Received: 8 December 2023 | Revised: 24 January 2024 | Accepted: 24 January 2024

DOI: 10.1002/wcms.1706

ADVANCED REVIEW

WIREs COMPUTATIONAL MOLECULAR SCIENCE WILEY

Making quantum chemistry compressive and expressive: Toward practical ab-initio simulation

Jun Yang[1,2]

[1]Department of Chemistry and State Key Laboratory of Synthetic Chemistry, The University of Hong Kong, Hong Kong, China

[2]Hong Kong Quantum AI Lab Limited, Hong Kong, China

Correspondence
Jun Yang, Department of Chemistry, The University of Hong Kong, Pokfulam Road, Hong Kong, China.
Email: juny@hku.hk

Funding information
Hong Kong Research Grant Council, Grant/Award Numbers: GRF17310922, ECS27307517, GRF17309020; University Research Committee, University of Hong Kong; Hong Kong Quantum AI Lab Limited through the AIR@InnoHK program of the Hong Kong SAR government; Faculty Computational Initiative Program; Chemistry Department Hui's Fund; University Postgraduate Fellowship

Edited by: Jinlong Yang, Associate Editor and Peter R. Schreiner, Editor-in-Chief

Abstract

Ab-initio quantum chemistry simulations are essential for understanding electronic structure of molecules and materials in almost all areas of chemistry. A broad variety of electronic structure theories and implementations has been developed in the past decades to hopefully solve the many-body Schrödinger equation in an approximate manner on modern computers. In this review, we present recent progress in advancing low-rank electronic structure methodologies that rely on the wavefunction sparsity and compressibility to select the important subset of electronic configurations for both weakly and strongly correlated molecules. Representative chemistry applications that require the many-body treatment beyond traditional density functional approximations are discussed. The low-rank electronic structure theories have further prompted us to highlight compressive and expressive principles that are useful to catalyze idea of quantum learning models. The intersection of the low-rank correlated feature design and the modern deep neural network learning provides new feasibilities to predict chemically accurate correlation energies of unknown molecules that are not represented in the training dataset. The results by others and us are discussed to reveal that the electronic feature sets from an extremely low-rank correlation representation, which is very poor for explicit energy computation, are however sufficiently expressive for capturing and transferring electron correlation patterns across distinct molecular compositions, bond types and geometries.

This article is categorized under:

Electronic Structure Theory > Ab Initio Electronic Structure Methods
Software > Quantum Chemistry
Software > Simulation Methods

KEYWORDS

electron correlation, low-rank wavefunction, quantum machine learning

1 | INTRODUCTION

Ab-initio electronic structure theories have long been developed to solve approximate Schrödinger equations. A number of standard blackbox many-body tools including perturbation theory (PT), coupled cluster (CC) and configuration interaction (CI) methods have been devised to predict electron correlation energies. By managing the particle (p)–hole (h) excitation hierarchy, the original exponential complexity is lowered to polynomial costs associated with a much reduced Hilbert space in which the wavefunction is represented. Among these methods, the exact MP2 (second-order Møller–Plesset perturbation theory) forms the simplest wavefunction-based method which utilizes the perturbative 2p–2h excitations relative to a single dominant reference (e.g., usually the Hartree–Fock Slater determinant) and scales as $\mathcal{O}(N^5)$ where N is a general measure of system sizes. The canonical CCSD(T) ansatz with the 1p–1h excitations for singles, 2p–2h excitations for doubles and perturbative 3p–3h excitations for triples is the minimum wavefunction model to obtain formally size-extensive and chemically accurate energies for close-shell molecules around equilibrium geometry, which however scales unfavorably as $\mathcal{O}(N^7)$.

In the past several decades, the steep computational scalings have been suggested to be unphysical and must be drastically reduced by exploiting the fundamental hypotheses for electron correlation,[1] assuming that (1) molecular properties are governed by the one-electron Hartree–Fock (HF) state; (2) correlation energies are additive to electron pairs; (3) correlation energies are insensitive to the long-range environment. This milestone idea has been implemented in many different schemes based on one-electron localization, including fragmentation methods,[2–4] local correlation methods,[5–28] and many others by combining subsystem and local correlation approaches.[29–47] Fragmentation methods, which divide macromolecules of interest into fragments based on the atomistic locality adhere to a group of atoms, solve all fragment problems separately, and combine fragment solutions to predict the macromolecular properties. Instead, the electronic locality[48,49] can be also exploited within a full system when the pair-electron operators are written in various compact forms. However, the prediction accuracy of both methods is usually drawn from benchmark systems and lacks direct validation for nonstandard macromolecules. Despite of many demonstrative applications to complex systems including biomolecules,[23,26,50–54] macro-clusters and liquids,[55–62] as well as condensed states,[44,63–71] the post-HF treatment through fragmentations and local correlations remains generally underutilized.

The aforementioned hypotheses may be broken when a single-reference state contributes insignificantly to strongly correlated systems, typically containing long π-conjugated carbon rings,[72] transition-metal elements,[73] and homolytically broken bonds.[74] Important contributions higher than 3p–3h excitation may arise with a prohibitively long CC or CI expansion that hinders quantitative computation of electron correlation energies and other properties. For one notorious example of Cr_2 molecule containing a sextuple metal–metal bond, the single-reference CCSD(T) prediction of Cr–Cr bond length[75] severely deviates from the experimental value; when surveying the predicted Cr_2 binding energies, the multireference computations yield a drastic range of the energy disagreements; however, it is possible to reproduce a more quantitative Cr_2 potential energy curve with more rigorous computations that include a large number of electronic configurations.[76]

One of the major challenges has been to design practical schemes for incorporating only important electronic configurations (e.g., an array of determinants) into the wavefunction, which aim to avoid little contributions from configuration components and reduce the computational cost. This idea conceptually derives from the nature of real physical interactions in molecular Hamiltonians,[77] and leads to a number of compressive and selective wavefunction representations implemented in state-of-the-art strong correlation methods for ab-initio quantum chemistry, including the density matrix renormalization group (DMRG),[78–83] the selected or adaptive CI,[84–92] the many-body expansion full CI,[93,94] the heat-bath CI,[95,96] the downfolded CI,[97] the CC reduction,[98–101] and selective high-level CC methods,[102–104] a variety of stochastic quantum Monte Carlo CI approaches,[105–109] and quasiparticle-based geminal wavefunction methods.[110–114] Recent benchmark studies of the non-relativistic frozen-core correlation energy of the benzene ground state[91,115] indicate a promising performance across a set of these methods yielding an energy deviation of about sub-kcal/mol, albeit with a considerable amount of computational resources using double-ζ basis set. The next hurdle is to find an efficient way of computing strongly correlated states of larger molecules with larger basis sets at comparable accuracy to benzene.

The difficulties facing these methods for efficiently treating complex systems are almost all attributed to the rapid expansion of the many-electron basis in which the wavefunction is represented with the increase of atom numbers. The underlying origin of such problems is due to the profound area law that many-electron states do not follow[116] in the presence of long-range interactions between subsystems of 2D and 3D macromolecules, that is, the correlation length increases with the increase of the system size.[117] It has been realized that the range of the interaction terms in a

Hamiltonian depends on the compactness of the many-electron basis in which the Hamiltonian is expressed. As a common practice, the one-electron localization has been widely used to retain only important determinants associated with the short-range interactions for dynamic correlation. The computational efficiency is thus significantly gained by either discarding long-range interactions or approximating them with classical or low-level couplings.[35,118–121] However, even weak interactions will mix many electronic configurations in largely unexplored subspaces necessary to determine the physical states to chemical accuracy. Moreover, systems exhibiting near degeneracy of valence orbitals must entangle the long-range electronic configurations that are almost entirely localized at intermediate distances,[122] as shown in the H_n bond breaking process.[123,124] Despite of the notable success of these selective methods, the search of weakly interacting configurations remains a daunting task for many electrons owning to the large size of Hilbert space.

The origin of the cost scalings of these high-level CC and CI methods is fundamentally traced to large systems comprising many atoms. This problem has been alternatively tackled with machine learning (ML) algorithms, with the ability of handling high-dimensional data structure in cheaper surrogate models than directly solving many-body electronic states. Mathematically, there exists a nonlinear network model that can universally represent any smooth and continuous multivariate function with sufficient neurons[125,126]; in practice, it has been shown to 1D or 2D Heisenberg models that simple deep network models can represent their quantum many-body states with a much reduced number of the hidden network parameters (e.g., neurons) compared with the original dimensionality of the Hilbert space, and the prediction fidelity can be systematically improved by increasing the hidden variables[127,128] For quantum chemistry Hamiltonian, machine learned configuration selection has demonstrated an efficient representability of learning networks for expressing important Slater determinants.[129,130] However, as the many-electron wavefunction does not change smoothly as a function of atomic positions, especially when state degeneracies and crossings occur, there may be a considerable challenge in directly predicting wavefunctions with ML models. A plethora of alternative machine learning approaches in the past decade[131–144] utilize the atomistic and electronic localities to further enhance the feature expressibility based on physically relevant knowledge that can be extracted directly from local environments within atomic subsystems or configuration subspaces.

In this advanced review, we will discuss the general idea, critical components of scale-up algorithm and applications of various compression-based quantum chemistry methods to both weakly and strongly correlated molecules. It becomes increasingly possible to handle previously difficult systems near chemical accuracy at reasonable costs, owning to the algorithmic advancement of systematically improvable low-rank representations. These techniques not only considerably shorten the many-body wavefunction parameters, but also facilitate quantum feature design for building an effective mapping from atomistic/electronic attributes to differences in molecular properties.[145] We will therefore also discuss the relevant development of expression-based quantum chemical neural network models to highlight the importance of physically motivated low-rank information that needs to be properly formulated to account for the transferable environment for atoms or electrons in molecule. Illustrative chemical applications will be demonstrated to molecular systems and processes that are controversial to traditional density functional theory (DFT) and generic post-HF computations.

2 | THEORY

2.1 | Low-rank wavefunction

In principle, a low-rank representation of any many-body electronic wavefunction exists for systems according to Schmidt decomposition[146]: as a linear algebra result from $|\Psi\rangle_{AB} = \sum_{ij} C_{ij} |i\rangle_A \bigotimes |j\rangle_B$ for arbitrary bipartite subsystems A enclosing states $|i\rangle_A$ and B enclosing $|j\rangle_B$, the wavefunction is expressed equally well in the low-rank orthonormal basis states $|\alpha_i\rangle_A$ for A and $|\beta_i\rangle_B$ for B,

$$|\Psi_{AB}\rangle = \sum_i \lambda_i |\alpha_i\rangle_A \bigotimes |\beta_i\rangle_B, \tag{1}$$

where $\sum_i \lambda_i^2 = 1$. If the full composite Hilbert space $\mathscr{H} = \mathscr{H}_A \bigotimes \mathscr{H}_B$ would be partitioned into small and large subspaces $\mathscr{H}_A$ and $\mathscr{H}_B$ of $\dim(\mathscr{H}_A) << \dim(\mathscr{H}_B)$, a rather small number of low-rank states can be utilized to formulate the wavefunction in the small subspace of the dimension $\dim(\mathscr{H}_A)$. However, the exact Schmidt decomposition is not practically feasible as this would require the full system solution which is unknown prior to basis rotations from original

$|i\rangle_A$ and $|j\rangle_B$ to Schmidt basis $|\alpha_i\rangle_A$ and $\left|\beta_j\right\rangle_B$, respectively. Nevertheless, it is possible to start from a trial full system wavefunction at low-level theory, typically uncorrelated mean-field wavefunction, as developed in the density matrix embedding theory (DMET) for lattice models,[147,148] quantum chemistry Hamiltonian[149,150] and periodic solids.[151,152] The DMET embedding theory is a reminiscent of low-rank Hamiltonian ($\overline{H}$) construction of $\overline{H}=\mathrm{PHP}$ in the exact Schmidt basis of $P=\sum_{ij\in \dim\mathscr{H}_A}|\alpha_i\rangle_A\left|\beta_j\right\rangle_B\left\langle\beta_j\right|_B\langle\alpha_i|_A$ resulting from many-particle rotations. In practice, the rotation can be approximated as single-particle rotations from single-particle objects, such as overlap or one-body reduced density matrix (1RDM) for designated fragments/baths of a trial wavefunction. This strategy has been discussed in various schemes including bootstrap embedding,[153–157] incremental embedding,[158] correlated bath states,[159] and complete active space (CAS) DMET.[160] Another interesting exploration to the low-rank wavefunction representation has been directed to the ab-initio quantum embedding scheme for the full system in which important many-body effects are systematically described by explicitly correlating the product state of the fragment electronic and environmental Drude oscillator wavefunctions variationally optimized in quantum Monte Carlo.[161] However, it still remains an open question to ensure the consistent fragment and bath description between the low- and high-level wavefunctions.

2.2 | Single-particle transformation

Another strategy is to set up the low-rank Hamiltonian by adopting a single-particle rotation that encodes the coarse knowledge of dynamic electron correlations. Such methods were developed as early as the 1950s[162] with introducing the natural orbital (NO) set through the low-rank tunable operator

$$\hat{a}_\chi^\dagger=\sum_p^{N_{\mathrm{orb}}}\hat{a}_p^\dagger U_{p\chi}, \tag{2}$$

with N_{orb} the number of molecular orbitals, by diagonalizing correlated 1RDM of the full system, which reduces the determinant space. Important NOs are selected according to their ordered eigenvalues $n_\chi=U^\dagger\chi U$ to lower the Hamiltonian complexity in $\overline{H}=\sum_{\chi\gamma}h_{\chi\gamma}\hat{a}_\chi^\dagger\hat{a}_\gamma+\sum_{\chi\gamma\delta\xi}g_{\chi\gamma\delta\xi}\hat{a}_\chi^\dagger\hat{a}_\gamma^\dagger\hat{a}_\xi\hat{a}_\delta$ in an approximate form. This type of NOs has further promoted other related approximations to the single-particle rotation U, including frozen NOs,[163,164] pair-natural orbitals (PNO, equally termed as pseudo-natural orbital in the early days)[19,165–168] and optimized correlating orbitals.[169,170] For efficiently recovering the CI solution, the PNO makes a separate basis rotation $\hat{a}_{\chi_{ij}}^\dagger=\sum_p^{N_{\mathrm{vir}}}\hat{a}_p^\dagger U_{p\chi_{ij}}$ (N_{vir}: the number of virtual orbitals) for each bonding or non-bonding electron pair ij, leading to a very compact correlating subspace for the pair. By combining PNOs with the electronic locality,[49] the local PNO schemes have been extensively developed in the recent decade for single-reference CC and PT theories[21,25,26,32,119,171–178] as well as their multireference variants.[179–185] An intermediate scheme between the full system NO and the pairwise PNO has been proposed to assign each single-particle rotation to an electronic orbital by $\hat{a}_{\chi_i}^\dagger=\sum_p^{N_{\mathrm{vir}}}\hat{a}_p^\dagger U_{p\chi_i}$, which makes the orbital-specific virtual (OSV) approximation of $U_{p\chi_i}$ and significantly reduces the complexities (see Section 3.2.1) arising for computing, storing and manipulating electronic repulsion integrals in PNO-based methods. The OSV scheme has been developed to MP2, CCSD and CCSD(T) theories,[22–24,34,35] which have been further implemented in OSV-PNO hybrid ansätz.[10,186–191] The direction determination of $\mathbf{U}_{ij}$ for PNOs[192,193] and $\mathbf{U}_i$ for OSVs[194] has been also developed to improve the correlation convergence.

It is clear that the rotation by $\mathbf{U}_{ij}$ and $\mathbf{U}_i$ generates partially non-orthogonal PNOs/OSVs between different pairs (ij and kl)/orbitals (i and j) due to the non-vanishing $\mathbf{U}_{ij}^\dagger\mathbf{U}_{kl}\neq 0$ and $\mathbf{U}_i^\dagger\mathbf{U}_j\neq 0$. As a result, the PNO non-orthogonality causes complications in carrying out extra linear algebraic operations associated with tremendous amounts of pairwise PNOs and repulsion integrals if all interacting pairs are included, and thus severe memory issues occur for very large molecules. The non-orthogonal OSVs considerably lower such complexities due to much fewer orbital-wise rotations of $U_{p\chi_i}$, despite that each $\mathbf{U}_i$ is less compact than $\mathbf{U}_{ij}$. Moreover, this complication can be effectively removed when the PNOs/OSVs are combined with the composite methods in various forms.[31,35,40,69,71,195–199] It must be pointed out that none of these techniques would be possible to compute large molecules at high efficiency without carefully handling the subtle complexities of rotated integrals, the sorting of local interacting pairs or orbitals, the retention of important low-rank rotations and the screening of long-range correlations, all of which are equally important in practical

computations. We have found that the computational cost-accuracy balance is different between different implementation schemes of these approximations. At similar accuracy level, the approximation parameters and the energy convergence performance are not directly comparable.[200] The predefined parameters that are optimally tuned against the benchmark results of small and medium molecules may not be easily validated for real large molecules, when assuming size-extensive correlation errors. Nevertheless, these low-rank methods combined with local correlations have been demonstrated to be capable of considerably shortening computations of the energies, structures and processes at different post-HF levels of theory[34,35,62,69,198,199,201–203] that were previously difficult.

2.3 | Many-particle transformation

For strong correlations, it is an attractive idea to be able to accurately express and solve model Hamiltonians in a subspace of much reduced dimensionality through the many-particle rotation of important electronic configurations, analogous to low-rank single-reference methods. We consider the m lowest eigenstates of an $N \times N$ Hamiltonian $H = H_A \bigotimes H_B$ for a bipartite system which is divided into the system A described by an $m \times m$ Hamiltonian H_A and the remaining environment B by H_B. Turning back to the exact Schmidt basis representation, the projection $P = \sum_{ij \in \dim \mathscr{H}_A} |\alpha_i\rangle_A |\beta_i\rangle_B \langle \beta_j|_B \langle \alpha_j|_A$ maps the original problem into $\overline{H} = \mathrm{PHP}$ requiring at most m^2 eigenvectors for exactly describing the m lowest states. On the other hand, under the complete decoupling limit, that is, when $H = H_A \bigotimes I_B \bigoplus I_A \bigotimes H_B$ is exactly block-diagonal between A and B subsystems, or alternatively when a many-particle decoupling rotation U is found such that the $U^{-1}HU$ is completely decoupled between A and B, a model Hamiltonian in the subspace $\mathscr{H}_A$ for A can be formulated as $\overline{H}_A = \sum_{i \in \dim \mathscr{H}_A} |\alpha_i\rangle_A H_A \langle \alpha_j|_A$ which requires m eigenvectors for solving the m states. Our analysis here, which does not provide a practical computational simplification, indicates that it is necessary to express a low-dimensional Hamiltonian in an $n \times n$ model subspace with $m < n < m^2$ in the presence of an intermediate system-environment coupling for describing m low-lying states sufficiently accurately.

Since the 1950s, different old flavors of the many-particle decoupling operator U[204–207] have been attempted to make H block-diagonal by $U^{-1}HU$, which were approximately solved by developing quasi-degenerate perturbation theory[204,207,208] and iterative determination techniques[209–211] without inputting a priori knowledge of the exact solutions. However, the convergence of this type of model Hamiltonians has been proved qualitatively poor in the presence of intruder states that are nearly degenerate with some states in the preselected model subspace,[211] which is a familiar problem facing modern MS-CASPT2 (multistate complete active space second-order perturbation theory) method.[212] The intruder-state problem is entirely avoidable in the driven similarity renormalization group approach,[213–217] which drives a continuous flow and decoupling of the Hamiltonian, practically truncated to one- and two-body operators toward the limit of linearized canonical transformation theory.[218] Although the choice of the decoupling rotations is arbitrary, in rigorous computations, an accurate model Hamiltonian must be defined in a subspace that contains sufficient components which largely overlap with a low-lying target state. An improved version of a more practical model Hamiltonian is to incorporate an intermediate region between the system and environment subspaces, establishing an intermediate effective Hamiltonian approach,[219,220] which underlies a similar spirit that the DMET resurrects for introducing quantum bath states to buffer the system–environment interaction. The principle of intermediate effective Hamiltonians has been applied to improve the selected CI (sCI) approach[221] for self-consistently selecting important configurations[222] toward the exact results of small polyatomic molecules, and recently invigorated to dress zero-order model Hamiltonians by perturbation for direct diagonalization in small CAS-CI spaces.[223]

The model Hamiltonian methods discussed above enforce a severe requirement $\Psi_P = P\Psi$, indicating that the model subspace projection (P) of an exact eigenstate (Ψ) of the original Hamiltonian must be a corresponding low-lying root (Ψ_P) of the model Hamiltonian, which is rather difficult for the Bloch wave operator theory.[210] While examining the seminal idea of Löwdin partitioning technique,[224,225] we consider the partitioning strategy that also fulfills this condition but conceptually different, leading to an effective Hermitian Hamiltonian $\overline{H}^P$ in the model subspace P upon a many-particle transformation $\Omega H \Omega$.

$$\overline{H}^P \Psi_P = E\Psi_P, \overline{H}^P = P(\Omega H \Omega)P, \tag{3}$$

$$\Psi = \Omega\Psi_P,\ \Omega = P + \frac{1}{E - QHQ} QHP. \tag{4}$$

Here, $Q = 1 - P$ is the outer subspace complementary to P. Ω behaves as an effective wave operator in another form that restores the exact eigenstate from the model subspace and holds relations $\Omega P = \Omega$, $P\Omega = P$, and $\Omega[H, \Omega] = 0$ similar to the Bloch wave equation.[210] This has instigated developments of the perturbative Ak and Bk methods,[226] the latter of which partitions the CI matrix with dominant configurations by including only the diagonal QHQ as the first-order perturbation. Further improved Bk variants include the state-specific shifted-Bk method[227–231] and the multistate shifted-Bk method combined with large sCI model subspaces,[90] which shifts the energy by the second-order perturbation in the denominator of Equation (4).

The super-operator Ω can be viewed as an external contraction and plays a key role in compressing critical outer configurations in the complementary Q subspace. This suggests an attractive feature that the resulting dimensionality of the model Hamiltonian $\overline{H}^P$ is not affected by the number of configurations in the outer Q subspace. But there are notable difficulties arising from the unknown energy and the need of inverting an exceedingly large matrix in the Ω denominator. Iterative diagonalization methods have been introduced in connection to different PT functions, including Rayleigh–Schrödinger PT,[224,232] Brillouin–Wigner PT,[233–236] and van Vleck PT,[237] which exploit several tractable perturbative functions of Ω toward the FCI convergence of small problems.

For relatively large-scale FCI matrices, we have recently developed a direct iterative method for determining the model Hamiltonian $\overline{H}^P$ at near-FCI accuracy by compressing selected outer determinants,[97] which is termed downfolded CI (dCI). This dCI algorithm relies on recursive formula of $\Omega_{ij} H \Omega_{ij}$ for a cluster-pair subspace $P_{ij} = P_i + P_j$, where each cluster P_i (or P_j) contains a small number of energetically close determinants. The complementary subspace Q_{ij} for each P_{ij} is refined by taking important determinants, and then compressed into P_{ij} in which $\overline{H}^P$ is expressed. We have found that many small interactions via $Q_{ij} H Q'_{ij} \neq 0$ from the outer subspace Q'_{ij} disconnected to P_{ij} turn out to aggregate into important contribution through the Ω_{ij} denominator. We therefore include them to modulate the connected $P_{ij} H Q_{ij}$ couplings. In all early PT versions of Löwdin partitioning, these minor contributions inevitably lead to a very large model subspace to couple with sufficient outer determinants for retrieving accurate dynamic correlations.

2.4 | Quantum machine learning

The difficulty associated with many electronic configurations has been largely resolved by the aforementioned theory developments advancing sophisticated algorithms and implementations. The computational scope has been significantly expanded for large systems and processes with greater predictive power. Despite of these developments, the resulting computations would be still limited to merely single point computations, and difficult for problems necessary to sample many atomic configurations, such as large-scale chemical space search and long-time ab-initio molecular dynamics (MD) simulations that would solve the Schrödinger equation repeatedly. The quantum machine learning (QML) provides another paradigm for making direct prediction of target molecules by learning the known solutions of other molecules that can be readily obtained, as a substitute to the difficult solution of the target molecule. The QML architecture attempts to encode high dimensional data structure with complex hidden patterns which bridge expressive atomistic or electronic features with a variety of target properties by training a pool of known molecules. A notable example is the QML determination of an end-to-end mapping between the electron density and external potential by employing highly nonlinear data structure beyond traditional functional forms of DFT.[133,238–242] It is important to emphasize that it becomes increasingly difficult to generate the reference datasets containing many molecules, since the QML training costs would be very demanding and even unfeasible as the molecule size grows across a threshold. It is of crucial importance to develop QML methods that efficiently harvest the universal and transferable knowledge of small molecules in reduced training datasets for expediting chemistry discovery of unknown complexes with predictive insights.

Although there are many aspects to consider for refining QML architectures, physically motivated representations of a chemical system are of critical importance to better data efficiency and transferability. The nearsightedness of an electronic matter[243] has become the core idea in driving QML developments closer to this goal. There have been two main alternatives in implementing the nearsightedness: the local atomistic environments and the local electronic environments. The former atomistic scheme has led to various kernel[244–247] and neural network[131,132,137,248,249] QML models. A notable and indeed very successful strategy for enhancing transferability has been to approximate extensive properties (e.g., the molecular total energy) with additive symmetry-constraint atom-centered functions which are learned separately for each local atomistic environment. These models design real-space atom-specific pairwise descriptors for 1- and 2-body interactions, and physically augmented many-body descriptors can be also added for

non-pairwise interactions[138,250] or via message passing neural networks.[251] In a similar spirit to aforementioned post-HF subsystem methods, the local atomic environments are used for either including only a limited range of atoms within a predefined real-space cutoff once and for all[131,132,136,249,252] which neglects long-range interatomic interactions, or augmenting molecular fragments as necessary on the fly.[140] It is obvious that there is an issue regarding the sufficiency of the local environments, which depends on the nature of the system extension. However, an inclusion of the longer-range interactions[134,253] may deteriorate the resulting QML transferability with poor energy prediction of large molecules if the training molecules are not large enough to represent the target structure and its chemical environment. We have to point out that the tradeoff between non-locality and transferability is a delicate issue and has to be carefully managed, since it is difficult to determine the *prior* importance of interatomic interactions.

The electronic locality explicitly accounts for electronic structures and interactions, which represents molecules by electronic orbitals or densities starting from inexpensive low-level (e.g., mean-field) wavefunction properties. The long-range electrostatic interactions are automatically included in a mean-field manner, leaving relatively short-range interactions (e.g., dispersions) to be accounted in electronic QML models. Moreover, the intermediate electronic descriptors from the low-level electronic structure instruct a coarse mapping from an atomistic geometry to an electronic distribution, which respects the quantum nature of electrons. However, the canonical HF or DFT orbitals are delocalized over molecules and no longer transferable due to their orbital orthogonalization components containing remote chemical environments. The nearsightedness of electronic interactions is ubiquitously resembled by localizing canonical orbitals. An explicit example shows the reliable transferability of localized orbitals through the determination of the electron densities obtained from the transfer of localized orbitals on molecular subunits (atoms, bonds or functional groups) that turn out to be very similar to the exact HF ones.[254] Recently developed electronic QML models learn and predict ab-initio properties (e.g., electron correlation energy) with various baseline descriptors derived from the correlation electron densities,[255,256] local electronic orbitals,[135,141,142,257–261] and post-HF wavefunction amplitudes/density tensors,[262–265] and so on. These proof-of-principle examples demonstrate an improved flexibility and transferability of QML models, for which the prediction errors are relatively less dependent on the range of chemical systems, albeit still limited to small organic molecules. It is not unexpected that the generation of orbital-based ab-initio descriptors may be prevented by prohibitive computational costs for a large number of training molecules.

At the intersection between the low-rank post-HF theory and electronic QML surrogate scheme, there are several important key advantages which mitigate these complexities, while still reserving chemically transferable and accurate prediction. The basic idea is that, when the orbital-based ab-initio descriptors are expressed in sufficiently reduced low-rank basis via single- (Section 2.2) or many-particle (Section 2.3) rotations, the expense of computing these descriptors can be significantly lowered and the model transferability can be enhanced by learning simple molecules. Hence, we have developed a transferable deep neural network (T-dNN) model[143] using OSV-based descriptors for predicting chemically accurate MP2 and CCSD correlation energies from a small training set containing small molecules. The low-rank OSV algorithm of these descriptors virtually compresses a global correlating environment for all electrons in the molecule into many local correlating environments, each for one electron, which simultaneously encodes the long-range correlation and retains the transferable feature. One appealing aspect of this method is that the balance between the non-locality and transferability can be systematically and automatically managed by tuning the compactness of ordered OSV-based descriptors (see Section 3.4). Most importantly, the intrinsically low-dimensional structure of the compressive input space may increase its inhomogeneity, favor better feature classification and selection, and prevent the ever-growing scale of the electronic T-dNN model.[266] We have provided a comprehensive study[143] and shown that the T-dNN prediction demonstrates an excellent transferability and data efficiency for a broad range of chemical systems, including alkanes, organic molecule and biomolecular interactions, and water clusters of various sizes and morphologies.

3 | SCALE-UP ALGORITHM

We present several low-rank algorithms that scale up various ideas as discussed above for weak and strong correlations. For expediting post-HF computations, the algorithms for small molecules need revision toward large molecules, for which the computational cost shifts to operations associated with electronic repulsion integrals that determine the important wavefunction components in both single- and multireference cases.

3.1 | Approximate single-particle rotation

The low-rank post-HF methods as described in Section 2.2 require the computation of single-particle rotation. The operational costs formally scale as $\mathcal{O}(N^3)$, $\mathcal{O}(N^4)$, and $\mathcal{O}(N^5)$ for frozen NOs, OSVs and PNOs, respectively, by their genuine definitions. A demonstrative example shows that this expense cannot be ignored for large molecules, for instance, the generation of OSVs is expensive for $(H_2O)_{190}$/cc-pVTZ water cluster, which takes about 400 min (24 CPU cores, 2.30 GHz) even using well optimized parallel implementation, two orders of magnitude slower than solving the OSV-MP2 residual equations (Table 1). This results from large virtual blocks of 1RDM in the canonical MO basis, and can be significantly lowered when the 1RDM is obtained in a reduced 2p–2h double excitation space in the projected atomic orbital (PAO) basis.[108,186] Similarly, this and similar complexities are even more severe for PNO-based methods, further due to the $\mathcal{O}(N^2)$ pair growth, and alleviated by approximating the 1RDM in a prior truncated OSV basis,[187] or in a hierarchical PAO → OSV → PNO treatment by combining both,[10,25] which also necessitates a rough estimate of MP2 pair screening before the PNOs are obtained. Apparently, the PNO accuracy is bounded to incomplete prior PAO domains, and the PNO-MP2 pair distribution may be significantly different from initial estimates. In what follows, we describe a low-rank one-off generation of OSVs which do not need estimated pair screening, which also makes it suitable for developing analytic gradients.

For each electronic spin-orbital i, we consider the OSV rotation $\mathbf{U}_i$ which diagonalizes the semi-canonical MP2 amplitude matrix $\mathbf{T}_{ii}$ with elements $[\mathbf{T}_{ii}]_{ab} = [ia|ib]/(f_{aa}+f_{bb}-2f_{ii})$, where $a,b,\cdots$ and $i,j,\cdots$ denote the virtual and occupied spin-orbitals, respectively. The matrix $\mathbf{T}_{ii}$ has the rank k_{osv} for measuring the intrinsic sparsity that determines the low-rank efficiency of OSV-based post-HF methods. We find that, in most cases for correlation energies, the OSV vector $\overline{\mathbf{U}}_i$ from a low-rank amplitude $\overline{\mathbf{T}}_{ii}$ is sufficiently accurate, instead of using the exact $\mathbf{T}_{ii}$, and the generation of OSVs is much more efficient from $\overline{\mathbf{T}}_{ii}$.[35] Here, we use the $N \times k_{\mathrm{osv}}$ ($k_{\mathrm{osv}} << N$) subset amplitude $\overline{\mathbf{T}}_{ii}$ as the basis to expand the remaining $N - k_{\mathrm{osv}}$ columns of the exact $\mathbf{T}_{ii}$ of the dimension $N \times N$,

$$\mathbf{T}_{ii} \approx \overline{\mathbf{T}}_{ii}\mathbf{C}_i, \ \left\|\overline{\mathbf{T}}_{ii}\mathbf{C}_i - \mathbf{T}_{ii}\right\| \le \delta, \tag{5}$$

where $\overline{\mathbf{T}}_{ii}$ has k_{osv} columns of the exact amplitude $\mathbf{T}_{ii}$, and the unknown $k_{\mathrm{osv}} \times N$ interpolative vector $\mathbf{C}_i$ must contain $k_{\mathrm{osv}} \times k_{\mathrm{osv}}$ identity submatrix for keeping selected k_{osv} columns. In fact, Equation (5) resembles exactly the interpolative decomposition that has been previously applied to localized Wannier function[267] and electron repulsion integral compression.[268] Thus, the norm discrepancy can be minimized up to the precision δ.

However, the direct application of Equation (5) is expensive due to the large $N \times N$ amplitude $\mathbf{T}_{ii}$. We sort to a randomized algorithm,[269] in which by acting an $n \times N$ matrix $\mathbf{R}_i$ ($n << N$ and n is slightly greater than the rank k_{osv}) to $\mathbf{T}_{ii}$ for a randomized fast Fourier transformation, the following minimization is carried out,

$$\left\|\mathbf{R}_i\overline{\mathbf{T}}_{ii}\mathbf{C}_i - \mathbf{R}\mathbf{T}_{ii}\right\| \le \sigma_{k+1}, \tag{6}$$

TABLE 1 Comparison of the average dimension (k_{osv}) and sparsity ($k_{\mathrm{osv}}/N \times 100\%$) of the low-rank semi-canonical amplitude $\overline{\mathbf{T}}_{ii}$, the timing ($t_{\mathrm{osv}}$) of OSV generation, and the accuracy of OSV-MP2 correlation energy (δE, relative to the result with exact OSVs of medium $l_{\mathrm{osv}} = 10^{-4}$) by tuning the interpolative decomposition rank (σ_{k+1}) for $C_{40}H_{64}O_{12}$ and $(H_2O)_{190}$.

	$C_{40}H_{64}O_{12}$/def2-TZVP			$(H_2O)_{190}$/cc-pVTZ		
Rank threshold (σ_{k+1})	k_{osv} ($k_{\mathrm{osv}}/N \times 100\%$)	t_{osv} (s)	δE (au)	k_{osv} ($k_{\mathrm{osv}}/N \times 100\%$)	t_{osv} (min)	δE (au)
10^{-4}	182 (10.1%)	4.6	7.3×10^{-6}	200 (2.0%)	22.5	2.3×10^{-5}
10^{-5}	303 (16.9%)	10.0	2.4×10^{-7}			
10^{-6}	430 (23.9%)	18.7	2.1×10^{-9}			
10^{-7}	564 (31.4%)	30.2	4.0×10^{-10}			
10^{-8}	699 (38.9%)	41.1	$< 10^{-10}$			
Exact OSV		28.9			402.9	

Note: Adapted with permission from Ref. [35]. Copyright 2021 American Chemical Society.

where the error is bounded to the $(k+1)$th greatest singular value σ_{k+1} of the projected $n \times N$ matrix $\mathbf{R}_i\mathbf{T}_{ii}$, and $\mathbf{R}_i\overline{\mathbf{T}}_{ii}$ collects the k_{osv} columns of $\mathbf{R}_i\mathbf{T}_{ii}$. Solving Equation (6) normally costs $\mathcal{O}(k_{\mathrm{osv}}nN)$ for each electronic orbital, much faster than $\mathcal{O}(k_{\mathrm{osv}}N^2)$ solving Equation (5). With the interpolative vector $\mathbf{C}_i$ identified, the approximate single-particle rotation $\mathbf{U}_i$ for OSVs can be computed via the QR decomposition of $\mathbf{C}_i$[35] at the reduced cost of $\mathcal{O}\left(2Nk_{\mathrm{osv}}^2\right)$. Hence, the interpolative decomposition significantly lowers the overall operational complexity of OSV generation from $\mathcal{O}\left(ON^3\right)$ to $\mathcal{O}(ON)$ for all occupied orbitals (measured by the O number of occupied orbitals). As seen in Table 1, it is clear that a sparse subspace $\overline{\mathbf{T}}_{ii}$ of the tunable rank $k_{\mathrm{osv}} << N$ can well approximate the OSVs with negligible error to correlation energies, as compared with results in the OSVs obtained from the exact amplitude $\mathbf{T}_{ii}$. For example, for a medium molecule $C_{40}H_{64}O_{12}$ with def2-TZVP basis, the interpolative decomposition accelerates the OSV generation by nearly seven folds, as compared with the exact OSVs, causing only a minor correlation energy loss of 7.3×10^{-6} au. For large $(H_2O)_{190}$/cc-pVTZ computation, the timing is reduced from 400 to 22 min.

3.2 | Weak electronic configuration

3.2.1 | Single-reference wavefunction

Another important part is to efficiently identify and treat important weak electronic configurations, such as the long-range electronic correlation at a single-reference post-HF level similar to the short-range one. The total contribution from the long-range correlation can become substantial and difficult to handle for macromolecules, due to a large amount of such weak electron pairs and relevant electronic repulsion integrals which increases rapidly with the inter-electronic distance. For example, the total PNO-MP2 contribution to the Auamin reaction energy from the long-range pairs amounts to 22 kJ/mol[118] which cannot be neglected, although each of the weak pairs has very small correlation below 10^{-5} au. The long-range pairs are normally identified based on the spatial criterion ($|r_i - r_j| > R_{\mathrm{long}}$) using the distance $|r_i - r_j|$ of two local orbitals, which can be further refined by screening the estimated energy magnitude via dipole–dipole interactions[10,25] or high-order multipole.[118] We note that the value of a proper spatial cutoff is largely affected by the nature of molecules, and varies from one molecule to another. Hence the precise determination of important long-range pairs is difficult from real space measurements alone, especially for molecules with extended π conjugation.

To avoid caveats from real space selection, we have developed an algorithmic metric[34,35] which assigns the long-range pairs according to the intrinsic compactness of the OSV orbital-domain overlap $\langle\overline{\mu}_i|\overline{\nu}_j\rangle$ that is capable of discerning the weak interaction strength between remote electron pairs,

$$s_{ij} = \frac{\sum_{\overline{\mu\nu}}\langle\overline{\mu}_i|\overline{\nu}_j\rangle^2}{\sqrt{n_i n_j}}, \quad n_i = \sum_{\overline{\mu\nu}}\langle\overline{\mu}_i|\overline{\nu}_i\rangle^2, \tag{7}$$

where n_i is the total number of OSVs for the ith LMO. We apply the Cauchy-Schwarz inequality to s_{ij} which leads to $0 < s_{ij} < 1$ for $i \neq j$ and $s_{ii} = 1$. Apparently, the magnitude of s_{ij} is closely related the nature of OSVs, which is adaptive to molecular attributes. As such, for long π molecules which yield greater s_{ij} from more delocalized OSVs, more long-range pairs emerge and can be included. This ensures that important long-range interactions can be adaptively, consistently and automatically identified. For example, when the long-range pairs are assigned by $10^{-7} < s_{ij}^{\mathbf{lr}} < 10^{-2}$ using triple-ζ basis sets, 25,179 long-range pairs out of 32,385 pairs are identified for extended C_{60}@catcher, 144,247 long-range pairs out of 289,180 pairs for $(H_2O)_{190}$, and only 13,870 long-range pairs out of 98,790 pairs for $(Gly)_{40}$. As seen in Table 2, the accuracy loss of the binding energy by discarding extremely remote pairs for which $s_{ij}^{\mathbf{dd}} \leq 10^{-7}$ is negligible. We have found (tab. S4 in Ref. [35]) that the basis set diffusion function does not necessarily lead to the inclusion of more long-range pairs using the pair classification in Equation (7), as opposed to real space selections. Interestingly, when adding diffuse basis functions, the amount of short-range pairs slightly decreases by 3%–4% for both C_{60}@catcher and $(H_2O)_{32}$, as a tradeoff with more OSVs.

The rapid computation of long-range pairs commonly invokes the semi-canonical formulation of $E_c^{\mathrm{long}} \approx \sum_{ijab}[ia|jb]^2/\left(f_{aa}+f_{bb}-f_{ii}-f_{jj}\right)$ without exchange terms,[10,25,119] or iteratively solves the long-range amplitude equations in the 2p–2h orbital-specific excitation subspace.[35,118] However, one has to mention that the semi-canonical

TABLE 2 The impact of long-range pairs on the total and binding energies (au) for 190 $H_2O \rightarrow (H_2O)_{190}$ with cc-pVTZ. The binding energy is given as $E_{\text{bind}} = E(H_2O)_{190} - \sum_i^{190} E_i(H_2O)$ (not assuming identical water molecules).

$s_{ij}^{\mathbf{lr}}$	sr pairs	lr pairs	dd pairs	$E(H_2O)_{190}$	$\sum_i^{190} E_i(H_2O)$	$\lvert E_{\text{bind}} \rvert$	$\lvert \delta E_{\text{bind}} \rvert$	$\lvert \delta E_{\text{bind}}/E_{\text{bind}} \rvert$
0	13,019	276,161	0	−51.476377	−50.182196	1.294182		
10^{-7}	13,019	144,247	131,914	−51.475779	−50.182196	1.293583	0.000598	0.05%

Note: The short-range (sr), long-range (lr), and discarded (dd) pairs are defined by $10^{-2} \le s_{ij}^{\mathbf{sr}} \le 1$ and $10^{-7} < s_{ij}^{\mathbf{lr}} < 10^{-2}$ and $s_{ij}^{\mathbf{dd}} \le 10^{-7}$, respectively. Adapted with permission from Ref. [35]. Copyright 2021 American Chemical Society.

approximation is prone to large energy error and also problematic to energy gradients due to the contribution from the response of long-range amplitudes; the iterative scheme accounts for only the genuine dispersion correlation for boosting computational efficiency, which may be a source of important errors. It is therefore critical for the iterative scheme to treat only the long-range pairs that truly contain negligible charge transfer and exchange correlation components. In our work, we make efficient use of the OSV domain overlap which modulates the couplings between different correlation components and is systematically tunable via $s_{ij}^{\mathbf{lr}}$ in Equation (7). We have found that within the range $10^{-7} < s_{ij}^{\mathbf{lr}} < 10^{-2}$, the resulting long-range pairs are dominated by only the dispersion correlations via $\{i \rightarrow \overline{\mu}_i, j \rightarrow \overline{\nu}_i\}$ 2p-2 h excitations for which the iterative scheme is sufficiently accurate. In the OSV basis, for solving the long-range $\overline{\mathbf{T}}_{ij}^{\mathbf{lr}}$, the dispersion-dominant amplitude equations are first projected out from the exact OSV-MP2 amplitude equations, and then further reduced by coupling each long-range pair ($\overline{\mathbf{T}}_{ij}^{\mathbf{lr}}$) with only short-range diagonal pair amplitudes ($\overline{\mathbf{T}}_{ii}^{\mathbf{sr}}$ and $\overline{\mathbf{T}}_{jj}^{\mathbf{sr}}$). Most importantly, the expense for computing the long-range repulsion integral $\left[i\mu_i|j\nu_j\right]^{\mathbf{lr}}$ is also much reduced as the 3-center-2-electron integrals $[i\mu_i|P]$ on the fitting basis P are readily available from short-range pairs. This and other costs arising from long-range correlations are significantly lowered in OSV-MP2 analytical energy theory.[35] Overall, the OSV-based iterative scheme sets a promising stage for computing the long-range correlation and analytical energy gradients at comparatively negligible cost, relative to these for solving the short-range contributions.

3.2.2 | Multireference wavefunction

The similar weak configuration prescreening from spatial or dipole–dipole interaction criteria has been introduced to identify the long-range dynamic correlations arising from the "inactive → external" and "active → external" subclasses of the excitation for PNO-based multireference CASPT2 and NEVPT2 approaches. Their correlation energies are estimated with the cheap semi-canonical[182] and high-order multipole approximations.[184] For large molecules and basis sets, this importance measure conveniently ranks a set of weakly coupling determinants that are connected to the active space through first-order perturbative couplings. These advances substantially reduce the expenses for computing the long-range dynamic correlations due to a large number of external orbitals, and shift the major computational bottleneck to the optimization of the reference CASSCF wavefunction which is limited to small active space. However, small active space computations are prone to large error arising from an enormous amount of disconnected external determinants, which are accumulated to make an important contribution to electron correlation through at least the third-order energy perturbation. Therefore, modern multireference approaches, such as the variants of sCI, attempt to gradually enlarge the active space by selecting non-negligible determinants and then perform the second-order PT correction on the resulting variational reference wavefunction. Notably, the importance of weak determinants is measured by several alternative metrics, including an estimate of the first-order wavefunction amplitudes ($\left|\frac{\sum_i H_{ji}c_i}{E_0 - H_{jj}}\right|$) for all determinants in the CI by perturbatively selecting iteratively[84] and a subset of them in the adaptive sampling CI,[86,89] a simple measure by selecting the maximal $\max_i\left(\left|H_{ji}c_i\right|\right)$ in the heat-bath CI,[95,96] as well as adaptive selection of CC configurations by moment expansion.[102–104] One has to be aware that, by iteratively selecting the tremendous number of weakly coupled determinants toward the near-exact solution, the Hamiltonian matrix drastically expands and ultimately exceeds the limit that modern computer resources can offer even for small molecules, such as benzene assessed with cc-pVDZ basis in various state-of-the-art approaches.[91,115] As a result, the application of these methods to more realistic molecules and basis sets still faces great challenges.

We consider our recent dCI Hamiltonian representation[97] as an alternative toward alleviating this challenge, as introduced with the tactics of simultaneously selecting and compressing weak configurations in Section 2.3. The dCI algorithm recursively builds up a very compact effective Hamiltonian for enabling direct diagonalization in a small model subspace P that is composite of cluster pairs $P_{ij} = \sum_{p \in P_{ij}} |\Phi_p\rangle\langle\Phi_p|$ containing a small number of determinants. Disconnected determinants to the local cluster pair P_{ij} are manifested and collected through the outer interactions $Q_{ij}H\overline{Q}_{ij}$ that attenuate the coupling magnitudes. Taking each determinant Φ_p from a local cluster subspace P_{ij}, its numerically connected outer subspace $Q_{ij} = \sum_{q \in Q_{ij}} |\Phi_q\rangle\langle\Phi_q|$ and fully disconnected subspace $\overline{Q}_{ij} = \sum_{q \in \overline{Q}_{ij}} |\overline{\Phi}_q\rangle\langle\overline{\Phi}_q|$ are identified as follows, respectively, according to the expectation value thresholds $\theta_1 = 10^{-8}$ and $\theta_2 = 10^{-6}$,

$$\sum_p \left|\langle\Phi_p|H|\Phi_q\rangle\right| > \theta_1, \tag{8}$$

$$\left|\left\langle\overline{\Phi}_q|V_{ij}^{(1)}|\overline{\Phi}_q\right\rangle\right| > \theta_2, \tag{9}$$

$$V_{ij}^{(1)} = \overline{Q}_{ij}HQ_{ij}\frac{1}{E - Q_{ij}HQ_{ij}}Q_{ij}H\overline{Q}_{ij}, \tag{10}$$

where $V_{ij}^{(1)}$ is the local screening potential that determines the selection of disconnected weak configurations specific to each cluster pair ij. As clearly revealed in Figure 1, the dCI correlation energies are improved by adding essential weak determinants in the outer subspace, and converged to chemical (~1 kcal/mol) and near-FCI (~0.05 kcal/mol) accuracy with the $N_Q = 64$ k and $N_Q = 460$ k determinants, respectively. The resulting effective Hamiltonian can be well represented in a small 2200×2200 model subspace enabling simple diagonalization, which is much smaller than the FCI problem. Apparently, the dCI selection and compression scheme is based on the quadratic outer coupling which is bounded by $\left|\left\langle\overline{\Phi}_q|V_{ij}^{(1)}|\overline{\Phi}_q\right\rangle\right| \sim \left|\overline{Q}_{ij}HQ_{ij}\right|^2$, as revealed in Equation (10). As the dCI search for important determinants is carried out separately within local cluster-based subspaces, a simple parallelization scheme can be implemented for dCI computations by distributing all local clusters.

3.3 | Low-rank analytical gradient theory

Another important aspect is to enable efficient post-HF molecular geometry and dynamics simulations for complex molecules. In principle, the low-rank post-HF formulation as discussed in Section 2.2 requires the response

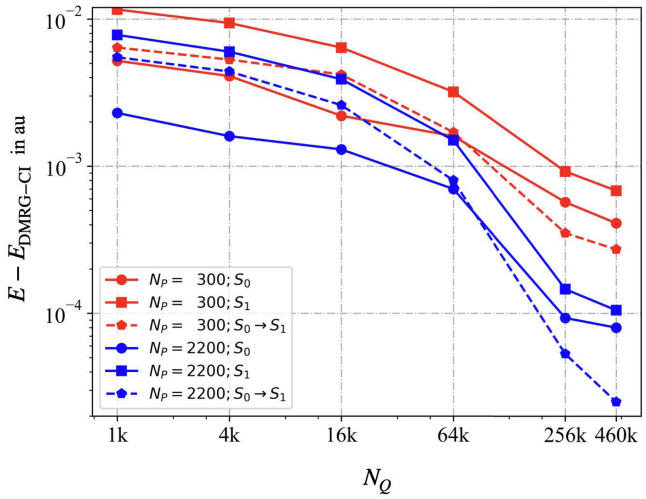

FIGURE 1 Convergence of the dCI energy errors for S_0 ($X^1\Sigma_g^+$) and S_1 ($B^1\Delta_g$) states of C_2 ($d_{C-C} = =1.24253$ Å) with respect to the outer subspace dimension N_Q in the cc-pVTZ basis for $N_P = 300$ and $N_P = 2200$ of the model subspace. Reprinted with permission from Ref. [97]. Copyright 2022 American Chemical Society.

contribution ($E_c^{\{\lambda\}}$) from the single-particle transformation $\mathbf{U}$ to obtain correlation energy gradients (E_c^{λ}), in addition to the relaxation of molecular ($E_c^{[\lambda]}$) and atomic ($E_c^{(\lambda)}$) orbitals.

$$E_c^{\lambda} = \frac{dE_c}{d\lambda} = E_c^{\{\lambda\}} + E_c^{[\lambda]} + E_c^{(\lambda)}, \tag{11}$$

$$\mathbf{U}(\lambda) = \mathbf{U}^{(0)}\mathbf{O}(\lambda), \tag{12}$$

$$\mathbf{U}^{\{\lambda\}} = \frac{d\mathbf{U}(\lambda)}{d\lambda} = \mathbf{U}^{(0)}\mathbf{O}^{\{\lambda\}}, \tag{13}$$

where E_c is the post-HF correlation energy and λ denotes a perturbation from, for example, atomic position displacements or an external field. It has been found that the absence of $E_c^{\{\lambda\}}$ results in significant error in predicted molecular structures.[34,270]

When a perturbation λ is applied to the system, the single-particle rotation $\mathbf{U}(\lambda)$ is perturbed and λ-dependent, which is represented exactly in a linear combination of the complete unperturbed rotation vectors $\mathbf{U}^{(0)}$ with unknown combination coefficients $\mathbf{O}(\lambda)$. The response $\mathbf{U}^{\{\lambda\}}$ of the perturbation-dependent rotation $\mathbf{U}(\lambda)$, as given in Equation (13), is obtained if $\mathbf{O}^{\{\lambda\}}$ is solved. In the NO-based schemes, the low-rank components of the reference $\mathbf{U}^{(0)} = \left[\mathbf{U}_l^{(0)}, \mathbf{U}_h^{(0)}\right]$ are normally selected within a kept eigenvector subspace ($\mathbf{U}_l^{(0)}$) of the virtual density matrix which is decoupled from the discarded complementary subspace ($\mathbf{U}_h^{(0)}$). The modern PNOs and OSVs are also similarly generated from the pair virtual density matrix and orbital-specific wavefunction amplitudes, respectively. This condition ensures that the density matrix expressed in the $\mathbf{U}^{(0)}$ basis is block-diagonal and the correlation energy is invariant to the rotation within either the kept or the discarded subspace. Hence, the response vector $\mathbf{U}^{\{\lambda\}}$ must be obtained via the rotation between the kept and discarded subspaces, which introduces for analytical theory new repulsion integrals in the long discarded basis that are not present in the energy computation. In PNO-based MP2 method,[25,202] this problem is handled by computing the response vector $\mathbf{U}^{\{\lambda\}}$ from the hierarchical relaxations of both PAOs and PNOs based on relatively compact discarded $\mathbf{U}_h^{(0)}$; in the OSV-MP2 gradient theory,[34,35] a compact $\mathbf{U}_h^{(0)}$ is automatically identified from the interpolative decomposition OSVs by tuning the rank of the semi-canonical amplitude matrix, as discussed in Section 3.1 and Table 1.

Another hurdle is the computation of the gradient contribution from weak electron correlations. For OSV-MP2 wavefunction, the number of weak electron pairs has been dramatically reduced according the OSV overlap criterion in Equation (7), and it is sufficient to consider only dispersion correlation (see Section 3.2.1) for which the cost is negligible, compared with strong electron pairs, since the computation of long-range integrals is avoided. According to our experimental evaluation to OSV-MP2 energy gradients, we invoke another approximation to avoid the expensive computation of the response of the one-electron part in the OSV-MP2 amplitude equations for weak pairs, which contribute little to the final gradients. For example, the nonactin molecule contains 8214 weak pairs out of 11,026 total pairs. The OSV-MP2 gradient error with def2-TZVP basis is only 1.2×10^{-5} au between the gradients with and without one-electron contribution and the maximum deviation only 8.3×10^{-5} au. Overall, we considerably boost the CPU, memory, and I/O efficiency for OSV-MP2 gradient evaluation without affecting accuracy.

3.4 | Cheap neural network learning

The low-rank post-HF methods discussed above are very useful for efficiently generating a large amount of almost noise-free training data of many atomic configurations within reasonable computational time and resources, rather than using cumbersome generic computations. However, the deep neural network training has to extract feature characters of large-scale and often redundant datasets for a learned mapping from input to output vectors. At a more fundamental level, the efficiency and learnability of the neural network hinges on how relevant physical properties are translated into the learning architecture.[271] As the low-rank data representation encodes certain physics including symmetry and locality, conceptually, we argue that the systematically tuned low-dimensional data structure increases the hierarchy and inhomogeneity of hidden features by removing the redundancy, and perhaps noise, in the generic wavefunction. This facilitates refactoring the outstanding interaction feature for electron correlations and improves the prediction transferability of the energy model between molecules of different size and geometry. Hence, the low-rank

operators, in which the original raw information of electron correlations is compressed through a lossy transformation, can are more expressive objects than handcrafted atom-based descriptors.

Bearing this in mind, in our electronic T-dNN model, we express the correlated descriptors[143] taking only a few OSVs, and we have found that the electron correlation characters are well reserved for making transferable prediction. Based on numerical experimentation, we define the feature amplitudes that respect the unique physical nature of the electron correlations according to the 2p–2h excitation patterns: vertical (vt, $\widetilde{\mathbf{T}}^{(\mathrm{vt})}_{ij,\overline{\mu}\overline{\nu}} = \frac{[i\overline{\mu}_i|j\overline{\nu}_j]}{\varepsilon_{ij}}$), exchange (ex, $\widetilde{\mathbf{T}}^{(\mathrm{ex})}_{ij,\overline{\mu}\overline{\nu}} = \frac{[i\overline{\mu}_j|j\overline{\nu}_i]}{\varepsilon_{ij}}$) and charge transfer (ct1, $\widetilde{\mathbf{T}}^{(\mathrm{ct1})}_{ij,\overline{\mu}\overline{\nu}} = \frac{[i\overline{\mu}_i|j\overline{\nu}_i]}{\varepsilon_{ij}}$ for type 1 and ct2, $\widetilde{\mathbf{T}}^{(\mathrm{ct2})}_{ij,\overline{\mu}\overline{\nu}} = \frac{[i\overline{\mu}_j|j\overline{\nu}_j]}{\varepsilon_{ij}}$ for type 2) correlations, which exhibit different attenuation dependence on the ij pair separation. Here $\varepsilon_{ij} = f_{ii} + f_{jj} - f_{\overline{\mu}_i\overline{\mu}_i} - f_{\overline{\nu}_j\overline{\nu}_j}$ with $f_{\overline{\mu}_i\overline{\nu}_j}$ the elements of virtual–virtual Fock matrix in the OSV basis, and there need only 8 OSVs $\{\overline{\mu}_1, \overline{\mu}_2, \cdots, \overline{\mu}_8\}$ automatically selected according to the most important singular values. One has to note that these feature amplitudes are extremely poor for directly computing correlation energies, but they describe the near-sighted limit of the amplitude equations for guiding the neural network learning of both MP2 and CCSD correlations. The feature amplitudes are further preprocessed to produce pseudo-energy inputs to the network,

$$\widetilde{\mathbf{e}}^{(\mathrm{vt})}_{ij,\overline{\mu}\overline{\nu}} = \widetilde{\mathbf{T}}^{(\mathrm{vt})}_{ij,\overline{\mu}\overline{\nu}}[i\overline{\mu}_i|j\overline{\nu}_j], \quad \widetilde{\mathbf{e}}^{(\mathrm{ex})}_{ij,\overline{\mu}\overline{\nu}} = \widetilde{\mathbf{T}}^{(\mathrm{ex})}_{ij,\overline{\mu}\overline{\nu}}\left[i\overline{\mu}_j|j\overline{\nu}_i\right], \quad \widetilde{\mathbf{e}}^{(\mathrm{ct1})}_{ij,\overline{\mu}\overline{\nu}} = \widetilde{\mathbf{T}}^{(\mathrm{ct1})}_{ij,\overline{\mu}\overline{\nu}}[i\overline{\mu}_i|j\overline{\nu}_i], \quad \widetilde{\mathbf{e}}^{(\mathrm{ct2})}_{ij,\overline{\mu}\overline{\nu}} = \widetilde{\mathbf{T}}^{(\mathrm{ct2})}_{ij,\overline{\mu}\overline{\nu}}\left[i\overline{\mu}_j|j\overline{\nu}_j\right]. \tag{14}$$

For each pair, the pseudo-energy tensor $\widetilde{\mathbf{e}}^{(\mathrm{X})}_{ij}$ for each type X uses only 64 exchange integrals. Our T-dNN model does not require an MP2 computation and needs considerably fewer repulsion integrals than what were reported.[257,260,265] The computation of these inputs is cheap and the main cost is dominated by the baseline HF with $\mathcal{O}(N^4)$. The scaling comparison is given in Table 3.The MPI-based parallel computation of the feature sets has been implemented by distributing LMO pairs over available processor cores, and nearly linear scaling computations per task can be carried out in many steps.

We briefly summarize the exceptional transferability of the T-dNN surrogate model in various aspects across different molecular sizes, datasets and conformations. For double-ζ basis sets, our results reveal that it is sufficient to predict chemically accurate correlation energies by training the T-dNN model with only a small dataset containing a few hundred molecules. For example, by only 100 training molecules randomly selected from QM9 dataset (including total 133,885 organic molecules), the mean absolute error (MAE) for CCSD/6–31g* correlation energies is about 1.05 kcal/mol for predicting the remaining 99.925% QM9 molecules, and is further reduced to 0.58, 0.52, 0.46, and 0.45 kcal/mol with 500, 1000, 2000, and 3000 training molecules, respectively. Moreover, by training only 100 QM9 molecular monomers, the predicted CCSD/6–31g* interaction energies of selected dimer complexes in ACONF, PCONF, S66, BBI, and SSI datasets are also accurate with MAEs <1 kcal/mol, including bimolecular interactions and non-covalent interactions. For triple-ζ basis sets, the T-dNN model trained on 100 QM7b-T molecules makes OSV-MP2/cc-pVTZ prediction

TABLE 3 Computational costs of all major steps for feature generation with the numbers of occupied LMOs (O) and atoms (N).

Computational steps	Asymptotic costs	Asymptotic costs per MO/pair
RHF energy	$\mathcal{O}(N^4)$	
Boys localization	$\mathcal{O}(N^3)$	
$[ii\|jj]$ and $[ij\|ij]$ integrals	$\mathcal{O}(O^2N^2)$	$\mathcal{O}(N^2)$
OSV generation	$\mathcal{O}(ON)$	$\mathcal{O}(N)$
OSV overlap $\langle\mu_i\|\nu_j\rangle$	$\mathcal{O}(O^2N)$	$\mathcal{O}(N)$
OSV Fock $\langle\mu_i\|\mathbf{F}\|\nu_j\rangle$	$\mathcal{O}(N^2)$	$\mathcal{O}(N)$
OSV exchange integral	$\mathcal{O}(N^2)$	$\mathcal{O}(N)$
Feature amplitudes $\widetilde{\mathbf{T}}_{ij}$	$\mathcal{O}(N)$	Constant
Pseudo-energy input $\widetilde{\mathbf{e}}_{ij}$	$\mathcal{O}(N)$	Constant

Note: Eight OSVs form the low-rank basis for descriptors. The asymptotic costs are estimated according to the linear growth of LMO pairs with N, the integral sparse-fitting implementation, and the interpolative decomposition for generating OSVs.[35] Adapted with permission from Ref. [143]. Copyright 2023 American Chemical Society.

with the MAE of 1.03 kcal/mol for all remaining QM7b-T molecules and 1.57 kcal/mol for molecules in another GDB13-T datasets, and the prediction MAEs are lowered to 0.49 and 0.89 kcal/mol for QM7b-T and GDB13-T molecules by training 800 QM7b-T molecules, respectively. The results suggest that more training molecules are necessary using larger basis sets.

Interestingly, the T-dNN model trained on small 800 $(H_2O)_{16}$ clusters exhibits systematic parallel deviations from the exact OSV-MP2/cc-pVTZ correlation energies by a few kcal/mol for $(H_2O)_{128}$ cluster of various conformations, as revealed in Figure 2. Nonetheless, this kind of near-constant errors does not affect the curvature of the potential energy surface by a constant global shift and thus we expect accurate geometry optimization and molecular dynamics simulations from T-dNN prediction. Although the total prediction error grows from $(H_2O)_{32}$ to $(H_2O)_{128}$ with the number of water molecules, the errors are systematic. When applying a global shift by the magnitude of the mean signed error, the prediction mean absolute relative error (MARE) is only 0.77 kcal/mol for $(H_2O)_{128}$. These positive results indicate that the electronic T-dNN model can transfer the underlying mapping between electron pairs and correlations toward larger molecules.

4 | ILLUSTRATIVE APPLICATIONS

We applied the low-rank correlated methods to several controversial molecular phenomena to which traditional DFT and post-HF methods are problematic. The development versions of OSV-MP2 and dCI programs[35,97] were employed. In the following illustration, we show that the scale-up algorithms of the low-rank OSV-MP2 analytical theory discussed above enable practical and accurate molecular structure optimization and Born–Oppenheimer molecular dynamics simulation that are difficult to generic methods for relatively complex systems and chemical processes. The first example devotes to the study of Cu-coordination structures (see Figure 3) of the interlocking Cu(I)–catenane supramolecule[272] that are managed by the ligand topologies and peripheral lengths, showing catalytic implications opposite to DFT results. The second example reveals that the MP2 electron correlation effects, drawn from the 10 ps classical-nuclei MD/NVE simulation driven by OSV-MP2 forces, retrieve the experimental broadening signature of the N–H vibration associated with intramolecular double hydrogen transfer in porphycene complex, which may not be attributed exclusively to proton quantum effects. The third application performs the hybrid OSV-MP2 and molecular mechanics (MM) MD simulation of a water microdroplet and reveals a substantial water–water autoionization on the microdroplet surface, which creates interfacial H_2O^+/H_2O^- radical pairs to catalyze an on-water reaction with two-carbon Criegee intermediate on the air/water surface. In the last example, we turn to strongly correlated excited states, and show that the low-rank dCI selection of important determinants is sufficient to recover near-exact excitation energies in both organic and transition metal compounds.

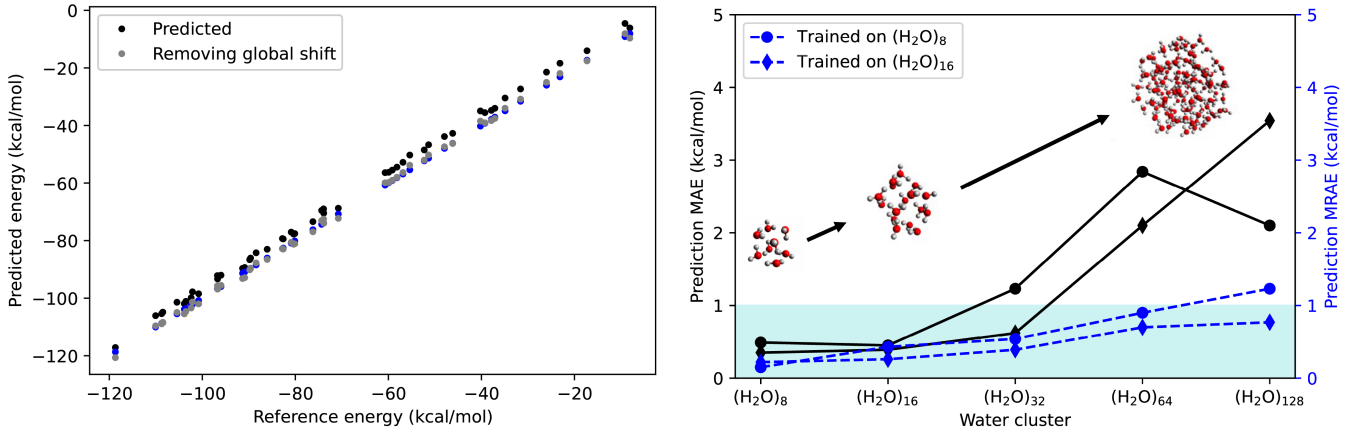

FIGURE 2 Comparison of the transferable predictions for MP2/cc-pVTZ energies from T-dNN model trained on small molecules. Left: The systematic shift of the predicted energies for $(H_2O)_{128}$ with a T-dNN model trained on 800 $(H_2O)_{16}$. Blue dots represent the energies from direct explicit computations. Right: The prediction errors for spherical water clusters of different sizes sampled from the molecular dynamics NVT trajectories. Adapted with permission from Ref. [143]. Copyright 2023 American Chemical Society.

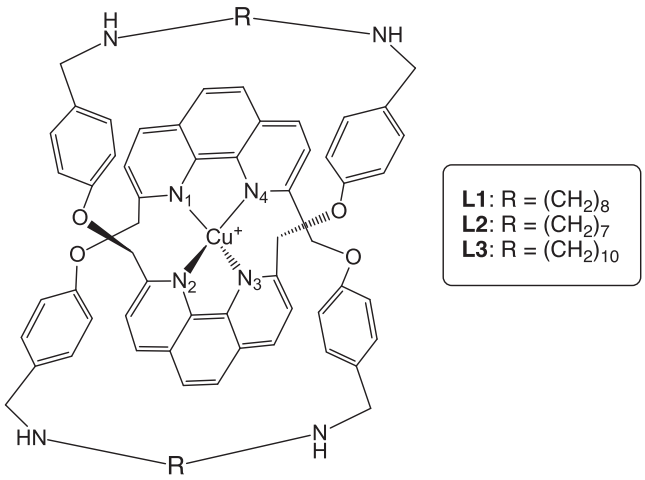

FIGURE 3 Chemical formula of Cu(I)–catenane complex. Adapted with permission from Ref. [35]. Copyright 2021 American Chemical Society.

TABLE 4 Comparison of the optimized –$Cu(N)_4$– coordination structures for $[Cu(L1)]PF_6$, $[Cu(L2)]PF_6$, and $[Cu(L3)]PF_6$ (Figure 3) between OSV-MP2/def2-TZVP (all electrons) and B3LYP-D3BJ/Lanl2dz/6-31 g(d,p) levels of theory.

Method		$[Cu(L1)]PF_6$	$[Cu(L2)]PF_6$	$[Cu(L3)]PF_6$
OSV-MP2	$d(N_1\text{–Cu})$ (pm)	202.22	200.22	201.75
	$d(N_2\text{-Cu})$ (pm)	197.53	197.44	198.26
	$d(N_3\text{-Cu})$ (pm)	202.30	200.51	201.39
	$d(N_4\text{-Cu})$ (pm)	197.47	197.12	198.23
	V_{coor} (pm^3)	3,377,975.78	3,346,818.43	3,366,231.91
	ΔV_{coor} (pm^3)	0.00	−31,157.35	−11,743.87
B3LYP-D3BJ	$d(N_1\text{-Cu})$ (pm)	205.48	204.50	205.52
	$d(N_2\text{-Cu})$ (pm)	204.99	208.76	208.21
	$d(N_3\text{-Cu})$ (pm)	205.48	204.01	205.58
	$d(N_4\text{-Cu})$ (pm)	205.00	209.68	207.53
	d (pm^3)	3,595,533.80	3,655,158.99	3,626,158.36
	ΔV_{coor} (pm^3)	0.00	59,625.19	30,624.56

Note: V_{coor} is the volume of the –$Cu(N)_4$– polyhedron. Adapted with permission from Ref. [35]. Copyright 2021 American Chemical Society.

4.1 | OSV-MP2 structure solver: Interlocking Cu–catenane supramolecular geometries

The mechanically interlocking tetradentate Cu(I)–catenane supramolecule (Figure 3) exhibits selective catalysis to $C(sp^3)$–O dehydrogenative reactions between phenol and bromodicarbonyl.[272] The catalytic activity for a broad scope of substrates can be managed by varying the ligand catenane topologies and peripheral lengths, which effectively adjusts the Cu(I)–catenane bonds and hence the Cu(I) coordination environment. Experimentally, our collaborators have shown that the Cu(L1) and Cu(L3) complexes in relatively loose mechanical bonds with long L1 ($R = (CH_2)_8$) and L3 ($R = (CH_2)_{10}$) ligands have a high catalytic yield of nearly 77%–80%, while the Cu(L2) complex in the tight Cu-N coordination considerably reduces the product generation at a yield of only 52%.

We attempt to examine whether the tetradentate –$Cu(N)_4$– coordination structure may be correlated with such catalytic activities shown in experiments. To this end, the OSV-MP2 structures of $[Cu(L1)]PF_6$, $[Cu(L2)]PF_6$, and $[Cu(L3)]PF_6$ were optimized with the def2-TZVP basis and compared with traditional DFT B3LYP-D3BJ optimizations. As seen in Table 4, the catenane ligand length in the number of methylene groups does not make a large impact on the Cu–N distances from OSV-MP2 prediction, causing <2 pm variation in all Cu–N bond lengths. However, the B3LYP-D3BJ computations lead to significantly longer Cu-N bond lengths and larger –$Cu(N)_4$– coordination volume V_{coor} than

OSV-MP2, in particular with large bond elongations of Cu–N_2 and Cu–N_4 from medium $[Cu(L1)]PF_6$ to either short $[Cu(L2)]PF_6$ or long $[Cu(L3)]PF_6$. Moreover, the OSV-MP2 predicts the smallest V_{coor} and hence strong interlocking mechanical bonds for $[Cu(L2)]P_6$ and larger V_{coor} for $[Cu(L1)]P_6$ and $[Cu(L3)]P_6$, which is well aligned with the catalytic efficacy ranking of $[Cu(L2)]PF_6 << [Cu(L1)]P_6 \sim [Cu(L3)]P_6$ given in experiments. However, the B3LYP-D3BJ computation results in the $-Cu(N)_4-$ volume change by $[Cu(L2)]P_6 > [Cu(L3)]P_6 > [Cu(L1)]P_6$, opposite to OSV-MP2 results and experiments. This study suggests that an ab-initio correlated model for post-HF energy and structure computations is critical for discerning the delicate response of the coordination environment to ligand changes, which is further implicated in the supramolecular catalytic efficiency.

4.2 | OSV-MP2 MD simulation: Tautomeric broadening of N–H vibrations in porphycene

Porphycene (Pc, $C_{20}H_{14}N_4$) provides a channel for fast double hydrogen transfer, resulting in tautomerization reactions at room temperature along the $N-H\cdots N$ in the molecular cavity formed by four nitrogens.[273] The hydrogen transfer leads to different tautomers: *cis*-Pc tautomer with two hydrogens bonded to nitrogens on the same side and *trans*-Pc tautomer with two hydrogens connected to nitrogens on the other side. However, the static computation of harmonic frequencies predicts only a single strong N–H stretching vibration at around 2900 cm^{-1}, while the experimental infrared spectrum shows a significant N–H band broadening over 2000–3000 cm^{-1}. The standard harmonic computation is flawed in the absence of vibrational anharmonicity and intermode couplings, which turn out to be significant in porphycene due to hydrogen transfer. The DFT-based ring-polymer path integral MD simulations ascribe the broadened N–H vibrational bands around 2200–3200 cm^{-1} to the nuclear quantum effect of transferred protons.[274] However, the appearance of the N–H stretching signature is highly sensitive to the chosen DFT functionals.

We probed the origin of the broad N-H vibrational peak by performing the 10 ps classical-nuclei ab-initio MD/NVE simulation using OSV-MP2 correlated model at a time step of 0.5 fs.[35] Our computed OSV-MP2 vibrational density of states (VDOS) retrieves both broadened low- and high-energy N–H stretching bands centered at 2600 and 3000 cm^{-1} (Figure 4), respectively, by propagating classical protons. However, the VDOS from the uncorrelated RHF MD simulation does not yield any band signature in 2000–3200 cm^{-1}, which indicates the importance of electron correlations. The lower N–H band at around 2400–2600 cm^{-1} is weak and assigned to *cis*-Pc tautomer, and the relatively strong band at 2800–3300 cm^{-1} originates from the *trans*-Pc tautomer, showing more *trans*-Pc tautomers than *cis*-Pc due to fast hydrogen transfer. This is in contrast to the literature ring-polymer B3LYP-vdW/MD infrared spectrum[274] which concludes a larger portion of *cis*-Pc tautomer than *trans*-Pc and points to the effect of quantal protons. The results by us and others imply that the origin of the broad N–H stretching bands is controversial between electron correlation and protonic quantum effects. Further studies are needed to investigate the impact of the proton-coupled correlated electrons.

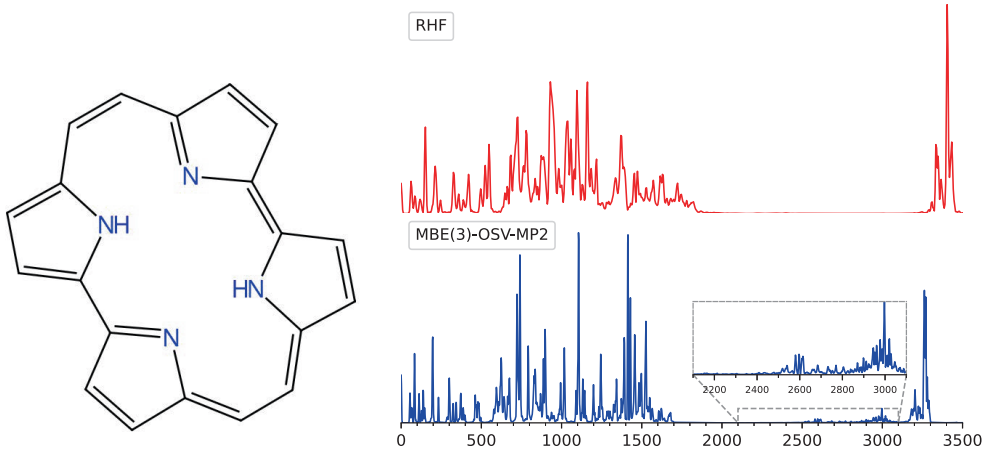

FIGURE 4 Left: Porphycene formula. Right: The VDOS spectra from the 10 ps MD/NVT simulation driven by RHF/6–31g* (red) and OSV-MP2/6–31g* (blue) forces at T = 291.9 K, followed by another 10 ps NVE equilibration. The OSV-MP2 MD simulation was carried out on 96 CPU cores (IntelXeon Platinum 9242@2.30 GHz). Adapted with permission from Ref. [35]. Copyright 2021 American Chemical Society.

4.3 | Multiscale OSV-MP2 QM/MM method: Surface water charge transfer and reactivity

We implemented the multiscale OSV-MP2/MM method to drive long-time ab-initio MD simulations with the dynamically adaptive QM-MM boundary for containing the constant number of QM molecules, including, for example, both the reactive center and a sufficient amount of surrounding environment molecules. This method features the low-rank OSV-MP2 electronic structure computation for explicitly treating QM molecules as well as the number-adaptive scheme[62,275] for updating the QM region along the time evolution. For explicitly including long-distance processes in the QM region, such as proton translocation and charge transfer, we further implemented a reallocation scheme of flexible QM center from which the numbers of QM and MM molecules are both constant on all full time-dependent trajectories.

We next address an interesting aspect of molecular reactivities on the water microdroplet[62] which exhibits a profound difference of the air-water interfacial environment from the bulk water for expediting certain reactions. Regarding the rich interfacial dangling protons near the air-water surface,[276] reactants and intermediate states may be favorably aligned to promote fast "on-water" reaction processes. High-level electronic structure methods are needed for precisely distinguishing the electronic structures between interfacial and bulk waters. Hence, we first performed the single-point OSV-MP2 computation of a large $(H_2O)_{190}$ microdroplet of the diameter of 22 Å with aug-cc-pVTZ basis, and analyzed the ab-initio water charge distribution. The microdroplet surface undergoes spontaneous water–water charge exchange that is highly inhomogeneous: a surprising large amount of charges up to ±0.20 e per water are present and create H_2O^+/H_2O^- radical pairs in the surface layer of 1–2 Å depth, and the waters are neutral as usual with only minor charge separations by 0.005 e per water near the microdroplet center.

The significance of the surface water radical pairs was investigated on an important atmospheric Criegee–water reaction, as often encountered within aerosols and clouds formed in the troposphere that impacts the global climate. Recent DFT-based QM/MM Born–Oppenheimer MD results suggest that CH_2OO–water reaction may rapidly proceed via several pathways,[277,278] and *anti*-CH_3CHOO is highly stable.[279] However, we do not fully understand the Criegee reactivities at the aqueous interface due to the presence of spontaneously charged waters. In our ab-initio simulation, the QM and MM subsystems were defined by *anti*-CH_3CHOO–$(H_2O)_{15}$ with OSV-MP2 and 1097 explicit water solvents with TIP3P model, respectively. The hybrid OSV-MP2 QM/MM MD simulation reveals that the two-carbon *anti*-CH_3CHOO Criegee molecule moving closer to water on the surface, for example, the Criegee–water distance falls within 1.5 Å, induces a conformation reorientation of 8 surface H_2O in the QM region, yielding water–water charge transfer on a negatively charged water by $\delta^- = -0.20$ near the surface and $\delta^+ = +0.39$ on the remaining 7 H_2O^+ penetrating the water microdroplet. The resulting H_2O^+/H_2O^- pairs build up a local electric field pointing to the surface (Figure 5a), which electrically reorients *anti*-CH_3CHOO by pulling the positively charged Criegee C_c atom toward

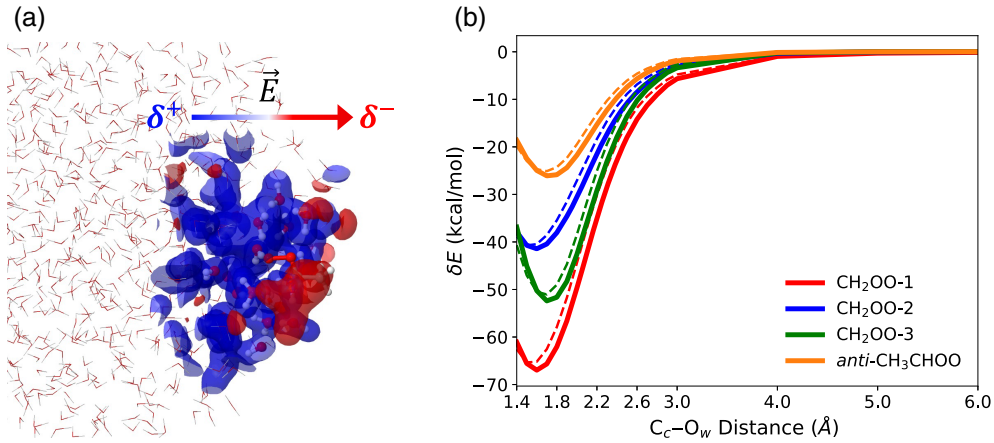

FIGURE 5 Water charge transfer creates the surface H_2O^+/H_2O^- radical pairs. The QM and MM subsystems are defined by *anti*-CH_3CHOO–$(H_2O)_{15}$ with OSV-MP2 and 1097 explicit water solvents with TIP3P model, respectively. (a) The resulting electrostatic potential of the QM water molecules nearby the air-water interface. (b) The interfacial stabilization energies (δE, solid) and the electrostatic contributions (dashed) when moving CI from air across the water microdroplet surface. Reprinted with permission from Ref. [62]. Copyright 2023 American Chemical Society.

H_2O^- near the surface, and the local field pushes an intramolecular electron transfer from the Criegee C_c to the terminal O_t atom to facilitate the reaction.

The OSV-MP2 QM/MM MD simulation captures Criegee(H_2O^-) intermediate state stabilized by an interfacial stabilization energy (Figure 5b, with DLPNO-CCSD(T)/aug-cc-pVTZ energy model) due to Criegee–water electrostatic attraction, for example, $\delta E \approx -28$ kcal/mol for *anti*-CH_3CHOO estimated from one QM/MM MD trajectory at a short C_c–O_w distance. δE is found to mainly arise from the interfacial stabilization of the *anti*-CH_3CHOO–H_2O^- electrostatic interactions. As seen in Figure 5b, compared with the gas phase reaction, the interfacial stabilization energy of the transition state (TS) corresponding to $d(C_c–O_w) = 1.9–2.0$ Å is estimated to be -22 kcal/mol, much greater in magnitude than the stabilization of -5 kcal/mol for the reactant state identified at $d(C_c–O_w) = 2.5–2.6$ Å. This suggests that the interfacial *anti*-CH_3CHOO reaction with water becomes barrierless with an activation energy $E_{TS,surface} = 0$. By assuming an activation energy of $E_{TS,gas} = 7 \sim 9$ kcal/mol and $E_{TS,gas} = 5 \sim 6$ kcal/mol for the gas phase *anti*-CH_3CHOO reaction with water monomer and dimer, respectively, we estimated an enhancement of the reaction rate coefficients by 5–6 and 3–4 orders of magnitude for the barrierless *anti*-CH_3CHOO/water reaction at the air-water interface at $T = 300$ K, according to a simple rate model ($e^{-\left(E_{TS,surface}-E_{TS,gas}\right)/k_BT}$).

4.4 | Low-rank dCI effective Hamiltonian: Low-lying excited states

We finally turn to showcase the possibility of setting up dCI effective Hamiltonians for accurately computing correlated molecular states that traditionally require a prohibitively long wavefunction expansion in the determinant basis. While there is much room to improve the algorithmic efficiency for selecting important state-specific determinants, our dCI assessment on low-lying states of several organic and Cu-ligand coordination compounds[97] clearly reveals that the molecular Hamiltonian is compressive and can be represented in compact dCI subspaces that are systematically tunable toward chemical accuracy. These molecular excitation energies for various low-lying singlet and doublet states of different character, shown in Table 5, agree to the reference DMRG-CI benchmark with deviation of 0.04–0.05 eV for which the FCI expansion typically demands about $10^{13} - 10^{18}$ determinants with practical cc-pVTZ basis set. For achieving chemical accuracy, the dCI recursive selection yields about $N_P = 650–2000$ and $N_Q = 250{,}000–550{,}000$ most important determinants. These results clearly demonstrate the advantage of the dCI selection algorithm: despite of the broad scale of FCI determinants across 5–6 orders of magnitude due to the diverse multireference character of these molecules, the dimensions of the selected model and outer subspaces are drastically narrowed down to a range differing by less than a factor of three among these molecules.

It is noted that the basis set impact leads to an increase of the dCI subspace dimensionality.[97] For C_2 molecule in Table 5, by using the consistently augmented basis sets of cc-pVDZ, cc-pVTZ, and cc-pVQZ, the ground state energies computed in a small $N_P = 2200$ model subspace deviate from the DMRG-CI reference values by 0.08, 0.7, and 3.6 mau with $N_Q = 64{,}000$ outer determinants, respectively, and are reduced to 0.02, 0.09, and 0.8 mau with $N_Q = 400{,}000$ outer determinants, respectively. The increasing number of outer determinants necessary for reaching chemical accuracy is however rather moderate, as compared with the blast increase of the FCI dimensions from 1.4×10^{11} for cc-pVDZ to 4.6×10^{18} for cc-pVTZ.

5 | CONCLUSION AND OUTLOOK

We have discussed the developments and results of various recent low-dimensional representations from large-scale MP2 and CC to CI types of wavefunction. These methods make broad explorations in either single-reference or multireference algorithms that attempt to automatically and systematically identify important wavefunction components which dominantly contribute to electron correlations in a promising cost-accuracy balanced fashion. However, Given the diverse correlation character of molecules, there is presently no single method that can be successfully applied to treat all many-body problems. Although they differ from one another in the variety of wavefunction formulations, implementation schemes and technical details, these approaches commonly feature a proper single-particle or many-particle wavefunction transformation to rank the significance of electronic configurations replying on the wavefunction sparsity or compressibility. It is important to point out that the transformation and ranking cause extra operations bearing non-negligible costs that are not present in conventional correlated methods, and extensive efforts have been made to reduce the ad-hoc impacts to the overall computational efficiency. In particular, for retrieving

TABLE 5 Error comparison of low-lying excitation energies ($|\Delta\omega|$) between dCI and reference DMRG-CI ($M = 4000$) results with respect to the numbers of complete (N_{Full}), model (N_P) and outer (N_Q) determinants for organic and transition metal compounds.

Molecules	Characters	Active space	N_{Full}	N_P (×1000)	N_Q (×1000)	DMRG-CI (eV)	dCI (eV)	$\lvert\Delta\omega\rvert$ (eV)
6–31g								
$NH_3 \rightarrow F_2$	Charge transfer	(30o, 22e)	3.0×10^{15}	0.91	400	9.26	9.31	0.05
$N_2 \rightarrow CH_2$	Charge transfer	(28o, 16e)	9.7×10^{12}	0.67	250	15.32	15.36	0.04
cc-pVTZ								
C_2H_4	$1^1A_g \rightarrow 2^1A_g$	(114o, 12e)	7.1×10^{18}	1.22	550	13.07	13.02	0.05
C_2	$X^1\Sigma_g^+ \rightarrow B^1\Delta_g$	(60o, 12e)	3.4×10^{15}	0.30	64	2.18	2.22	0.04
C_3	$1^1\Sigma_g^+ \rightarrow 1^1\Delta_g$	(87o, 12e)	2.6×10^{17}	0.95	450	5.22	5.18	0.05
HNO	$1^1A' \rightarrow 2^1A'$	(72o, 12e)	2.4×10^{16}	0.82	350	4.33	4.37	0.04
H_2S	$S_0 \rightarrow S_1$	(61o, 16e)	8.7×10^{18}	1.14	500	6.95	6.91	0.04
HCHO	$S_0 \rightarrow S_1$	(86o, 12e)	2.2×10^{17}	0.98	450	4.15	4.21	0.05
def2-TZVP								
$[CuN_6C_{20}H_{18}]^+$	$S_0 \rightarrow S_1$	(30o, 30e)	2.4×10^{16}	1.62	440	6.28	6.32	0.04
	$S_0 \rightarrow S_2$			1.73	470	6.77	6.81	0.04
$[CuN_7C_{22}H_{21}]^{2+}$	$D_0 \rightarrow D_1$	(30o, 29e)	2.3×10^{16}	1.45	440	5.19	5.24	0.05
	$D_0 \rightarrow D_2$			1.75	460	4.31	4.35	0.04
$[CuN_7C_{22}H_{21}]^{3+}$	$S_0 \rightarrow S_1$	(30o, 30e)	2.4×10^{16}	1.64	480	4.38	4.43	0.05
	$S_0 \rightarrow S_2$			1.84	470	5.24	5.29	0.05

Note: All valence electrons are correlated in organic molecules. For $[Cu(NHC)_2(pyridine)_2]^+$ ($x = 1, 2, 3$) complexes of different Cu oxidation state, the active space contains 30 molecular orbitals for 30 or 29 valence electrons for valency $x = 1, 3$ and $x = 2$, respectively, with predominant atomic orbitals of Cu/3d, C/2p and N/2p. The dCI energy is converged with 10^{-4} au with Pipek–Mezey localized orbitals. The source data is available in Ref. [97].

dynamic correlation from low-rank post-HF methods, a prior estimate of important wavefunction amplitudes from single-particle objects is usually sufficiently accurate. However, an iterative augmentation of the sCI wavefunction is normally necessary by selecting subclasses of configurations according to state-specific heuristic solutions, which is still prone to an exponential scaling with the system size when exploring the configurational space.

The compressive design of correlated electronic structure methods offers exciting feasibilities to solve problems that are traditionally difficult and controversial to generic DFT (with poor computational reliability) and post-HF (with poor computational scalability) methods, which we briefly illustrate and discuss in this review. The weakly correlated ground states can be now routinely handled by PNO- or OSV-based MP2 and CC methods for large molecules. In particular, the low-rank MP2 analytical gradient theories offer an alternative to DFT for optimizing ab-initio structures of complex molecule. For strongly correlated molecules, the variety of sCI implementations has been improved to afford highly accurate computation of polyatomic molecules containing several non-hydrogen atoms, albeit consuming considerable computational resources.

We have provided a snapshot of the current state-of-art in ML surrogate models for substituting explicit electronic structure computations aiming for energy chemical accuracy. The low-rank technologies provide new idea for engineering electron-based transferrable and expressive feature sets, more than being used as toolkits for expediting correlated many-body computation. An ongoing grand challenge, which is to reliably predict molecular energies of complex molecules that are not well represented in ML training datasets, can be now clearly addressed when employing cheap low-rank electronic descriptors that respect the electron correlated characters. We have demonstrated that accurate MP2 and CCSD correlation energies can be predicted by learning small molecules in small datasets. The data efficiency and transferable learnability are validated across alkanes, organic molecules, biomolecular interactions, and water clusters of various sizes and morphologies.

There are several areas that need continued efforts to make improvements of these theoretical models. One important issue is to make implementation progress on many graphical processing units (GPU) in light of the low-dimensional data objects needed in low-rank post-HF computations are very suitable for efficient instruction on GPU threads. Another pressing theme is to develop low-rank post-HF methods for evaluating the electronic and geometric structures of periodic solids. The intersection of period low-rank correlated methods and transferable electronic ML models is promising to offer new opportunities for analyzing the energy thermodynamic limit and tackling a range of condensed matter phenomena with predictive power. One valuable observation is that practical low-rank correlation computations show superiority in run-time efficiency to generic Hartree-Fock for medium and large molecules, the latter of which forms the next hurdle to remove. We expect that an efficient combination of the tunable low-rank correlation method for more expressive feature extraction, better transferable low-data quantum ML model and the hybrid CPU/GPU platform will be developed in the near future for simulating macromolecules and complex processes.

AUTHOR CONTRIBUTIONS

Jun Yang: Conceptualization (lead); data curation (lead); formal analysis (lead); funding acquisition (lead); methodology (lead); project administration (lead); resources (lead); writing – original draft (lead); writing – review and editing (lead).

ACKNOWLEDGMENTS

Jun Yang thanks the funding supports from Hong Kong Research Grant Council (ECS27307517, GRF17309020, GRF17310922), the Hong Kong Quantum AI Lab Limited, the Computational Initiative Program of the Faculty of Science and the Hui's fund of the Department of Chemistry at the University of Hong Kong. Computational works were partially performed on the National Supercomputer Center in Guangzhou of China and the research computing facilities offered by Information Technology Services at the University of Hong Kong.

FUNDING INFORMATION

Hong Kong Research Grant Council under Grant No. ECS27307517, GRF17309020 and GRF17310922; the Hong Kong Quantum AI Lab Limited through the AIR@InnoHK program of the Hong Kong SAR government; the Faculty Computational Initiative Program; the Chemistry Department Hui's Fund; the University Postgraduate Fellowship.

CONFLICTS OF INTEREST STATEMENT

The author declares no potential conflict of interests for this article.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

RELATED WIREs ARTICLES

Periodic and fragment models based on the local correlation approach

Chemical transformations and transport phenomena at interfaces

Atomistic neural network representations for chemical dynamics simulations of molecular, condensed phase, and interfacial systems: Efficiency, representability, and generalization

Recent advances in quantum fragmentation approaches to complex molecular and condensed-phase systems

ORCID

Jun Yang https://orcid.org/0000-0001-8701-9297

REFERENCES

1. Nesbet RK. Electronic correlation in atoms and molecules. Adv Chem Phys. 1965;9:321–63.
2. Gordon MS, Fedorov DG, Pruitt SR, Slipchenko LV. Fragmentation methods: a route to accurate calculations on large systems. Chem Rev. 2012;112(1):632–72.
3. Fang T, Li Y, Li S. Generalized energy-based fragmentation approach for modeling condensed phase systems. Wiley Interdiscip Rev Comput Mol Sci. 2017;7(2):e1297.
4. Liu J, He X. Recent advances in quantum fragmentation approaches to complex molecular and condensed-phase systems. Wiley Interdiscip Rev Comput Mol Sci. 2023;13(3):e1650.

5. Maslen PE, Head-Gordon M. Non-iterative local second order Møller–Plesset theory. Chem Phys Lett. 1998;283:102–8.
6. Ayala PY, Scuseria GE. Linear scaling second-order Moller–Plesset theory in the atomic orbital basis for large molecular systems. J Chem Phys. 1999;110(8):3660–71.
7. Lee MS, Maslen PE, Head-Gordon M. Closely approximating second-order Moller–Plesset perturbation theory with a local triatomics in molecules model. J Chem Phys. 2000;112(8):3592–601.
8. Doser B, Lambrecht DS, Kussmann J, Ochsenfeld C. Linear-scaling atomic orbital-based second-order Møller–Plesset perturbation theory by rigorous integral screening criteria. J Chem Phys. 2009;130(6):064107.
9. Hampel C, Werner H-J. Local treatment of electron correlation in coupled cluster theory. J Chem Phys. 1996;104(16):6286–97.
10. Werner H-J, Knizia G, Krause C, Schwilk M, Dornbach M. Scalable electron correlation methods I.: PNO-LMP2 with linear scaling in the molecular size and near-inverse-linear scaling in the number of processors. J Chem Theory Comput. 2015;11:484–507.
11. Schütz M, Werner H-J. Local perturbative triples correction (T) with linear cost scaling. Chem Phys Lett. 2000;318(4-5):370–8.
12. Schütz M, Werner H-J. Low-order scaling local electron correlation methods. IV. Linear scaling local coupled-cluster (LCCSD). J Chem Phys. 2001;114(2):661–81.
13. Schütz M. Low-order scaling local electron correlation methods. V. Connected triples beyond (T): linear scaling local CCSDT-1b. J Chem Phys. 2002;116(20):8772–85.
14. Schütz M. A new, fast, semi-direct implementation of linear scaling local coupled cluster theory. Phys Chem Chem Phys. 2002;4(16):3941–7.
15. Werner H-J, Schütz M. An efficient local coupled cluster method for accurate thermochemistry of large systems. J Chem Phys. 2011;135(14):144116.
16. Subotnik JE, Head-Gordon M. A local correlation model that yields intrinsically smooth potential-energy surfaces. J Chem Phys. 2005;123(6):064108.
17. Auer AA, Nooijen M. Dynamically screened local correlation method using enveloping localized orbitals. J Chem Phys. 2006;125(2):024104.
18. Subotnik JE, Sodt A, Head-Gordon M. The limits of local correlation theory: electronic delocalization and chemically smooth potential energy surfaces. J Chem Phys. 2008;128(3):034103.
19. Meyer W. Ionization energies of water from PNO-CI calculations. Int J Quantum Chem. 1971;5(S5):341–8.
20. Ahlrichs R, Lischka H, Staemmler V, Kutzelnigg W. PNO–CI (pair natural orbital configuration interaction) and CEPA–PNO (coupled electron pair approximation with pair natural orbitals) calculations of molecular systems. I. Outline of the method for closed-shell states. J Chem Phys. 1975;62(4):1225–34.
21. Neese F, Hansen A, Liakos DG. Efficient and accurate approximations to the local coupled cluster singles doubles method using a truncated pair natural orbital basis. J Chem Phys. 2009;131(6):064103.
22. Yang J, Kurashige Y, Manby FR, Chan GKL. Tensor factorizations of local second-order Møller-Plesset theory. J Chem Phys. 2011;134(4):044123.
23. Yang J, Chan GK-L, Manby FR, Schütz M, Werner H-J. The orbital-specific-virtual local coupled cluster singles and doubles method. J Chem Phys. 2012;136(14):144105.
24. Schütz M, Yang J, Chan GK-L, Manby FR, Werner H-J. The orbital-specific virtual local triples correction: OSV-L(T). J Chem Phys. 2013;138(5):054109.
25. Riplinger C, Neese F. An efficient and near linear scaling pair natural orbital based local coupled cluster method. J Chem Phys. 2013;138(3):034106.
26. Riplinger C, Sandhoefer B, Hansen A, Neese F. Natural triple excitations in local coupled cluster calculations with pair natural orbitals. J Chem Phys. 2013;139(13):134101.
27. Flocke N, Bartlett RJ. A natural linear scaling coupled-cluster method. J Chem Phys. 2004;121(22):10935–44.
28. Hughes TF, Flocke N, Bartlett RJ. Natural linear-scaled coupled-cluster theory with local transferable triple excitations: applications to peptides. J Phys Chem A. 2008;112(26):5994–6003.
29. Stoll H. Correlation energy of diamond. Phys Rev B. 1992;46(11):6700–4.
30. Friedrich J, Hanrath M, Dolg M. Fully automated implementation of the incremental scheme: application to CCSD energies for hydrocarbons and transition metal compounds. J Chem Phys. 2007;126(15):154110.
31. Nagy PR, Samu G, Kállay M. An integral-direct linear-scaling second-order Møller-Plesset approach. J Chem Theory Comput. 2016;12(10):4897–914.
32. Haldar S, Riplinger C, Demoulin B, Neese F, Izsak R, Dutta AK. Multilayer approach to the IP-EOM-DLPNO-CCSD method: theory, implementation, and application. J Chem Theory Comput. 2019;15(4):2265–77.
33. Ghosh S, Neese F, Izsák R, Bistoni G. Fragment-based local coupled cluster embedding approach for the quantification and analysis of noncovalent interactions: exploring the many-body expansion of the local coupled cluster energy. J Chem Theory Comput. 2021;17(6):3348–59.
34. Zhou R, Liang Q, Yang J. Complete OSV-MP2 analytical gradient theory for molecular structure and dynamics simulations. J Chem Theory Comput. 2019;16(1):196–210.
35. Liang Q, Yang J. Third-order many-body expansion of OSV-MP2 wave function for low-order scaling analytical gradient computation. J Chem Theory Comput. 2021;17(11):6841–60.

36. Li S, Ma J, Jiang Y. Linear scaling local correlation approach for solving the coupled cluster equations of large systems. J Comput Chem. 2002;23(2):237–44.
37. Li W, Li S. Divide-and-conquer local correlation approach to the correlation energy of large molecules. J Chem Phys. 2004;121(14):6649–57.
38. Li S, Shen J, Li W, Jiang Y. An efficient implementation of the "cluster-in-molecule" approach for local electron correlation calculations. J Chem Phys. 2006;125(7):074109.
39. Li W, Piecuch P, Gour JR, Li S. Local correlation calculations using standard and renormalized coupled-cluster approaches. J Chem Phys. 2009;131(11):114109.
40. Rolik Z, Kállay M. A general-order local coupled-cluster method based on the cluster-in-molecule approach. J Chem Phys. 2011;135(10):104111.
41. Rolik Z, Szegedy L, Ladjánszki I, Ladóczki B, Kállay M. An efficient linear-scaling CCSD(T) method based on local natural orbitals. J Chem Phys. 2013;139(9):094105.
42. Hirata S, Valiev M, Dupuis M, Xantheas SS, Sugiki S, Sekino H. Fast electron correlation methods for molecular clusters in the ground and excited states. Mol Phys. 2005;103(15-16):2255–65.
43. Dahlke EE, Truhlar DG. Electrostatically embedded many-body expansion for large systems, with applications to water clusters. J Chem Theory Comput. 2007;3(1):46–53.
44. Hirata S. Fast electron-correlation methods for molecular crystals: an application to the α, β1, and β2 modifications of solid formic acid. J Chem Phys. 2008;129(20):204104.
45. Bygrave PJ, Allan NL, Manby FR. The embedded many-body expansion for energetics of molecular crystals. J Chem Phys. 2012;137(16):164102.
46. Liu K-Y, Herbert JM. Energy-screened many-body expansion: a practical yet accurate fragmentation method for quantum chemistry. J Chem Theory Comput. 2019;16(1):475–87.
47. Herbert JM. Fantasy versus reality in fragment-based quantum chemistry. J Chem Phys. 2019;151(17):170901.
48. Adams WH. On the solution of the Hartree-Fock equation in terms of localized orbitals. J Chem Phys. 1961;34(1):89–102.
49. Pulay P. Localizability of dynamic electron correlation. Chem Phys Lett. 1983;100(2):151–4.
50. Komeiji Y, Ishida T, Fedorov DG, Kitaura K. Change in a protein's electronic structure induced by an explicit solvent: an ab initio fragment molecular orbital study of ubiquitin. J Comput Chem. 2007;28(10):1750–62.
51. Fedorov DG, Jensen JH, Deka RC, Kitaura K. Covalent bond fragmentation suitable to describe solids in the fragment molecular orbital method. J Phys Chem A. 2008;112(46):11808–16.
52. Duan LL, Tong Y, Mei Y, Zhang QG, Zhang JZH. Quantum study of HIV-1 protease-bridge water interaction. J Chem Phys. 2007;127(14):145101.
53. He X, Fusti-Molnar L, Cui G, Merz KM. Importance of dispersion and electron correlation in ab initio protein folding. J Phys Chem B. 2009;113(15):5290–300.
54. Bernát Szabó P, Csóka J, Kállay M, Nagy PR. Linear-scaling open-Shell MP2 approach: algorithm, benchmarks, and large-scale applications. J Chem Theory Comput. 2021;17(5):2886–905.
55. Willow SY, Salim MA, Kim KS, Hirata S. Ab initio molecular dynamics of liquid water using embedded-fragment second-order many-body perturbation theory towards its accurate property prediction. Sci Rep. 2015;5(1):14358.
56. Liu J, He X, Zhang JZH. Structure of liquid water – a dynamical mixture of tetrahedral and 'ring-and-chain' like structures. Phys Chem Chem Phys. 2017;19(19):11931–6.
57. Conrad JA, Kim S, Gordon MS. Ionic liquids from a fragmented perspective. Phys Chem Chem Phys. 2019;21(31):16878–88.
58. Liu J, Yang J, Zeng XC, Xantheas SS, Yagi K, He X. Towards complete assignment of the infrared spectrum of the protonated water cluster $H^+(H_2O)_{21}$. Nat Commun. 2021;12(1):6141.
59. Garcia-Ratés M, Becker U, Neese F. Implicit solvation in domain based pair natural orbital coupled cluster (DLPNO-CCSD) theory. J Comput Chem. 2021;42(27):1959–73.
60. Liu J, Lan J, He X. Toward high-level machine learning potential for water based on quantum fragmentation and neural networks. J Phys Chem A. 2022;126(24):3926–36.
61. Seeger ZL, Izgorodina EI. A DLPNO-CCSD(T) benchmarking study of intermolecular interactions of ionic liquids. J Comput Chem. 2022;43(2):106–20.
62. Liang Q, Zhu C, Yang J. Water charge transfer accelerates Criegee intermediate reaction with H_2O–radical anion at the aqueous Interface. J Am Chem Soc. 2023;145(18):10159–66.
63. Fujita T, Nakano T, Tanaka S. Fragment molecular orbital calculations under periodic boundary condition. Chem Phys Lett. 2011;506(1-3):112–6.
64. Wen S, Beran GJO. Accurate molecular crystal lattice energies from a fragment QM/MM approach with on-the-fly ab initio force field parametrization. J Chem Theory Comput. 2011;7(11):3733–42.
65. Friedrich J, Perlt E, Roatsch M, Spickermann C, Kirchner B. Coupled CLUSTER IN CONDENSED Phase. Part I: Static quantum chemical calculations of hydrogen fluoride clusters. J Chem Theory Comput. 2011;7(4):843–51.
66. Pisani C, Schütz M, Casassa S, Usvyat D, Maschio L, Lorenz M, et al. Cryscor: a program for the post-Hartree–Fock treatment of periodic systems. Phys Chem Chem Phys. 2012;14(21):7615–28.
67. Usvyat D, Maschio L, Schütz M. Periodic local MP2 method employing orbital specific virtuals. J Chem Phys. 2015;143:10.

68. Usvyat D, Maschio L, Schütz M. Periodic and fragment models based on the local correlation approach. Wiley Interdiscip Rev Comput Mol Sci. 2018;8(4):e1357.
69. Yang J, Weifeng H, Usvyat D, Matthews D, Schütz M, Chan GK-L. Theoretical chemistry. Ab initio determination of the crystalline benzene lattice energy to sub-kilojoule/mole accuracy. Science. 2014;345(6197):640–3.
70. Fang T, Li W, Fangwei G, Li S. Accurate prediction of lattice energies and structures of molecular crystals with molecular quantum chemistry methods. J Chem Theory Comput. 2015;11(1):91–8.
71. Wang Y, Ni Z, Neese F, Li W, Guo Y, Li S. Cluster-in-molecule method combined with the domain-based local pair natural orbital approach for electron correlation calculations of periodic systems. J Chem Theory Comput. 2022;18(11):6510–21.
72. Trinquier G, Malrieu J-P. Kekulé versus Lewis: when aromaticity prevents electron pairing and imposes polyradical character. Chemistry. 2015;21(2):814–28.
73. Li X, Paldus J. Binding in transition metal complexes: reduced multireference coupled-cluster study of the MCH_{2+} (M = Sc to Cu) compounds. J Chem Phys. 2007;126(23):234303.
74. Musia M, Bartlett RJ. Critical comparison of various connected quadruple excitation approximations in the coupled-cluster treatment of bond breaking. J Chem Phys. 2005;122(22):224102.
75. Bauschlicher CW, Partridge H. Cr2 revisited. Chem Phys Lett. 1994;231(2):277–82.
76. Larsson HR, Zhai H, Umrigar CJ, Chan GK-L. The chromium dimer: closing a chapter of quantum chemistry. J Am Chem Soc. 2022;144(35):15932–7.
77. Motta M, Ye E, McClean JR, Li Z, Minnich AJ, Babbush R, et al. Low rank representations for quantum simulation of electronic structure. npj Quantum Inf. 2021;7(1):83.
78. White SR. Density matrix formulation for quantum renormalization groups. Phys Rev Lett. 1992;69(19):2863–6.
79. White SR. Density-matrix algorithms for quantum renormalization groups. Phys Rev B. 1993;48(14):10345–56.
80. White SR, Martin RL. Ab initio quantum chemistry using the density matrix renormalization group. J Chem Phys. 1999;110(9):4127–30.
81. Chan GK-L, Head-Gordon M. Highly correlated calculations with a polynomial cost algorithm: A study of the density matrix renormalization group. J Chem Phys. 2002;116(11):4462–76.
82. Chan GK-L, Sharma S. The density matrix renormalization group in quantum chemistry. Annu Rev Phys Chem. 2011;62:465–81.
83. Baiardi A, Reiher M. The density matrix renormalization group in chemistry and molecular physics: recent developments and new challenges. J Chem Phys. 2020;152:4.
84. Huron B, Malrieu JP, Rancurel P. Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wavefunctions. J Chem Phys. 1973;58(12):5745–59.
85. Evangelista FA. Adaptive multiconfigurational wave functions. J Chem Phys. 2014;140(12):124114.
86. Tubman NM, Lee J, Takeshita TY, Martin Head-Gordon K, Whaley B. A deterministic alternative to the full configuration interaction quantum Monte Carlo method. J Chem Phys. 2016;145(4):044112.
87. Zhang N, Liu W, Hoffmann MR. Iterative configuration interaction with selection. J Chem Theory Comput. 2020;16(4):2296–316.
88. Zhang N, Liu W, Hoffmann MR. Further development of $iCIPT_2$ for strongly correlated electrons. J Chem Theory Comput. 2021;17(2):949–64.
89. Tubman NM, Freeman CD, Levine DS, Hait D, Head-Gordon M, Whaley KB. Modern approaches to exact diagonalization and selected configuration interaction with the adaptive sampling CI method. J Chem Theory Comput. 2020;16(4):2139–59.
90. Garniron Y, Scemama A, Giner E, Caffarel M, Loos P-F. Selected configuration interaction dressed by perturbation. J Chem Phys. 2018;149(6):064103.
91. Loos P-F, Damour Y, Scemama A. The performance of CIPSI on the ground state electronic energy of benzene. J Chem Phys. 2020;153(17):176101.
92. Damour Y, Véril M, Kossoski F, Caffarel M, Jacquemin D, Scemama A, et al. Accurate full configuration interaction correlation energy estimates for five- and six-membered rings. J Chem Phys. 2021;155(13):134104.
93. Eriksen JJ, Lipparini F, Gauss J. Virtual orbital many-body expansions: a possible route towards the full configuration interaction limit. J Phys Chem Lett. 2017;8(18):4633–9.
94. Eriksen JJ, Gauss J. Many-body expanded full configuration interaction. II. Strongly correlated regime. J Chem Theory Comput. 2019;15(9):4873–84.
95. Holmes AA, Changlani HJ, Umrigar CJ. Efficient heat-bath sampling in fock space. J Chem Theory Comput. 2016;12(4):1561–71.
96. Holmes AA, Tubman NM, Umrigar CJ. Heat-bath configuration interaction: an efficient selected configuration interaction algorithm inspired by heat-bath sampling. J Chem Theory Comput. 2016;12(8):3674–80.
97. Li J, Yang J. Downfolded configuration interaction for chemically accurate electron correlation. J Phys Chem Lett. 2022;13(43):10042–7.
98. Li X, Paldus J. Reduced multireference CCSD method: an effective approach to quasidegenerate states. J Chem Phys. 1997;107(16):6257–69.
99. Li X, Paldus J. Reduced multireference coupled cluster method with singles and doubles: perturbative corrections for triples. J Chem Phys. 2006;124(17):174101.
100. Enhua X, Uejima M, Ten-no SL. Full coupled-cluster reduction for accurate description of strong electron correlation. Phys Rev Lett. 2018;121(11):113001.

101. Enhua X, Uejima M, Ten-no SL. Towards near-exact solutions of molecular electronic structure: full coupled-cluster reduction with a second-order perturbative correction. J Phys Chem Lett. 2020;11(22):9775–80.
102. Emiliano Deustua J, Shen J, Piecuch P. Converging high-level coupled-cluster energetics by Monte Carlo sampling and moment expansions. Phys Rev Lett. 2017;119(22):223003.
103. Emiliano Deustua J, Magoulas I, Shen J, Piecuch P. Communication: approaching exact quantum chemistry by cluster analysis of full configuration interaction quantum Monte Carlo wave functions. J Chem Phys. 2018;149(15):151101.
104. Gururangan K, Piecuch P. Converging high-level coupled-cluster energetics via adaptive selection of excitation manifolds driven by moment expansions. J Chem Phys. 2023;159(8):084108.
105. Silvestrelli PL, Baroni S, Car R. Auxiliary-field quantum Monte Carlo calculations for systems with long-range repulsive interactions. Phys Rev Lett. 1993;71(8):1148–51.
106. Zhang S, Krakauer H. Quantum Monte Carlo method using phase-free random walks with slater determinants. Phys Rev Lett. 2003;90(13):136401.
107. Booth GH, Thom AJW, Alavi A. Fermion Monte Carlo without fixed nodes: a game of life, death, and annihilation in slater determinant space. J Chem Phys. 2009;131(5):054106.
108. Petruzielo FR, Holmes AA, Changlani HJ, Nightingale MP, Umrigar CJ. Semistochastic projector Monte Carlo method. Phys Rev Lett. 2012;109(23):230201.
109. Motta M, Zhang S. Ab initio computations of molecular systems by the auxiliary-field quantum Monte Carlo method. Wiley Interdiscip Rev Comput Mol Sci. 2018;8(5):e1364.
110. Surján PR, Szabados Á, Jeszenszki P, Zoboki T. Strongly orthogonal geminals: size-extensive and variational reference states. J Math Chem. 2012;50:534–51.
111. Johnson PA, Ayers PW, Limacher PA, de Baerdemacker S, van Neck D, Bultinck P. A size-consistent approach to strongly correlated systems using a generalized antisymmetrized product of nonorthogonal geminals. Comput Theor Chem. 2013;1003:101–13.
112. Tecmer P, Boguslawski K, Johnson PA, Limacher PA, Chan M, Verstraelen T, et al. Assessing the accuracy of new geminal-based approaches. J Phys Chem A. 2014;118(39):9058–68.
113. Henderson TM, Bulik IW, Stein T, Scuseria GE. Seniority-based coupled cluster theory. J Chem Phys. 2014;141(24):244104.
114. Kim TD, Miranda-Quintana RA, Richer M, Ayers PW. Flexible ansatz for N-body configuration interaction. Comput Theor Chem. 2021;1202:113187.
115. Eriksen JJ, Anderson TA, Emiliano Deustua J, Ghanem K, Hait D, Hoffmann MR, et al. The ground state electronic energy of benzene. J Phys Chem Lett. 2020;11(20):8922–9.
116. Hastings MB. Quasi-adiabatic continuation in gapped spin and fermion systems: Goldstone's theorem and flux periodicity. J Stat Mech Theory Exp. 2007;2007(5):P05010.
117. Eisert J, Cramer M, Plenio MB. *Colloquium*: Area laws for the entanglement entropy. Rev Mod Phys. 2010;82(1):277–306.
118. Werner H-J. Communication: multipole approximations of distant pair energies in local correlation methods with pair natural orbitals. J Chem Phys. 2016;145(20):201101.
119. Riplinger C, Pinski P, Becker U, Valeev EF, Neese F. Sparse maps—a systematic infrastructure for reduced-scaling electronic structure methods. II. Linear scaling domain based pair natural orbital coupled cluster theory. J Chem Phys. 2016;144(2):024109.
120. Bertoni C, Slipchenko LV, Misquitta AJ, Gordon MS. Multipole moments in the effective fragment potential method. J Phys Chem A. 2017;121(9):2056–67.
121. Mitrushchenkov AO, Fano G, Linguerri R, Palmieri P. On the importance of orbital localization in QC-DMRG calculations. Int J Quantum Chem. 2012;112(6):1606–19.
122. Izsák R, Ivanov AV, Blunt NS, Holzmann N, Neese F. Measuring electron correlation: the impact of symmetry and orbital transformations. J Chem Theory Comput. 2023;19(10):2703–20.
123. Tsuchimochi T, Scuseria GE. Strong correlations via constrained-pairing mean-field theory. J Chem Phys. 2009;131(12):121102.
124. Scuseria GE, Jiménez-Hoyos CA, Henderson TM, Samanta K, Ellis JK. Projected quasiparticle theory for molecular electronic structure. J Chem Phys. 2011;135(12):124108.
125. Kolmogorov AN. On the representation of continuous functions of several variables by superpositions of continuous functions of a smaller number of variables. Dokl. Akad. Nauk SSSR; 1956;108:179–82.
126. Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. Neural Netw. 1989;2(5):359–66.
127. Carleo G, Troyer M. Solving the quantum many-body problem with artificial neural networks. Science. 2017;355(6325):602–6.
128. Cai Z, Liu J. Approximating quantum many-body wave functions using artificial neural networks. Phys Rev B. 2018;97(3):035116.
129. Coe JP. Machine learning configuration interaction. J Chem Theory Comput. 2018;14(11):5739–49.
130. Coe JP. Machine learning configuration interaction for ab initio potential energy curves. J Chem Theory Comput. 2019;15(11):6179–89.
131. Behler J, Parrinello M. Generalized neural-network representation of high-dimensional potential-energy surfaces. Phys Rev Lett. 2007;98(14):146401.
132. Schütt KT, Arbabzadah F, Chmiela S, Müller KR, Tkatchenko A. Quantum-chemical insights from deep tensor neural networks. Nat Commun. 2017;8(1):13890.
133. Brockherde F, Vogt L, Li L, Tuckerman ME, Burke K, Müller K-R. Bypassing the Kohn-Sham equations with machine learning. Nat Commun. 2017;8(1):872.

134. Chmiela S, Tkatchenko A, Sauceda HE, Poltavsky I, Schütt KT, Müller K-R. Machine learning of accurate energy-conserving molecular force fields. Sci Adv. 2017;3(5):e1603015.
135. Welborn M, Cheng L, Miller TF. transferability in machine learning for electronic structure via the molecular orbital basis. J Chem Theory Comput. 2018;14(9):4772–9.
136. Unke OT, Meuwly M. A reactive, scalable, and transferable model for molecular energies from a neural network approach based on local information. J Chem Phys. 2018;148(24):241708.
137. Zhang L, Han J, Wang H, Car R, Weinan EJPRL. Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics. Phys Rev Lett. 2018;120(14):143001.
138. Unke OT, Meuwly M. PhysNet: a neural network for predicting energies, forces, dipole moments, and partial charges. J Chem Theory Comput. 2019;15(6):3678–93.
139. Qiao Z, Welborn M, Anandkumar A, Manby FR, Miller TF. OrbNet: deep learning for quantum chemistry using symmetry-adapted atomic-orbital features. J Chem Phys. 2020;153(12):124111.
140. Huang B, von Lilienfeld OA. Quantum machine learning using atom-in-molecule-based fragments selected on-the-fly. Nat Chem. 2020; 12(10):945–51.
141. Cheng L, Sun J, Miller TF. Accurate molecular-orbital-based machine learning energies via unsupervised clustering of chemical space. J Chem Theory Comput. 2022;18(8):4826–35.
142. Qiao Z, Christensen AS, Welborn M, Manby FR, Anandkumar A, Miller TF. Informing geometric deep learning with electronic interactions to accelerate quantum chemistry. Proc Natl Acad Sci U S A. 2022;119(31):e2205221119.
143. Ng W-P, Liang Q, Yang J. Low-data deep quantum chemical learning for accurate MP2 and coupled-cluster correlations. J Chem Theory Comput. 2023;19(16):5439–49.
144. Ko TW, Finkler JA, Goedecker S, Behler J. Accurate fourth-generation machine learning potentials by electrostatic embedding. J Chem Theory Comput. 2023;19(12):3567–79.
145. Huang B, von Lilienfeld OA. Ab initio machine learning in chemical compound space. Chem Rev. 2021;121(16):10001–36.
146. Peschel I. Special review: entanglement in solvable many-particle models. Braz J Phys. 2012;42:267–91.
147. Knizia G, Chan GK-L. Density matrix embedding: a simple alternative to dynamical mean-field theory. Phys Rev Lett. 2012;109(18): 186404.
148. Zheng B-X, Chan GK-L. Ground-state phase diagram of the square lattice Hubbard model from density matrix embedding theory. Phys Rev B. 2016;93(3):035126.
149. Knizia G, Chan GK-L. Density matrix embedding: a strong-coupling quantum embedding theory. J Chem Theory Comput. 2013;9(3): 1428–32.
150. Wouters S, Jiménez-Hoyos CA, Sun Q, Chan GK-L. A practical guide to density matrix embedding theory in quantum chemistry. J Chem Theory Comput. 2016;12(6):2706–19.
151. Pham HQ, Hermes MR, Gagliardi L. Periodic electronic structure calculations with the density matrix embedding theory. J Chem Theory Comput. 2019;16(1):130–40.
152. Cui Z-H, Zhu T, Chan GK-L. Efficient implementation of ab initio quantum embedding in periodic systems: density matrix embedding theory. J Chem Theory Comput. 2019;16(1):119–29.
153. Welborn M, Tsuchimochi T, van Voorhis T. Bootstrap embedding: an internally consistent fragment-based method. J Chem Phys. 2016; 145:7.
154. Ye H-Z, Ricke ND, Tran HK, van Voorhis T. Bootstrap embedding for molecules. J Chem Theory Comput. 2019;15(8):4497–506.
155. Ye H-Z, van Voorhis T. Atom-based bootstrap embedding for molecules. J Phys Chem Lett. 2019;10(20):6368–74.
156. Ye H-Z, Tran HK, van Voorhis T. Bootstrap embedding for large molecular systems. J Chem Theory Comput. 2020;16(8):5035–46.
157. Ye H-Z, Tran HK, van Voorhis T. Accurate electronic excitation energies in full-valence active space via bootstrap embedding. J Chem Theory Comput. 2021;17(6):3335–47.
158. Ye H-Z, Welborn M, Ricke ND, van Voorhis T. Incremental embedding: a density matrix embedding scheme for molecules. J Chem Phys. 2018;149(19):194108.
159. Tsuchimochi T, Welborn M, van Voorhis T. Density matrix embedding in an antisymmetrized geminal power bath. J Chem Phys. 2015; 143:2.
160. Hermes MR, Gagliardi L. Multiconfigurational self-consistent field theory with density matrix embedding: the localized active space self-consistent field method. J Chem Theory Comput. 2019;15(2):972–86.
161. Ditte M, Barborini M, Sandonas LM, Tkatchenko A. Molecules in environments: toward systematic quantum embedding of electrons and drude oscillators. Phys Rev Lett. 2023;131:228001.
162. Löwdin P-O. Quantum theory of many-particle systems. I. Physical interpretations by means of density matrices, natural spin-orbitals, and convergence problems in the method of configurational interaction. Phys Rev. 1955;97(6):1474–89.
163. Barr TL, Davidson ER. Nature of the configuration-interaction method inAb InitioCalculations. I. Ne ground state. Phys Rev A. 1970; 1(3):644–58.
164. Sosa C, Geertsen J, Trucks GW, Bartlett RJ, Franz JA. Selection of the reduced virtual space for correlated calculations. An application to the energy and dipole moment of H_2O. Chem Phys Lett. 1989;159(2-3):148–54.
165. Edmiston C, Krauss M. Pseudonatural orbitals as a basis for the superposition of configurations. I. He^{2+}. J Chem Phys. 1966;45(5): 1833–9.

166. Meyer W. PNO–CI studies of electron correlation effects. I. Configuration expansion by means of nonorthogonal orbitals, and application to the ground state and ionized states of methane. J Chem Phys. 1973;58(3):1017–35.
167. Taylor PR, Bacskay GB, Hush NS, Hurley AC. The coupled-pair approximation in a basis of independent-pair natural orbitals. Chem Phys Lett. 1976;41(3):444–9.
168. Ahlrichs R, Driessler F, Lischka H, Staemmler V, Kutzelnigg W. PNO–CI (pair natural orbital configuration interaction) and CEPA–PNO (coupled electron pair approximation with pair natural orbitals) calculations of molecular systems. II. The molecules BeH_2, BH, BH_3, CH_4, CH_{-3}, NH_3 (planar and pyramidal), H_2O, OH_{+3}, HF and the Ne atom. J Chem Phys. 1975;62(4):1235–47.
169. Ahlrichs R, Driessler F. Direct determination of pair natural orbitals. Theor Chim Acta. 1975;36:275–87.
170. Adamowicz L, Bartlett RJ. Optimized virtual orbital space for high-level correlated calculations. J Chem Phys. 1987;86(11):6314–24.
171. Pinski P, Riplinger C, Valeev EF, Neese F. Sparse maps—a systematic infrastructure for reduced-scaling electronic structure methods. I. An efficient and simple linear scaling local MP2 method that uses an intermediate basis of pair natural orbitals. J Chem Phys. 2015;143(3):034108.
172. Pavoševi'c F, Pinski P, Riplinger C, Neese F, Valeev EF. SparseMaps—A systematic infrastructure for reduced-scaling electronic structure methods. IV. Linear-scaling second-order explicitly correlated energy with pair natural orbitals. J Chem Phys. 2016;144(14):144109.
173. Pavoševi'c F, Peng C, Pinski P, Riplinger C, Neese F, Valeev EF. SparseMaps—A systematic infrastructure for reduced scaling electronic structure methods. V. Linear scaling explicitly correlated coupled-cluster method with pair natural orbitals. J Chem Phys. 2017;146(17):174108.
174. Saitow M, Becker U, Riplinger C, Valeev EF, Neese F. First UHF implementation of the incremental scheme for open-shell systems. J Chem Phys. 2017;146(16):164105.
175. Helmich B, Hättig C. Local pair natural orbitals for excited states. J Chem Phys. 2011;135(21):214106.
176. Helmich B, Hättig C. A pair natural orbital implementation of the coupled cluster model CC2 for excitation energies. J Chem Phys. 2013;139(8):084114.
177. Frank MS, Hättig C. A pair natural orbital based implementation of CCSD excitation energies within the framework of linear response theory. J Chem Phys. 2018;148(13):134102.
178. Peng C, Clement MC, Valeev EF. State-averaged pair natural orbitals for excited states: a route toward efficient equation of motion coupled-cluster. J Chem Theory Comput. 2018;14(11):5597–607.
179. Demel O, Pittner J, Neese F. A local pair natural orbital-based multireference mukherjee's coupled cluster method. J Chem Theory Comput. 2015;11(7):3104–14.
180. Brabec J, Lang J, Saitow M, Pittner J, Neese F, Demel O. Domain-based local pair natural orbital version of Mukherjee's state-specific coupled cluster method. J Chem Theory Comput. 2018;14(3):1370–82.
181. Lang J, Brabec J, Saitow M, Pittner J, Neese F, Demel O. Perturbative triples correction to domain-based local pair natural orbital variants of Mukherjee's state specific coupled cluster method. Phys Chem Chem Phys. 2019;21:5022–38.
182. Guo Y, Sivalingam K, Valeev EF, Neese F. SparseMaps—A systematic infrastructure for reduced-scaling electronic structure methods. III. Linear-scaling multireference domain-based pair natural orbital N-electron valence perturbation theory. J Chem Phys. 2016;144(9):094111.
183. Guo Y, Pavošević F, Sivalingam K, Becker U, Valeev EF, Neese F. SparseMaps—A systematic infrastructure for reduced-scaling electronic structure methods. VI. Linear-scaling explicitly correlated N-electron valence state perturbation theory with pair natural orbital. J Chem Phys. 2023;158(12):124120.
184. Menezes F, Kats D, Werner H-J. Local complete active space second-order perturbation theory using pair natural orbitals (PNO-CASPT2). J Chem Phys. 2016;145(12):124115.
185. Kats D, Werner H-J. Multi-state local complete active space second-order perturbation theory using pair natural orbitals (PNO-MS-CASPT2). J Chem Phys. 2019;150:21.
186. Krause C, Werner H-J. Comparison of explicitly correlated local coupled-cluster methods with various choices of virtual orbitals. Phys Chem Chem Phys. 2012;14:7591–604.
187. Schmitz G, Helmich B, Hättig C. A scaling PNO–MP2 method using a hybrid OSV–PNO approach with an iterative direct generation of OSVs†. Mol Phys. 2013;111(16-17):2463–76.
188. Ma Q, Werner H-J. Scalable electron correlation methods. 2. Parallel PNO-LMP_2-F_{12} with near linear scaling in the molecular size. J Chem Theory Comput. 2015;11(11):5291–304.
189. Schwilk M, Ma Q, Køppl C, Werner H-J. Scalable electron correlation methods. 3. Efficient and accurate parallel local coupled cluster with pair natural orbitals (PNO-LCCSD). J Chem Theory Comput. 2017;13(8):3650–75.
190. Ma Q, Schwilk M, Köppl C, Werner H-J. Scalable electron correlation methods. 4. Parallel explicitly correlated local coupled cluster with pair natural orbitals (PNO-LCCSD-F_{12}). J Chem Theory Comput. 2017;13(10):4871–96.
191. Ma Q, Werner H-J. Scalable electron correlation methods. 5. Parallel perturbative triples correction for explicitly correlated local coupled cluster with pair natural orbitals. J Chem Theory Comput. 2018;14(1):198–215.
192. Clement MC, Zhang J, Lewis CA, Yang C, Valeev EF. Optimized pair natural orbitals for the coupled cluster methods. J Chem Theory Comput. 2018;14(9):4581–9.
193. Kottmann JS, Bischoff FA, Valeev EF. Direct determination of optimal pair-natural orbitals in a real-space representation: the second-order Moller–Plesset energy. J Chem Phys. 2020;152(7):124120.

194. Kurashige Y, Yang J, Chan GK-L, Manby FR. Optimization of orbital-specific virtuals in local Møller-Plesset perturbation theory. J Chem Phys. 2012;136(12):124106.
195. Kállay M. Linear-scaling implementation of the direct random-phase approximation. J Chem Phys. 2015;142:20.
196. Nagy PR, Kállay M. Optimization of the linear-scaling local natural orbital CCSD(T) method: redundancy-free triples correction using Laplace transform. J Chem Phys. 2017;146(21).
197. Guo Y, Becker U, Neese F. Comparison and combination of "direct" and fragment based local correlation methods: cluster in molecules and domain based local pair natural orbital perturbation and coupled cluster theories. J Chem Phys. 2018;148:12.
198. Nagy PR, Samu G, Kállay M. Optimization of the linear-scaling local natural orbital CCSD(T) method: improved algorithm and benchmark applications. J Chem Theory Comput. 2018;14(8):4193–215.
199. Ni Z, Guo Y, Neese F, Li W, Li S. Cluster-in-molecule local correlation method with an accurate distant pair correction for large systems. J Chem Theory Comput. 2021;17(2):756–66.
200. Werner H-J, Hansen A. Accurate calculation of isomerization and conformational energies of larger molecules using explicitly correlated local coupled cluster methods in Molpro and ORCA. J Chem Theory Comput. 2023;19(20):7007–30.
201. Ni Z, Wang Y, Li W, Pulay P, Li S. Analytical energy gradients for the cluster-in-molecule MP_2 method and its application to geometry optimizations of large systems. J Chem Theory Comput. 2019;15(6):3623–34.
202. Pinski P, Neese F. Analytical gradient for the domain-based local pair natural orbital second order Møller-Plesset perturbation theory method (DLPNO-MP2). J Chem Phys. 2019;150:16.
203. Sharma B, Tran VA, Pongratz T, Galazzo L, Zhurko I, Bordignon E, et al. A joint venture of ab initio molecular dynamics, coupled cluster electronic structure methods, and liquid-state theory to compute accurate isotropic hyperfine constants of nitroxide probes in water. J Chem Theory Comput. 2021;17(10):6366–86.
204. Bloch C. Sur la théorie des perturbations des états liés. Nucl Phys. 1958;6:329–47.
205. Okubo S. Diagonalization of Hamiltonian and Tamm-Dancoff equation. Prog Theor Phys. 1954;12(5):603–22.
206. Cloizeaux JD. Extension d'une formule de Lagrange à des problèmes de valeurs propres. Nucl Phys. 1960;20:321–46.
207. Shavitt I, Redmon LT. Quasidegenerate perturbation theories. A canonical van Vleck formalism and its relationship to other approaches. J Chem Phys. 1980;73(11):5711–7.
208. Brandow BH. Linked-cluster expansions for the nuclear many-body problem. Rev Mod Phys. 1967;39(4):771–828.
209. Martensson AM. An iterative, numeric procedure to obtain pair functions applied to two-electron systems. J Phys B. 1979;12(24):3995–4012.
210. Durand P. Direct determination of effective Hamiltonians by wave-operator methods. I. General formalism. Phys Rev A. 1983;28(6):3184–92.
211. Maynau D, Ph Durand JP, Daudey JPM. Direct determination of effective Hamiltonians by wave-operator methods. II. Application to effective-spin interactions inπ-electron systems. Phys Rev A. 1983;28(6):3193–206.
212. Finley J, Malmqvist P-Å, Roos BO, Serrano-Andrés L. The multi-state CASPT2 method. Chem Phys Lett. 1998;288(2-4):299–306.
213. Evangelista FA. A driven similarity renormalization group approach to quantum many-body problems. J Chem Phys. 2014;141(5):054109.
214. Li C, Evangelista FA. Multireference driven similarity renormalization group: a second-order perturbative analysis. J Chem Theory Comput. 2015;11(5):2097–108.
215. Li C, Evangelista FA. Towards numerically robust multireference theories: The driven similarity renormalization group truncated to one- and two-body operators. J Chem Phys. 2016;144(16):164114.
216. Zhang T, Li C, Evangelista FA. Improving the efficiency of the multireference driven similarity renormalization group via sequential transformation, density fitting, and the noninteracting virtual orbital approximation. J Chem Theory Comput. 2019;15(8):4399–414.
217. Li C, Evangelista FA. Multireference theories of electron correlation based on the driven similarity renormalization group. Annu Rev Phys Chem. 2019;70:245–73.
218. Yanai T, Chan GK. Canonical transformation theory from extended normal ordering. J Chem Phys. 2007;127(10):104107.
219. Malrieu JP, Durand P, Daudey JP. Intermediate Hamiltonians as a new class of effective Hamiltonians. J Phys A: Math Gen. 1985;18(5):809–26.
220. Heully J-L, Malrieu J-P. Exploiting the flexibility of intermediate effective Hamiltonians. Chem Phys. 2009;356(1-3):76–85.
221. Caballol R, Malrieu J-P. Direct selected configuration interaction using a hole-particle formalism. Chem Phys Lett. 1992;188(5-6):543–9.
222. Daudey J-P, Heully J-L, Malrieu J-P. Size-consistent self-consistent truncated or selected configuration interaction. J Chem Phys. 1993;99(2):1240–54.
223. Pathak S, Lang L, Neese F. A dynamic correlation dressed complete active space method: theory, implementation, and preliminary applications. J Chem Phys. 2017;147:23.
224. Löwdin P-O. A note on the quantum-mechanical perturbation theory. J Chem Phys. 1951;19(11):1396–401.
225. Löwdin P-O. Studies in perturbation theory. IV. Solution of eigenvalue problem by projection operator formalism. J Math Phys. 1962;3(5):969–82.
226. Gershgorn Z, Shavitt I. An application of perturbation theory ideas in configuration interaction calculations. Int J Quantum Chem. 1968;2(6):751–9.
227. Nitzsche LE, Davidson ER. A perturbation theory calculation on the $\sim$* state of formamide. J Chem Phys. 1978;68(7):3103–9.

228. Davidson ER, McMurchie LE, Day SJ. The BK method: application to methylene. J Chem Phys. 1981;74(10):5491–6.
229. Rawlings DC, Davidson ER. The Rayleigh—Schrödinger BK method applied to the lower electronic states of pyrrole. Chem Phys Lett. 1983;98(5):424–7.
230. Kozlowski PM, Davidson ER. Construction of open shell perturbation theory invariant with respect to orbital degeneracy. Chem Phys Lett. 1994;226(5-6):440–6.
231. Staroverov VN, Davidson ER. The reduced model space method in multireference second-order perturbation theory. Chem Phys Lett. 1998;296(5-6):435–44.
232. Roos B. A new method for large-scale Cl calculations. Chem Phys Lett. 1972;15(2):153–9.
233. Löwdin P-O. Studies in perturbation theory. J Mol Spectrosc. 1964;13(1-4):326–37.
234. Bartlett RJ, Brändas EJ. Reduced partitioning procedure in configuration interaction studies. I. Ground states. J Chem Phys. 1972; 56(11):5467–77.
235. Bartlett RJ, Brändas EJ. Reduced partitioning procedure in configuration interaction studies. II. Excited states. J Chem Phys. 1973; 59(4):2032–42.
236. Windom ZW, Bartlett RJ. On the iterative diagonalization of matrices in quantum chemistry: reconciling preconditioner design with Brillouin–Wigner perturbation theory. J Chem Phys. 2023;158(13):134107.
237. Choi JH. Partitioning method and Van Vleck's perturbation theory. Prog Theor Phys. 1975;53(6):1641–51.
238. Rupp M, Tkatchenko A, Müller K-R, Von Lilienfeld OA. Fast and accurate modeling of molecular atomization energies with machine learning. Phys Rev Lett. 2012;108(5):058301.
239. Tsubaki M, Mizoguchi T. Quantum deep descriptor: physically informed transfer learning from small molecules to polymers. J Chem Theory Comput. 2021;17(12):7814–21.
240. Margraf JT, Reuter K. Pure non-local machine-learned density functional theory for electron correlation. Nat Commun. 2021;12(1):1–7.
241. Dick S, Fernandez-Serra M. Machine learning accurate exchange and correlation functionals of the electronic density. Nat Commun. 2020;11(1):1–10.
242. Kirkpatrick J, McMorrow B, Turban DHP, Gaunt AL, Spencer JS, Matthews AGDG, et al. Pushing the frontiers of density functionals by solving the fractional electron problem. Science. 2021;374(6573):1385–9.
243. Fias S, Heidar-Zadeh F, Geerlings P, Ayers PW. Chemical transferability of functional groups follows from the nearsightedness of electronic matter. Proc Natl Acad Sci U S A. 2017;114(44):11633–8.
244. Bartók AP, Payne MC, Kondor R, Csányi G. Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons. Phys Rev Lett. 2010;104(13):136403.
245. Bartók AP, Kondor R, Csányi G. On representing chemical environments. Phys Rev B. 2013;87(18):184115.
246. Hansen K, Biegler F, Ramakrishnan R, Pronobis W, von Lilienfeld OA, Muller K-R, et al. Machine learning predictions of molecular properties: accurate many-body potentials and nonlocality in chemical space. J Phys Chem Lett. 2015;6(12):2326–31.
247. Faber FA, Christensen AS, Bing Huang O, von Lilienfeld A. Alchemical and structural distribution based representation for universal quantum machine learning. J Chem Phys. 2018;148(24):241717.
248. Zhang L, Han J, Wang H, Saidi W, Car R, Weinan E. End-to-end symmetry preserving inter-atomic potential energy model for finite and extended systems. Adv Neural Inf Process Sys. 2018;31;4436−4446.
249. Schütt KT, Sauceda HE, Kindermans P-J, Tkatchenko A, Müller K-R. SchNet – a deep learning architecture for molecules and materials. J Chem Phys. 2018;148(24):241722.
250. Pronobis W, Tkatchenko A, Müller K-R. Many-body descriptors for predicting molecular properties with machine learning: analysis of pairwise and three-body interactions in molecules. J Chem Theory Comput. 2018;14(6):2991–3003.
251. Gilmer J, Schoenholz SS, Riley PF, Vinyals O, Dahl GE. Neural message passing for quantum chemistry. In International conference on machine learning, PMLR; 2017. p. 1263–1272.
252. Grisafi A, Fabrizio A, Meyer B, Wilkins DM, Corminboeuf C, Ceriotti M. Transferable machine-learning model of the electron density. ACS Cent Sci. 2018;5(1):57–64.
253. Chmiela S, Sauceda HE, Müller K-R, Tkatchenko A. Towards exact molecular dynamics simulations with machine-learned force fields. Nat Commun. 2018;9(1):1–10.
254. Meyer B, Guillot B, Ruiz-Lopez MF, Genoni A. Libraries of extremely localized molecular orbitals. 1. Model molecules approximation and molecular orbitals transferability. J Chem Theory Comput. 2016;12(3):1052–67.
255. Nudejima T, Ikabata Y, Seino J, Yoshikawa T, Nakai H. Machine-learned electron correlation model based on correlation energy density at complete basis set limit. J Chem Phys. 2019;151(2):024104.
256. Han R, Rodríguez-Mayorga M, Luber S. A machine learning approach for MP_2 correlation energies and its application to organic compounds. J Chem Theory Comput. 2021;17(2):777–90.
257. Cheng L, Welborn M, Christensen AS, Miller TF. Interaction of hydrogen with actinide dioxide (111) surfaces. J Chem Phys. 2019; 150(13):131103.
258. Cheng L, Kovachki NB, Welborn M, Miller TF. Regression clustering for improved accuracy and training costs with molecular-orbital-based machine learning. J Chem Theory Comput. 2019;15(12):6668–77.
259. Chen Y, Zhang L, Wang H, Weinan E. Ground state energy functional with Hartree-Fock efficiency and chemical accuracy. J Phys Chem A. 2020;124(35):7155–65.

260. Husch T, Sun J, Cheng L, Lee SJR, Miller TF III. Improved accuracy and transferability of molecular-orbital-based machine learning: Organics, transition-metal complexes, non-covalent interactions, and transition states. J Chem Phys. 2021;154(6):064108.
261. Karandashev K, von Lilienfeld OA. An orbital-based representation for accurate quantum machine learning. J Chem Phys. 2022; 156(11):114101.
262. Margraf JT, Reuter K. Making the coupled cluster correlation energy machine-learnable. J Phys Chem A. 2018;122(30):6343–8.
263. Townsend J, Vogiatzis KD. Data-driven acceleration of the coupled-cluster singles and doubles iterative solver. J Phys Chem A. 2019; 10(14):4129–35.
264. Peyton BG, Briggs C, DCunha R, Margraf JT, Crawford TD. Machine-learning coupled cluster properties through a density tensor representation. J Phys Chem A. 2020;124(23):4861–71.
265. Townsend J, Vogiatzis KD. Transferable MP_2-based machine learning for accurate coupled-cluster energies. J Chem Theory Comput. 2020;16(12):7453–61.
266. Ma Y, Tsao D, Shum H-Y. On the principles of parsimony and self-consistency for the emergence of intelligence. Front Inform Technol Electron Eng. 2022;23(9):1298–323.
267. Damle A, Lin L, Ying L. Compressed representation of Kohn-Sham orbitals via selected columns of the density matrix. J Chem Theory Comput. 2015;11(4):1463–9.
268. Jianfeng L, Ying L. Compression of the electron repulsion integral tensor in tensor hypercontraction format with cubic scaling cost. J Comput Phys. 2015;302:329–35.
269. Woolfe F, Liberty E, Rokhlin V, Tygert M. A fast randomized algorithm for the approximation of matrices. Appl Comput Harmon Anal. 2008;25(3):335–66.
270. Frank MS, Schmitz G, Hättig C. The PNO–MP gradient and its application to molecular geometry optimisations. Mol Phys. 2017; 115(3):343–56.
271. Lin HW, Tegmark M, Rolnick D. Why Does Deep and Cheap Learning Work So Well? J Stat Phys. 2017;168:1223–47.
272. Zhu L, Li J, Yang J, Au-Yeung HY. Cross dehydrogenative C-O coupling catalysed by a catenane-coordinated copper(i). Chem Sci. 2020;11(48):13008–14.
273. Gawinkowski S, Walewski Ł, Vdovin A, Slenczka A, Rols S, Johnson MR, et al. Vibrations and hydrogen bonding in porphycene. Phys Chem Chem Phys. 2012;14(16):5489–503.
274. Litman Y, Richardson JO, Kumagai T, Rossi M. Elucidating the nuclear quantum dynamics of intramolecular double hydrogen transfer in porphycene. J Am Chem Soc. 2019;141(6):2526–34.
275. Takenaka N, Kitamura Y, Koyano Y, Nagaoka M. The number-adaptive multiscale QM/MM molecular dynamics simulation: application to liquid water. Chem Phys Lett. 2012;524:56–61.
276. Hao H, Pestana LR, Qian J, Liu M, Xu Q, Head-Gordon T. Chemical transformations and transport phenomena at interfaces. Wiley Interdiscip Rev Comput Mol Sci. 2022;13(2):e1639.
277. Zhu C, Kumar M, Zhong J, Li L, Francisco JS, Zeng XC. New mechanistic pathways for Criegee-Water chemistry at the air/water interface. J Am Chem Soc. 2016;138(35):11164–9.
278. Liu J, Liu Y, Yang J, Zeng XC, He X. Directional proton transfer in the reaction of the simplest Criegee intermediate with water involving the formation of transient H_3O. J Phys Chem Lett. 2021;12(13):3379–86.
279. Zhong J, Kumar M, Zhu CQ, Francisco JS, Zeng XC. Surprising stability of larger criegee intermediates on aqueous interfaces. Angew Chem Int Ed. 2017;56(27):7740–4.

How to cite this article: Yang J. Making quantum chemistry compressive and expressive: Toward practical ab-initio simulation. WIREs Comput Mol Sci. 2024;14(2):e1706. https://doi.org/10.1002/wcms.1706