

Delta-Augmented Subsystem Density Functional Theory: A Study Across Diverse Systems

Michela Pauletti^{a,b}, Marcella Iannuzzi^a, and Vladimir V. Rybkin^{a,c,*}

Abstract: In this study, we expand upon and benchmark the Kim–Gordon method (KG), a subsystem density functional theory (DFT) approach appended with a machine-learned correction to compensate for errors in the kinetic energy term and thereby match Kohn–Sham (KS) DFT accuracy. This correction is obtained through ‘delta-learning’ based on KS-DFT data. The method promises sampling of configurations for condensed molecular systems at the Kohn–Sham DFT level of accuracy at a fraction of the computational cost. Despite encouraging results for liquid water, it was not obvious whether the scheme had more general appeal. In this work, we show that the approach allows for a broad range of applications. In particular, we successfully apply it to complex molecular liquids, such as bulk ammonia and methanol. As a bonus, the correction trained on the bulk KS data is applicable to clusters, illustrating its transferability. By focusing on ‘delta-learning’—predicting small corrections rather than full Kohn–Sham (KS) energies and forces—we significantly reduce the required training data. This approach, especially when combined with linear-scaling self-consistent field (LS-SCF) techniques, establishes the method as a highly efficient computational tool for molecular dynamics.

Keywords: Delta-learning · Kim–Gordon method · Machine-learning potentials · Molecular liquids · Subsystem DFT



Michela Pauletti is currently an AI Software Engineer at the University Hospital of Würzburg. She is a chemist and holds a PhD in Quantum Chemistry from the University of Zurich. She has previously conducted research at ETH Zurich and the University of Basel in the field of mathematical modeling in biology.



Marcella Iannuzzi is an adjoint Professor and group leader in computational Materials Science at the department of Chemistry of the University of Zurich. Her research centres on *ab initio* simulations to unravel fundamental processes at interfaces, in condensed phases, and under excitation, with applications to sustainable energy and materials. She is one of the core developers of the CP2K, an open source package for atomistic simulations for extended systems.



Vladimir V. Rybkin is the team leader of the *ab initio* spectroscopy team at HQS Quantum Simulations GmbH, Karlsruhe, Germany. He holds a PhD in Theoretical Chemistry from the University of Oslo. After being a postdoc at ETH Zurich and the University of Zurich and holding an SNSF Ambizione Fellowship at the latter he joined the quantum computing industry.

His research is focused on *ab initio* molecular dynamic simulations, many body electronic structure theory and quantum computing use cases. He has contributed to the CP2K and Dalton software packages.

1. Introduction

Classical force fields are often the method of choice for long-time and large-scale molecular dynamics (MD) simulations.^[1–4] Their computational efficiency allows for extensive exploration of phase space and the study of structural, thermodynamic, and dynamical properties in multiatomic systems. For liquids, this includes diffusion and transport coefficients, solvation dynamics, and interfacial phenomena such as wetting, friction, and surface tension, which are crucial for understanding both equilibrium and nonequilibrium behaviour of matter.^[5–9] Nevertheless, most force fields remain empirical: they are typically parametrized for specific systems and may perform poorly when electronic effects such as polarization, charge transfer, or bond rearrangements play a significant role.

Electronic-structure methods based on quantum mechanics provide a more fundamental description, as they include the relevant interactions from first principles. In principle, these methods are preferable whenever transferable accuracy is required, for example in reactive liquids, complex mixtures, or charged systems. However, the computational cost of approaches such as density functional theory (DFT) severely limits the accessible length and time scales. To address this challenge, a wide range of approximate electronic-structure schemes and algorithmic strategies have been proposed, including linear-scaling techniques that exploit the locality of electronic structures.^[10–12]

*Correspondence: V. V. Rybkin, E-mail: vladimir.rybkin@quantumsimulations.de

^aPhysical Chemistry Institute, University of Zurich, Winterthurerstrasse 190, Zurich, Switzerland; ^bCurrent address: Interventional and Experimental Endoscopy (InExEn), Department of Internal Medicine 2, University Hospital Würzburg, Würzburg, Germany; ^cCurrent address: HQS Quantum Simulations GmbH, Rintheimer Strasse 23, 76131 Karlsruhe, Germany

Subsystem DFT offers an interesting strategy for complexity reduction by partitioning the electronic-structure problem into smaller subsystems that can be treated independently.^[13,14] This idea is particularly natural for molecular liquids, where the system can be divided into molecular fragments. Among the various formulations of subsystem DFT, the Kim–Gordon (KG) scheme^[15–18] has been studied. In the original formulation,^[19] the energies of subsystems are minimized independently, with a correction term for the non-additive exchange–correlation (XC) and kinetic energy contributions.

Implementations such as that in the CP2K program package,^[20,21] which employ the Gaussian and plane-wave (GPW) formalism,^[22] can be coupled with linear-scaling algorithms and thereby treat large systems. Despite these advantages, the applicability of the KG scheme to *ab initio* molecular dynamics (AIMD) remains limited. First, technical bottlenecks arise because the kinetic energy functional and other density-dependent terms must be represented on real-space grids and Fourier-transformed for each subsystem, a step that becomes prohibitively costly in very large systems. Second, at the level of orbital-free density functionals, KG performs adequately for weakly interacting systems such as gases, but fails to capture the correct intermolecular interactions in condensed phases, leading to incorrect sampling.^[14,23] These limitations are both technical and fundamental: the former concerns scaling with respect to grid operations, while the latter reflects the difficulty of constructing universally accurate non-additive kinetic energy functionals. Several remedies have been proposed, ranging from non-decomposable corrections^[24] to orbital-dependent functionals.^[25]

A pragmatic alternative is to introduce data-driven corrections that can approximate the missing contributions at low cost. Recent progress in machine learning (ML) has shown that potential energy surfaces can be accurately represented by models trained on large datasets of electronic-structure calculations.^[26] While ML force fields are typically designed to reproduce the entire potential energy surface, in the KG context the correction can be restricted to the small but critical non-additive contributions. This retains the molecular electronic structure of the subsystems while recovering Kohn–Sham accuracy for energies and forces at a fraction of the computational price. Despite recent progress^[27] in universal^[28] and liquid-specific ML-based force fields (such as MB-Pol^[29]), we argue that direct electronic-structure information remains essential because it provides access to properties that cannot be captured by purely mechanistic force-field descriptions. Further opportunities are opened by using ML density functionals^[30] within subsystem DFT frameworks, such as the one developed in this work. In this work, we extend the ML-based KG approach to molecular liquids. Our aim is to assess its performance across different condensed-phase systems and to compare it with established subsystem methods. In particular, we benchmark against the absolutely localized molecular orbital (ALMO) scheme,^[31–33] which provides a controlled approximation to the full-system electronic structure at the DFT or Hartree–Fock level and represents a well-established $O(N)$ method for large-scale simulations.^[21,34] By combining linear-scaling algorithms with ML corrections, we seek to demonstrate a route towards accurate and efficient quantum-based molecular dynamics of liquids.

2. Subsystem DFT: the Kim–Gordon Method

The starting point of orbital-free density functional theory (OF-DFT)^[35,36] is the Kohn–Sham (KS) energy functional:

$$E[\rho] = T_s[\rho] + \int \rho(\mathbf{r}) v_{\text{ext}}(\mathbf{r}) d\mathbf{r} + E_H[\rho] + E_{xc}[\rho], \quad (1)$$

where ρ is the electron density, v_{ext} is the external potential, $E_{xc}[\rho]$ is the exchange–correlation energy functional, and $E_H[\rho]$ is the Hartree functional:

$$E_H[\rho] = \frac{1}{2} \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'. \quad (2)$$

In conventional Kohn–Sham DFT, the non-interacting kinetic energy functional $T_s[\rho]$ is evaluated from auxiliary orbitals. By contrast, in OF-DFT, this quantity is approximated directly as an explicit functional of the electron density,

$$T_s[\rho] \approx T_s^{\text{OF}}[\rho], \quad (3)$$

so that the total energy becomes

$$E[\rho] \approx T_s^{\text{OF}}[\rho] + \int \rho(\mathbf{r}) v_{\text{ext}}(\mathbf{r}) d\mathbf{r} + E_H[\rho] + E_{xc}[\rho]. \quad (4)$$

Thus, a key aspect of the OF-DFT scheme is the choice of the kinetic energy functional, in particular its explicit dependence on the electron density ρ .^[37,38] Accurate results are, however, often hampered by the lack of sufficiently accurate approximations for this quantity.^[39–41] $E_{xc}[\rho]$ is also always approximated and is another major error source in practical DFT; this issue is not being addressed here.

The Kim–Gordon (KG) theory combines these two formalisms and exploits fragmentation into subsystems. The total electronic density ρ_{tot} is computed as the sum over the subsystem densities ρ_A , which are constructed by restricting the expansion to the basis-set functions $\{\phi_{\alpha,A}(r)\}$ belonging to the subsystem itself:

$$\rho_{\text{tot}}(r) = \sum_A \rho_A(r) = \sum_{\alpha_A\beta_A} P_{\alpha_A\beta_A} \phi_{\alpha_A}(r)\phi_{\beta_A}(r), \quad (5)$$

where \mathbf{P} is a one-electron reduced density matrix.

The total energy is then obtained as

$$E_{\text{tot}} = E_{\text{OF}}[\rho_{\text{tot}}] - \sum_A E_{\text{OF}}[\rho_A] + \sum_A E_{\text{KS}}[\rho_A], \quad (6)$$

where the orbital-free functional is defined as

$$E_{\text{OF}}[\rho] = T_s^{\text{OF}}[\rho] + E_{\text{ext}}[\rho] + E_{\text{hxc}}[\rho], \quad (7)$$

with

$$E_{\text{hxc}}[\rho] = E_H[\rho] + E_{xc}[\rho]. \quad (8)$$

Due to the partitioning, the density matrix \mathbf{P} is by construction block-diagonal, and the same holds for the KS matrix, since only matrix elements of the form $P_{\alpha_A\beta_A}$ are taken into account. Consequently, the orbital kinetic energy contribution in the subsystem KS part is additive,

$$T_s[P] = \sum_A T_s[P_A]. \quad (9)$$

Assuming that the external energy functional is linear in the density,

$$E_{\text{ext}}[\rho_{\text{tot}}] = \sum_A E_{\text{ext}}[\rho_A], \quad (10)$$

which holds for local potentials and can be extended to pseudopotentials with non-local parts,^[19] the total energy becomes

$$E_{\text{tot}} = \sum_A (T_s[P_A] + E_{\text{ext}}[\rho_A]) + E_{\text{hxc}}[\rho_{\text{tot}}] + T_s^{\text{OF,nadd}}[\rho_{\text{tot}}, \{\rho_A\}], \quad (11)$$

where the non-additive kinetic energy is defined as

$$T_s^{\text{OF,nadd}}[\rho_{\text{tot}}, \{\rho_A\}] = T_s^{\text{OF}}[\rho_{\text{tot}}] - \sum_A T_s^{\text{OF}}[\rho_A]. \quad (12)$$

Within the GPW formalism, the evaluation of $T_s^{\text{OF}}[\rho_A]$ and the associated contribution to the KS matrix is performed by locating the density of each subsystem on the real-space grid and carrying out numerical integration to obtain $[T_{s,A}^{\text{OF}}]_{\alpha\beta}$. This operation is approximately N_A times more expensive than locating ρ_{tot} as a whole and integrating the corresponding terms.

2.1 Potentials

To avoid integrating the kinetic energy functional for each subsystem, the non-additive term can be replaced by a potential form. Introducing the local kinetic energy density functional $\mu[\rho]$, the non-additive term (for simplicity denoted as T_{nadd}) can be rewritten as

$$\begin{aligned} T_{\text{nadd}} &= \int \rho_{\text{tot}}(r) \mu[\rho_{\text{tot}}] dr - \sum_A \int \rho_A(r) \mu[\rho_A] dr \\ &= \sum_A \left[\int \rho_A(r) (\mu[\rho_{\text{tot}}] - \mu[\rho_A]) dr \right]. \end{aligned} \quad (13)$$

Expanding in a Taylor series truncated at the linear term, we obtain the approximation

$$\mu[\rho_{\text{tot}}] - \mu[\rho_A] \approx \frac{\partial \mu[\rho_A]}{\partial \rho} \sum_{B \neq A} \rho_B, \quad (14)$$

which gives

$$T_{\text{nadd}} = \sum_A \sum_{B \neq A} \int \mu'[\rho_A] \rho_A(r) \rho_B(r) dr. \quad (15)$$

A further approximation of the derivative functional in atomic contributions is

$$\mu'[\rho_A] \rho_A = V^K[\rho_A] \approx \sum_{a \in A} V_a^K(R_a). \quad (16)$$

In the simplest Thomas–Fermi approximation, which is exact for homogeneous systems, the kinetic energy is proportional to $\rho^{5/3}$, which leads to a model for atomic local potentials of the form

$$V_a^K(R_a) = N_a \rho_a^{2/3}, \quad (17)$$

where ρ_a is the model atomic density and N_a is a fitted parameter. In this case the atomic density is approximated with a Gaussian. The potential has been obtained directly using the CP2K code,^[42] starting from the Goedecker–Teter–Hutter pseudopotentials,^[43,44] and the fitting region has been confined within the Bohr radius for each atom type.

As discussed above, our goal is to apply this method to molecular liquids, taking advantage of the clear definition of subsystems and the localization of electronic density. Intermolecular interactions and the resulting polarization or charge-transfer effects can be particularly significant in hydrogen-bonded networks, such as liquid water, or in other polar liquids. To gain a deeper understanding of the kinetic energy-functional term and to explore how it can be described more accurately for optimal agreement with the Kohn–Sham reference, we performed a series of tests on small molecular systems.

2.2 Delta-Learning Training

The missing term in the KG scheme arises mainly from the repulsive forces resulting from the orthogonalization of wave functions on neighbouring subsystems. The relatively short range of this interaction makes it suitable for description using a ML potential.^[45–49] By interpolating energies and forces over a set of reference structures, it is possible to generate interatomic potentials that accurately reproduce the potential energy surfaces (PESs) obtained from various quantum-chemistry methods. A key requirement is thorough exploration of phase space to ensure representative sampling of configuration space, which in turn guarantees the accuracy of the resulting potential. One well-established ML approach for potential energy surfaces is the Behler–Parrinello neural-network (BPNNP) scheme,^[50,51] which maps atomic coordinates onto a set of two- and three-body symmetry functions^[51] that capture correlations between atoms in the neighbourhood of a central atom. The more optimized these symmetry functions are, the more accurately energies and forces are reproduced compared with the reference data. The two-body term is represented by

$$G_2^i = \sum_j e^{-\eta(R_{ij}-R_s)^2} f_c(R_{ij}), \quad (18)$$

where η and R_s define the width and position of the Gaussian with respect to the atom centre and $f_c(R_{ij})$ is the switching function that ensures a smooth drop to zero at the cutoff value r_c . The three-body term is

$$\begin{aligned} G_3^i &= 2^{1-\zeta} \sum_j \sum_{k \neq j} (1 + \lambda \cos \theta_{ijk})^\zeta \\ &\times e^{-\eta(R_{ij}^2 + R_{ik}^2 + R_{jk}^2)} f_c(R_{ij}) f_c(R_{ik}) f_c(R_{jk}), \end{aligned} \quad (19)$$

where ζ , η , and λ define the shape of the angular part of the functions. Lastly, the cutoff is defined as

$$f_c(R_{ij}) = \tanh^3 \left(1 - \frac{R_{ij}}{r_c} \right) \quad (20)$$

for $R_{ij} \geq r_c$, and

$$f_c(R_{ij}) = 0 \quad (21)$$

for $R_{ij} \geq r_c$.

In some cases, it is possible to add a potential that reproduces only a fraction of the total energy of the system. For instance, a specific term in the energy expression can be captured using the so-called delta-learning scheme. The training set for this approach is represented by

$$E_{\text{delta}} = E_{\text{target}} - E_{\text{starting}} \quad (22)$$

and

$$F_{i,\text{delta}} = F_{i,\text{target}} - F_{i,\text{starting}}, \quad (23)$$

where i runs over all atoms, ‘target’ denotes the reference method to be reproduced, and ‘starting’ denotes the method to be corrected. From these differences, the algorithm can generate an empirical potential that is computationally inexpensive. When added to P-KG, this potential effectively approximates the KS potential.

The machine-learning potentials are trained, in this work, using configurations extracted from PBE-KS simulations, with varying system sizes and temperatures to ensure broad phase-space coverage. In some cases, the initially generated potentials fail to reproduce the KS results, and high-energy configurations from these attempts are incorporated to help the model recognize unfeasible states. Typically, 10% of the total structures are reserved for testing, while the remaining data are used for training. This procedure is applied consistently throughout this work.

3. Evaluating the Kim–Gordon Method: Application Across Multiple Systems and Accuracy Analysis

3.1 Methods and Computational Setup

All calculations in this study were performed using the CP2K program package.^[21,42] The Kohn–Sham calculations were performed in combination with a generalized-gradient-approximation (GGA) exchange–correlation functional, supplemented by dispersion corrections.^[52,53] Core electrons were treated using norm-conserving pseudopotentials of the Goedecker–Teter–Hutter type,^[43] and valence electrons were expanded in atom-centred Gaussian basis sets optimized for condensed-phase simulations.^[54] The GPW formalism was used to efficiently solve the Kohn–Sham DFT equations, with plane waves serving as an auxiliary basis. Different energy cutoffs were used for each system; the specific values are provided in the corresponding sections. Periodic boundary conditions were applied for bulk phases, while finite systems such as molecular clusters were treated in vacuum without periodicity. In periodic calculations, only the Γ -point was considered within the supercell approach.

ML corrections were implemented using a delta-learning scheme, primarily in combination with the Kim–Gordon method and with linear-scaling SCF techniques, including purification-based approaches.^[55] Molecular-dynamic simulations were used to generate training and validation data in the relevant ensembles, with temperature controlled by Nosé–Hoover thermostats. Delta-learning models were trained on data derived from DFT trajectories, with systematic variation of numerical settings

such as basis-set completeness, plane-wave cutoffs, and SCF accuracy to assess robustness.

The delta-learning KG approach was evaluated through comparisons with reference KS calculations using root mean square errors (RMSEs) in energies and forces, as well as structural metrics such as radial distribution functions (RDFs). Transferability of models was assessed by applying bulk-trained models to molecular clusters of varying size and topology. Additional validations were conducted on randomly sampled geometries to ensure generalizability.

Simulations and analyses were performed using the trained potentials, based on the available n2p2 outputs^[56,57] (the original training configurations were not accessible to the authors at the time of the manuscript preparation, but all relevant log files and reproducibility documentation were available), together with the cross-validation data and symmetry-function definitions. Complete benchmarking data are provided in the Supporting Information.

3.2 Bulk Ammonia

We begin by describing a molecular-dynamics simulation of liquid ammonia at 230 K. Our goal is to investigate how different factors, including training and linear-scaling setup, can impact performance. We performed a 4 ps MD simulation in the NPT (constant-number, pressure, and temperature) ensemble at 230 K and at atmospheric pressure, using full Kohn–Sham DFT on a periodic cubic cell containing 64 ammonia molecules. The system was equilibrated to a cell size of 13.6830 Å. In our delta-learning approach, we have two main components: one represented by the DFT setup and the other by the potential correction. It is important to consider the impact of the calculation settings on training. To address this question, we used a dataset comprising 500 configurations from a 20 ps simulation at 230 K, 80 configurations from a simulation using an initial ML potential, and 180 configurations from two simulations at 210 K and 240 K. All calculations were performed at the Kohn–Sham level. Approximately 10% of the total number of structures were used for testing after training. We conducted five training sessions: first, we increased the cutoff energy from 800 Ry to 1000 Ry; second, we tested various linear-scaling configurations by adjusting the filter settings (loose, tight, and very tight); and finally, we completely removed the linear-scaling. Fig. 1 shows that the training does not appear to be

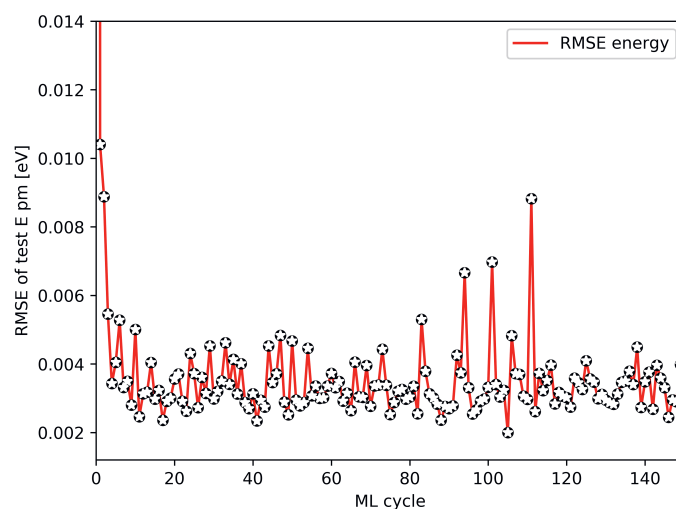


Fig. 1. Learning curve for the training of 64 ammonia molecules. The y-axis reports the RMSE of the energy per molecule. Three setups were used on the same dataset: the red line represents delta training performed with a cutoff of 800 Ry and a linear-scaling filter of 10^{-8} ; the black dots correspond to delta training with a cutoff of 1000 Ry and a linear-scaling filter of 10^{-8} ; and the white stars indicate delta training with a cutoff of 1000 Ry and a linear-scaling filter of 10^{-10} .

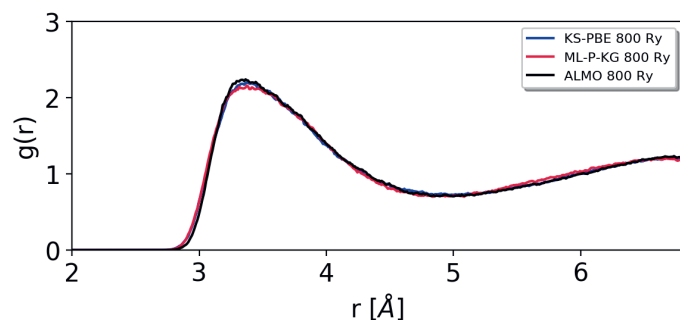


Fig. 2. N–N pair distribution function calculated for KS with a cutoff 800 Ry (blue line), ML–P–KG with a cutoff equal to 800 Ry (red line), and ALMO–SCF with a cutoff equal to 800 Ry (black line). The MD sampling has been carried out at 230 K.

significantly sensitive to small changes in forces and energies due to different numerical accuracy. This suggests that delta-learning potentials can be applied to various DFT setups without introducing significant errors.

To ensure consistency in this study, we simulated each system using its respective potential and obtained 50 ps trajectories for each. All calculations were carried out with the CP2K program package.^[21,42] We used dual-space Goedecker–Teter–Hutter pseudopotentials^[43,44] and MOLOPT TZV2P basis sets.^[54] Periodic boundary conditions were applied to the cell mentioned above. We used the PBE exchange–correlation functional^[52] augmented by Grimme’s D3 dispersion corrections^[53] for all the KS, ALMO, and KG calculations. Molecular-dynamic simulations were carried out within the canonical ensemble (constant number of particles, volume, and temperature; NVT) at 230 K with a time step of 0.5 fs using the Nosé–Hoover thermostat. The KG method was always applied together with a linear-scaling scheme based on trace-resetting purification of an effective Hamiltonian (TRS4).^[55]

The N–N ammonia peak is reported in Fig. 2. Here we compare KS–PBE (blue line), ML–P–KG (red line), and ALMO–SCF (black line). All curves show good agreement; however, the KG method shows a slightly lower intensity, while ALMO shows a slightly higher intensity in the first peak compared with KS. In the Supporting Information (SI), we also present the other RDFs from all our simulations, for example those obtained using higher cutoffs and different least-squares algorithms, as well as the cross-validation results related to the training of this system. In the final section, we examine the accuracy of the calculations, discussing the factors that influenced reliability and the implications for the results presented above. It is known from previous literature that linear-scaling applied to molecular dynamics can lead to conservative energy drifts.^[58,59] This behaviour depends solely on the linear-scaling parameters and is not related to the KG method or the training procedure. In fact, we observed that without linear-scaling, no drift was detected, and modifications to the potential did not affect the energy trend.

Fig. 3 displays the conservative energy monitored during the simulations with different calculation setups. It is clear that linear-scaling SCF always leads to a drift that depends on how loose the filters are. However, the orange line shows that without LS there are no evident energy drifts. It is important to note that reference scaling tests for the KG method, including linear-scaling behaviour, can be found in Ref. [17]. Its computational cost depends on the number of molecules, the number of processors, and the chosen DFT-level accuracy (*e.g.* basis set and cutoff).

3.3 Bulk Methanol

It is desirable to extend the method to more complex liquids. In this section, we performed full machine learning and KG del-

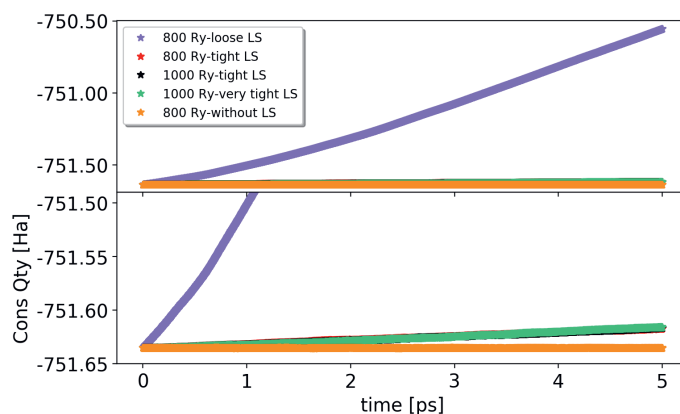


Fig. 3. Conservative energy plotted during the simulations with different calculation setups. Linear-scaling SCF always leads to a drift that depends on how loose the filters are. The orange line shows that without LS there are no evident energy drifts.

ta-learning for a system containing 32 methanol molecules and compared the results. We aim to understand whether the delta-learning scheme provides a meaningful advantage over full-range training. As in the previous case, the program package CP2K^[21,42] was used, employing dual-space Goedecker–Teter–Hutter pseudopotentials^[43,44] and MOLOPT TZV2P basis sets.^[54] Periodic boundary conditions were always applied to the cell with side length 13.0562 Å. We used the PBE XC functional^[52] augmented by Grimme’s D3 dispersion corrections^[53] for all the KS, ALMO, and KG calculations. Molecular-dynamic simulations were carried out within the NVT ensemble at 298.15 K with a time step of 0.5 fs using the Nosé–Hoover thermostat. The KG method was always applied together with a linear-scaling scheme based on trace-resetting purification of an effective Hamiltonian (TRS4).^[55] We extracted 600 configurations from a 30 ps KS–PBE simulation to perform the machine-learning part, and approximately 10% of them were used for testing after training.

Fig. 4 shows that lower values of the root mean square error are reached in the case of delta-learning compared with full machine-learning training. This dataset can be considered relatively small in size, as it comprises only 600 configurations obtained from a single temperature. It was specifically chosen to test the efficiency of the delta-learning approach. Despite both training

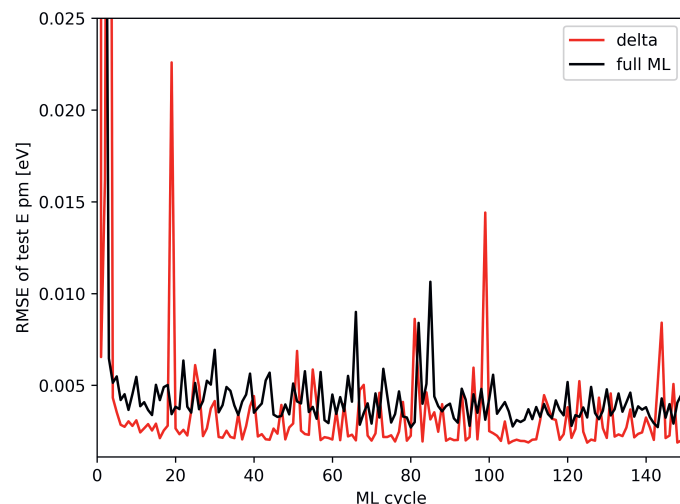


Fig. 4. Learning curve relative to the training of 32 methanol molecules. The RMSE of the energy per molecule is shown; the red line represents the delta training, and the black line displays the full ML training.

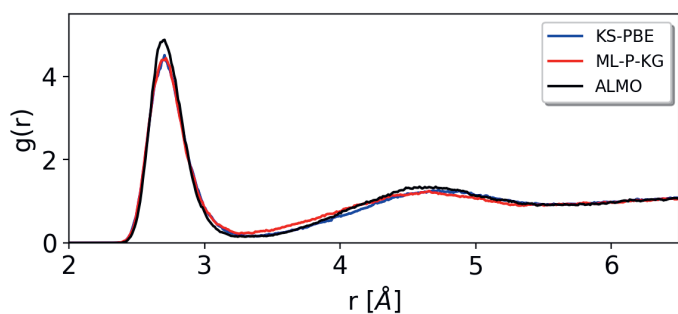


Fig. 5. O–O pair distribution function calculated for KS (blue line), ML–P–KG (red line), and ALMO–SCF (black line). The MD sampling has been carried out at 298.15 K.

approaches yielding acceptable RMSE values, we observed that the full machine-learning potential was not yet stable enough for running MD simulations, likely due to the limited phase space explored. However, employing the exact same training parameters in our delta scheme, the resulting potential was very stable and suitable for performing simulations.

The radial distribution function can be calculated from the MD simulation, and the O–O curve is reported in Fig. 5. All the methods are in agreement, although the first peak is slightly overestimated by the ALMO method and there are some small differences between the first minimum and the second shell for all the methods compared with KS. Despite the limited size of the dataset, the results are consistent and demonstrate satisfactory accuracy. The Supporting Information provides the machine-learning inputs and cross-validation details associated with the training of this system.

3.4 Ammonia and Methanol Clusters

Lastly, we want to assess to what extent the potential generated for the bulk (see Fig. 1) can be transferred to clusters and whether the resulting errors remain within an acceptable range. To achieve this, we started from four ammonia structures: dimer, trimer, tetramer, and pentamer. The DFT calculation setup was the same as described in the previous section, except that in this case

periodic boundary conditions were not applied. We generated a ‘cluster potential’ starting from 500 ammonia tetramer configurations extracted randomly from bulk ammonia simulations. Here, we compare the RMSEs of energies and forces across the different systems.

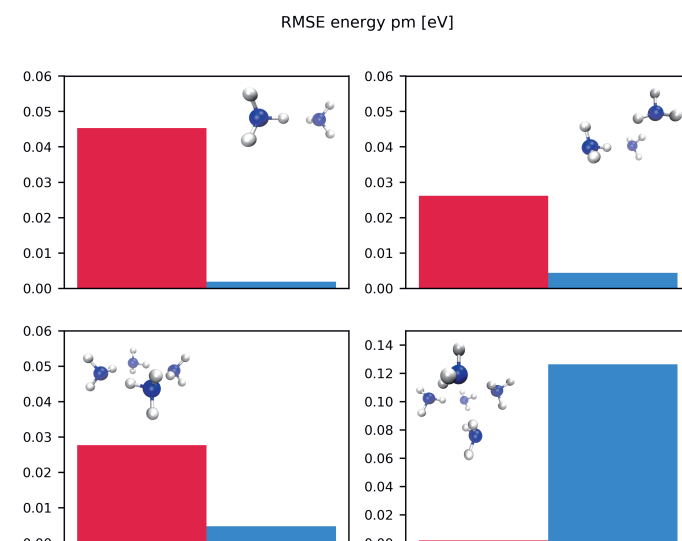


Fig. 6. RMSE of energy per molecule in eV for four different cluster systems (dimer, trimer, tetramer, and pentamer). All structures were previously optimized at the Kohn–Sham level. The red bars represent the RMSE calculated by comparing KS and KG corrected by the bulk potential, whereas the blue bars compare KS and KG corrected by the cluster potential.

Figs. 6 and 7 report the RMSE of energy and forces for ammonia clusters. The potential trained on tetramers (blue bars) shows better performance than the bulk potential (red bars) for small clusters (2–3 molecules); however, as expected, it is less accurate than the bulk potential for larger clusters such as the pentamer. Because the bulk potential was trained on many molecules under varying conditions, it can reproduce a wider range of systems compared with the cluster potential. Even in the case of dimers and trimers, the resulting RMSE remains within an acceptable range.

RMSE forces [eV/Å]

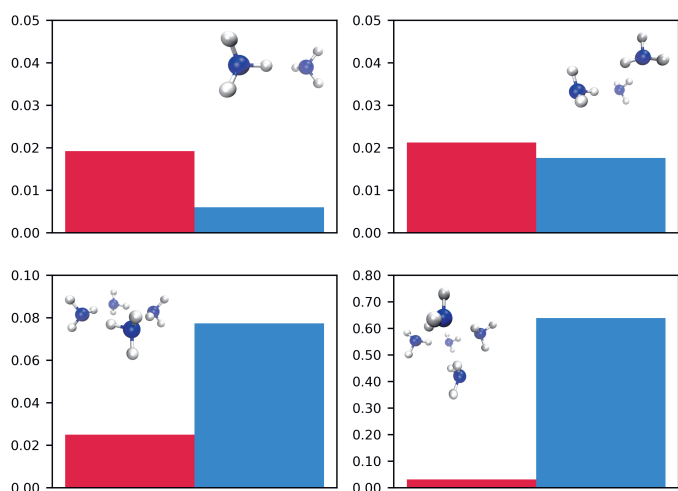


Fig. 7. RMSE of forces per atom in eV/Å for four different cluster systems (dimer, trimer, tetramer, and pentamer). All structures were previously optimized at the Kohn–Sham level. The red bars represent the RMSE calculated by comparing KS and KG corrected by the bulk potential, whereas the blue bars compare KS and KG corrected by the cluster potential.

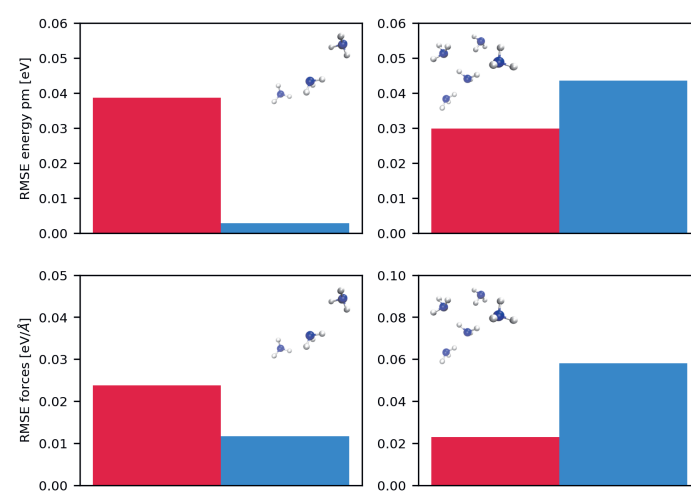


Fig. 8. RMSE of energy per molecule in eV (top line) and RMSE of forces per atom in eV/Å (bottom line) for trimers and pentamers extracted randomly from bulk ammonia. The total number of configurations is 100. The red bars represent the RMSE calculated by comparing KS and KG corrected by the bulk potential, whereas the blue bars compare KS and KG corrected by the cluster potential.

Table 1. RMSE of energy and forces for methanol trimers. 'pm' indicates that the values are given per molecule.

Quantity	RMSE value
Energy (pm) [eV]	0.0352(7)
Forces x [eV/Å]	0.0278(5)
Forces y [eV/Å]	0.0307(7)
Forces z [eV/Å]	0.0341(9)

Because the analysis based on a single equilibrium configuration for each structure is not statistically representative, we extended it to a further 100 trimer and 100 pentamer configurations extracted randomly from a KS-PBE trajectory of bulk ammonia. Fig. 8 shows in the top line the RMSE related to energies and in the bottom line the RMSE related to the forces of trimers and pentamers. The tested potentials are the same as those discussed in Figs. 6 and 7. The inaccuracy of the cluster potential (blue bars) in predicting forces becomes increasingly evident as the number of molecules increases. In contrast, the bulk potential (red bars) shows lower accuracy in smaller systems. As expected, even with an increase in the number of configurations, the results reported in Figs. 6, 7, and 8 remain in agreement. The same procedure was repeated for the methanol trimers extracted from the bulk system, in this case employing only the bulk potential described previously (see Fig. 5).

The RMSE can be considered to be in the same range as in the case of ammonia clusters.

4. Conclusions

In this work, we explored the efficiency and accuracy of the Kim–Gordon method within the framework of subsystem DFT, with a focus on the application of a delta-learning correction for molecular liquids. The KG method is an alternative to conventional Kohn–Sham DFT because it leverages subsystem partitioning and local kinetic energy approximations to reduce computational costs while maintaining accuracy.

We demonstrated that the application of delta-learning significantly improves the accuracy of KG calculations by effectively correcting for missing energy contributions due to subsystem partitioning. The training strategy used for the delta-learning part proved to be robust in different DFT setups, as shown by the benchmark results on bulk ammonia and bulk methanol. In particular, our results indicate that delta-learning-based corrections can successfully capture short-range repulsive interactions, enhancing the accuracy of the KG method, originally applied to the gas phase, in liquid-phase simulations.

Through molecular-dynamic simulations, we established that the KG approach yields structural and dynamical properties that are in good agreement with full Kohn–Sham DFT calculations. The radial distribution functions obtained for bulk ammonia and methanol highlighted the validity of the KG method. Additionally, our analysis of conservative energy drifts in linear-scaling calculations shows the importance of fine-tuning the linear-scaling filter threshold to ensure stability and reliability in long-time molecular-dynamics simulations.

In general, the KG method combined with delta-learning remains a machine-learning-based approach that requires training. However, it offers a significantly more computationally efficient alternative to traditional machine-learning potentials in quantum chemistry. This makes it well suited for large-scale simulations of molecular liquids. Future work will focus on further refining the

kinetic energy functionals used in the subsystem approach and exploring its applicability to more complex chemical systems, including heterogeneous interfaces and reactive environments.

Supporting Information

Additional details, including symmetry-function definitions, cross-validation data, and derived molecular-dynamics simulations and computed properties, are provided in the Supporting Information.

Acknowledgements

The authors thank Prof. Dr. Jürg Hutter (University of Zürich) for support and valuable advice and Prof. Rustam Z. Khaliullin for introducing the codes and algorithms used for ALMO. The authors thank the Swiss National Science Foundation (SNF) for financial support under grant No. 200021_162432 (MP). This work has also been supported by the Swiss National Science Foundation (SNSF) in the form of Ambizione grant No. PZ00P2_174227 (VVR). The authors thank the University of Zürich for providing generous computational resources under Project ID uzh1 at the Swiss National Supercomputing Centre.

Received: March 8, 2026

- [1] B. R. Brooks, R. E. Brucocoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus, *J. Comput. Chem.* **1983**, *4*, 187, <https://doi.org/10.1002/jcc.540040211>.
- [2] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179, <https://doi.org/10.1021/ja00124a002>.
- [3] C. Oostenbrink, A. Villa, A. E. Mark, W. F. van Gunsteren, *J. Comput. Chem.* **2004**, *25*, 1656, <https://doi.org/10.1002/jcc.20090>.
- [4] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc.* **1996**, *118*, 11225, <https://doi.org/10.1021/ja9621760>.
- [5] A. Torres-Carbajal, V. M. Trejos, L. A. Nicasio-Collazo, *J. Chem. Phys.* **2018**, *149*, 144104, <https://doi.org/10.1063/1.5031132>.
- [6] D. Keffer, University of Tennessee **2001**.
- [7] A. P. Willard, D. Chandler, *J. Phys. Chem. B* **2010**, *114*, 1954, <https://doi.org/10.1021/jp909219k>.
- [8] Y. Nagata, T. Ohto, M. Bonn, T. D. Kühne, *J. Chem. Phys.* **2016**, *144*, 204705, <https://doi.org/10.1063/1.4951710>.
- [9] M. Pauletti, V. V. Rybkin, M. Iannuzzi, *J. Phys.: Condens. Matter* **2021**, *34*, 044003, <https://doi.org/10.1088/1361-648X/ac2e8f>.
- [10] J. VandeVondele, U. Borstnik, J. Hutter, *J. Chem. Theory Comput.* **2012**, *8*, 3565, <https://doi.org/10.1021/ct200897x>.
- [11] S. Goedecker, *Rev. Mod. Phys.* **1999**, *71*, 1085, <https://doi.org/10.1103/RevModPhys.71.1085>.
- [12] S. Goedecker, G. E. Scuseria, *Comput. Sci. Eng.* **2003**, *5*, 14, <https://doi.org/10.1109/MCISE.2003.1208637>.
- [13] P. Cortona, *Phys. Rev. B* **1991**, *44*, 8454, <https://doi.org/10.1103/PhysRevB.44.8454>.
- [14] C. R. Jacob, J. Neugebauer, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2024**, *14*, e1700, <https://doi.org/10.1002/wcms.1700>.
- [15] R. G. Gordon, Y. S. Kim, *J. Chem. Phys.* **1972**, *56*, 3122, <https://doi.org/10.1063/1.1677649>.
- [16] Y. S. Kim, R. G. Gordon, *J. Chem. Phys.* **1974**, *60*, 1842, <https://doi.org/10.1063/1.1681283>.
- [17] M. Pauletti, V. V. Rybkin, M. Iannuzzi, *J. Chem. Theory Comput.* **2021**, *17*, 6423, <https://doi.org/10.1021/acs.jctc.1c00592>.
- [18] C. Herrero, M. Pauletti, G. Tocci, M. Iannuzzi, L. Joly, *Proc. Natl. Acad. Sci. U. S. A.* **2022**, *119*, e2121641119, <https://doi.org/10.1073/pnas.2121641119>.
- [19] M. Iannuzzi, B. Kirchner, J. Hutter, *Chem. Phys. Lett.* **2006**, *421*, 16, <https://doi.org/10.1016/j.cplett.2005.09.113>.
- [20] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, J. Hutter, *Comput. Phys. Commun.* **2005**, *167*, 103, <https://doi.org/10.1016/j.cpc.2004.12.014>.
- [21] T. D. Kühne, M. Iannuzzi, M. Del Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann, D. Golze, J. Wilhelm, S. Chulkov, M. H. Bani-Hashemian, V. Weber, U. Borstnik, M. Taillefumier, A. S. Jakobovits, A. Lazzaro, H. Pabst, T. Müller, R. Schade, M. Guidon, S. Andermatt, N. Holmberg, G. K. Schenter, A. Hehn, A. Bussy, F. Belleflamme, G. Tabacchi, A. Glöß, M. Lass, I. Bethune, C. J. Mundy, C. Plessl, M. Watkins, J. VandeVondele, M. Krack, J. Hutter, *J. Chem. Phys.* **2020**, *152*, 194103, <https://doi.org/10.1063/5.0007045>.
- [22] G. Lippert, J. Hutter, M. Parrinello, *Mol. Phys.* **1997**, *92*, 477, <https://doi.org/10.1080/00268979709482119>.
- [23] V. L. Lignères, E. A. Carter, 'Handbook of Materials Modeling' **2005**, 137, https://doi.org/10.1007/978-1-4020-3286-8_9.

- [24] C. R. Jacob, S. M. Beyhan, L. Visscher, *J. Chem. Phys.* **2007**, *126*, 234116, <https://doi.org/10.1063/1.2743013>.
- [25] L. S. Eitelhuber, D. G. Artiukhin, *J. Chem. Phys.* **2025**, *162*, 5, <https://doi.org/10.1063/5.0241361>.
- [26] J. Behler, *J. Chem. Phys.* **2016**, *145*, 170901, <https://doi.org/10.1063/1.4971792>.
- [27] J. Xia, Y. Zhang, B. Jiang, *Chem. Soc. Rev.* **2025**, *54*, 4790, <https://doi.org/10.1039/D5CS00104H>.
- [28] D. P. Kovács, I. Batatia, E. S. Arany, G. Csányi, *J. Chem. Phys.* **2023**, *159*, 044109, <https://doi.org/10.1063/5.0155322>.
- [29] X. Zhu, M. Riera, E. F. Bull-Vulpe, F. Paesani, *J. Chem. Theory Comput.* **2023**, *19*, 3551, <https://doi.org/10.1021/acs.jctc.3c00326>.
- [30] G. Luise, C.-W. Huang, T. Vogels, D. P. Kooi, S. Ehlert, S. Lanius, K. J. H. Giesbertz, A. Karton, D. Gunceler, M. Stanley, W. P. Bruinsma, L. Huang, X. Wei, J. Garrido Torres, A. Katbashev, R. Chavez Zavaleta, B. Máté, S.-O. Kaba, R. Sordillo, Y. Chen, D. B. Williams-Young, C. M. Bishop, J. Hermann, R. van den Berg, P. Gori-Giorgi, *arXiv* **2025**, 2506.14665, <https://doi.org/10.48550/arXiv.2506.14665>
- [31] R. Z. Khaliullin, M. Head-Gordon, A. T. Bell, *J. Chem. Phys.* **2006**, *124*, 204105, <https://doi.org/10.1063/1.2191500>.
- [32] R. Z. Khaliullin, E. A. Cobar, R. C. Lochan, A. T. Bell, M. Head-Gordon, *J. Phys. Chem. A* **2007**, *111*, 8753, <https://doi.org/10.1021/jp073685z>.
- [33] R. Z. Khaliullin, A. T. Bell, M. Head-Gordon, *J. Chem. Phys.* **2008**, *128*, 184112, <https://doi.org/10.1063/1.2912041>.
- [34] G. Galli, *Curr. Opin. Solid State Mater. Sci.* **1996**, *1*, 864, [https://doi.org/10.1016/S1359-0286\(96\)80114-8](https://doi.org/10.1016/S1359-0286(96)80114-8).
- [35] E. Smargiassi, P. A. Madden, *Phys. Rev. B* **1994**, *49*, 5220, <https://doi.org/10.1103/PhysRevB.49.5220>.
- [36] V. V. Karasiev, S. B. Trickey, *Adv. Quantum Chem.* **2015**, *71*, 221, <https://doi.org/10.1016/bs.aiq.2015.02.004>.
- [37] T. A. Wesolowski, *J. Chem. Phys.* **1997**, *106*, 8516, <https://doi.org/10.1063/1.473907>.
- [38] T. A. Wesolowski, H. Chermette, J. Weber, *J. Chem. Phys.* **1996**, *105*, 9182, <https://doi.org/10.1063/1.472823>.
- [39] E. Sim, J. Larkin, K. Burke, C. W. Bock, *J. Chem. Phys.* **2003**, *118*, 8140, <https://doi.org/10.1063/1.1565316>.
- [40] T. Pope, W. Hofer, *arXiv* **2019**, 1907.06433.
- [41] D. G. Artiukhin, C. R. Jacob, J. Neugebauer, *J. Chem. Phys.* **2015**, *142*, 234101, <https://doi.org/10.1063/1.4922429>.
- [42] J. Hutter, M. Iannuzzi, F. Schiffmann, J. VandeVondele, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 15, <https://doi.org/10.1002/wcms.1159>.
- [43] S. Goedecker, M. Teter, J. Hutter, *Phys. Rev. B* **1996**, *54*, 1703, <https://doi.org/10.1103/PhysRevB.54.1703>.
- [44] C. Hartwigsen, S. Goedecker, J. Hutter, *Phys. Rev. B* **1998**, *58*, 3641, <https://doi.org/10.1103/PhysRevB.58.3641>.
- [45] P. O. Dral, *J. Phys. Chem. Lett.* **2020**, *11*, 2336, <https://doi.org/10.1021/acs.jpclett.9b03664>.
- [46] N. Artrith, K. T. Butler, F.-X. Coudert, S. Han, O. Isayev, A. Jain, A. Walsh, *Nat. Chem.* **2021**, *13*, 505, <https://doi.org/10.1038/s41557-021-00716-z>.
- [47] P. M. Pflüger, F. Glorius, *Angew. Chem. Int. Ed.* **2020**, *59*, 18860, <https://doi.org/10.1002/anie.202008366>.
- [48] H. Wang, L. Zhang, J. Han, E. Weinan, *Comput. Phys. Commun.* **2018**, *228*, 178, <https://doi.org/10.1016/j.cpc.2018.03.016>.
- [49] L. Zhang, J. Han, H. Wang, R. Car, E. Weinan, *Phys. Rev. Lett.* **2018**, *120*, 143001, <https://doi.org/10.1103/PhysRevLett.120.143001>.
- [50] J. Behler, M. Parrinello, *Phys. Rev. Lett.* **2007**, *98*, 146401, <https://doi.org/10.1103/PhysRevLett.98.146401>.
- [51] J. Behler, *J. Chem. Phys.* **2011**, *134*, 074106, <https://doi.org/10.1063/1.3553717>.
- [52] M. Ernzerhof, G. E. Scuseria, *J. Chem. Phys.* **1999**, *110*, 5029, <https://doi.org/10.1063/1.478401>.
- [53] E. Caldeweyher, C. Bannwarth, S. Grimme, *J. Chem. Phys.* **2017**, *147*, 034112, <https://doi.org/10.1063/1.4993215>.
- [54] J. VandeVondele, J. Hutter, *J. Chem. Phys.* **2007**, *127*, 114105, <https://doi.org/10.1063/1.2770708>.
- [55] A. M. N. Niklasson, C. J. Tymczak, M. Challacombe, *J. Chem. Phys.* **2003**, *118*, 8611, <https://doi.org/10.1063/1.1559913>.
- [56] A. Singraber, T. Morawietz, J. Behler, C. Dellago, *J. Chem. Theory Comput.* **2019**, *15*, 3075, <https://doi.org/10.1021/acs.jctc.8b01092>.
- [57] A. Singraber, J. Behler, C. Dellago, *J. Chem. Theory Comput.* **2019**, *15*, 1827, <https://doi.org/10.1021/acs.jctc.8b00770>.
- [58] J.-L. Fattebert, F. Gygi, *Phys. Rev. B* **2006**, *73*, 115124, <https://doi.org/10.1103/PhysRevB.73.115124>.
- [59] M. J. Cawkwell, A. M. N. Niklasson, *J. Chem. Phys.* **2012**, *137*, 134105, <https://doi.org/10.1063/1.4755991>.

License and Terms



This is an Open Access article under the terms of the Creative Commons Attribution License CC BY 4.0. The material may not be used for commercial purposes.

The license is subject to the CHIMIA terms and conditions: (<https://chimia.ch/chimia/about>).

The definitive version of this article is the electronic one that can be found at <https://doi.org/10.2533/chimia.2026.319>