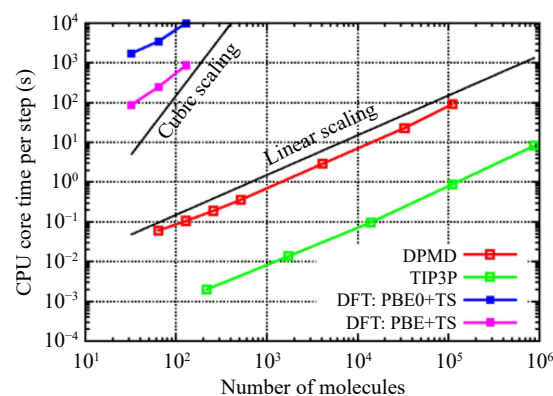




chemical reactions.

Computational simulations have become indispensable for probing atomic-scale details. Classical molecular dynamics (MD), which uses empirical force fields, can simulate large systems over long timescales, but it often lacks accuracy because empirical force fields frequently fail to capture quantum effects. In aqueous systems, for instance, MD may poorly describe the structure and dynamics of the hydrogen-bond network [9], or proton transfer processes [10, 11], or electronic structure [12], leading to systematic errors in predicting properties like ion hydration free energies and reaction pathways. In contrast, *ab initio* molecular dynamics (AIMD) relies on quantum mechanical principles to describe electronic interactions with high accuracy. Based on AIMD, researchers have achieved accurate predictions and gained deeper insights into fundamental aqueous properties, such as the solvation structure of ion [12–16], self-diffusion of water [17–19], and proton transfer dynamics in aqueous solutions [20, 21]. Despite its precision, computational cost of AIMD is extremely high, typically limiting simulations to small systems ( $\sim 100$  water molecules) and short timescales (picoseconds), which is insufficient for studying rare events or large-scale collective dynamics. This long-standing accuracy-efficiency trade-off has been a major obstacle to reliable simulations of complex aqueous systems.

The emergence of machine learning potentials (MLPs) has revolutionized molecular simulations by bridging the gap between quantum accuracy and classical efficiency. MLPs employ machine learning algorithms, primarily neural networks, to construct high-dimensional potential energy surfaces (PES) from *ab initio* data. A fundamental requirement for MLPs is invariance under translation, rotation, and permutation of identical atoms, achieved via rotationally invariant descriptors (e.g., atom-centered symmetry functions or bispectrum components) [22]. Another key is decomposing the energy into local atomic contributions, enabling near-linear scaling. The Behler-Parrinello neural network (BPNN) [23] pioneered the use of atom-centered symmetry functions to achieve rotational invariance and locality, enabling linear-scaling simulations of materials such as bulk silicon. Building on this, the Gaussian Approximation Potential (GAP) [24] introduced Gaussian process regression with Smooth Overlap of Atomic Positions (SOAP) descriptors, improving the faithfulness of the MLP [25]. To further enhance accuracy and automation, the Deep Potential Molecular Dynamics (DPMD) model [26] employed deep neural networks to automatically learn optimal features from DFT data, achieving near-linear scaling and DFT-level accuracy for large systems. Around the same time, SchNet [27] introduced continuous-filter convolutional networks on molecular graphs, demonstrating that graph-based representations could learn molecular properties with high accuracy and transferability. More recently,



**Fig. 1** Computational cost per MD step as a function of Number of molecules. DPMD (deep potential) exhibits near-linear scaling  $O(N)$ , in contrast to the cubic scaling  $O(N^3)$  of DFT-based methods (PBE+TS and PBE0+TS) and the constant cost of classical TIP3P. Reproduced with permission from Ref. [26]. Copyright © 2018 American Physical Society.

equivariant architectures have improved data efficiency and accuracy by incorporating physical symmetries directly into the network structure: NequIP [28] uses E(3)-equivariant graph neural networks, while MACE [29] introduces higher-order message passing network to capture complex many-body interactions with fewer parameters. These models, together with advances in handling long-range electrostatics (e.g., DPLR), provide a versatile toolkit for simulating aqueous systems with *ab initio* accuracy over extended scales.

The computational advantage of MLPs over AIMD is most evident in their scaling behavior. AIMD, typically based on density functional theory (DFT), exhibits a cubic scaling  $O(N^3)$  with number of molecules  $N$  due to the electronic structure problem, while well-designed MLPs achieve near-linear scaling  $O(N)$  (see Fig. 1) [26]. This reduction is achieved because the MLP inference step involves evaluating local atomic environments within a cutoff radius, bypassing the expensive global electronic minimization. Once trained, MLPs can predict energies and forces with *ab initio* accuracy at a computational cost that is orders of magnitude less expensive than direct AIMD, enabling microsecond-scale simulations of systems containing tens of thousands of atoms. Techniques like active learning further allow MLPs to autonomously explore configuration space [30, 31], enhancing their robustness and generalizability. Due to the high accuracy and efficiency, MLPs have emerged as a powerful and efficient tool for simulating complex aqueous systems, effectively overcoming the longstanding accuracy-efficiency trade-off.

This review aims to provide a comprehensive overview of the applications of MLPs in complicated aqueous systems, highlighting recent breakthroughs, methodological innovations, and persistent challenges. We will begin by exploring how MLPs have advanced our understanding

of structural and dynamical properties in bulk water and aqueous solutions, such as ion hydration and water diffusion. Then we will discuss insights gained from machine learning molecular dynamics (MLMD) on confined aqueous systems, including water in confined systems and at interfaces, where finite-size effects and surface interactions dominate. Finally, we will examine the role of MLPs in probing reactive processes, such as proton transport, acid-base reactions, and catalytic mechanisms at interfaces. Although DFT serves as the primary reference for most MLPs, its intrinsic accuracy limitations — discussed further in the Outlook — must be acknowledged, with higher-level methods like quantum Monte Carlo (QMC) and CCSD(T) providing benchmark references.

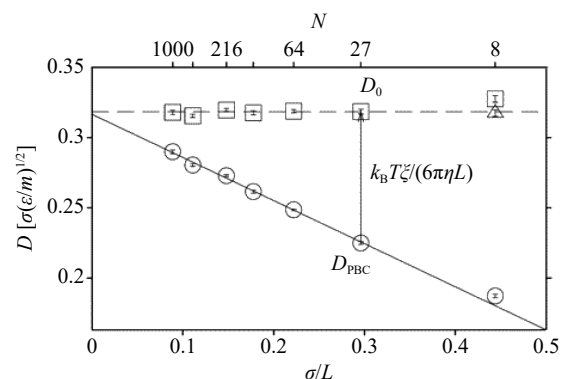
## 2 Structural and dynamical properties of bulk systems

Building on the motivation for MLPs in aqueous systems, this section delves into the fundamental structural and dynamical properties of bulk water and aqueous solutions, which serve as a baseline for understanding complex interactions. MLP methods like DPMD enable *ab initio* level insights into hydrogen bonding networks, reorientation dynamics, and solute-solvent effects over nanosecond timescales, overcoming limitations of traditional simulations.

### 2.1 Bulk water

Bulk water constitutes the fundamental system for understanding aqueous environments, where its structural and dynamical properties govern behaviors in more complex solutions. Prior to the advent of MLPs, AIMD established a foundational understanding of the microscopic structure of liquid water. AIMD studies demonstrated that advanced density functionals, including hybrid functionals with exact exchange and van der Waals corrections, as well as the meta-GGA, could accurately predict the structure, thermodynamics, and electronic properties of water [17–19].

However, constrained by computational cost, AIMD simulations are typically limited to small system sizes and short timescales. An insufficiently large system size may give rise to finite-size effects, while an insufficiently long simulation time results in poor statistical convergence of dynamical convergence. Early benchmark studies revealed that for basic structural properties such as the O-O radial distribution function (RDF), simulations with about 64 water molecules were sufficient to eliminate finite-size effect [32, 32]. While modeling the X-ray absorption spectra (XAS) of pure water requires a system size of 128 water molecules to achieve numerically converged results [33]. For dynamic properties, the limi-

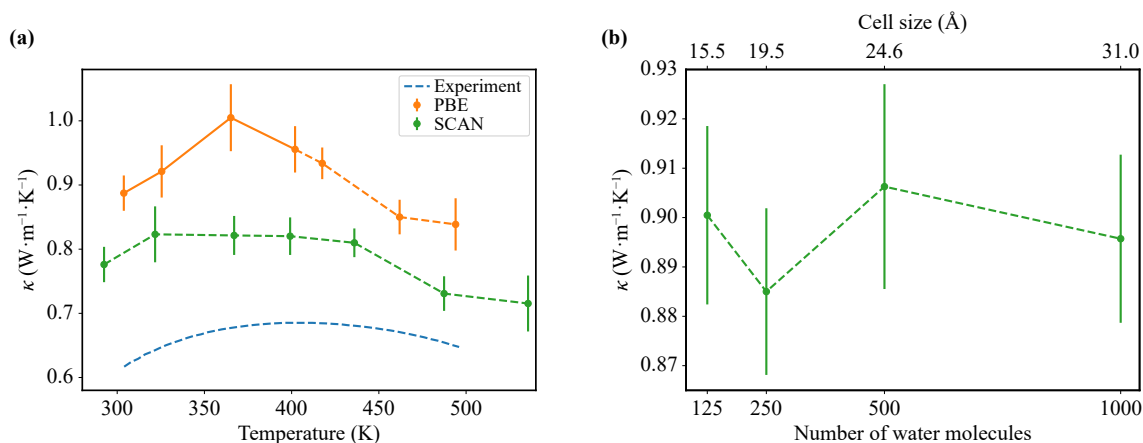


**Fig. 2** System size dependence of the self-diffusion coefficient in a Lennard–Jones fluid. The diffusion coefficient converges slowly with system size; even at  $N \approx 2000$ , it remains  $\sim 10\%$  below the thermodynamic limit, highlighting the need for finite-size corrections in dynamic properties. Reproduced with permission from Ref. [34]. Copyright © 2004 American Chemical Society.

tation of finite-size effects is more severe. Classical MD studies quantified that the self-diffusion coefficient of water exhibits significant finite-size dependence, underestimating values by  $\sim 10\%$  even for systems of  $\sim 2000$  molecules and necessitating hydrodynamic corrections (see Fig. 2) [34]. Properties such as dielectric response, solvation free energies, and collective dynamics in electrolyte solutions also require very large simulation cells for numerical convergence [35, 36].

For dynamical and time-correlation properties, achieving statistical convergence requires not only the mitigation of finite-size effects but also extensive sampling over long simulation times. In a system of 64 water molecules, obtaining a statistically converged diffusion coefficient necessitates at least 20–30 ps of trajectory [17–19], while properties involving rotational motions, which require well-averaged time-correlation functions, demand even longer simulation durations [37, 38]. These requirements for large-scale, long-timescale sampling exceeded the practical limits of conventional AIMD, thus highlighting the need for a simulation method that could simultaneously deliver both electronic-structure accuracy and high computational efficiency.

Recent years, MLPs have emerged as a transformative solution, bridging the accuracy-sampling gap by extending the scope and depth of first-principles simulations in two key dimensions: (i) overcoming the severe sampling limitation, (ii) mapping atomic configurations to electronic and spectroscopic observables. MLPs overcome the sampling limitations of AIMD by enabling large-scale, long-time simulations, thereby allowing for the rigorous convergence of properties sensitive to system size and trajectory length. State-of-the-art MLP-driven simulations have provided definitive evidence for a dynamic, “dancing” distorted Eigen cation ( $\text{H}_9\text{O}_4^+$ ) as the prevalent



**Fig. 3** Thermal conductivity of liquid water from DPMD simulations. **(a)** Temperature dependence of thermal conductivity  $\kappa$  from 300 K to 500 K, comparing DPMD results trained on PBE (orange) and SCAN (green) functionals. The SCAN-based DPMD results show significantly improved agreement with experimental measurements. **(b)** System-size dependence of  $\kappa$  from DPMD with the SCAN functional, confirming convergence. Direct AIMD cannot easily access such data due to the analytical complexity of the energy flux. Reproduced with permission from Ref. [40]. Copyright © 2021 American Physical Society.

species, resolving a long-standing structural debate through computationally intensive sampling that would be prohibitive for direct AIMD [39]. In the study of heat transport, a deep neural network potential trained on DFT data enabled the calculation of thermal conductivity using the nonempirical SCAN meta-GGA functional, which significantly improved agreement with experiment — a task exceedingly difficult with direct AIMD due to the analytical complexity of the required energy flux (see Fig. 3) [40].

Beyond addressing sampling limitations, MLPs dramatically broaden the spectrum of accessible properties by directly mapping atomic configurations to electronic and spectroscopic observables, moving far beyond the traditional output of forces and energies. This capability manifests on two interconnected fronts. First, MLPs enable the efficient prediction of electronic response properties. For instance, deep neural networks trained to predict molecular multipole moments from electron density allow for the calculation of dielectric constants and other susceptibilities from long trajectories, creating a direct bridge from structure to electronic response [41]. Second, MLPs provide a framework for *ab initio*-quality spectroscopy at molecular dynamics scales. To predict such spectroscopic observables, MLPs are trained not only on energies and forces but also on response properties computed from DFT (e.g., via finite-field methods). These quantities are local functions of atomic configurations, making them learnable using the same descriptors. A dual-network framework is often employed: one network learns the potential energy, while a separate network learns the dipole or polarizability. Once trained, the MLP can compute time-correlation functions over microsecond-long trajectories — a task far beyond direct AIMD. This yields statistically converged spectral line

shapes and relaxation times, enabling direct comparison with experiment. For instance, such a dual-network framework modeling both the potential energy surface and electronic polarizability has enabled the calculation of Raman spectra over nanoseconds [42]. This spectroscopic power extends across multiple techniques: MLPs have been used to compute infrared (IR), Raman, and sum-frequency generation spectra for direct comparison with experiment [43]; to reveal how isotopic substitution ( $\text{D}_2\text{O}$  vs.  $\text{H}_2\text{O}$ ) blue-shifts the X-ray absorption (XAS) edge due to a more ordered hydrogen-bond network [44]. Most recently, by analyzing shifts in the OH-stretching bands, researchers successfully linked these spectral changes to the degree of ordering within hydrogen-bond chains [45]. Collectively, these advancements transform MLP simulations into a versatile spectroscopic platform, yielding direct structural and dynamical insights that both validate and refine our understanding of the hydrogen-bonding network.

The ability of MLPs to deliver first-principles accuracy across extended scales not only solves sampling problems but also creates a new pathway to efficiently incorporate nuclear quantum effects (NQEs), which are central to water's behavior. NQEs, arising from the quantum nature of atomic nuclei, are central to the behavior of water and far from minor perturbations. The importance of NQEs is unambiguously evidenced by isotopic differences, such as the  $\sim 20\%$  higher self-diffusion coefficient of  $\text{H}_2\text{O}$  versus  $\text{D}_2\text{O}$  [46], and by the NQE-induced softening of the oxygen-oxygen radial distribution function observed in neutron scattering [47]. Accurately capturing these effects requires both an extremely precise potential energy surface and computationally feasible access to quantum statistical mechanics. Prior to the advent of MLPs, path-integral AIMD (PI-AIMD) was widely used

to provide foundational insights of NQEs [48]. However, the computational cost of PI-AIMD is very expensive. This stems from the path-integral formalism, which maps each quantum nucleus onto a ring of  $P$  classical replicas (“beads”), discretizing its Feynman path. Since each bead requires an independent electronic structure calculation, the cost per MD step scales linearly with  $P$ , which is typically 16–32 for aqueous systems. Even though applying colored-noise thermostat could enhance sampling efficiency [49–51], when combined with the necessity for a smaller integration timestep, the total computational effort escalates by two to three orders of magnitude compared to an AIMD simulation of the same physical system and duration [52].

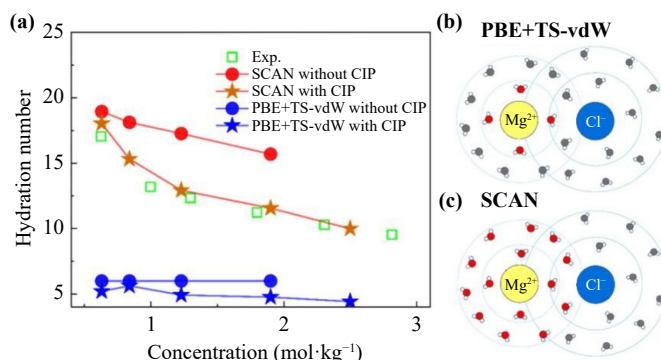
The integration of machine learning potentials with path-integral molecular dynamics (MLP-PIMD) has emerged as a transformative strategy to overcome this barrier. A landmark demonstration of this power is the work of Cheng *et al.* [53], which employed an MLP trained on the SCAN functional to perform extensive PIMD simulations. This study achieved an *ab initio* thermodynamic description of water, accurately predicting its phase diagram while quantitatively capturing the isotopic shift between  $\text{H}_2\text{O}$  and  $\text{D}_2\text{O}$  — a direct and stringent test of NQEs. This exemplifies how MLP-PIMD enables quantum-accurate simulations at scales previously inaccessible, fundamentally extending the scope of first-principles investigation into nuclear quantum phenomena.

This capability to reliably incorporate NQEs has opened new avenues for probing their influence on dynamical properties. For instance, a detailed analysis across a temperature range shows that NQEs slow down the self-diffusion and rotational motion of water molecules, with intermolecular effects diminishing as temperature rises while intramolecular effects persist [54]. Further dynamical insights emerge from studies of the dielectric response of water under pressure, where a nonlinear increase of the dielectric constant is attributed to pressure-induced strengthening of the hydrogen bond network, enhancing collective dipole fluctuations [55].

This systematic progression — from establishing structural benchmarks and hydrogen-bonding networks, through elucidating NQE dynamics and pressure responses, to spectroscopic validations — demonstrates how MLPs provide a unified framework for unraveling the complexity of bulk water. These foundational insights create the necessary basis for exploring more intricate aqueous solutions in the following subsection.

## 2.2 Aqueous solutions

MLPs extend their predictive power to aqueous solutions, unraveling how ions reshape structural and dynamical behavior of water. This section summarizes key insights across diverse systems, organized by solute-

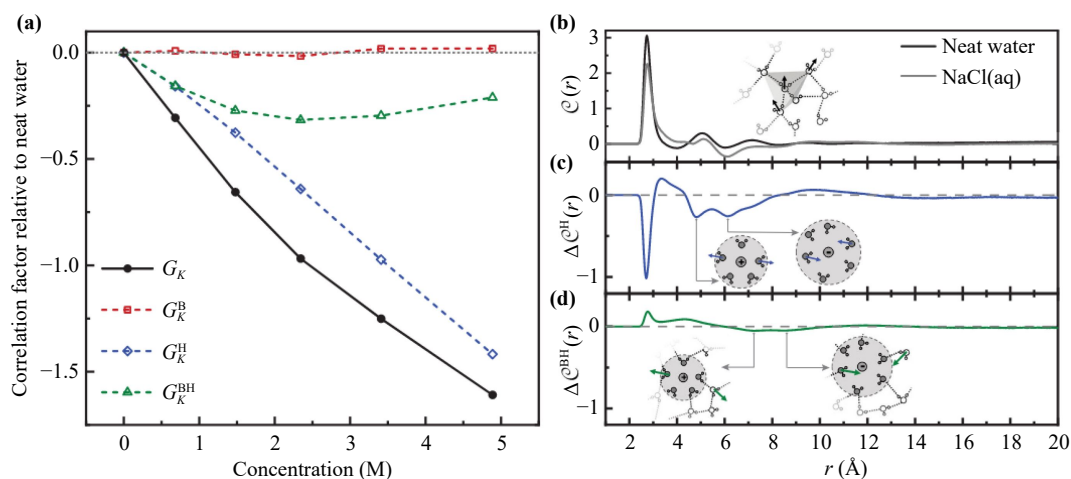


**Fig. 4** The concentration dependence of hydration number in the  $\text{MgCl}_2$  solution, obtained by SCAN without contact ion pairs (CIPs), SCAN with CIPs, PBE+TS-vdW without CIPs, PBE+TS-vdW with CIPs, and the experiment [38]. Reproduced from Ref. [63].

induced perturbations.

In this review we mainly focus on two fundamental and widely debated questions regarding aqueous solutions: (i) do ions influence water’s structure and dynamics through short-range or long-range effects? (ii) how does concentration effect influence the structural and dynamical properties? The first concerns the spatial range of ion-induced perturbations. Since the early 2000s, experimental findings have presented a stark contrast. Evidence supporting a long-range influence comes from neutron diffraction studies, dielectric spectroscopy, and femtosecond elastic second-harmonic scattering experiments [56–58]. Conversely, femtosecond infrared spectroscopy and X-ray absorption experiments have concluded that ion-induced perturbations are confined to the first solvation shell [59, 60]. Classical MD simulations, whose outcomes are sensitive to the empirical force field employed [61], fail to resolve this debate. Meanwhile, AIMD, while providing first-principles accuracy, lacks the computational throughput to simulate the large systems and long timescales necessary to gather statistically robust evidence for long-range correlations [9]. Notably, Gaiduk *et al.* [62] highlighted that addressing this question may require moving beyond single-ion models to consider ion-pair effects. This context is ideally suited for MLPs, which uniquely combine AIMD-level accuracy with the computational efficiency needed to tackle this long-standing puzzle.

To address this question, Li *et al.* [63] investigated intermediate-range of  $\text{Mg}^{2+}$  ions. By employing an MLP-based approach, they compared the properties of  $\text{MgCl}_2$  and  $\text{MgBr}_2$  solutions at varying concentrations. Their simulations revealed that the retarding effect of  $\text{Mg}^{2+}$  on water dynamics extends beyond the first hydration shell. This point was further proved by the results of hydration number, which is consistent with the experiment (see Fig. 4) [38]. This medium-range effect was attributed to a competition between ion-water and hydrogen-bond



**Fig. 5** The contribution to dipolar correlation function of different components. (a) The dependence of correlation factor relative to neat water on concentration in NaCl solutions. The contributions are divided into correlations between bulk water  $G_K^B$ , hydration water  $G_K^H$ , and cross-correlations between bulk water and hydration water  $G_K^{BH}$ . (b) The dipolar correlation function of neat water and NaCl solution at 4.9 M. (c) The dipolar correlation function between hydration water. (d) The dipolar cross-correlation function between bulk water and hydration water. Reproduced with permission from Ref. [69]. Copyright © 2023 American Physical Society.

interactions. This finding provides a new perspective for understanding the dichotomy between short-range and long-range perturbation theories. Furthermore, their work demonstrated that the choice of the underlying density functional significantly influences the predicted dynamical properties [63].

We now turn to the second question regarding concentration effects. Experimental observations dating back nearly a century have shown that the static dielectric constant of aqueous NaCl solutions decreases with increasing salt concentration [64]. Subsequent quantitative measurements further revealed that this decrease follows a distinct nonlinear trend [65–67]. However, the molecular origin of this nonlinear behavior has remained elusive. To resolve this, Zhang and colleagues leveraged an MLP approach to simulate NaCl solutions across a wide concentration range. Predicting the dielectric response of aqueous solutions demands both large system sizes (hundreds of molecules) and nanosecond-scale trajectories to get numerically and statistically converged dipolar correlations, yet AIMD simulations of such scale are computationally prohibitive. Moreover, standard MLPs with local cutoffs cannot intrinsically capture the long-range  $1/r$  tail of dipole–dipole and ion–water interactions, which are critical for accurate dielectric constants. To overcome these limitations, the authors employed the Deep Potential Long-Range (DPLR) framework [68], which augments a local deep potential with a physics-based electrostatic term that explicitly treats long-range interactions beyond the cutoff. By classifying water molecules into distinct subpopulations (e.g., bulk and hydration water) and quantitatively deconvoluting their contributions to the total system dipole, their work

demonstrated the nonlinearity to the evolving coupling in dipole correlations between bulk water and hydration water (see Fig. 5) [69].

In addition to probing concentration effects, the capability of MLPs to efficiently perform simulations across multiple concentrations enables the routine screening of solution properties, thereby facilitating the investigation of a broader spectrum of problems in aqueous chemistry. For NaCl solutions, MLP simulations has shown that the structure perturbation originates from concentration effects is not equivalent to that originates from applying pressure, demonstrating distinct hydrogen-bonding network disruptions that cannot be mapped to pressure-induced compression [70].

In recent years, water-in-salt electrolytes (WISEs) have attracted significant research interest due to their promising applications. Wang *et al.* [71] employed MLP-based molecular dynamics to demonstrate that the high reductive stability stems from a fundamental switching of redox levels between anions and water. This critical inversion, driven by the isolation of water molecules, suppresses the hydrogen evolution reaction and promotes the formation of a protective, anion-derived solid-electrolyte interphase [71]. Beyond thermodynamic stability, MLPs also elucidate the unique transport dynamics in WISEs. In concentrated  $Zn^{2+}$  systems, a subsequent study revealed a sophisticated three-stage carrier transport mechanism, progressing from independent ion diffusion to strongly correlated motion, and finally to the diffusion of small positive charges through anionic polymeric clusters [72]. Together, these works elucidate the power of MLPs in unraveling the coupled thermodynamic and kinetic phenomena that govern the performance

of next-generation electrolytes.

Beyond these equilibrium properties, MLPs provide profound insights into the dynamical behavior and transport mechanisms governing ionic solutions. Utilizing DPMD, Avula *et al.* [73] successfully simulated the anomalous diffusion of water in CsI and NaCl solutions, obtaining results that are qualitatively consistent with experimental observations across a range of concentrations. The researchers employed DPMD to elucidate the competitive mechanism by which different ions exert opposing effects on water diffusion. Long-timescale MLP simulations further resolve NaCl dissociation kinetics, revealing a stochastic pathway involving dynamic intermediates [74], while potential of mean force (PMF) calculations show minimal energy differences between contact ion pairs and solvent-shared ion pairs [75]. Similar PMF-based analyses have been extended to divalent metal ions interacting with biologically relevant ligands. Using neural network potentials, recent studies systematically characterized the binding of  $Mg^{2+}$ ,  $Ca^{2+}$ , and  $Zn^{2+}$  to amino acid side-chain analogues and the  $Mg^{2+}$ -acetate ion pair in water, revealing ion-specific coordination structures and a  $\sim 5.8$  kcal/mol barrier for monodentate-bidentate conversion, while also providing benchmarks for classical force fields [76, 77].

### 3 Properties of confined and interfacial aqueous systems

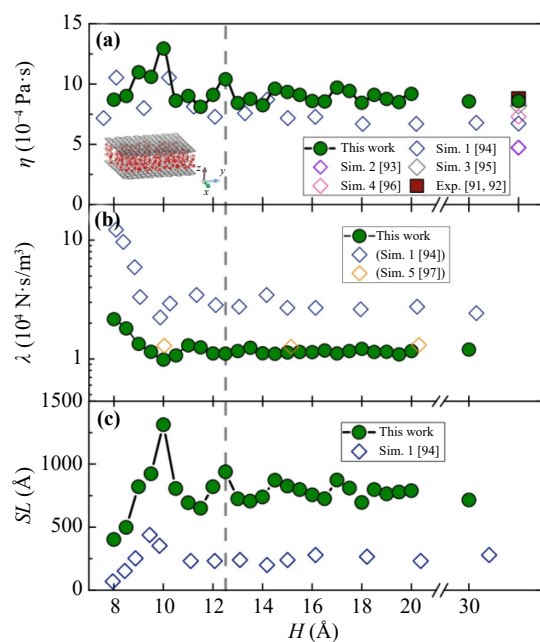
The exploration of water and electrolytes under spatial confinement and at solid interfaces has long been hindered by the dual challenges of experimental characterization and computational cost. The advent of MLPs is now overcoming these barriers, enabling a precise and comprehensive dissection of the unique physical chemistry that emerges when water is taken out of its bulk environment.

Water in nanoscale confinement is ubiquitous in nature. Such extreme spatial confinement can profoundly alter the dynamic transport behavior of liquid water, with significant implications for a wide range of biological and industrial processes [78, 79]. Over the past decade, confinement effects have been recognized as a key factor modulating liquid transport. Experiments have demonstrated that water can flow through nanochannels at ultrafast speeds. This enhanced flow surpasses the predictions of classical Navier-Stokes hydrodynamics by over an order of magnitude. This ultrafast transport is highly dependent on the confinement dimension and the material of the channel walls, being particularly pronounced within carbon-based nanostructures [80, 81]. In contrast, it is markedly less evident in systems such as those confined by  $MoS_2$  or hexagonal boron nitride (hBN) [82, 83]. The prevailing experimental interpretation of this phenomenon has largely relied on

macroscopic hydrodynamic arguments [84], attributing the differences to wall hydrophobicity and atomic-scale smoothness, which collectively determine the interfacial friction and thus the transport properties [85]. However, these continuum models fail to account for the anomalous water transport observed specifically in graphene-based channels. While classical MD simulations have been employed to probe the role of interfacial friction, their results are notably sensitive to the empirical force field chosen, often leading to significant discrepancies with experimental measurements [86]. To systematically unravel the underlying confinement effects, an *ab initio* accurate description of the water-confining material interaction is essential [87]. Herein lies the key strength of the MLP approach, which provides a powerful and accurate means to investigate how nanoconfinement influences water transport properties.

Recent advancements leveraging MLPs have clarified the dependence of nanofluidic transport on channel material and confinement geometry. First, addressing the long-standing puzzle of material specificity, a seminal MLP study directly compared water flow in carbon nanotubes (CNTs) and boron nitride nanotubes (BNNTs) [88]. It quantified  $a \approx 5$  times lower friction on carbon surfaces, tracing the origin to atomic-scale mechanisms: facile oxygen motion in CNTs versus specific hydrogen-nitrogen interactions in BNNTs. This moves the explanation beyond macroscopic hydrophobicity. At a more fundamental level, MLPs are quantifying how confinement distorts the hydrogen-bond network itself. Simulations reveal that hydrogen bond orient preferentially parallel to the interface, imparting a “two-dimensional” character that significantly reduces the water layers needed for network percolation compared to bulk water [89]. This structural insight directly accounts for the geometric dependence of transport. Another pivotal study on graphene nanoslits revealed that the interlayer distance ( $H$ ) dictates dynamics through a competition between parallel flow and perpendicular fluctuations, with a critical transition at  $H \approx 12.5$  Å (see Fig. 6). Importantly, this mechanistic understanding yields a rational design principle: transport can be enhanced several-fold by optimizing friction at specific confinement sizes [90].

MLPs have uniquely enabled the exploration of the novel physical properties of confined water. Zhao, Qiu, and Guo [98] developed a deep neural network potential for water confined in graphene nanocapillaries, trained on DFT calculations, which achieves near-DFT accuracy at substantially reduced computational cost and accurately reproduces a wide range of structural and dynamical properties of nanoconfined water. A landmark study revealed that water confined to a monolayer exhibits a rich phase diagram, including a superionic phase with proton conductivity exceeding that of battery materials — a phenomenon dictated by nanoscale pressure and



**Fig. 6** Transport properties of confined water from MLP-based simulations. The variations of (a) shear viscosity, (b) friction coefficient and (c) slip length of confined water obtained by MLPs methods. Some results from previous experiment [91, 92] and simulation [93–97] are also plotted. Reproduced from Ref. [90].

inaccessible to conventional sampling [99]. Furthermore, MLPs have elucidated how such confinement fundamentally alters collective electronic responses. Simulations of water between graphene and hBN directly connected the spontaneous polarization of interfacial water to the system’s anomalously low dielectric constant, providing a mechanistic link between subnanometer structure and macroscopic dielectric properties [100].

MLPs are uniquely suited to address the fundamental challenges of understanding aqueous solutions in confinement. For NaCl solutions in graphene slits, MLP simulations have demonstrated that increasing confinement not only reduces ionic conductivity but also triggers a fundamental shift in the ion transport mechanism from a collective process to vehicular motion [101]. Furthermore, the effective interactions between ions are profoundly altered. PMF calculations, now feasible with *ab initio* accuracy over relevant distances, reveal that ion-ion interactions in nanoconfinement deviate significantly from bulk solutions, explaining anomalous phase behavior and ion selectivity [102].

Among interfacial aqueous systems, one of the most widely studied systems is the electric double layer (EDL). The EDL plays a central role in electrochemical energy storage, catalysis, and corrosion, yet its atomic-scale structure and potential distribution have long remained elusive to direct experimental probing. Conventional AIMD is severely limited by its computa-

tional cost when simulating large, realistic electrode-electrolyte interfaces over relevant timescales. Recent MLP developments, particularly those incorporating long-range electrostatics, have opened new avenues for EDL research. For instance, Li *et al.* [103] combined AIMD with MLP simulations to reveal how alkali metal cations in base alter the hydrogen-bond network continuity of interfacial water within the EDL, thereby modulating the proton transfer kinetics in alkaline hydrogen electrocatalysis [103]. Zhang and coworkers employed a DFT-based MLP with explicit long-range electrostatic corrections to systematically investigate the anatase  $\text{TiO}_2$ -electrolyte interface under different pH conditions. Their large-scale simulations captured water dissociation/recombination and proton transport, showing that the higher affinity of cations for the oxide surface under basic conditions leads to an increased differential capacitance, in quantitative agreement with experiment [104]. Zhu and Cheng further developed a hybrid MLP framework that simultaneously treats local electronic polarization in the electrolyte and non-local charge transfer in the metal electrode. This approach successfully reproduced the bell-shaped differential Helmholtz capacitance at the Pt(111)-electrolyte interface and provided new insights into the dielectric response of the EDL [105]. Collectively, these studies demonstrate that MLPs can effectively unravel the complex behavior of EDLs at atomic resolution.

At solid–liquid interfaces, the subtle interplay between water and the atomic specifics of the substrate dictates structure and dynamics. While AIMD has been limited to simulating small patches for short durations, MLPs now enable direct comparison with experimental observables across large, realistic interfaces. For example, MLPs combined with experimental atomic force microscopy can differentiate the distinct hydrogen-bonding structures of water confined between graphene and hexagonal boron nitride surfaces — a task beyond the reach of conventional force-field molecular dynamics [106]. Beyond static structure, MLPs reveal how mechanical coupling governs transport. Simulations of flexible channels show that the solid–liquid friction in materials like graphene scales inversely with lateral size — a key relationship absent in rigid models and one that opens new avenues for flow control [107]. Similarly, MLP simulations have resolved the long-standing question of hydronium ( $\text{H}_3\text{O}^+$ ) and hydroxide ( $\text{OH}^-$ ) ion distribution at the graphene-water interface, linking it to interfacial polarization and charge rearrangement effects [108]. These studies exemplify how MLPs bridge the gap between nanoscale simulation and macroscopic experiment, providing mechanistic insights that are both atomically precise and directly relevant to observable interfacial phenomena.

In summary, MLP-based simulations are not merely incremental improvements but represent a paradigm shift in the study of confined and interfacial aqueous

systems. By bridging the gap between the accuracy of quantum mechanics and the scale of classical molecular dynamics, they are systematically unraveling the complex, atomic-level physics that govern these technologically critical environments, from energy storage and catalysis to nanofluidics and desalination.

## 4 Reactions in aqueous systems

The accurate simulation of chemical reactions in aqueous environments represents one of the most demanding challenges in computational chemistry. These processes, which include proton transfer (PT), acid dissociation, and interfacial catalysis, are characterized by concerted atomic motions, significant solvent reorganization, and the making and breaking of chemical bonds. Capturing these processes requires both the electronic-level accuracy necessary to describe bond rearrangements and the ability to sample rare events and achieve statistical convergence for reliable kinetics. MLPs enable long-timescale simulations at *ab initio* accuracy. This capability allows for the rigorous mapping of multi-step reaction pathways, the quantification of free energy landscapes, and the direct connection of simulated dynamics to experimental observables. This section showcases the transformative impact of MLPs across a spectrum of aqueous reactivity, from elucidating the intricate mechanism of proton transport to unraveling complex reaction dynamics at interfaces and in bulk solution.

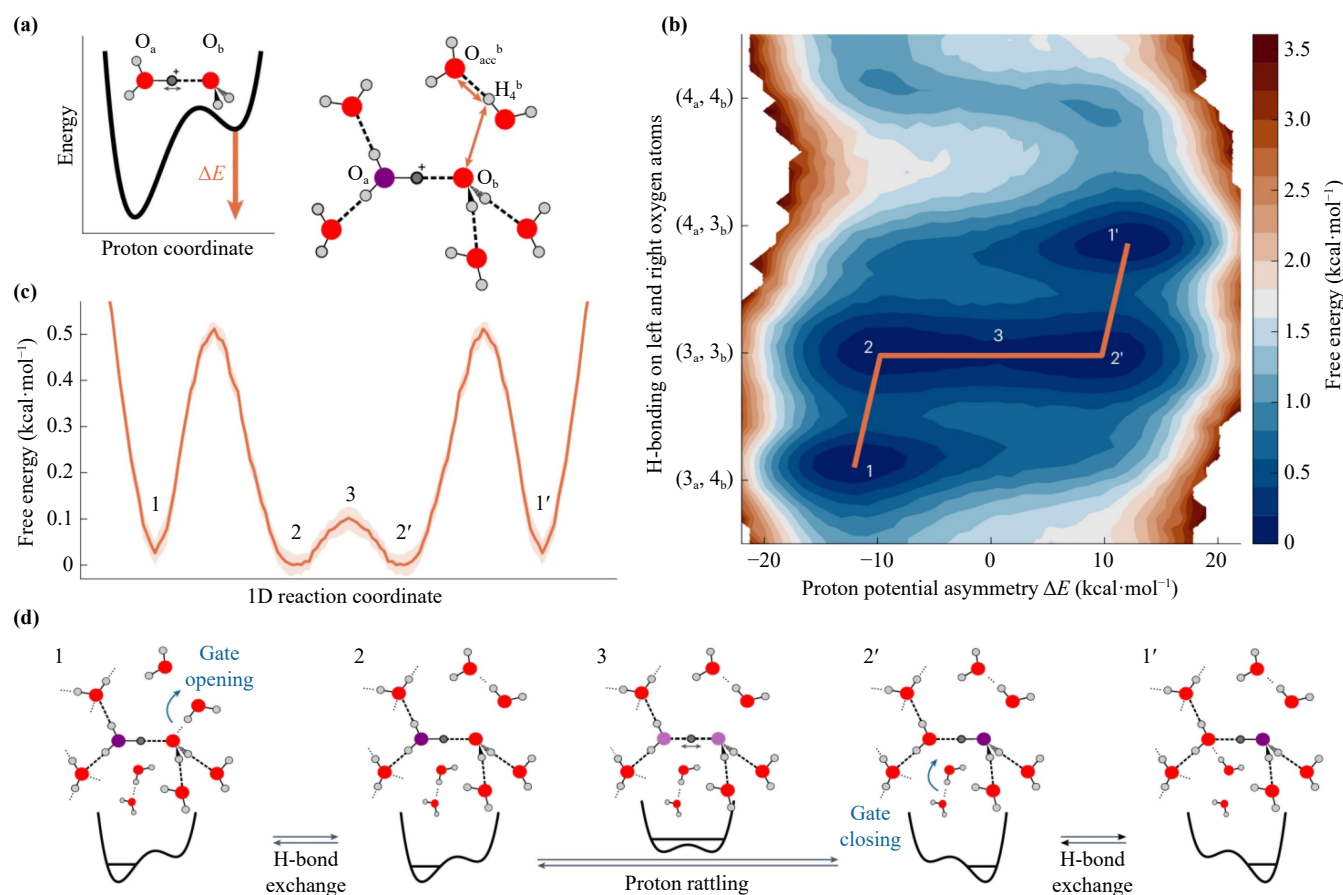
AIMD has laid the foundation for understanding PT in aqueous systems, capturing key aspects of the Grotthuss mechanism through visualization of hydrogen-bond network rearrangements and the dynamics of hydrated proton complexes [21]. These studies have provided insights into proton hopping mechanisms and the quantification of free-energy barriers [20]. However, AIMD's severe computational constraints prevent adequate sampling of complete PT coordinates and reaction kinetics. A critical illustration of this limitation is the finding that reliable estimation of proton and hydroxide ( $\text{OH}^-$ ) ion diffusion coefficients requires simulations spanning hundreds of picoseconds — a timescale far beyond routine AIMD capabilities [109]. Therefore, there is a pressing need for efficient and highly accurate sampling methods, enabled by MLPs, to investigate PT events.

MLPs enable the precise quantification of PT mechanisms in bulk systems. A seminal study leveraged DPMD to unravel a three-step mechanism for proton transport in bulk water, which is governed by two consecutive hydrogen-bond exchanges (see Fig. 7) [110]. This finding provides unprecedented statistical confidence in the detailed reaction coordinate of the Grotthuss mechanism. Furthermore, MLPs have proven crucial for mapping PT pathways at complex interfaces. For instance, DPMD simulations of the air-water interface

enabled the computation of the SFG spectrum of hydronium ( $\text{H}_3\text{O}^+$ ) ions, directly connecting simulated interfacial structures with experimental spectroscopic signatures [111]. Researchers have further extended their investigations of proton transfer to aqueous solutions. Using DPMD, Zhao *et al.* [112] has found that the dissociation kinetics of  $\text{H}_2\text{CO}_3$  are governed not by the inherent stability of carbonic acid, but by the hydrogen-bonding ring structure of specific conformational isomers, underscoring the critical role of solvation topology (see Fig. 8). Collectively, these studies have elucidated the PT mechanism.

While proton transfer remains foundational in bulk systems, it is the processes at surfaces and interfaces, most notably the air-water interface, that are often more critical in practical scenarios and represent a domain well-suited to the strengths of MLP simulations. DPMD simulations of the air-water interface revealed a decoupling between water and proton dynamics: while water molecular diffusion accelerates, proton diffusion slows down considerably [113]. This contrast stems from reduced hydrogen bond coordination at the interface, which facilitates water mobility but impedes the stable proton hopping central to the Grotthuss mechanism. MLP-based reactive force fields have also resolved long-standing debates about interfacial acidity. Studies on nitric and formic acids at the air-water interface showed that the measured degree of dissociation is highly sensitive to the experimental probing depth and system size [114]. The simulations revealed that acidity drops at the interface due to desolvation but can be enhanced just below it, offering a molecular explanation for contrasting experimental results. This concept extends to water droplets, where MLP simulations showed that although pH is uniform in the core, substantial concentration gradients of  $\text{H}_3\text{O}^+$  and  $\text{OH}^-$  ions across the outermost molecular layers lead to increased acidity at the interface [115].

Moving beyond proton hopping, MLPs have enabled a deep dive into the mechanisms of complex chemical reactions across diverse aqueous environments. At fluid interfaces, MLP simulations have uncovered unconventional reaction paradigms, benefiting from their high-efficiency sampling of molecular configurations [116]. A prime example is the discovery, via MACE-based MD simulations, of an “in-and-out” mechanism for  $\text{CO}_2$  hydration at the air-water interface, wherein  $\text{CO}_2$  diffuses into the aqueous surface layer to react and the product is subsequently expelled; this confirms a significant contribution from surface-adsorbed  $\text{CO}_2$  to the overall acidification rate [117]. Similarly, the dissociation dynamics of  $\text{N}_2\text{O}_5$  on water surfaces are dominated by interfacial hydrolysis and competitive re-evaporation rather than bulk-solution processes [118]. In bulk solution, MLPs coupled with enhanced sampling provide quantitative insights into reaction free energies and mechanisms. This is exemplified by the study of weak



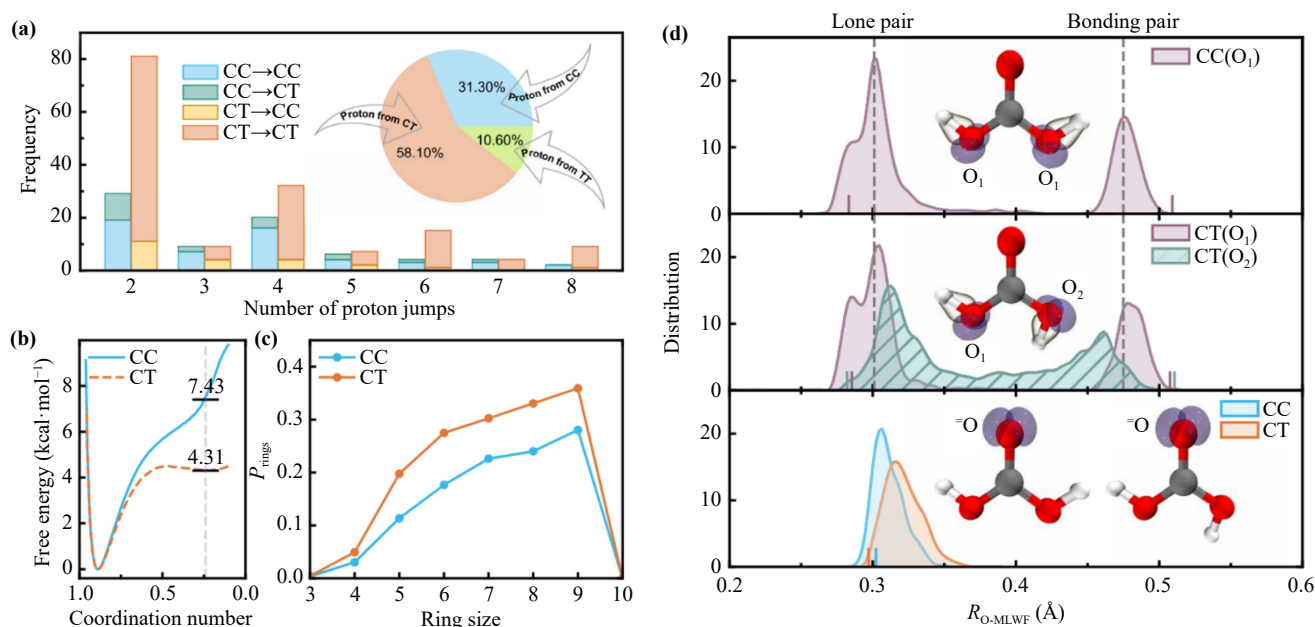
**Fig. 7** Free-energy landscape and mechanism of proton transfer. (a) Definition of proton coordinates: proton potential asymmetry  $\Delta E$  (left) and periodic hydrogen bond coordinate probing the donor and acceptor numbers of hydrogen bonded partners (right). (b) Free energy surface along the proton potential asymmetry and hydrogen bond coordinate denoted  $(n_a, n_b)$ , where  $n_{a,b}$  is the number of hydrogen bond partners for  $O_{a,b}$ . (c) Free energy profile along the pathway shown in (b). (d) Schematic PT mechanism with typical hydrogen bond arrangements and proton potential energy profiles for key locations on the surface. Reproduced from Ref. [110].

carboxylic acids, where MLPs successfully captured the concerted proton dissociation of formic and acetic acid to yield accurate  $pK_a$  estimates and characterized their dimerization processes [119]. MLPs also illuminate reactions of fundamental importance in synthesis, as demonstrated by the exploration of the amide bond formation mechanism in solution [120].

## 5 Outlooks

This review highlights the versatility of MLPs and their importance across multiple aqueous systems — including bulk water, aqueous solutions, confined systems, and interfaces — due to their ability to capture water–water, ion–water, and water–interface interactions with *ab initio* accuracy. Although the MLP approach has provided solutions to long-standing problems concerning accuracy and efficiency, fundamental questions persist regarding its applicability to the most complex aqueous scenarios.

A fundamental open question remains: do ions exert genuine long-range effects, and if so, how do these effects influence the structure and dynamics of water? Resolving this issue crucially depends on the development of MLP approaches capable of capturing long-range interatomic interactions, which still remains difficult. To date, most models utilize local descriptors that only capture the atomic environment within a cutoff radius, implicitly attributing all interactions to atoms within this sphere. Due to computational constraints, this cutoff cannot be extended arbitrarily, thus precluding the modeling of true long-range physical effects. Recent efforts fall into several categories. One strategy involves incorporating physics-based corrective terms for specific interactions beyond the cutoff, such as electrostatics (e.g., DPLR or CACE). However, this approach requires prior knowledge of the functional form of the long-range interactions. Another strategy employs message-passing neural networks to propagate information over longer ranges (e.g., SO3krates, MACE) [29, 121]. Yet, the



**Fig. 8** Conformer-dependent proton-transfer dynamics. **(a)** Distribution of proton-hop counts for the cis-cis (CC) and cis-trans (CT) conformers between dissociation and recombination events. **(b)** Free-energy landscape for dissociation, projected onto the hydrogen coordination number of the deprotonated oxygen, comparing the CC and CT conformers. **(c)** Distribution of ring size that CC or CT participate in. **(d)** The distribution of the distance between the centers of the corresponding maximally localized Wannier functions and the intramolecular oxygen atoms in H<sub>2</sub>CO<sub>3</sub> of CC (top), CT (middle) and the carbonyl oxygens in both CC and CT conformers (bottom). Reproduced from Ref. [112].

computational cost of these models escalates significantly with the number of message-passing layers. While the latest research explores capturing information in Fourier or spherical harmonic space [122], a general and efficient model for long-range interactions remains elusive.

As mentioned in the introduction, the accuracy of MLPs is inherently bounded by the DFT reference data on which they are trained. While DFT remains the most practical electronic structure method for generating large training sets, its well-known limitations — including self-interaction error, incomplete treatment of van der Waals dispersion, and approximate exchange-correlation functionals — can propagate into MLP predictions for hydrogen-bonding networks, proton transfer barriers, and interfacial charge distributions. If the reference functional fails to accurately describe these key interactions, MLPs trained on such data will inevitably inherit the same deficiencies, regardless of their numerical efficiency. To overcome this limitation, the community is increasingly turning to higher-accuracy benchmarks such as quantum Monte Carlo (QMC) and CCSD(T) [123–128]. Both methods provide near-exact energies and forces for small water clusters and reaction intermediates, albeit at high computational cost. Although still too demanding for direct dynamics, QMC/CCSD(T) data can serve as a reference for training or correcting MLPs, as discussed below.

Beyond the quest for higher-accuracy reference data, another major challenge lies in the efficient generation of

training datasets that comprehensively sample the complex configurational space of aqueous systems. Practical applications of aqueous solutions often involve multi-element systems, solid-liquid or gas-liquid interfaces, and long-range interactions, all of which make it prohibitively expensive to acquire sufficient training data using traditional DFT methods. To address this, current strategies focus on improving data efficiency. One powerful approach is active learning, which iteratively explores configuration space using ML-driven MD rather than *ab initio* MD, identifies regions where the MLP has high uncertainty, and selectively adds new training points from those regions [30]. This minimizes the amount of DFT data required while maximizing model robustness. Other methods include  $\Delta$ -learning, where a low-cost MLP learns the correction between a cheaper baseline method and a higher-level theory (e.g., DFT vs. QMC/CCSD(T)) [129], and multi-fidelity approaches that combine data from multiple levels of theory [130]. Additionally, the emergence of universal foundation models pretrained on extensive datasets offers a promising route: such models can be fine-tuned for specific aqueous systems with limited additional data [131, 132]. Despite these advances, a systematic and robust solution for generating training sets for highly complex solutions and chemical reactions remains an open challenge.

Once the aforementioned challenges are overcome, MLPs will evolve into a transformative tool, not only deepening our fundamental understanding of the mecha-

nisms of aqueous systems but also directly contributing to the rational design of materials and processes to address critical challenges in environmental science and industrial chemistry [133].

**Declarations** The authors declare that they have no competing interests and there are no conflicts.

**Acknowledgements** This work was supported by the Shanghai Rising-Star Program (No. 23QA1406800) and ShanghaiTech AI Initiative (Grant No. AI2026A09).

## References

- J. P. McEvoy and G. W. Brudvig, Water-splitting chemistry of photosystem II, *Chem. Rev.* 106(11), 4455 (2006)
- S. Cukierman, Et tu, Grotthuss! and other unfinished stories, *Biochim. Biophys. Acta Bioenerg.* 1757(8), 876 (2006)
- C. A. Wraight, Chance and design — Proton transfer in water, channels and bioenergetic proteins, *Biochim. Biophys. Acta Bioenerg.* 1757(8), 886 (2006)
- L. Vrbka, P. Jungwirth, P. Bauduin, D. Touraud, and W. Kunz, Specific ion effects at protein surfaces: A molecular dynamics study of bovine pancreatic trypsin inhibitor and horseradish peroxidase in selected salt solutions, *J. Phys. Chem. B* 110(13), 7036 (2006)
- M. C. Pinna, P. Bauduin, D. Touraud, M. Monduzzi, B. W. Ninham, and W. Kunz, Hofmeister effects in biology: Effect of choline addition on the salt-induced super activity of horseradish peroxidase and its implication for salt resistance of plants, *J. Phys. Chem. B* 109(34), 16511 (2005)
- P. Jungwirth and D. J. Tobias, Specific ion effects at the air/water interface, *Chem. Rev.* 106(4), 1259 (2006)
- A. Sanna, M. Uibu, G. Caramanna, R. Kuusik, and M. M. Maroto-Valer, A review of mineral carbonation technologies to sequester CO<sub>2</sub>, *Chem. Soc. Rev.* 43(23), 8049 (2014)
- L. Wang, J. Yan, Y. Hong, Z. Yu, J. Chen, and J. Zheng, Ultrahigh-rate and ultralong-life aqueous batteries enabled by special pair-dancing proton transfer, *Sci. Adv.* 9(18), eadf4589 (2023)
- Y. Ding, A. A. Hassanali, and M. Parrinello, Anomalous water diffusion in salt solutions, *Proc. Natl. Acad. Sci. USA* 111(9), 3310 (2014)
- D. Marx, M. E. Tuckerman, J. Hutter, and M. Parrinello, The nature of the hydrated excess proton in water, *Nature* 397(6720), 601 (1999)
- M. Tuckerman, K. Laasonen, M. Sprik, and M. Parrinello, *Ab initio* molecular dynamics simulation of the solvation and transport of hydronium and hydroxyl ions in water, *J. Chem. Phys.* 103(1), 150 (1995)
- M. DelloStritto, J. Xu, X. Wu, and M. L. Klein, Aqueous solvation of the chloride ion revisited with density functional theory: impact of correlation and exchange approximations, *Phys. Chem. Chem. Phys.* 22(19), 10666 (2020)
- C. Zhang, T. A. Pham, F. Gygi, and G. Galli, Communication: Electronic structure of the solvated chloride anion from first principles molecular dynamics, *J. Chem. Phys.* 138(18), 181102 (2013)
- A. Bankura, V. Carnevale, and M. L. Klein, Hydration structure of salt solutions from *ab initio* molecular dynamics, *J. Chem. Phys.* 138(1), 014501 (2013)
- T. T. Duignan, G. K. Schenter, J. L. Fulton, T. Huthweller, M. Balasubramanian, M. Galib, M. D. Baer, J. Wilhelm, J. Hutter, M. Del Ben, X. S. Zhao, and C. J. Mundy, Quantifying the hydration structure of sodium and potassium ions: Taking additional steps on Jacob's ladder, *Phys. Chem. Chem. Phys.* 22(19), 10641 (2020)
- F. C. Lightstone, E. Schwegler, R. Q. Hood, F. Gygi, and G. Galli, A first principles molecular dynamics simulation of the hydrated magnesium ion, *Chem. Phys. Lett.* 343(5–6), 549 (2001)
- R. A. Jr DiStasio, B. Santra, Z. Li, X. Wu, and R. Car, The individual and collective effects of exact exchange and dispersion interactions on the *ab initio* structure of liquid water, *J. Chem. Phys.* 141(8), 084502 (2014)
- M. Chen, H. Y. Ko, R. C. Remsing, M. F. Calegari Andrade, B. Santra, Z. Sun, A. Selloni, R. Car, M. L. Klein, J. P. Perdew, and X. Wu, *Ab initio* theory and modeling of water, *Proc. Natl. Acad. Sci. USA* 114(41), 10846 (2017)
- L. Zheng, M. Chen, Z. Sun, H. Y. Ko, B. Santra, P. Dhuvad, and X. Wu, Structural, electronic, and dynamical properties of liquid water by *ab initio* molecular dynamics based on SCAN functional within the canonical ensemble, *J. Chem. Phys.* 148(16), 164505 (2018)
- M. Chen, L. Zheng, B. Santra, H. Y. Ko, M. L. Jr DiStasio, R. Klein, Car, and X. Wu, Hydroxide diffuses slower than hydronium in water because its solvated structure inhibits correlated proton transfer, *Nat. Chem.* 10(4), 413 (2018)
- A. Hassanali, F. Giberti, J. Cuny, T. D. Kühne, and M. Parrinello, Proton transfer through the water gossamer, *Proc. Natl. Acad. Sci. U. S. A.* 110(34), 13723 (2013)
- J. Behler, Atom-centered symmetry functions for constructing high-dimensional neural network potentials, *J. Chem. Phys.* 134(7), 074106 (2011)
- J. Behler and M. Parrinello, Generalized neural-network representation of high-dimensional potential-energy surfaces, *Phys. Rev. Lett.* 98(14), 146401 (2007)
- A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons, *Phys. Rev. Lett.* 104(13), 136403 (2010)
- A. P. Bartók, R. Kondor, and G. Csányi, On representing chemical environments, *Phys. Rev. B* 87(18), 184115 (2013)
- L. Zhang, J. Han, H. Wang, and R. Car, Deep potential molecular dynamics: A scalable model with the accuracy of quantum mechanics, *Phys. Rev. Lett.* 120(14), 143001 (2018)



27. K. Schütt, P. J. Kindermans, H. E. Saucedo Felix, S. Chmiela, A. Tkatchenko, and K. R. Müller, SchNet: A continuous-filter convolutional neural network for modeling quantum interactions, in: *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2017, Vol. 30
28. S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt, and B. Kozinsky,  $E(3)$ -equivariant graph neural networks for data-efficient and accurate interatomic potentials, *Nat. Commun.* 13(1), 2453 (2022)
29. I. Batatia, D. P. Kovacs, G. Simm, C. Ortner, and G. Csanyi, MACE: Higher order equivariant message passing neural networks for fast and accurate force fields, *Adv. Neural Inf. Process. Syst.* 35, 11423 (2022)
30. L. Zhang, D. Y. Lin, H. Wang, and R. Car, Active learning of uniformly accurate interatomic potentials for materials simulation, *Phys. Rev. Mater.* 3(2), 023804 (2019)
31. Y. Zhang, H. Wang, W. Chen, J. Zeng, L. Zhang, and H. Wang, DP-GEN: A concurrent learning platform for the generation of reliable deep learning based potential energy models, *Comput. Phys. Commun.* 253, 107206 (2020)
32. E. Schwegler, J. C. Grossman, F. Gygi, and G. Galli, Towards an assessment of the accuracy of density functional theory for first principles simulations of water. II, *J. Chem. Phys.* 121(11), 5400 (2004)
33. Z. Sun, M. Chen, L. Zheng, J. Wang, B. Santra, H. Shen, L. Xu, W. Kang, M. L. Klein, and X. Wu, X-ray absorption of liquid water by advanced *ab initio* methods, *Phys. Rev. B* 96(10), 104202 (2017)
34. I. C. Yeh and G. Hummer, System-size dependence of diffusion coefficients and viscosities from molecular dynamics simulations with periodic boundary conditions, *J. Phys. Chem. B* 108(40), 15873 (2004)
35. A. Torres, L. S. Pedroza, M. Fernandez-Serra, and A. R. Rocha, Using neural network force fields to ascertain the quality of *ab initio* simulations of liquid water, *J. Phys. Chem. B* 125(38), 10772 (2021)
36. A. Z. Panagiotopoulos and S. Yue, Dynamics of aqueous electrolyte solutions: Challenges for simulations, *J. Phys. Chem. B* 127(2), 430 (2023)
37. G. Bekçioğlu-Neff, C. Allolio, Y. S. Desmukh, M. R. Hansen, and D. Sebastiani, Dynamical dimension to the Hofmeister series: insights from first-principles simulations, *ChemPhysChem* 17(8), 1166 (2016)
38. S. Kim, X. Wang, J. Jang, K. Eom, S. L. Clegg, G. Park, and D. Di Tommaso, Hydrogen-bond structure and low-frequency dynamics of electrolyte solutions: hydration numbers from *ab initio* water reorientation dynamics and dielectric relaxation spectroscopy, *ChemPhysChem* 21(20), 2334 (2020)
39. P. B. Calio, C. Li, and G. A. Voth, Resolving the structural debate for the hydrated excess proton in water, *J. Am. Chem. Soc.* 143(44), 18672 (2021)
40. D. Tisi, L. Zhang, R. Bertossa, H. Wang, R. Car, and S. Baroni, Heat transport in liquid water from first-principles and deep neural network simulations, *Phys. Rev. B* 104(22), 224202 (2021)
41. Y. Shi, C. C. Doyle, and T. L. Beck, Condensed phase water molecular multipole moments from deep neural network models trained on *ab initio* simulation data, *J. Phys. Chem. Lett.* 12(42), 10310 (2021)
42. G. M. Sommers, M. F. Calegari Andrade, L. Zhang, H. Wang, and R. Car, Raman spectrum and polarizability of liquid water from deep neural networks, *Phys. Chem. Chem. Phys.* 22(19), 10592 (2020)
43. V. Kapil, D. P. Kovács, G. Csányi, and A. Michaelides, First-principles spectroscopy of aqueous interfaces using machine-learned electronic and quantum nuclear effects, *Faraday Discuss.* 249, 50 (2024)
44. C. Zhang, L. Zhang, J. Xu, F. Tang, B. Santra, and X. Wu, Isotope effects in X-ray absorption spectra of liquid water, *Phys. Rev. B* 102(11), 115155 (2020)
45. F. Tang, D. Y. Qiu, and X. Wu, Optical absorption spectroscopy probes water wire and its ordering in a hydrogen-bond network, *Phys. Rev. X* 15(1), 011048 (2025)
46. R. Mills, Self-diffusion in normal and heavy water in the range 1-45 Deg, *J. Phys. Chem.* 77(5), 685 (1973)
47. A. K. Soper and C. J. Benmore, Quantum differences between heavy and light water, *Phys. Rev. Lett.* 101(6), 065502 (2008)
48. J. A. Morrone and R. Car, Nuclear quantum effects in water, *Phys. Rev. Lett.* 101(1), 017801 (2008)
49. M. Ceriotti, G. Bussi, and M. Parrinello, Nuclear quantum effects in solids using a colored-noise thermostat, *Phys. Rev. Lett.* 103(3), 030603 (2009)
50. M. Ceriotti, G. Bussi, and M. Parrinello, Colored-noise thermostats à La Carte, *J. Chem. Theory Comput.* 6(4), 1170 (2010)
51. M. Ceriotti, G. Bussi, and M. Parrinello, Langevin equation with colored noise for constant-temperature molecular dynamics simulations, *Phys. Rev. Lett.* 102(2), 020601 (2009)
52. M. Ceriotti, W. Fang, P. G. Kusalik, R. H. McKenzie, A. Michaelides, M. A. Morales, and T. E. Markland, Nuclear quantum effects in water and aqueous systems: Experiment, theory, and current challenges, *Chem. Rev.* 116(13), 7529 (2016)
53. B. Cheng, E. A. Engel, J. Behler, C. Dellago, and M. Ceriotti, *Ab initio* thermodynamics of liquid and solid water, *Proc. Natl. Acad. Sci. USA* 116(4), 1110 (2019)
54. Y. Yao and Y. Kanai, Temperature dependence of nuclear quantum effects on liquid water via artificial neural network model based on SCAN meta-GGA functional, *J. Chem. Phys.* 153(4), 044114 (2020)
55. Y. Song and X. Wu, Pressure-induced structural and dielectric changes in liquid water at room temperature, *J. Chem. Phys.* 162(11), 114508 (2025)
56. A. K. Soper and K. Weckström, Ion solvation and water structure in potassium halide aqueous solutions, *Biophys. Chem.* 124(3), 180 (2006)
57. K. J. Tielrooij, N. Garcia-Araez, M. Bonn, and H. J. Bakker, Cooperativity in ion hydration, *Science* 328(5981), 1006 (2010)
58. Y. Chen, H. I. Okur, N. Gomopoulos, C. Macias-Romero, P. S. Cremer, P. B. Petersen, G. Tocci, D. M. Wilkins, C. Liang, M. Ceriotti, and S. Roke, Electrolytes induce long-range orientational order and free energy changes in the h-bond network of bulk water, *Sci. Adv.*

- 2(4), e1501891 (2016)
59. A. W. Omta, M. F. Kropman, S. Woutersen, and H. J. Bakker, Negligible effect of ions on the hydrogen-bond structure in liquid water, *Science* 301(5631), 347 (2003)
  60. I. Juurinen, T. Pylkkänen, K. O. Ruotsalainen, C. J. Sahle, G. Monaco, K. Hämäläinen, S. Huotari, and M. Hakala, Saturation behavior in X-ray Raman scattering spectra of aqueous LiCl, *J. Phys. Chem. B* 117(51), 16506 (2013)
  61. S. Funkner, G. Niehues, D. A. Schmidt, M. Heyden, G. Schwaab, K. M. Callahan, D. J. Tobias, and M. Havenith, Watching the low-frequency motions in aqueous salt solutions: The terahertz vibrational signatures of hydrated ions, *J. Am. Chem. Soc.* 134(2), 1030 (2012)
  62. A. P. Gaiduk and G. Galli, Local and global effects of dissolved sodium chloride on the structure of water, *J. Phys. Chem. Lett.* 8(7), 1496 (2017)
  63. W. Li and Z. Sun, Second hydration shell of  $Mg^{2+}$ : Competition between ion–water interaction and hydrogen bonding interaction, *J. Phys. Chem. Lett.* 16(2), 503 (2025)
  64. O. Blüh, Dielektrizitätskonstanten von elektrolytlösungen, *Z. Phys.* 25(1), 220 (1924)
  65. J. B. Hasted, D. M. Ritson, and C. H. Collie, Dielectric properties of aqueous ionic solutions. Parts I and II, *J. Chem. Phys.* 16(1), 1 (1948)
  66. J. H. Christensen, A. J. Smith, R. B. Reed, and K. L. Elmore, Dielectric properties of phosphoric acid solutions at 25° C, *J. Chem. Eng. Data* 11(1), 60 (1966)
  67. R. Buchner, G. T. Hefter, and P. M. May, Dielectric relaxation of aqueous NaCl solutions, *J. Phys. Chem. A* 103(1), 1 (1999)
  68. L. Zhang, H. Wang, M. C. Muniz, A. Z. Panagiotopoulos, and R. Car, Deep potential model with long-range electrostatic interactions, *J. Chem. Phys.* 156(12), 124107 (2022)
  69. C. Zhang, S. Yue, A. Z. Panagiotopoulos, M. L. Klein, and X. Wu, Why dissolving salt in water decreases its dielectric permittivity, *Phys. Rev. Lett.* 131(7), 076801 (2023)
  70. C. Zhang, S. Yue, A. Z. Panagiotopoulos, M. L. Klein, and X. Wu, Dissolving salt is not equivalent to applying a pressure on water, *Nat. Commun.* 13(1), 822 (2022)
  71. F. Wang, Y. Sun, and J. Cheng, Switching of redox levels leads to high reductive stability in water-in-salt electrolytes, *J. Am. Chem. Soc.* 145(7), 4056 (2023)
  72. C. Cao, A. Kingan, R. C. Hill, J. Kuang, L. Wang, C. Zhang, M. R. Carbone, H. Van Dam, S. Yoo, S. Yan, E. S. Takeuchi, K. J. Takeuchi, X. Wu, A. M. Abeykoon, A. C. Marschilok, and D. Lu, Resolving the solvation structure and transport properties of aqueous zinc electrolytes from salt-in-water to water-in-salt using neural network potential, *PRX Energy* 4(2), 023004 (2025)
  73. N. V. S. Avula, M. L. Klein, and S. Balasubramanian, Understanding the anomalous diffusion of water in aqueous electrolytes using machine learned potentials, *J. Phys. Chem. Lett.* 14(42), 9500 (2023)
  74. N. O’Neill, C. Schran, S. J. Cox, and A. Michaelides, Crumbling crystals: On the dissolution mechanism of NaCl in water, *Phys. Chem. Chem. Phys.* 26(42), 26933 (2024)
  75. N. O’Neill, B. X. Shi, K. Fong, A. Michaelides, and C. Schran, To pair or not to pair? Machine-learned explicitly-correlated electronic structure for NaCl in water, *J. Phys. Chem. Lett.* 15(23), 6081 (2024)
  76. J. Gu, J. Xiao, X. Wu, X. Zhu, H. Chen, J. Z. H. Zhang, T. Zhu, Y. Gao, and Z. Yin, Interaction of magnesium ion and acetate anion in bulk water: Toward high-level machine learning potential, *Chin. J. Chem. Phys.* 38(1), 95 (2025)
  77. Q. Zhang and T. Zhu, Interaction of divalent cations and amino acids in bulk water: Molecular simulations with neural network potentials, *Chin. J. Chem. Phys.* 36(2), 162 (2023)
  78. B. E. Logan and M. Elimelech, Membrane-based processes for sustainable power generation using water, *Nature* 488(7411), 313 (2012)
  79. M. Elimelech and W. A. Phillip, The future of seawater desalination: Energy, technology, and the environment, *Science* 333(6043), 712 (2011)
  80. H. Huang, Z. Song, N. Wei, L. Shi, Y. Mao, Y. Ying, L. Sun, Z. Xu, and X. Peng, Ultrafast viscous water flow through nanostrand-channelled graphene oxide membranes, *Nat. Commun.* 4(1), 2979 (2013)
  81. E. Secchi, S. Marbach, A. Niguès, D. Stein, A. Siria, and L. Bocquet, Massive radius-dependent flow slippage in carbon nanotubes, *Nature* 537(7619), 210 (2016)
  82. B. Radha, A. Esfandiar, F. C. Wang, A. P. Rooney, K. Gopinadhan, A. Keerthi, A. Mishchenko, A. Janardan, P. Blake, L. Fumagalli, M. Lozada-Hidalgo, S. Garaj, S. J. Haigh, I. V. Grigorieva, H. A. Wu, and A. K. Geim, Molecular transport through capillaries made with atomic-scale precision, *Nature* 538(7624), 222 (2016)
  83. A. Keerthi, S. Goutham, Y. You, P. Iamprasertkun, R. A. W. Dryfe, A. K. Geim, and B. Radha, Water friction in nanofluidic channels made from two-dimensional crystals, *Nat. Commun.* 12(1), 3092 (2021)
  84. C. Lee, C. H. Choi, and C. J. C. J. Kim, Structured surfaces for a giant liquid slip, *Phys. Rev. Lett.* 101(6), 064501 (2008)
  85. M. Majumder, N. Chopra, R. Andrews, and B. J. Hinds, Enhanced flow in carbon nanotubes, *Nature* 438(7064), 44 (2005)
  86. S. K. Kannam, B. D. Todd, J. S. Hansen, and P. J. DAVIS, How fast does water flow in carbon nanotubes, *J. Chem. Phys.* 138(9), 094701 (2013)
  87. G. Cicero, J. C. Grossman, E. Schwegler, F. Gygi, and G. Galli, Water confined in nanotubes and between graphene sheets: A first principle study, *J. Am. Chem. Soc.* 130(6), 1871 (2008)
  88. F. L. Thiemann, C. Schran, P. Rowe, E. A. Müller, and A. Michaelides, Water flow in single-wall nanotubes: Oxygen makes it slip, hydrogen makes it stick, *ACS Nano* 16(7), 10775 (2022)
  89. R. Wang, M. DelloStritto, M. L. Klein, E. Borguet, and V. Carnevale, Topological properties of interfacial hydrogen bond networks, *Phys. Rev. B* 110(1), 014105 (2024)



90. W. Pu and Z. Sun, Synergistic effects of confinement size and interface on anomalous ultrafast transport in nanofluidics, *Commun. Phys.* 8(1), 167 (2025)
91. M. J. Holmes, N. G. Parker, and M. J. W. Povey, Temperature dependence of bulk viscosity in water using acoustic spectroscopy, *J. Phys. Conf. Ser.* 269, 012011 (2011)
92. P. Först, F. Werner, and A. Delgado, The viscosity of water at high pressures – especially at subzero degrees centigrade, *Rheol. Acta* 39(6), 566 (2000)
93. M. Neek-Amal, F. M. Peeters, I. V. Grigorieva, and A. K. Geim, Commensurability effects in viscosity of nanoconfined water, *ACS Nano* 10(3), 3685 (2016)
94. H. G. Kalashami, M. Neek-Amal, and F. M. Peeters, Slippage dynamics of confined water in graphene oxide capillaries, *Phys. Rev. Mater.* 2(7), 074004 (2018)
95. W. Xiong, J. Z. Liu, M. Ma, Z. Xu, J. Sheridan, and Q. Zheng, Strain engineering water transport in graphene nanochannels, *Phys. Rev. E* 84(5), 056329 (2011)
96. N. Wei, X. Peng, and Z. Xu, Breakdown of fast water transport in graphene oxides, *Phys. Rev. E* 89(1), 012113 (2014)
97. K. Falk, F. Sedlmeier, L. Joly, R. R. Netz, and L. Bocquet, Molecular origin of fast water transport in carbon nanotube membranes: Superlubricity versus curvature dependent friction, *Nano Lett.* 10(10), 4067 (2010)
98. W. Zhao, H. Qiu, and W. Guo, A deep neural network potential for water confined in graphene nanocapillaries, *J. Phys. Chem. C Nanomater. Interfaces* 126(25), 10546 (2022)
99. V. Kapil, C. Schran, A. Zen, J. Chen, C. J. Pickard, and A. Michaelides, The first-principles phase diagram of monolayer nanoconfined water, *Nature* 609(7927), 512 (2022)
100. T. Dufils, C. Schran, J. Chen, A. K. Geim, L. Fumagalli, and A. Michaelides, Origin of dielectric polarization suppression in confined water from first principles, *Chem. Sci. (Camb.)* 15(2), 516 (2024)
101. K. D. Fong, C. P. Grey, and A. Michaelides, On the physical origins of reduced ionic conductivity in nanoconfined electrolytes, *ACS Nano* 19(13), 13191 (2025)
102. K. D. Fong, B. Sumić, N. O'Neill, C. Schran, C. P. Grey, and A. Michaelides, The interplay of solvation and polarization effects on ion pairing in nanoconfined electrolytes, *Nano Lett.* 24(16), 5024 (2024)
103. P. Li, Y. L. Jiang, Y. Men, Y. Z. Jiao, and S. Chen, Kinetic cation effect in alkaline hydrogen electrocatalysis and double layer proton transfer, *Nat. Commun.* 16(1), 1844 (2025)
104. C. Zhang, M. F. Calegari Andrade, Z. K. Goldsmith, A. S. Raman, Y. Li, P. M. Piaggi, X. Wu, R. Car, and A. Selloni, Molecular-scale insights into the electrical double layer at oxide-electrolyte interfaces, *Nat. Commun.* 15(1), 10270 (2024)
105. J. X. Zhu and J. Cheng, Machine learning potential for electrochemical interfaces with hybrid representation of dielectric response, *Phys. Rev. Lett.* 135(1), 018003 (2025)
106. Y. Wang, H. Luo, X. R. Advincula, Z. Zhao, A. Esfandiari, D. Wu, K. D. Fong, L. Gao, A. S. Hazrah, T. Taniguchi, C. Schran, Y. Nagata, L. Bocquet, M. L. Bocquet, Y. Jiang, A. Michaelides, and M. Bonn, Spontaneous surface charging and Janus nature of the hexagonal boron nitride–water interface, *J. Am. Chem. Soc.* 147(33), 30107 (2025)
107. J. Lian, S. Jiao, W. Yin, and K. Zhou, Machine-learning potential molecular dynamics reveals the critical role of flexibility in solid–liquid nanofluidic friction, *ACS Nano* 19(36), 32422 (2025)
108. X. R. Advincula, K. D. Fong, A. Michaelides, and C. Schran, Protons accumulate at the graphene–water interface, *ACS Nano* 19(18), 17728 (2025)
109. D. Muñoz-Santiburcio, Accurate diffusion coefficients of the excess proton and hydroxide in water via extensive *ab initio* simulations with different schemes, *J. Chem. Phys.* 157(2), 024504 (2022)
110. A. Gomez, W. H. Thompson, and D. Laage, Neural-network-based molecular dynamics simulations reveal that proton transport in water is doubly gated by sequential hydrogen-bond exchange, *Nat. Chem.* 16(11), 1838 (2024)
111. M. de la Puente, A. Gomez, and D. Laage, Neural network-based sum-frequency generation spectra of pure and acidified water interfaces with air, *J. Phys. Chem. Lett.* 15(11), 3096 (2024)
112. Y. Zhao, F. Tian, and Z. Sun, Ab initio deep neural network simulations reveal that carbonic acid dissociation is dominated by minority cis-trans conformers, *Sci. Adv.* 11(19), eadu6525 (2025)
113. M. de la Puente, A. Gomez, and D. Laage, Why proton grotthuss diffusion slows down at the air–water interface while water diffusion accelerates, *J. Phys. Chem. Lett.* 16(10), 2645 (2025)
114. M. de la Puente, R. David, A. Gomez, and D. Laage, Acids at the edge: Why nitric and formic acid dissociations at air–water interfaces depend on depth and on interface specific area, *J. Am. Chem. Soc.* 144(23), 10524 (2022)
115. M. de la Puente, and D. Laage, How the acidity of water droplets and films is controlled by the air–water interface, *J. Am. Chem. Soc.* 145(46), 25186 (2023)
116. M. Yang, L. Bonati, D. Polino, and M. Parrinello, Using metadynamics to build neural network potentials for reactive events: The case of urea decomposition in water, *Catal. Today* 387, 143 (2021)
117. S. G. H. Brookes, V. Kapil, A. Michaelides, and C. Schran, CO<sub>2</sub> hydration at the air–water interface: A surface-mediated “in-and-out” mechanism, *Proc. Natl. Acad. Sci. USA* 122(34), e2502684122 (2025)
118. M. Galib and D. T. Limmer, Reactive uptake of N<sub>2</sub>O<sub>5</sub> by atmospheric aerosol is dominated by interfacial processes, *Science* 371(6532), 921 (2021)
119. A. S. Raman and A. Selloni, Modeling the solvation and acidity of carboxylic acids using an *ab initio* deep neural network potential, *J. Phys. Chem. A* 126(40), 7283 (2022)
120. R. David, I. Tuñón, and D. Laage, Competing reaction mechanisms of peptide bond formation in water revealed by deep potential molecular dynamics and path sampling, *J. Am. Chem. Soc.* 146(20), 14213

- (2024)
121. T. Frank, K. R. Müller, and O. Unke, So3krates: Equivariant attention for interactions on arbitrary length-scales in molecular systems, in: *Advances in Neural Information Processing Systems 35*, Neural Information Processing Systems Foundation, Inc. (NeurIPS): New Orleans, Louisiana, USA, 2022, pp 29400–29413
  122. Y. Ji, J. Liang, and Z. Xu, Machine-learning interatomic potentials for long-range systems, *Phys. Rev. Lett.* 135(17), 178001 (2025)
  123. W. Ren, W. Fu, X. Wu, and J. Chen, Towards the ground state of molecules via diffusion Monte Carlo on neural networks, *Nat. Commun.* 14(1), 1860 (2023)
  124. Z. Tang, H. Chen, Y. Li, Y. Qian, Y. Wang, W. Fu, J. Li, C. Si, W. Duan, J. Chen, and Y. Xu, Deep-learning electronic structure calculations, *Nat. Comput. Sci.* 5(12), 1133 (2025)
  125. Z. Huang, Z. Guo, C. Cao, H. Q. Pham, X. Wen, G. H. Booth, J. Chen, and D. Lv, A multi-resolution systematically improvable quantum embedding scheme for large-scale surface chemistry calculations, *Nat. Commun.* 16(1), 9297 (2025)
  126. A. Mukhopadhyay, S. S. Xantheas, and R. J. Saykally, The water dimer II: Theoretical investigations, *Chem. Phys. Lett.* 700, 163 (2018)
  127. A. Zen, Y. Luo, G. Mazzola, L. Guidoni, and S. Sorella, *Ab initio* molecular dynamics simulation of liquid water by quantum Monte Carlo, *J. Chem. Phys.* 142(14), 144111 (2015)
  128. J. Guo, L. Zhou, A. Zen, A. Michaelides, X. Wu, E. Wang, L. Xu, and J. Chen, Hydration of  $\text{NH}_4^+$  in water: Bifurcated hydrogen bonding structures and fast rotational dynamics, *Phys. Rev. Lett.* 125(10), 106001 (2020)
  129. A. Nandi, P. Pandey, P. L. Houston, C. Qu, Q. Yu, R. Conte, A. Tkatchenko, and J. M. Bowman,  $\Delta$ -machine learning to elevate DFT-based potentials and a force field to the CCSD(*T*) level illustrated for ethanol, *J. Chem. Theory Comput.* 20(20), 8807 (2024)
  130. V. Vinod, S. Maity, P. Zaspel, and U. Kleinekathöfer, Multifidelity machine learning for molecular excitation energies, *J. Chem. Theory Comput.* 19(21), 7658 (2023)
  131. D. Zhang, X. Liu, X. Zhang, C. Zhang, C. Cai, H. Bi, Y. Du, X. Qin, A. Peng, J. Huang, B. Li, Y. Shan, J. Zeng, Y. Zhang, S. Liu, Y. Li, J. Chang, X. Wang, S. Zhou, J. Liu, X. Luo, Z. Wang, W. Jiang, J. Wu, Y. Yang, J. Yang, M. Yang, F. Q. Gong, L. Zhang, M. Shi, F. Z. Dai, D. M. York, S. Liu, T. Zhu, Z. Zhong, J. Lv, J. Cheng, W. Jia, M. Chen, and G. Ke, DPA-2: A large atomic model as a multi-task learner, *npj Comput. Mater.* 10(1), 293 (2024)
  132. T. Liang, K. Xu, E. Lindgren, Z. Chen, R. Zhao, J. Liu, E. Berger, B. Tang, B. Zhang, Y. Wang, K. Song, P. Ying, N. Xu, H. Dong, S. Chen, P. Erhart, Z. Fan, T. Ala-Nissila, and J. Xu, NEP89: Universal neuroevolution potential for inorganic and organic materials across 89 elements. arXiv: 2504.21286 (2025)
  133. S. S. Chong, Y. S. Ng, H. Q. Wang, and J. C. Zheng, Advances of machine learning in materials science: Ideas and techniques, *Front. Phys. (Beijing)* 19(1), 13501 (2024)