Exploring complex reaction networks using neural network-based molecular dynamics simulation

Qingzhao Chu[a,b], Kai H. Luo[c], Dongping Chen[a,b,*]

[a] *State Key Laboratory of Explosion Science and Technology, Beijing Institute of Technology, Beijing 100081, China*

[b] *Explosion Protection and Emergency Disposal Technology Engineering Research Center of the Ministry of Education, Beijing, 100081, China*

[c] *Department of Mechanical Engineering, University College London, Torrington Place, London WC1E 7JE, UK*

**Corresponding author: dc516@bit.edu.cn*

Abstract

Ab initio molecular dynamics (AIMD) is an established method to reveal the reactive dynamics of complex systems. However, the computational cost of AIMD restricts the explorable length and time scales to a great extent. Here, we develop a fundamentally different approach using molecular dynamics simulations powered by a neural network potential to investigate complex reaction networks. This potential is trained via a workflow combining AIMD and interactive molecular dynamics in virtual reality (VRMD) to accelerate the sampling of a rare reactive process. The capability of the methodology is demonstrated by achieving a panoramic visualization of the complex reaction networks for decomposition of a novel high explosive (ICM-102), without any predefined reaction coordinates. The study leads to the discovery of new pathways that would be difficult to uncover employing established methods. These results highlight the power of neural network-based molecular dynamics simulations for exploration of complex reaction mechanisms under extreme conditions at the *ab initio* level, pushing the limit of theoretical and computational chemistry towards the realism and fidelity of experiments.

Keywords: Energetic Material; Reaction Kinetics and Dynamics; Machine Learning; Ab initio Molecular Dynamics; Virtual Reality; Reaction Network

Introduction

High explosives (HEs) have contributed enormously to the prosperity of humankind since the invention of black powder in China and the era of nitroglycerin brought by Alfred Nobel[1,2]. In an HE, the chemical energy stored in the bonds of the explosive material is converted into kinetic energy of the gaseous products via chemical reactions. A quantitative as well as qualitative understanding of the complex reaction network is critical to the application of HE. Such a reaction network involves thousands of elementary reactions that take place over a wide range of time scales. In addition, the reactions usually occur in extreme conditions (*i.e.,* > 10 GPa and > 3000 K[3]), which makes it very difficult or even impossible to obtain the detailed reaction mechanisms of HE materials using experimental or conventional simulation approaches.

In recent decades, *ab initio* molecular dynamics (AIMD) simulations[4] have been applied to gain atomic insights into HE materials[4–10]. In an AIMD simulation, an HE material is represented by an atomic model with interatomic forces determined by electronic structure calculations, and the Newtonian equations of motion are solved to obtain the dynamic trajectories. Thus, AIMD allows chemical reactions (i.e. bond breaking and forming events) to occur and accounts for electronic polarization effects. The initiation reaction of HE involves coupled processes subject to mechanical, thermal, and chemical stimulation. With the AIMD method, the response of HE material to stimulation and the subsequent reaction mechanism can, in principle, be captured. AIMD has been applied to investigate the shock-induced mechanical response and

initiation of a detonation in numerous studies, providing fundamental insight into the microscopic mechanisms of such complex phenomena. However, the computational cost of AIMD is so high that the accessible system size and time scale in simulations are limited to several hundred atoms and dozens of picoseconds, respectively.

Recently, artificial neural networks (NNs) have been applied to construct potential energy surfaces (PESs) in a fully data-driven manner, where the PES is abstracted from a well-selected training dataset using suitable functional expressions[11]. Several formulations, such as DeepMD[12], GAP[13], sGDML[14], and SchNet[15], have been proposed to develop NN potentials and have achieved success in the modeling of water[16], small organic molecules[14], and metal materials[17]. The performance of NN potential depends on the completeness of the training dataset[18]. In other words, NN does well in finding solutions in the function space of the training dataset but might fail in configurations outside the dataset. Therefore, the quality of the dataset is critical to developing NN potentials, and it is recommended to include all the critical configurations during potential reaction processes[19]. MD sampling is the most straightforward way to construct the training dataset, and the evolution of energies and forces are recorded as training datasets[20]. Various bond dissociations and recombinations exist in a reactive system, e.g., the decomposition of HE materials. Such processes cannot be sampled appropriately in classical MD sampling due to the high energy barriers. Enhanced sampling techniques, such as metadynamics[21,22],[19], could improve the sampling quality in the potential reaction coordinates. The concept

of interactive molecular dynamics in virtual reality (VRMD) was first proposed by Glowacki et al. [23] and applied to sample the PES of a simple hydrogen abstraction reaction[24,25]. This method has been proven to accelerate the intelligent curation of high-quality datasets. In a VRMD simulation, VR forces are implemented on atoms to accelerate the reaction processes, similar to metadynamics.

In this paper, we develop a neural network-based molecular dynamics (NNMD) method for investigation into the reaction dynamics of a novel high explosive (ICM-102)[26]. The training dataset is derived from both AIMD and VRMD simulations. The configuration spaces of datasets from different sampling methods are compared to yield a good NN potential. A set of MD simulations is performed to explore the elementary reactions in the decomposition of ICM-102 molecules. A panoramic visualization of the complex reaction networks is abstracted, and new pathways are identified from atomic trajectories for the first time.

Computational Methods

Workflow of VR-enhanced NN potential

The full dataset consists of configurations sampled from AIMD and VRMD methods. A detailed description of the training dataset is shown in Table S1. As shown in Fig. 1, the following training procedures are performed:

(i) Performing AIMD simulation in the NVT ensemble at multiple temperatures (300, 1000, 2000, 3000, 4000 K);

(ii) Selecting additional configurations with an active learning strategy;

(iii) Extracting the configurations of potential chemical reactions with the VRMD method.

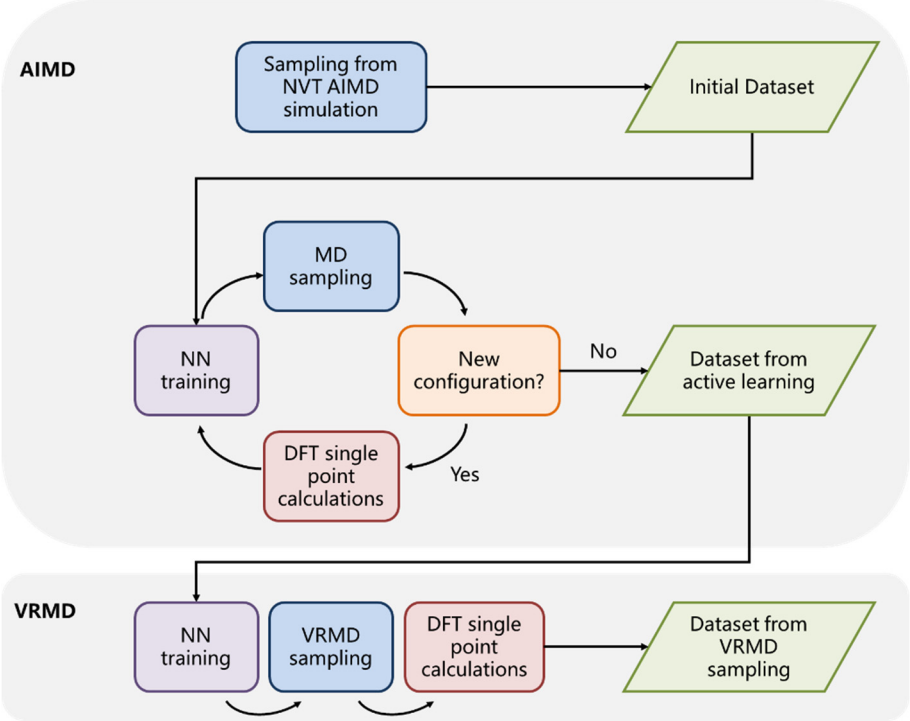


Fig. 1 Schematic illustration of the workflow for NN potential training.

DFT and AIMD Calculations

DFT calculations were performed using the CP2K package[27]. Core electrons were treated using Goedecker−Teter−Hutter (GTH) pseudopotentials and the Perdew Burke Ernzerhof generalized gradient approximation method[28,29]. The Grimme DFT-D3 method[30] was used to account for dispersion interactions. A double-zeta Gaussian basis set plus polarization (DZVP-MOLOPT)[31] was considered. AIMD calculations were performed for an ICM-102 system of 160 atoms using the Quickstep module in CP2K. In both cases, simulations were carried out at constant volume and temperature conditions with periodic boundary conditions. An integration time step of 0.5 fs was used.

NN descriptor and training

We trained NN potentials using the DeePMD-kit package[32,33]. The smooth version of the deep potential model is adopted with a cut-off radius of 6.0 Å[32]. To remove the discontinuity introduced by the cut-off, the *1/r* term in the network construction is smoothly switched off by a cosine shape function from 1.0 Å to 6.0 Å. The filter (embedding) network has three layers with (25, 50, 100) nodes/layer, and the fitting net is composed of three layers, with 240 nodes each. The network is trained with the ADAM optimizer, with an exponentially decaying learning rate from 1.0×10^{-3} to 5.0×10^{-8}. During the optimization process, the pre-factors in the loss function change from 1 to 10 and 1000 to 1 for the energy and force terms, respectively. The final NN model used for the production run was trained for 1.0×10^{6} steps. Additional configurations were obtained using an active learning sampling procedure as implemented in the DP-GEN package[34]. Four NN potentials were trained on the same dataset with different initializations of weights and biases. NVT-MD simulations were performed at multiple temperatures (300, 3000, and 4000 K) using the LAMMPS package[35]. By comparing the structures from MD simulations, the agreement on the force predictions made by these potentials is used to select new configurations. When the deviation of NN prediction for one configuration was in the range of [0.4, 0.8] eV/Å, the corresponding structure was labeled as a candidate for the training set. The upper limit is used to filter nonphysical configurations[19].

VRMD sampling with Manta

Figure 2 illustrates the basic architecture, including multiple VR clients, a network server, and an MD simulator. The VR client renders the virtual reality environment and captures the user's actions using the Unity3D engine. The network server enables a mode of cloud computing to control information flow between the local VR client and the MD simulator hosted on a high-performance cluster. The MD simulator empowers Manta to conduct the real-time calculations of NNMD and integrate the interactive forces from the VR clients. Details of Manta implementation can be found in supplementary materials (Fig. S1). With Manta, users could explore the evolution of molecular structures along their intuitionistic reaction pathways, and corresponding configurations are recorded for further analysis.

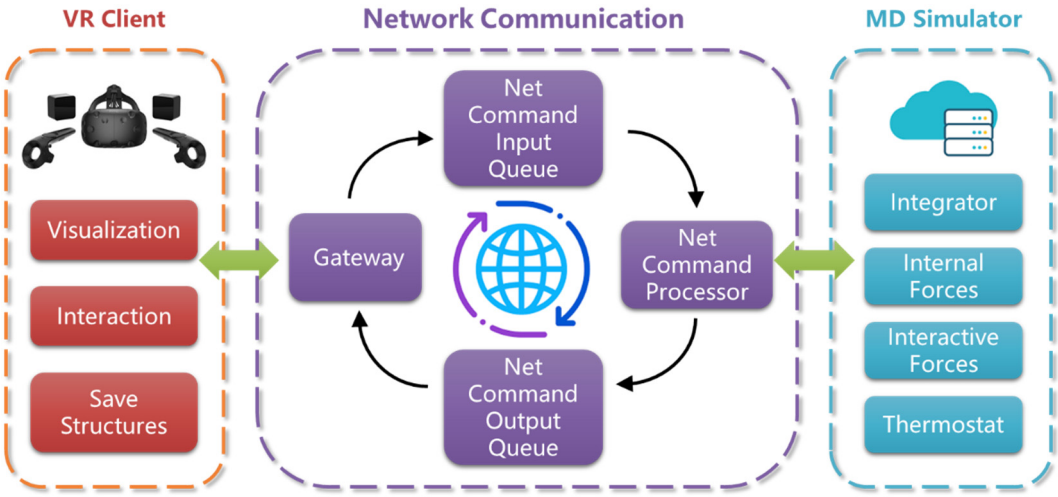


Fig. 2 General architecture of Manta. The core of Manta is built on an MD simulator solving the molecular interactions from internal and interactive forces. The front end of the VR client includes multiple VR helmets/handles allowing molecular visualization and interactions. A gateway and net command processor control the network

communication between the MD simulator and the VR client.

Reaction network

The reaction network was postprocessed using ReacNetGenerator by Zeng et al.[36] from the NNMD trajectories. All the species in the reaction network are clustered by their fingerprints using the scikit-learn package[37]. The fingerprints are defined by the molecular structure and calculated by RDKit[38]. The importance of each species is calculated according to the observed number that it occurs and is represented by the size of the dot.

Results and discussion

Sampling the PES with VRMD

The decomposition of ICM-102 involves complex PESs, and VRMD is applied as an enhanced sampling method to potential reactions. To illustrate the effect of VR forces on the sampling processes, we first sample the reaction of a single ICM-102 molecule. As shown in Fig. 3a and 1b, the ICM-102 molecule is visualized in the helmet screen. The user interacts with atoms using handles, where the movement of handles is translated as VR forces applied on atoms. In Fig. 3c, we select a reaction that involves the breakage of a C-N bond and the formation of a C-O bond. This reaction is a rare event hindered by the high energy barrier. As illustrated in Fig 3d, the PES from the initial state to the transition state is steep. With the aid of VR forces, the reactants overcome the high energy and undergo C-N bond scission. The full PES of this reaction

is plotted along the reaction coordinates in Fig. 3e; the lengths of the C-N and C-O bonds are represented by d1 and d2, respectively. There is a local minimum as the initial state in Fig. 3c at d1=1.40 Å and d2=2.28 Å. The other local minimum sits at d1=2.30 Å and d2=1.36 Å, corresponding to the final state (Fig. 3c). The energy barrier corresponds to 138 kcal/mol. Several AIMD simulations are further performed for 10 ps to sample the PES at 300-4000 K. It is found that most AIMD trajectories are distributed near the initial state. As the temperature increases, the distribution of configurations becomes wider. When the temperature exceeds 3000 K, a few configurations near the transition state and final state are sampled. The trajectories from VRMD are also mapped in Fig. 3e, which samples the whole reaction process. The above comparison illustrates that AIMD sampling is mainly trapped in low-energy regions, resulting in a limited number of reaction configurations. VRMD sampling could obtain the reaction event guided by human intuition that almost follows the minimum energy path (MEP).

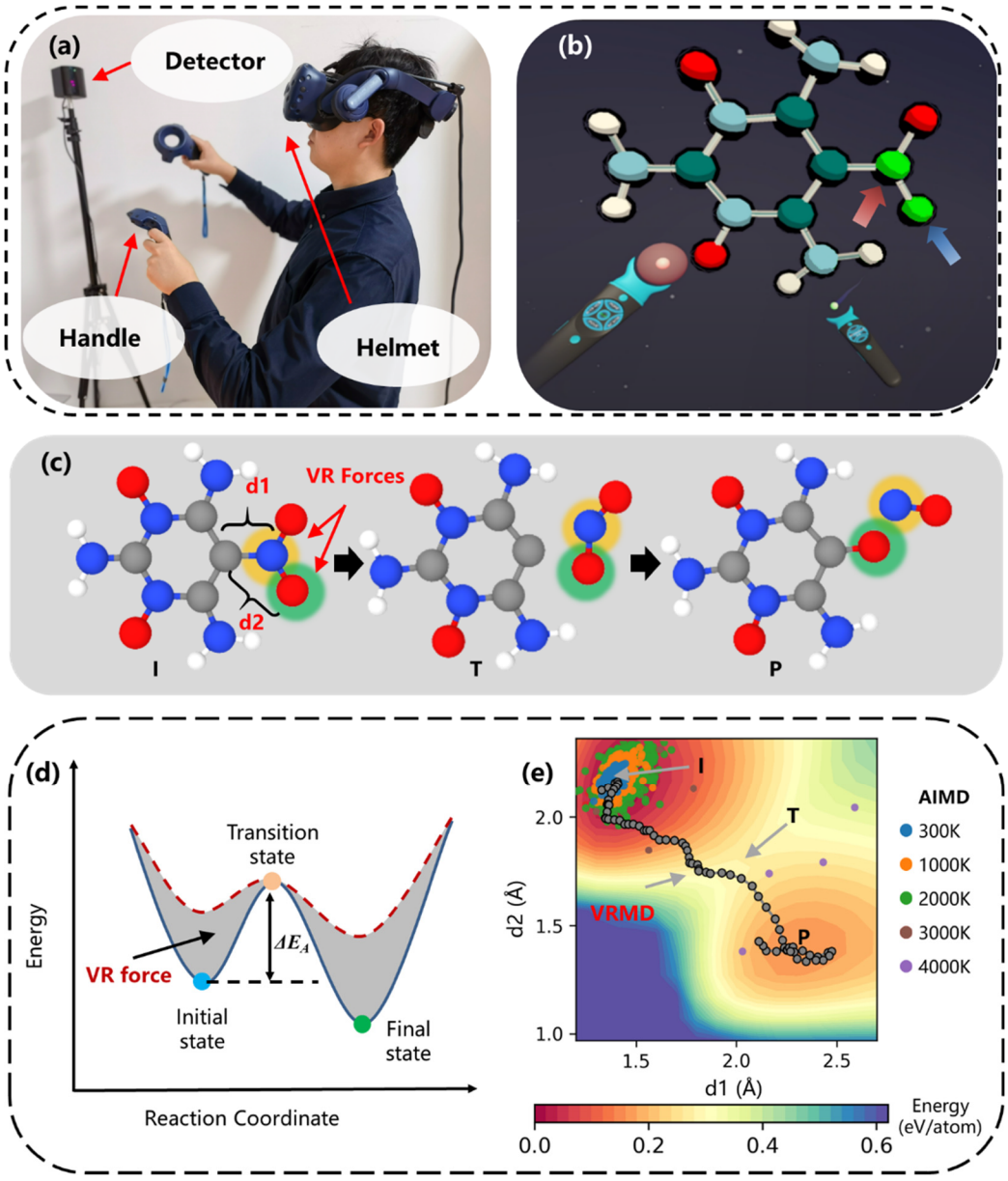


Fig. 3 The VRMD sampling process. (a) The hardware for VRMD, including one helmet, two handles, and two detectors. (b) Snapshots of a single ICM-102 molecule in a VRMD environment. Red and blue arrows represent the moving direction of handles. (c) The reaction pathway involves the breakage of the N-O bond and the formation of the CO bond. Yellow and green halos represent VR forces applied by handles. (d) Illustration of VR forces applied on the potential energy surface. (e) Trajectory visualization sampled by VRMD (gray dots). The background is the potential energy

surface projected on the reaction coordinates of d1 and d2. The AIMD trajectories at temperatures of 300, 1000, 2000, 3000 and 4000 K are highlighted by blue, orange, green, red, and purple dots, respectively.

Dataset exploration

The performance of NN potential depends on the quality of the training dataset, which describes the chemical space of PES[18,39]. The whole training dataset of bulk ICM-102 molecules is constructed using both AIMD and VRMD sampling methods, and the detailed configurations are listed in Table S1. The structural landscapes of ICM-102 decomposition are shown in Fig. 4a, where the molecular configurations are projected onto their first two principal components (PCs). The full configurations can be divided into crystal, partial decomposition, and gas species. The most common structure is the crystalline phase of ICM-102, represented by a sheet-like structure. As the decomposition reaction's starting point, the crystal group has the lowest potential energies, as expected. Along with the 1st principal component (PC1), crystalline ICM-102 gradually turns into states undergoing partial decomposition. The molecules become irregular under thermal stimulation, and the loss of hydrogen atoms is observed. Further examining the configurations along PC1, small gas molecules, *i.e.,* H_2O, NO, NH_2, HCN, are observed, suggesting that ICM-102 molecules have been entirely decomposed. Figure 4b shows the average molecular masses along PC1. When PC1 < -2, the average molecular mass is 202 a.u., corresponding to the mass of the ICM-102 molecule. As PC1 increases, the molecular mass gradually decreases, indicating the

start of the decomposition reaction. Therefore, we conclude that PC1 refers to the overall reaction coordinate, where the initial stable reactants gain energy to overcome energy barriers and reach a stable product state. Then, the distributions of subsets are compared. The first subset is, in fact, obtained from AIMD simulations at 300-4000 K, combined with an active learning sampling strategy[40]. An additional dataset of VRMD is constructed from the intuition of the first author using an in-house VRMD simulator (Manta). It allows researchers to explore the configuration space of chemical reactions and performs expert-biased enhanced sampling on the targeted PES. In the VRMD dataset, reaction pathways include the transfer of hydrogen atoms between ICM-102 molecules, ring-opening reactions and the formation of gas molecules (Fig. S2b). A detailed distribution of each subset along PC1 and energy is included in Fig. S2. Similar to Fig. 3e, the range of the AIMD dataset increases with temperature, and the VRMD datasets cover the full range of PC1. Figure 4c shows the distribution density of configurations along PC1. The configurations of the AIMD dataset cluster in the range of [-2, 0], and high-energy configurations are less sampled due to the energy barriers. In other words, most of the configurations sampled by AIMD simulations are nonreactive (~70%). In contrast, the distribution of the VRMD dataset has two peaks. In particular, it includes more structures with high energies, which largely extends the boundary of PES sampling. Then, NN potentials are trained on different dataset configurations. Fig. S3 shows that the NN potentials trained on the AIMD+VRMD datasets have lower mean absolute errors (MAEs) than potentials trained on only the

AIMD dataset. This demonstrates that the final NN potential can perfectly reproduce the energy and forces in AIMD trajectories (Fig. S4). The configurations sampled by VRMD can be a good supplement for AIMD samples to create a balanced training set for model training.

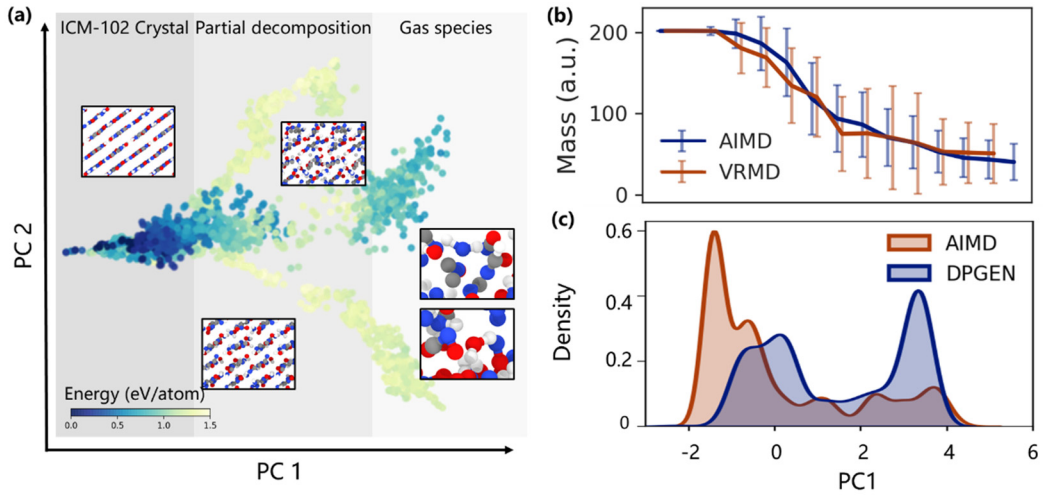


Fig. 4 Exploring the datasets of AIMD and VRMD. (a) Landscapes of the AIMD dataset for the decomposition of ICM-102 with principal component analysis (PCA). Each dot represents a configuration of 160 atoms. The colors refer to the atomic energy values. The inserts indicate representative snapshots of configurations. (b) The average molecular mass along the PC1. The vertical bars represent the standard deviations. (c) The distribution of AIMD and VRMD datasets.

Reaction network of ICM-102 decomposition

To capture the reactive dynamics of ICM-102 molecules, 100 ps MD simulations were performed using the new NN potential (*e.g.*, case 4 in Table S2) with 64 ICM-102 molecules at 3000 K. The species evolution during ICM-102 decomposition predicted by NN and AIMD are compared in Fig. S5. It is clear that NN agrees well with the

AIMD results, where ICM-102 molecules are rapidly consumed, and H_2O is the main product. With the application of NN potential, the system size could be extended to thousands or even millions of atoms[41,42]. Compared to the large oscillation in the predicted species evolution of AIMD, NNMD simulates an eight-times larger system, resulting in smooth species evolution. Such a large system allows studies on the macroscopic nature of a statistical ensemble, which is beyond the ability of the traditional AIMD method.

Figure 5 shows the decomposition of 64 ICM-102 molecules. The simulation reaches equilibrium after ~50 ps, where the main species include $C_4H_5N_6O_4$ (ICM-102), $C_4H_5N_6O_4$, H_2O, NO, N_2, CHNO, and CO_2. The atomic trajectory is visualized by identifying the new product and highlighting these molecules in molecule-specific colors. As shown in Fig. 5a, most ICM-102 molecules are consumed within the first 1 ps, and H_2O and NO are subsequently observed. These species can be produced by direct bond dissociation from ICM-102 molecules. The formation of N_2, CHNO, and CO_2 requires ~5 ps to occur, and this cannot be observed from the short AIMD simulations, suggesting that the kinetics derived from AIMD might not represent the full reactive image (Fig. S5).

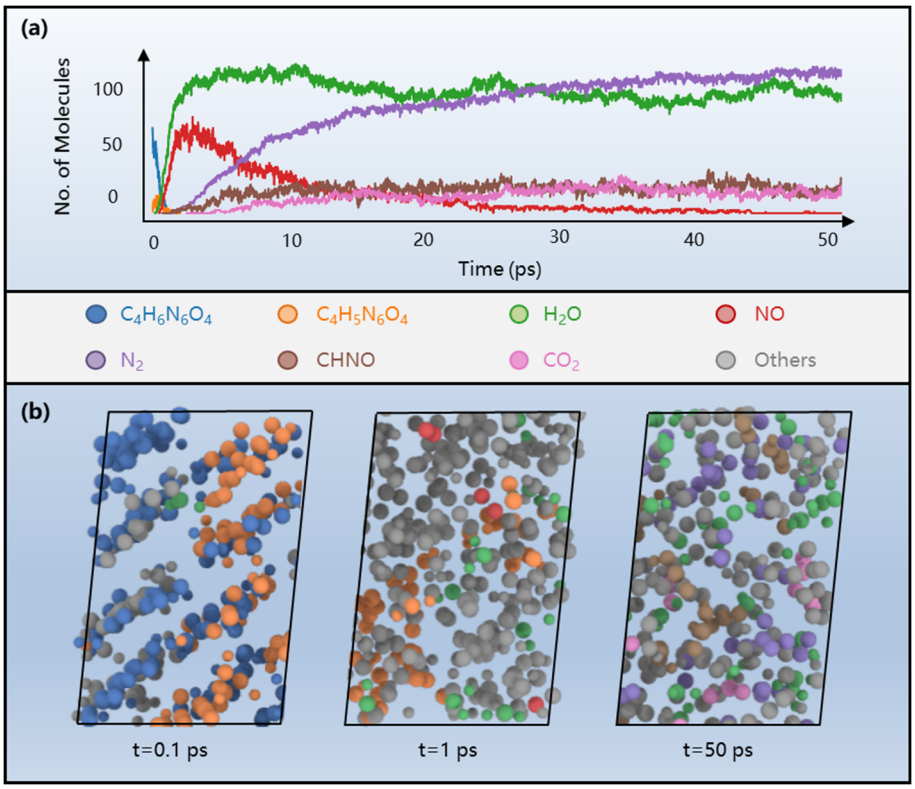


Fig. 5 Evolution of an NNMD simulation. (a) Evolution of major species. (b) Local snapshots of the formation of key species (only 1/8 of the system is shown). Left: The simulation begins with the hydrogen abstraction of ICM-102 molecules (blue) to form $C_4H_5N_6O_4$ (orange). Middle: small gas molecules exist (H_2O, green; NO, red). Right: At longer simulation times, other gas species are observed (N_2, purple; CHNO, brown; CO_2, pink).

In total, 5799 products are identified from the NNMD trajectories. Direct derivation of the reaction mechanism from thousands of products can be quite challenging. We propose an interactive reaction network to illustrate the detailed reaction mechanism. In Fig. 6, the full reaction network contains detailed pathways and intermediates. These intermediates are represented by fingerprints[38] and clustered into eight groups using the k-means clustering algorithm[43]. Groups 1 and 2 represent

ICM-102 and molecules produced by releasing H·, O· and OH· radicals, respectively. Groups 3 and 4 share similar elemental compositions as groups 1 and 2 but with broken C-N rings. The decomposition of groups 5, 6, and 7 produces intermediates with a low H/O ratio (see Figure S6). Group 8 represents small gas products, including CHNO, H_2O, OH·, NO, and N_2. These species are frequently involved in the reaction network as the product of many reaction pathways. There are two pathways for reactants to release the final gas products: group 1 => 2 => 8; group 1 => 3 and 4 => 5 and 6 => 7 => 8. The first pathway corresponds to reactions in early initiation (<1 ps) and produces molecules/radicals such as H_2O, NO, OH· and H·. The second pathway requires a higher energy barrier and mainly occurs during 5-50 ps.

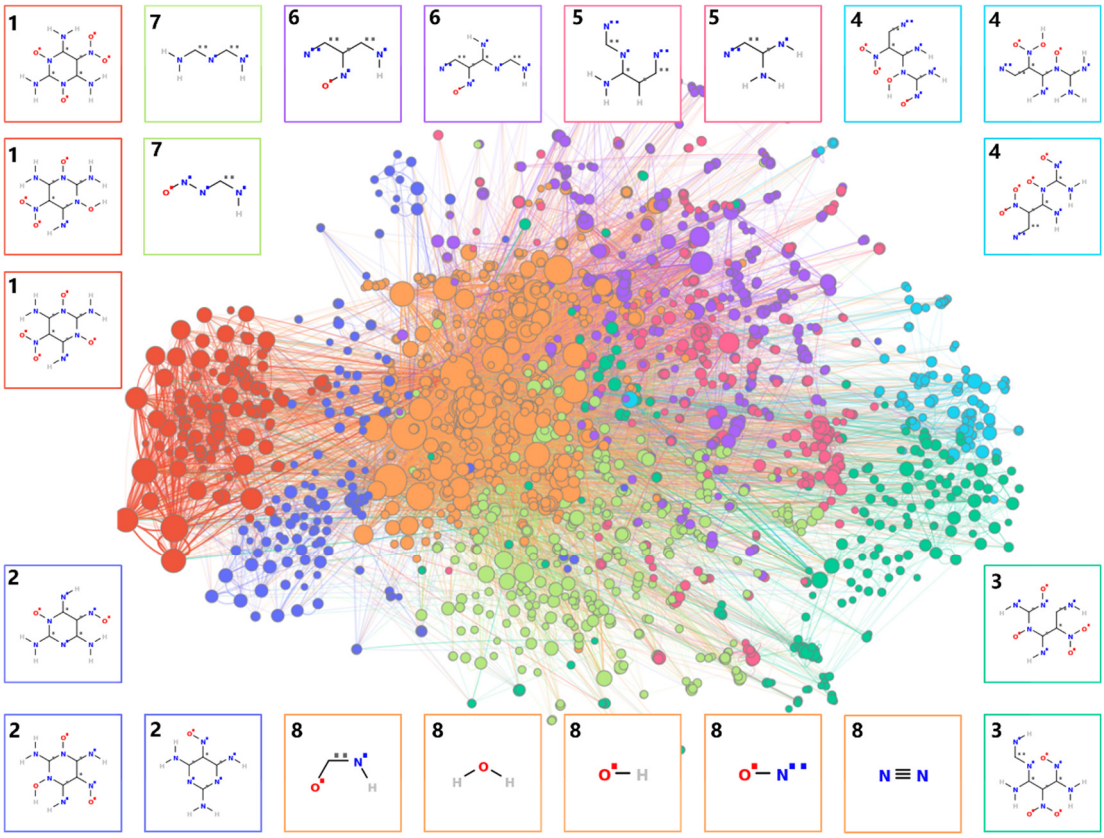


Fig. 6 A panoramic visualization of the ICM-102 decomposition. Each dot and line

represent a species and a reaction between species, respectively. The dot size is proportional to the observed frequency of species in the network. Only species with a frequency higher than five are shown. The inserts show the representative molecular structure in each group.

Derived from the overall ICM-102 decomposition (Fig. 6), the primary reaction pathways of ICM-102 and the formation of H_2O are constructed in Fig. 7. ICM-102 decomposition starts with intermolecular and intramolecular H transfer to form R-OH species (R1 and R7). These species further decompose into radicals such as OH· and H· (R2 and R10). The NO_2 molecule also abstracts a hydrogen atom from ICM-102 to form HNO_2 (R8), which further dissociates as an OH· radical (R9). The combination of OH· and H· radials forms water molecules (R6). The R-OH structures could also undergo ring-opening reactions by C-N bond scission (R3), and this reaction is not preferred due to the high bond energy (305 kJ/mol for C-N bonds[44]).

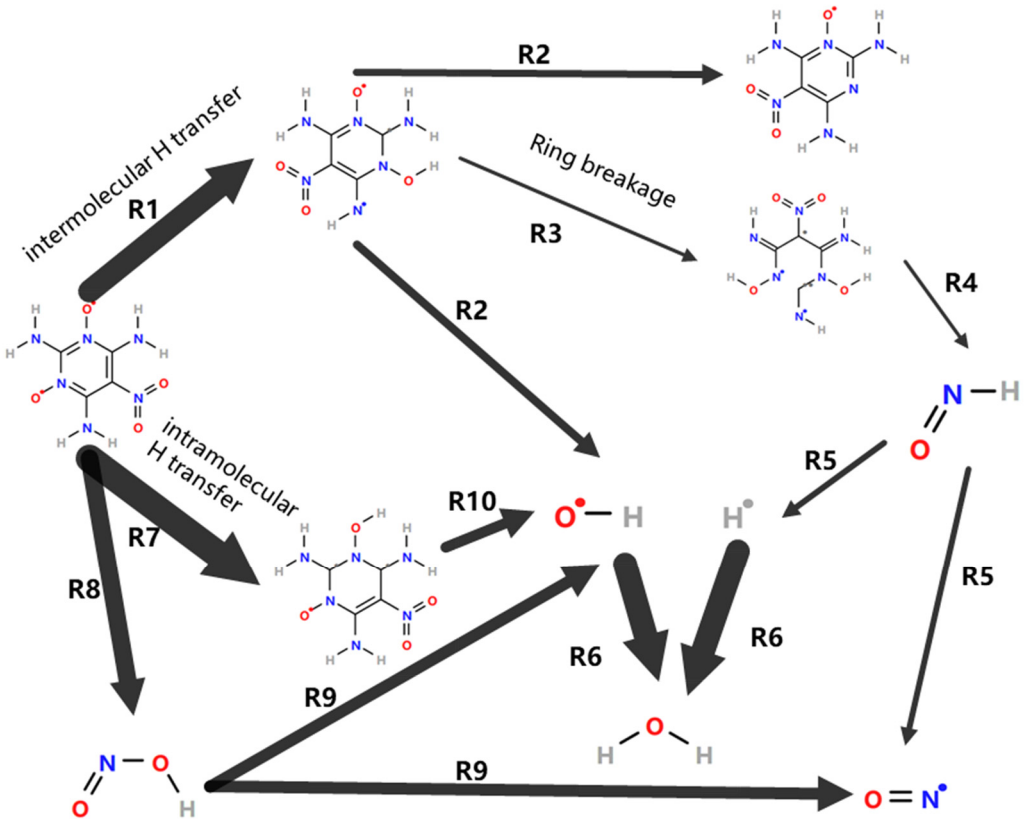


Fig. 7 Primary reaction pathways for the ICM-102 decomposition and formation of H_2O. The arrow width represents the observed number of reactions (*n*), where width = *ln* (n+1).

From the above discussion, the NNMD method breaks through the limitations of AIMD. On the one hand, it extends the timescale of simulation to allow observation of the reaction pathways related to N_2, CHNO, CO_2 and many other intermediates of ring-opening reactions; on the other hand, it expands the system size to thousands of atoms to obtain statistically significant results on the branching of detailed reaction pathways. In summary, ICM-102 decomposition reactions are explored using NN-based molecular dynamics simulations. This is the first study to achieve a panoramic view of the complex explosive reaction process with an *ab initio* level of accuracy. NNMD extends

simulation of high explosives to the sub-nanosecond scale, which helps the development of high accuracy kinetics models and promotes the design and application of high explosives.

Conclusion

This work evaluates the datasets generated from different sampling methods: AIMD and VR enhanced sampling. The complex system is abstracted by PCA to compare the similarity of configurations. Surprisingly, we find that PC1 is strongly correlated with the overall reaction coordinates of ICM-102 decomposition, where the explosive molecules decompose into gaseous species. According to the principal components, a prior estimation of the training set is achieved, where configurations of AIMD are concentrated in low-energy states. In contrast, VRMD provides more configurations at high-energy states corresponding to gaseous products. An NN potential is developed from the AIMD and VRMD datasets with superior accuracy. The dataset's quality is critical to the training of a neural network potential. The trained NN potentials will not capture the correct dynamics if the configuration space is not adequately sampled.

Although MD simulations can easily generate thousands of configurations, only a fraction of them is involved in chemical reactions due to the high energy barriers. With the help of VRMD sampling, the full PES of interest could be appropriately represented with a small number of additional configurations. We believe such a reduced dataset is

extremely valuable to the application of supervised learning in the field of computational chemistry because dataset labeling is the most time-consuming step in supervised learning, corresponding to electronic structure calculations in NN potential training. A small improvement in accuracy requires a great effort in computational resources. Recently, it has been found that the computational efforts at high-level *ab initio* calculations can be minimized by combining datasets of molecular forces from different levels of theory[14]. However, such a strategy is highly dependent on the well-selected training set of high-level *ab initio* calculations. We are working on integrating the VR-enhanced sampling algorithms to develop NN potentials with a high level of accuracy.

Gaps between the MD simulations and experiments still exist in explosive decomposition reactions. Although the landscape of ICM-102 molecules has been drawn *via* AIMD and VR-enhanced sampling, further work is required to assemble reaction networks into kinetics models to interpret macro-phenomena in diverse fields, such as combustion, catalysis, and atmospheric chemistry. There are many well-established methods, such as model reduction and sensitivity analysis, to construct kinetic models and describe the macroproperties. Importantly, the present work constructs a novel VR-enhanced sampling method to train NN potentials and enables MD simulations of a novel high explosive (*i.e.* ICM-102) with complex reaction networks at the level of ab initio calculations, providing atomic insights into the detailed reaction pathways that are otherwise difficult to uncover, such as ring-opening reactions

and the formation of N_2, CHNO, and CO_2. Thus, the study opens up new possibilities to build reaction kinetics models based on high-fidelity and low-cost AI algorithms.

Acknowledgements

This paper is supported by the State Key Laboratory of Explosion Science and Technology (Grant No. ZDKT21-01) and the National Natural Science Foundation of China (Grant No. 52106130). The authors also acknowledge the support from the Foundation of Science and Technology on Combustion and Explosion Laboratory.

Reference

[1] T.M. Klapötke, Chemistry of high-energy materials, de Gruyter, 2019.
[2] J.P. Agrawal, High energy materials: propellants, explosives and pyrotechnics, John Wiley & Sons, 2010.
[3] Y. Tan, Z. Yang, H. Wang, H. Li, F. Nie, Y. Liu, Y. Yu, High energy explosive with low sensitivity: a new energetic cocrystal based on CL-20 and 1, 4-DNI, Cryst. Growth Des. 19 (2019) 4476–4482.
[4] R. Iftimie, P. Minary, M.E. Tuckerman, *Ab initio* molecular dynamics: Concepts, recent developments, and future trends, Proc. Natl. Acad. Sci. 102 (2005) 6654–6659.
[5] S. Feng, F. Guo, C. Yuan, X. Cheng, Z. Li, L. Su, Ab-initio molecular dynamics study on chemical decomposition reaction of α-HMX, Chem. Phys. Lett. 748 (2020) 137394. https://doi.org/10.1016/j.cplett.2020.137394.
[6] V. Rizzi, D. Polino, E. Sicilia, N. Russo, M. Parrinello, The onset of dehydrogenation in solid ammonia borane: an *ab initio* metadynamics study, Angew. Chem. Int. Ed. 58 (2019) 3976–3980. https://doi.org/10.1002/anie.201900134.
[7] S.J. Klippenstein, C. Cavallotti, Ab initio kinetics for pyrolysis and combustion systems, in: Comput. Aided Chem. Eng., Elsevier, 2019: pp. 115–167. https://doi.org/10.1016/B978-0-444-64087-1.00002-4.
[8] Q. Wu, H. Chen, G. Xiong, W. Zhu, H. Xiao, Decomposition of a 1,3,5-triamino-2,4,6-trinitrobenzene crystal at decomposition temperature coupled with different pressures: an *ab initio* molecular dynamics study, J. Phys. Chem. C. 119 (2015)

16500–16506. https://doi.org/10.1021/acs.jpcc.5b05041.
[9] I.V. Schweigert, *Ab initio* molecular dynamics of high-temperature unimolecular dissociation of gas-phase RDX and its dissociation products, J. Phys. Chem. A. 119 (2015) 2747–2759. https://doi.org/10.1021/jp510034p.
[10] O. Isayev, L. Gorb, M. Qasim, J. Leszczynski, *Ab initio* molecular dynamics study on the initial chemical events in nitramines: thermal decomposition of CL-20, J. Phys. Chem. B. 112 (2008) 11005–11013.
[11] Y. Zuo, C. Chen, X. Li, Z. Deng, Y. Chen, J. Behler, G. Csányi, A.V. Shapeev, A.P. Thompson, M.A. Wood, S.P. Ong, Performance and cost assessment of machine learning interatomic potentials, J. Phys. Chem. A. 124 (2020) 731–745. https://doi.org/10.1021/acs.jpca.9b08723.
[12] L. Zhang, J. Han, H. Wang, R. Car, W. E, Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics, Phys. Rev. Lett. 120 (2018) 143001. https://doi.org/10.1103/PhysRevLett.120.143001.
[13] D. Dragoni, T.D. Daff, G. Csányi, N. Marzari, Achieving DFT accuracy with a machine-learning interatomic potential: Thermomechanics and defects in bcc ferromagnetic iron, Phys. Rev. Mater. 2 (2018) 013808.
[14] S. Chmiela, H.E. Sauceda, K.-R. Müller, A. Tkatchenko, Towards exact molecular dynamics simulations with machine-learned force fields, Nat. Commun. 9 (2018) 3887. https://doi.org/10.1038/s41467-018-06169-2.
[15] K.T. Schütt, P.-J. Kindermans, H.E. Sauceda, S. Chmiela, A. Tkatchenko, K.-R. Müller, SchNet: A continuous-filter convolutional neural network for modeling quantum interactions, in: Proc. 31st Int. Conf. Neural Inf. Process. Syst., Curran Associates Inc., Red Hook, NY, USA, 2017: pp. 992–1002.
[16] L. Zhang, H. Wang, R. Car, W. E, Phase diagram of a deep potential water model, Phys. Rev. Lett. 126 (2021) 236001. https://doi.org/10.1103/PhysRevLett.126.236001.
[17] W. Jiang, Y. Zhang, L. Zhang, H. Wang, Accurate deep potential model for the Al–Cu–Mg alloy in the full concentration space, Chin. Phys. B. 30 (2021) 050706. https://doi.org/10.1088/1674-1056/abf134.
[18] J.A. Keith, V. Vassilev-Galindo, B. Cheng, S. Chmiela, M. Gastegger, K.-R. Müller, A. Tkatchenko, Combining machine learning and computational chemistry for predictive insights into chemical systems, Chem. Rev. 121 (2021) 9816–9872. https://doi.org/10.1021/acs.chemrev.1c00107.
[19] M. Yang, L. Bonati, D. Polino, M. Parrinello, Using metadynamics to build neural network potentials for reactive events: the case of urea decomposition in water, Catal. Today. (2021). https://doi.org/10.1016/j.cattod.2021.03.018.
[20] J. Zeng, L. Cao, M. Xu, T. Zhu, J.Z.H. Zhang, Complex reaction processes in combustion unraveled by neural network-based molecular dynamics simulation, Nat. Commun. 11 (2020) 5713. https://doi.org/10.1038/s41467-020-19497-z.
[21] A. Laio, M. Parrinello, Escaping free-energy minima, Proc. Natl. Acad. Sci. 99 (2002) 12562–12566.

[22] O. Valsson, P. Tiwary, M. Parrinello, Enhancing important fluctuations: Rare events and metadynamics from a conceptual viewpoint, Annu. Rev. Phys. Chem. 67 (2016) 159–184.

[23] M. O'Connor, H.M. Deeks, E. Dawn, O. Metatla, A. Roudaut, M. Sutton, L.M. Thomas, B.R. Glowacki, R. Sage, P. Tew, M. Wonnacott, P. Bates, A.J. Mulholland, D.R. Glowacki, Sampling molecular conformations and dynamics in a multiuser virtual reality framework, Sci. Adv. 4 (2018) eaat2731. https://doi.org/10.1126/sciadv.aat2731.

[24] S. Amabilino, L.A. Bratholm, S.J. Bennie, A.C. Vaucher, M. Reiher, D.R. Glowacki, Training neural nets to learn reactive potential energy surfaces using interactive quantum chemistry in virtual reality, J. Phys. Chem. A. 123 (2019) 4486–4499. https://doi.org/10.1021/acs.jpca.9b01006.

[25] S. Amabilino, L.A. Bratholm, S.J. Bennie, M.B. O'Connor, D.R. Glowacki, Training atomic neural networks using fragment-based data generated in virtual reality, J. Chem. Phys. 153 (2020) 154105. https://doi.org/10.1063/5.0015950.

[26] Y. Wang, Y. Liu, S. Song, Z. Yang, X. Qi, K. Wang, Y. Liu, Q. Zhang, Y. Tian, Accelerating the discovery of insensitive high-energy-density materials by a materials genome approach, Nat. Commun. 9 (2018) 2444. https://doi.org/10.1038/s41467-018-04897-z.

[27] G. Lippert, J. Hutter, M. Parrinello, The Gaussian and augmented-plane-wave density functional method for ab initio molecular dynamics simulations, Theor. Chem. Acc. 103 (1999) 124–140.

[28] S. Goedecker, M. Teter, J. Hutter, Separable dual-space Gaussian pseudopotentials, Phys. Rev. B. 54 (1996) 1703.

[29] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77 (1996) 3865.

[30] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, J. Chem. Phys. 132 (2010) 154104.

[31] J. VandeVondele, J. Hutter, Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases, J. Chem. Phys. 127 (2007) 114105.

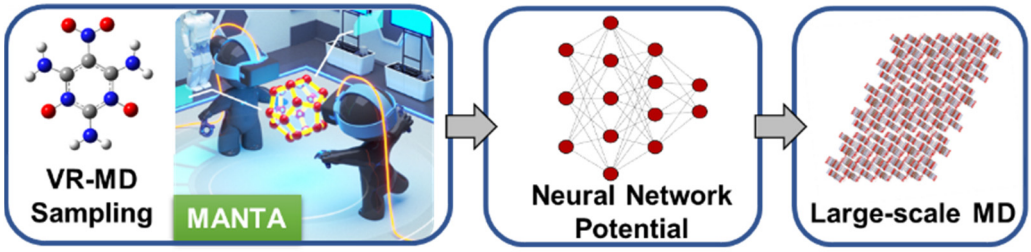
[32] L. Zhang, J. Han, H. Wang, W. Saidi, R. Car, W. E, End-to-end symmetry preserving inter-atomic potential energy model for finite and extended systems, in: Adv. Neural Inf. Process. Syst., Curran Associates, Inc., 2018. https://proceedings.neurips.cc/paper/2018/file/e2ad76f2326fbc6b56a45a56c59fafdb-Paper.pdf.

[33] H. Wang, L. Zhang, J. Han, W. E, DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics, Comput. Phys. Commun. 228 (2018) 178–184. https://doi.org/10.1016/j.cpc.2018.03.016.

[34] Y. Zhang, H. Wang, W. Chen, J. Zeng, L. Zhang, H. Wang, W. E, DP-GEN: A concurrent learning platform for the generation of reliable deep learning based potential energy models, Comput. Phys. Commun. 253 (2020) 107206.

https://doi.org/10.1016/j.cpc.2020.107206.
[35] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, J. Comput. Phys. 117 (1995) 1–19.
[36] J. Zeng, L. Cao, C.-H. Chin, H. Ren, J.Z.H. Zhang, T. Zhu, ReacNetGenerator: an automatic reaction network generator for reactive molecular dynamics simulations, Phys. Chem. Chem. Phys. 22 (2020) 683–691. https://doi.org/10.1039/C9CP05091D.
[37] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: machine learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.
[38] G. Landrum, RDKit: Open-source cheminformatics, (n.d.). http://www.rdkit.org.
[39] F. Musil, A. Grisafi, A.P. Bartók, C. Ortner, G. Csányi, M. Ceriotti, Physics-inspired structural representations for molecules and materials, Chem. Rev. 121 (2021) 9759–9815. https://doi.org/10.1021/acs.chemrev.1c00021.
[40] L. Zhang, D.-Y. Lin, H. Wang, R. Car, W. E, Active learning of uniformly accurate interatomic potentials for materials simulation, Phys. Rev. Mater. 3 (2019) 023804. https://doi.org/10.1103/PhysRevMaterials.3.023804.
[41] D. Lu, H. Wang, M. Chen, L. Lin, R. Car, W. E, W. Jia, L. Zhang, 86 PFLOPS deep potential molecular dynamics simulation of 100 million atoms with *ab initio* accuracy, Comput. Phys. Commun. 259 (2021) 107624. https://doi.org/10.1016/j.cpc.2020.107624.
[42] W. Jia, H. Wang, M. Chen, D. Lu, L. Lin, R. Car, W. E, L. Zhang, Pushing the limit of molecular dynamics with *ab initio* accuracy to 100 million atoms with machine learning, in: Proc. Int. Conf. High Perform. Comput. Netw. Storage Anal., IEEE Press, 2020.
[43] J. MacQueen, others, Some methods for classification and analysis of multivariate observations, in: Proc. Fifth Berkeley Symp. Math. Stat. Probab., Oakland, CA, USA, 1967: pp. 281–297.
[44] C. Qi, Q.-H. Lin, Y.-Y. Li, S.-P. Pang, R.-B. Zhang, C–N bond dissociation energies: An assessment of contemporary DFT methodologies, J. Mol. Struct. THEOCHEM. 961 (2010) 97–100.

Table of Contents Graphic



We develop a workflow combining ab initio molecular dynamics and interactive molecular dynamics in virtual reality (VRMD) to construct a neural network potential of a novel high explosive. The VRMD helps sample the reactive process and significantly improves the model performance. Our potential enables the large-scale MD simulation with an ab initio level of accuracy, achieving a panoramic view of the complex explosive reaction process.