

Universālo mašīnmācīšanās starpatomu potenciālu precizēšana un pārbaude, izmantojot volframa disulfīda piemēru

Pjotrs Žguns^{1,2}, Inga Pudža¹, Aleksejs Kuzmins¹

¹Latvijas Universitātes Cietvielu fizikas institūts

²Materiālzinātnes un inženierzinātņu nodaļa, Masačūsetsas Tehnoloģiju institūts, ASV

Universālie mašīnmācīšanās starpatomu potenciāli (uMLIP) nodrošina *ab initio* precizitāti enerģijas un spēku aprēķinos, vienlaikus saglabājot zemas skaitļošanas izmaksas, padarot tos par nenovērtējamu rīku materiālu modelēšanā. Lai gan uMLIP ir iepriekš apmācīti, izmantojot plašus *ab initio* datu kopumus, to plašākai izmantošanai ir nepieciešama rūpīga validācija. Šajā pētījumā mēs izmantojam CHGNet uMLIP [1], lai modelētu termisko nekārtību slāņveida WS₂, salīdzinot tā rezultātus ar *ab initio* datiem un paplašinātās rentģenabsorbcijas sīkstruktūras (EXAFS) spektru, kas fiksē termiskās izmaiņas saišu garumos un leņķos. CHGNet precizēšana WS₂ specifiskiem *ab initio* (DFT) datiem samazina uMLIP raksturīgo sistemātisko spēku nenovērtēšanu un vienlaikus uzlabo sakritību starp molekulārās dinamikas simulācijās iegūto un eksperimentālo EXAFS spektru. Lai gan precizēšana, izmantojot vienu DFT struktūru, ir iespējama, ieteicams izmantot aptuveni 100 struktūras, lai nodrošinātu precīzu teorētisko EXAFS spektru un DFT precizitāti. uMLIP salīdzinājums ar DFT un eksperimentālajiem EXAFS datiem stiprina pārliecību par tā veikspēju un sniedz vadlīnijas optimāla precizēšanas datu kopuma lieluma noteikšanai.

I.P. pateicas Latvijas Zinātnes padomes projektam Nr. LZP-2023/1-0528 par atbalstu.

Fine-tuning and validating universal machine learning interatomic potentials on the example of tungsten disulfide

Pjotrs Žguns^{1,2}, Inga Pudža¹, and Alexei Kuzmin¹

¹Institute of Solid State Physics, University of Latvia

²Department of Materials Science and Engineering, Massachusetts Institute of Technology, USA

Universal machine learning interatomic potentials (uMLIPs) offer *ab initio* accuracy in energy and force calculations at minimal computational cost, making them essential for materials modeling. Although uMLIPs are pre-trained on vast *ab initio* datasets, their validation is crucial for reliable practical applications. In this study, we employ the CHGNet uMLIP [1] to model thermal disorder in layered tungsten disulfide (WS₂), comparing its performance against *ab initio* data and extended X-ray absorption fine structure (EXAFS) spectrum, which captures thermal variations in bond lengths and angles. Fine-tuning CHGNet uMLIP with WS₂-specific *ab initio* (DFT) data corrects the systematic softening (force underestimation) typically observed in uMLIPs while also improving agreement between molecular dynamics-derived and experimental EXAFS spectra. We demonstrate that although a single DFT structure provides some improvement, approximately one hundred structures are recommended for generating accurate theoretical EXAFS spectra and achieving DFT-level precision. Benchmarking uMLIP against both DFT and experimental EXAFS data strengthens confidence in its reliability and helps determine the optimal fine-tuning dataset size.

I.P. thanks the support of the Latvian Council of Science project No. LZP-2023/1-0528.

[1] B. Deng, P. Zhong, K. Jun, J. Riebesell, K. Han, C.J. Bartel, G. Ceder, Nat. Mach. Intell. 5 (2023) 1031–1041.