Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supporting Information:

Prediction of Optoelectronic Properties of Cu_2O Using Neural Network Potentials

Balaranjan Selvaratnam, Ranjit T. Koodali, and Pere Miró*

Department of Chemistry, University of South Dakota, Vermillion, SD 57069

E-mail: pere.miro@usd.edu

List of Tables

List of Figures

Training Data

The training data for the Neural Network Potential (NNP) was gathered using an iterative process. Initially, a NVT MD simulation was performed at 1550 K with a time step of 1 fs for 250 steps. Berendsen thermostat with a coupling constant of 25 fs was used to control the temperature. For this MD, the energy and forces were calculated from VASP and the NVTBerendsen algorithm implemented in ASE was used to propagate the MD. All 250 snapshots from this initial MD was added to the training dataset and a NNP was trained using the training dataset. For the NNP, the default Gaussian parameters defined in the AMP package was used and they are presented in Table S1. Then, a random snapshot from the training dataset and a random temperature between 50 K and 1550 K (in 25 K intervals, i.e. 25K, 50K, 75K, ... 1525K) were selected and a MD was performed using the selected snapshot and temperature using same MD parameters described earlier. For this MD, NNP was used to calculate the energy and forces. The potential energies calculated by the NNP during this MD run were validated against DFT by single point calculations using VASP (see Methods in main text for computational details). Then all structures with an energy difference greater than or equal to 5 kcal were added to the training dataset. Then a random snapshot and a random temperature were selected, and the loop was continued until no structures were found with potential energy difference of 5 kcal or higher for three consecutive MD runs. The selected temperatures and the number of structures selected for training at each temperature are given in Table S2. The initial temperature 1550 K was selected to sample configurations that are far from equilibrium structure and the subsequent temperatures and starting configurations were selected randomly to explore the potential energy space. In this work, we ran all MD simulations for shorter time and used NVT-Berendsen thermostat since our main objective of this work is to test the Behler-Parrinello scheme to predict opto-electronic properties. However, for a production potential, MD sampling with different ensembles and longer simulations are necessary to sample the potential energy surface.

Table S1: Gaussian parameters used to create the atomic environment vector.

No	Type	Elements	η	ζ	λ
1	G2	Cu	0.05	-	-
2	G2	O	0.05	-	-
3	G2	Cu	0.23	-	-
4	G2	O	0.23	-	-
5	G2	Cu	1.08	-	-
6	G2	O	1.08	-	-
7	G2	Cu	5.00	-	-
8	G2	O	5.00	-	-
9	G4	(Cu, Cu)	0.005	1.0	1.0
10	G4	(Cu, O)	0.005	1.0	1.0
11	G4	(O, O)	0.005	1.0	1.0
12	G4	(Cu, Cu)	0.005	1.0	-1.0
13	G4	(Cu, O)	0.005	1.0	-1.0
14	G4	(O, O)	0.005	1.0	-1.0
15	G4	(Cu, Cu)	0.005	4.0	1.0
16	G4	(Cu, Cu)	0.005	4.0	1.0
17	G4	(O, O)	0.005	4.0	1.0
18	G4	(Cu, Cu)	0.005	4.0	-1.0
19	G4	(Cu, O)	0.005	4.0	-1.0
20	G4	(O, O)	0.005	4.0	-1.0

Table S2: Temperatures used to iterate the dataset aggregation cycle and number of structures added to training set at each temperatures.

Iteration No.	Temperature (K)	Number of selected structures
1	1550	250
2	1225	221
3	725	233
4	350	226
5	750	40
6	1475	161
7	200	233
8	1075	201
9	950	161
10	275	52
11	150	0
12	1275	91
13	1125	116
14	1450	218
15	700	183
16	1500	53
17	175	0
18	125	0
19	975	224
20	450	49
21	775	0
22	1325	0
23	75	0

MD Validation

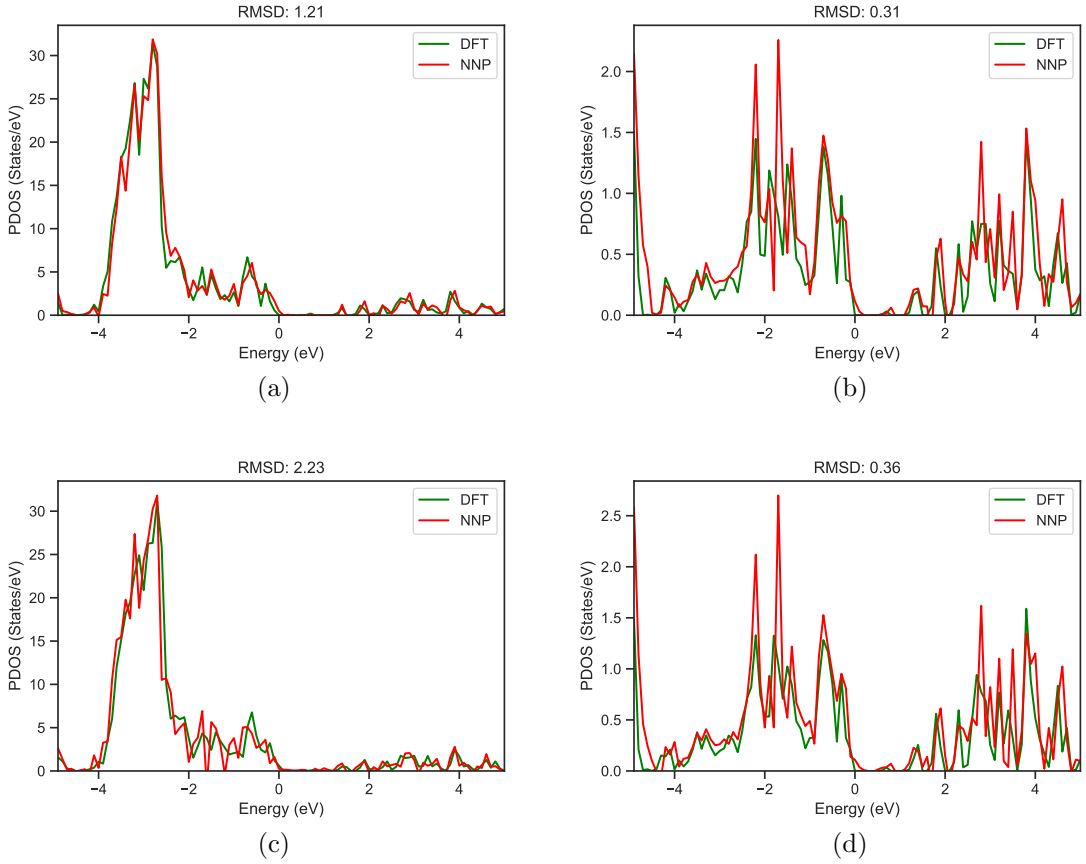


Figure S1: Comparison of raw PDOS of Cu (a, c), and O (b, d) sampled at two different RMSD windows.

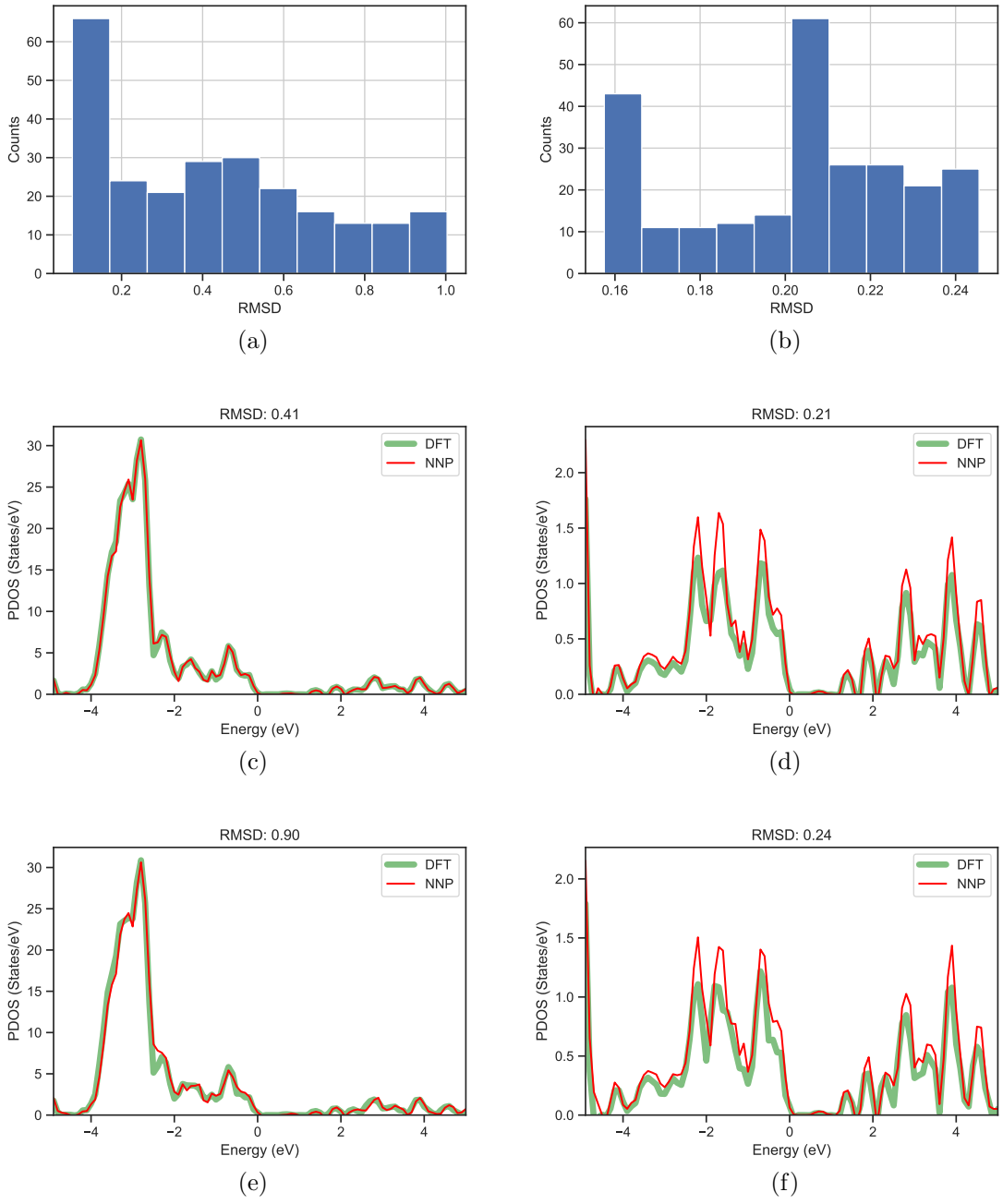
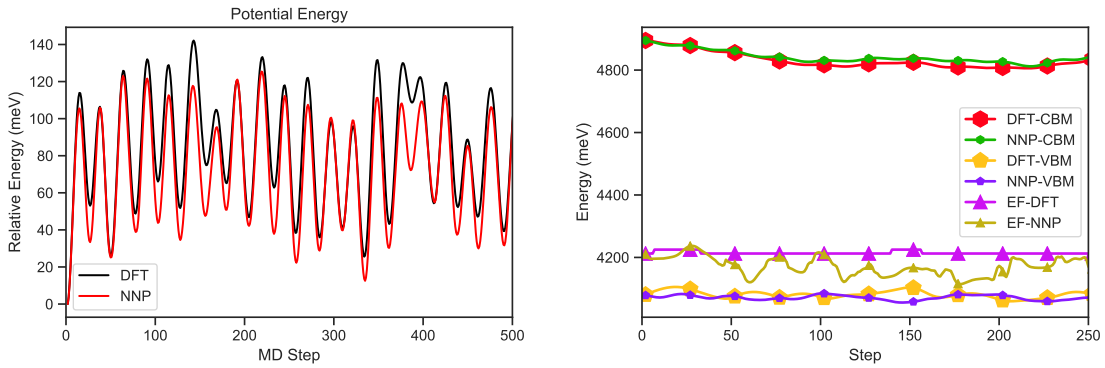


Figure S2: Histogram of RMSD for the PDOS prediction of (a) Cu and (b) O. (c), (e) comparison of DFT and NNP PDOS of Cu at two different RMSDs. (d), (f) comparison of DFT and NNP PDOS of O at two different RMSDs.



(a) (b)

Figure S3: comparison of DFT and NNP (a) potential energies, and (b) Fermi energy, CBM. and VBM along the molecular dynamics trajectory

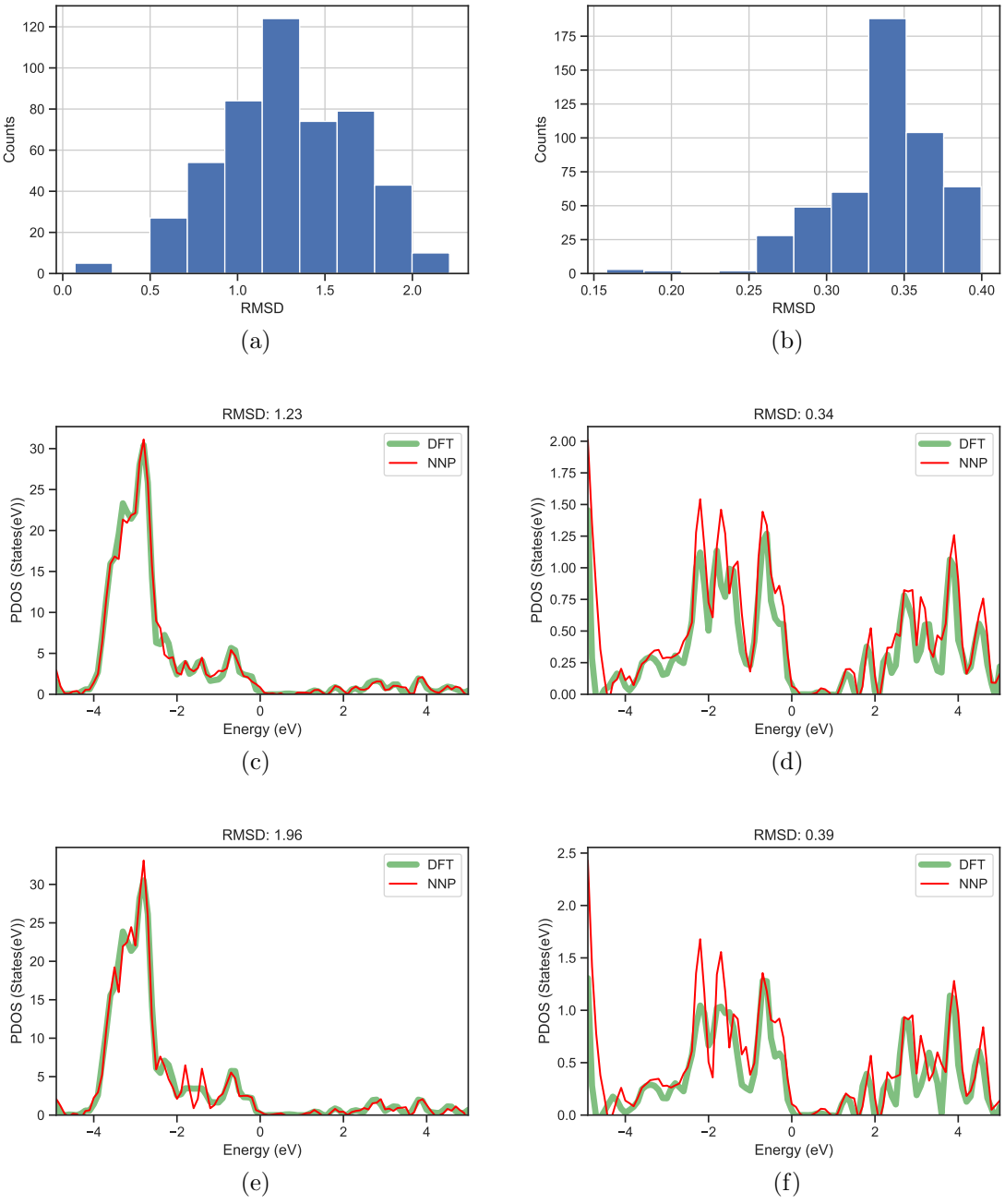


Figure S4: Histogram of RMSD for the PDOS prediction of (a) Cu and (b) O. (c), (e) comparison of DFT and NNP PDOS of Cu at two different RMSDs. (d), (f) comparison of DFT and NNP PDOS of O at two different RMSDs.

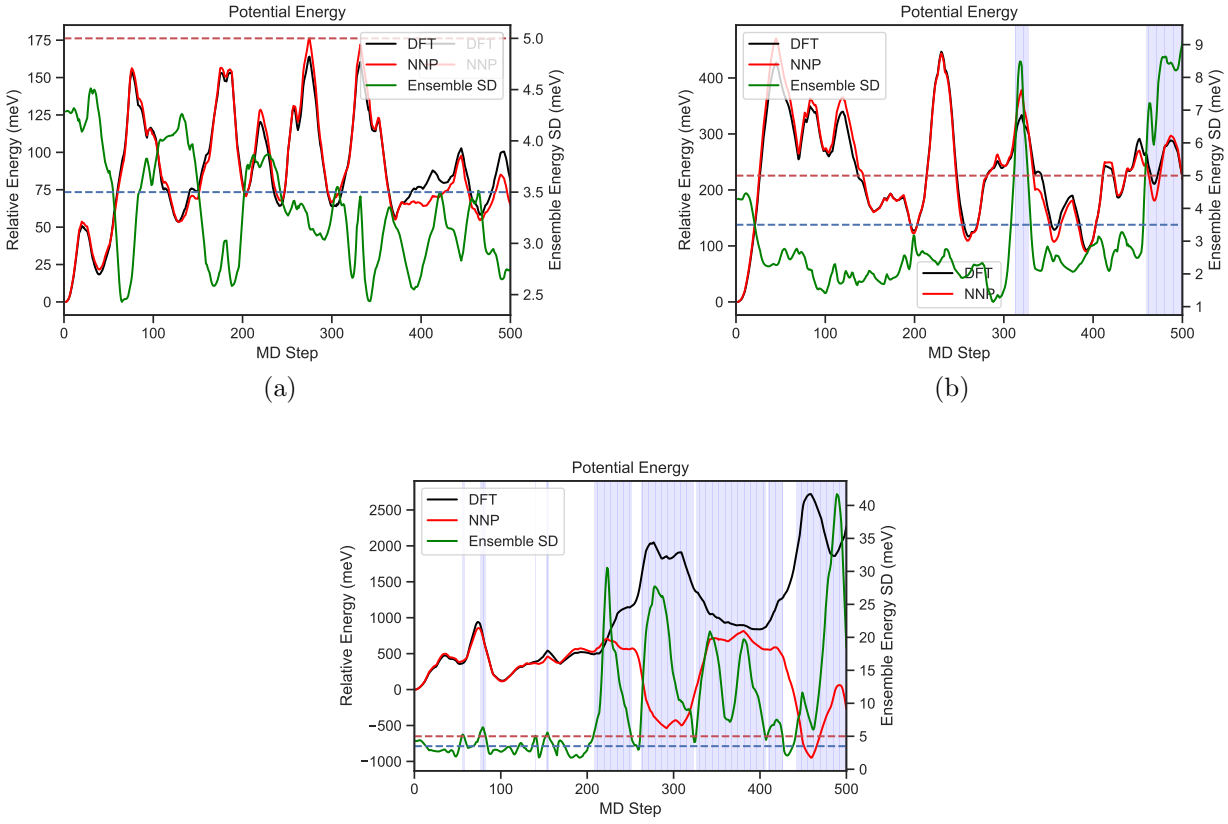


Figure S5: **comparison** of NNP and DFT energies for NVT MD simulations at (a) 300 K, (b) 600 K, and (c) 900 K. Horizontal dashed lines red - ensemble SD 5.0, blue - ensemble SD 3.5.