Evaluating Machine Learning Workflows for Structural Relaxation and Strained Bandgap Modeling in Monolayer Transition Metal Dichalcogenides

Edward Deleu

CBE423/MSE423: May 13, 2026

Abstract

This project explores machine learning workflows for structural relaxation and strained bandgap modeling in monolayer transition metal dichalcogenides (TMDs), across MoS_2, $MoSe_2$, $MoTe_2$, WS_2, and WSe_2. Various modern machine learning interatomic potentials (MLIPs) are benchmarked against density functional theory predictions and experimental measurements: all of the tested MLIPs achieve comparable structural relaxations to DFT, achieving average RMSDs below 0.1Å compared to experimental unit-cells. The most accurate MLIP tested, TensorNet-MatPES-r^{2}SCAN, is utilized to relax a large WS_2/WSe_2 lateral heterostructure, demonstrating the utility of MLIPs for complex TMD supercells. Machine learning workflows are also tested for bandgap prediction under biaxial strain using the PET-MAD-DOS model. While this model captures qualitative bandgap trends, it does not quantitatively agree with DFT results. In contrast, a mixed workflow using MLIPs for structural relaxation followed by static DFT bandgap calculations closely reproduces fully DFT-based strain-bandgap trends. In summary, MLIPs are highly promising for accelerating TMD structural modeling, but reliable electronic-property predictions still require either improved machine learning models or DFT-based post-processing.

1 Introduction

Transition metal dichalcogenides (TMDs) are a class of semiconducting materials that exhibit direct bandgap characteristics when isolated into monolayer form.[1] Monolayer group-VI TMDs follow an MX_2 structure (with M = Mo or W, X = S, Se, Te) and belong to space group 187, as shown in Fig. 1. They are direct-bandgap two-dimensional semiconductors, unlike graphene's zero-bandgap nature, making them highly attractive for transistors, photovoltaics, and sensing applications. Their electronic properties are highly sensitive to alloying, external electric fields, and mechanical strain, offering several pathways to precisely and continuously tune device characteristics.[1] The high strain limit of monolayer TMDs, reaching upwards of 12% in MoS_2, provides an effective means to precisely tune both the electronic and optical properties.[2,3]

Density functional theory (DFT) has shown success in modeling TMDs for various applications including strain engineering, alloying, and simulating band alignments in heterostructures.[1,4–6] While certain exchange-correlation functionals like PBE have known limitations underestimating properties such as electronic bandgaps, hybrid methods further improve predictions and have been well-studied for standard TMDs.[4,5] However, there is limited systematic study of TMDs across application areas and materials, since DFT becomes computationally expensive as the number of atoms per unit-cell increases or the number of required simulations rises. This is especially present in studies of strain engineering where simulations are typically only run for a few materials under a relatively coarse strain grid and limited strain configurations.[4] Furthermore, computational studies of TMD alloying and heterostructures have been limited to quite small unit-cells, ultimately limiting their robustness in comparison to experiment.[7,8]

As such, a fast and accurate replacement for DFT modeling would be highly useful for studying TMDs. Foundational machine learning interatomic potentials (MLIPs) have shown significant

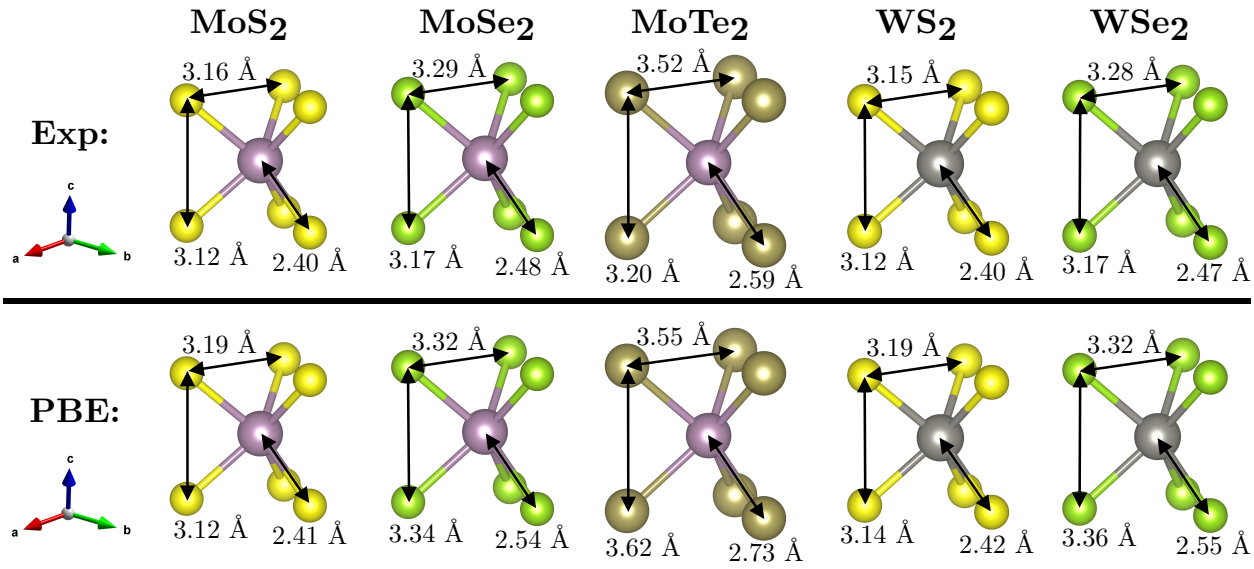


Figure 1: Comparison of monolayer TMD unit-cells constructed from experimental lattice parameters reported by Samadi *et al.*[1] and PBE-relaxed geometries for MoS_2, $MoSe_2$, $MoTe_2$, WS_2, and WSe_2. Key structural distances are annotated for each structure; figures are not precisely to scale.

promise in computational materials modeling, often offering several orders of magnitude speedups compared to DFT with comparable accuracy.[9,10] These models are trained on a wide range of materials data, allowing them to be used as a general drop-in replacement for materials modeling.

This project launches a two-pronged study applying machine learning workflows for monolayer TMD modeling, specifically looking at the five most common group-VI TMDs: MoS_2, $MoSe_2$, $MoTe_2$, WS_2, and WSe_2. First, MLIPs of various architectures and training datasets are compared to DFT and experimental measurements across group-VI TMDs. The most promising candidate is then applied to relax a large in-plane heterostructure as a toy example. Second, machine learning based workflows are evaluated for bandgap prediction in TMDs compared to DFT, across materials and biaxial strain levels. Thus, two generally applicable application areas are studied to determine the accuracy and efficiency of MLIPs for TMD modeling.

2 Methods

This work employs both density functional theory and machine learning based workflows for modeling TMDs. Structural files for the five monolayer TMDs under study are retrieved from the Materials Project database (mp-1023924, mp-1023934, mp-1023935, mp-1023937, mp-1023936) for subsequent DFT and MLIP relaxation.[11] Python is used for all data analysis and automation in this work, and VESTA is used for all visualizations. All MLIP computations are performed on an Apple M1 Pro processor and all DFT computations are performed on Princeton Research Computing's Della cluster.

2.1 Density Functional Theory

Density functional theory (DFT) simulations are the leading methodology for first-principles modeling of TMD behavior. VASP version 6.4.3[12–15] and the helper-software VASPKIT[16,17] are utilized for DFT computations in this work. VASP is a widely used first-principles electronic-structure soft-

ware based on density functional theory, employing plane-wave basis sets and projector-augmented-wave (PAW) potentials to efficiently treat core-valence interactions and solve the Kohn-Sham equations. DFT provides a quantum-mechanical description of materials using the electron density rather than complex many-body wavefunctions, allowing accurate predictions of structural and electronic properties without using system-specific empirical parameters. This makes DFT particularly suitable for systematically studying how external perturbations such as strain modify the electronic band structures of monolayer TMDs. The Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional is the standard level of theory for VASP-based DFT, with both fast and efficient results. The HSE06 hybrid functional is also used, providing more accurate predictions for quantities like bandgap by including a portion of the exact Hartree-Fock exchange energy.

Reference DFT computations are carried out in VASP with ENCUT= 700eV, a 15Å vacuum spacing, spin polarization disabled, and a $2\pi \times 0.03\text{Å}^{-1}$ reciprocal-space mesh. A majority of reference DFT computations have been carried out as part of a previous work and are hence not the main focus of this brief study. Still, these optimal parameters are chosen for the new DFT results presented in this study.

2.2 Machine Learning Interatomic Potentials

Machine learning interatomic potentials (MLIPs) are a promising DFT-surrogate, offering energy and force predictions utilizing large-scale machine learning models. These models are typically trained on large atomic datasets, accepting atomic structures as input and outputting the system's energy. Various MLIPs are explored in this study, implemented in the Atomic Simulation Environment (ASE) in Python. A maximum force convergence threshold of $|F_{max}| < 0.001\text{eV/Å}$ is implemented to ensure robust, optimal structures using the BFGS optimizer.

Table 1: Summary of machine learning interatomic potentials (MLIPs), their training datasets, and highest level of theory for their training data.

MLIP Name	**Training Dataset**	**Dataset Size**	**Level of Theory**	**Parameters**
uma-s-1p2-OMat[18]	OMat24[19]	118M	PBE	6.6M/150M
uma-m-1p1-OMat[18]	OMat24[19]	118M	PBE	50M/1.4B
TensorNet-MatPES-PBE[20]	MatPES[20]	434,712	PBE	838,000
TensorNet-MatPES-r2SCAN[20]	MatPES[20]	387,897	r^{2}SCAN	838,000
PET-OMat-PES-L[21]	OMat24[19] $\rightarrow$ MatPES[20]	118M + 387k	r^{2}SCAN	255M
PET-MAD-DOS[22]	MAD[23]	95,595	PBEsol	8,625,226

Table 1 outlines the various MLIPs and their training datasets used in this work. Notably, models of varying size and level of theory are tested to compare the accuracy of predictions under these varying factors. The Open Materials 2024 (OMat24) dataset is by far the largest solid-state material dataset, trained on PBE data, including a significant number of molecular dynamics and non-equilibrium structures.[19] Its structures are randomly sampled from the Alexandria bulk materials database which *does* contain the monolayer TMD structures under study: however, the OMat24 dataset is too large to query for this project, so it is uncertain whether the monolayer TMDs are contained in the OMat24 dataset. The Materials Potential Energy Surface (MatPES) dataset includes subsets trained on both PBE and r^{2}SCAN data, containing 62 instances of the monolayer TMDs as either equilibrium structures or as part of molecular dynamics trajectories.[20]

For the sake of brevity, the *TensorNet-MatPES-* models will be abbreviated as *TensorNet-PBE* and *TensorNet-r^{2}SCAN* in this report.

While traditional MLIPs are suitable for energy and force predictions, they cannot predict more complex quantities like bandgap. The PET-MAD-DOS model is not a traditional MLIP– it is a density-of-states prediction model that also has bandgap prediction capabilities.[22] The Massive Atomic Diversity (MAD) dataset includes several subsets, one of which is for two-dimensional materials (MC2D), which does include the monolayer TMDs in this study.[23] This model accepts a relaxed structure and is used for machine learning based bandgap prediction in this study.

3 Results and Discussion

3.1 Structural Results

Experimental unit-cells for the five TMD monolayers under study are shown in Fig. 1, created using experimental measurements,[1] in comparison to the PBE relaxed structures. This demonstrates that DFT very closely reproduces the experimental structures. These structures are very simple, with just 3 atoms per cell. While DFT is already quite efficient for these small cells, this study aims to benchmark differences in DFT, MLIPs, and experiment, to evaluate the potential of MLIPs for more complex TMD simulations with larger unit-cells.

Table 2: Comparison of structural prediction accuracy across DFT and MLIPs compared to reference experimental structures from Samadi *et al.*[1] Reported values are mean absolute percentage errors (MAE) for lattice constants, monolayer thickness, and transition metal to chalcogen bond-lengths, along with average RMSD in Å. Results are averaged over the five TMDs under study.

Metric	**PBE**	**uma-s-1p2**	**uma-m-1p1**	**TensorNet -PBE**	**TensorNet -r2SCAN**	**PET-OMat -PES-L**
Lattice Constant MAE (%)	1.031	1.227	1.009	0.629	0.729	1.008
Thickness MAE (%)	4.984	4.794	4.927	5.592	3.725	3.986
Bond-Length MAE (%)	2.626	2.660	2.591	2.649	1.926	2.212
Average RMSD (Å)	0.069	0.068	0.068	0.074	0.058	0.060

Table 2 quantifies the differences in structural relaxations for various DFT and MLIP results compared to the experimental unit-cells, averaged over the five monolayer TMDs. The HSE06 results are omitted, since the HSE06 functional relaxes the structures nearly identically to the PBE functional. TMD lattice constants may be considered as the transition metal to transition metal bond distances, or alternatively as the length of the in-plane lattice vectors (**a** and **b** in Fig. 1). All of the DFT and MLIP methods produce extremely similar results for the predicted lattice constants and metal to chalcogen bond lengths, all with mean absolute errors under 3%. The DFT and MLIP methods agree slightly worse with experiment in terms of monolayer thicknesses, with average errors closer to $4-5\%$. Finally, the average root mean square deviations (RMSDs) are computed using ASE; this indicates, on average, how far the atomic positions of the relaxed cells deviate from the corresponding positions in the experimental cells. Across methods, on average, the predicted atomic positions tend to be < 0.1Å away from the experimental locations. Both DFT and MLIP methods tend to consistently underbind atoms, resulting in measurements that are slightly too large compared to experiment. Further data comparing the exact structural parameters between materials and methods is presented in Appendix A.1 in Tables 4 to 7.

These results demonstrate that overall, all of the MLIPs tested produce comparable relaxed structures compared to DFT– no approach is significantly better in reproducing experimental struc-

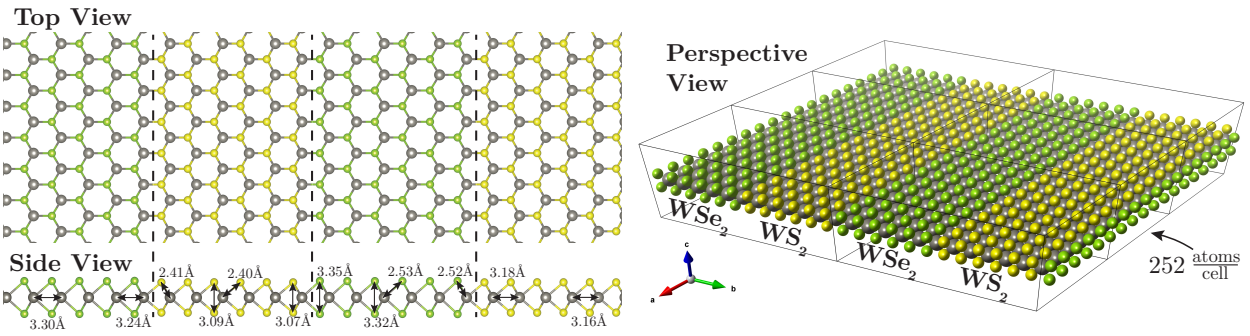


Figure 2: Relaxed lateral heterostructure of monolayer WS_2/WSe_2 shown in top, side, and perspective views. Tungsten atoms are colored gray, selenium green, and sulfur yellow. Annotated dimensions include local tungsten-chalcogen bond lengths, monolayer thicknesses, and in-plane W–W separations across distinct material regions. The perspective rendering displays the repeated supercell and is displayed in space-filling representation for enhanced visual clarity.

tures. Even the greatest differences compared to experiment are on the order of $1-2\%$. Likely, regardless of the MLIP and level of theory of its training set, most contemporary MLIPs will produce satisfactory structural relaxations for TMDs. Still, the two r^{2}SCAN MLIPs (TensorNet and PET-OMat-PES-L) match the experimental results more closely than any of the other methods, including PBE. This result may be expected, since r^{2}SCAN is a higher level of theory than PBE (GGA vs. meta-GGA). Still, it is surprising that its MLIPs outperform the HSE06 relaxed structures given that HSE06 tends to be a more accurate (yet more expensive) functional (meta-GGA vs. hybrid). This r^{2}SCAN improvement is more pronounced than varying the model size: the uma-s and uma-m models trained on PBE data produce almost identical results, indicating that the higher model complexity does not yield improved results for these TMD structural relaxations.

The TensorNet-r^{2}SCAN model produces results closest to experiment for all of the metrics shown in Table 2 (except lattice constants), despite its compact size compared to the other models– it is the smallest MLIP tested. This indicates that training the model on the MatPES r^{2}SCAN dataset may produce slightly enhanced results compared to the PET-OMat-PES-L MLIP which trains on the OMat24 (PBE) dataset and fine-tunes on MatPES. Of course, this difference could also be down to architectural differences. In general, for TMD structural relaxations, it appears that the level of theory of the MLIP has a more significant effect in modeling accuracy compared to model size.

Next, the TensorNet-r^{2}SCAN model is used to relax a large in-plane TMD heterostructure as a toy problem for relaxing a large TMD supercell. The heterostructure is generated between WS_2/WSe_2 with 252 atoms per cell, inspired by the superlattice created experimentally by Xie *et al.*[24] Because of the $\approx 4\%$ lattice-mismatch between materials, the heterostructure supercell is generated with lattice parameters averaged between the two materials to ensure periodicity, starting with the PBE-relaxed structures for simplicity; the average lattice constant is 3.25Å. The interface is created at a metal-terminated zig-zag edge, in an attempt to match the most favorable experimental conditions. Of course, many other viable heterostructure interfaces are plausible but are outside the scope of this experiment. The relaxed supercell along with annotated measurements is demonstrated in Fig. 2.

The heterostructure is successfully relaxed in just 55 optimization steps, producing a realistic and viable heterostructure. The lattice is distorted by non-uniform strain to accommodate for the differing lattice parameters of the two materials. Both materials relax to the same lattice constant of

3.23Å in the direction of the interface (the red **a** direction in Fig. 2), with varying lattice constants in the perpendicular direction. This is actually a slight decrease from the original average value. Near the interface, the WSe_2 and WS_2 lattice constants are 3.24Å and 3.18Å, respectively, compared to 3.30Å and 3.17Å from the TensorNet equilibrium structures. Away from the interface at the center of each material's region, the lattice constants are 3.30Å and 3.16Å, respectively, significantly closer to the equilibrium values. This indicates that strain is induced at the interface for a more favorable alignment of the lattices, bringing the lattices closer to the average, while the individual materials trend toward their equilibrium values away from the interface. Similar trends are observed with the transition metal to chalcogen bond lengths. A slightly different trend is present with monolayer thickness, with distortion away from the averages at the interface: near the interfaces the WSe_2 and WS_2 monolayer thicknesses are 3.35Å and 3.07Å, respectively, compared to 3.31Å and 3.11Å from the TensorNet equilibrium structures.

One surprising consequence of this relaxation is that the symmetry of the original unit-cells is not preserved, with differing lengths of the in-plane lattice vectors, demonstrating the distortion and non-local strain perpendicular to the interface. Overall, this experiment demonstrates the viability of MLIPs for structural relaxations of complex many-atom TMD supercells which may then be used as the starting point for static DFT calculations to compute more complex properties.

3.2 Strained Bandgap Modeling

Machine learning workflows are also evaluated for bandgap modeling in TMDs. While the MLIPs in this study cannot directly predict bandgaps, various workflows are tested, as presented in Table 3. To begin, the PBE and HSE06 bandgaps are approximately 32% and 15% underestimates on average across the TMD materials compared to experimental values. Next, the PET-MAD-DOS bandgap model is tested with PBE-relaxed structures, producing slightly worse predictions compared to PBE-predicted bandgaps on average. The PET-MAD-DOS model is also tested with various MLIP-relaxed structures, producing comparable results on average. This indicates that the small differences between MLIP-relaxed structures do not have a huge impact on predicted properties, and the limiting factor for PET-MAD-DOS prediction accuracy is its PBEsol level of theory. It does not come close to the improved accuracy of the HSE06 functional.

Table 3: Comparison of electronic bandgaps (in eV) for various TMDs obtained from experiment, DFT, and machine learning augmented approaches. Experimental values are from Samadi *et al.*.[1] The right half depicts bandgaps predicted using the PET-MAD-DOS model (abbreviated as "+ML") evaluated on geometries relaxed with PBE and various MLIPs. The final row reports the mean absolute percentage error (MAE) relative to experimental values.

Material	**Exp.**	**HSE06**	**PBE**	**PBE +ML**	**TensorNet-PBE+ML**	**TensorNet-r2SCAN+ML**	**PET-OMat-PES-L+ML**
MoS_2	2.5	2.147	1.644	1.369	1.563	1.795	1.778
$MoSe_2$	2.2	1.879	1.452	1.539	1.528	1.500	1.536
$MoTe_2$	1.6	1.457	1.080	1.089	0.990	1.017	1.017
WS_2	2.7	2.102	1.820	1.884	1.927	1.629	1.746
WSe_2	2.1	1.768	1.547	1.583	1.577	1.642	1.624
MAE (%)	–	-15.12%	-31.95%	-32.42%	-31.94%	-31.59%	-30.70%

Still, the PET-MAD-DOS and other bandgap workflows are tested for their capabilities in capturing trends of TMDs under biaxial strain. To apply biaxial strain, the in-plane lattice vectors

of the optimized unstrained structures are uniformly scaled according to the strain parameter ε such that

$$\mathbf{a}' = (1+\varepsilon)\mathbf{a}, \qquad \mathbf{b}' = (1+\varepsilon)\mathbf{b}, \qquad \mathbf{c}' = \mathbf{c},$$

where $\mathbf{a}$ and $\mathbf{b}$ are the original in-plane lattice vectors and ε is the applied biaxial strain. Positive values of ε correspond to tensile strain, while negative values correspond to compressive strain. After scaling the lattice vectors, the atomic positions are relaxed using either DFT or MLIP methods while keeping the strained lattice vectors fixed. The advantage of MLIPs is that they may relax significantly more strain configurations in a given period of time, allowing for a more precise understanding of strain-bandgap trends. The trends for DFT and various machine learning workflows for predicting strained bandgaps are presented in Fig. 3 for MoS_2. DFT bandgaps are computed in increments of 1% strain, while machine learning methods utilize smaller increments of 0.1% strain.

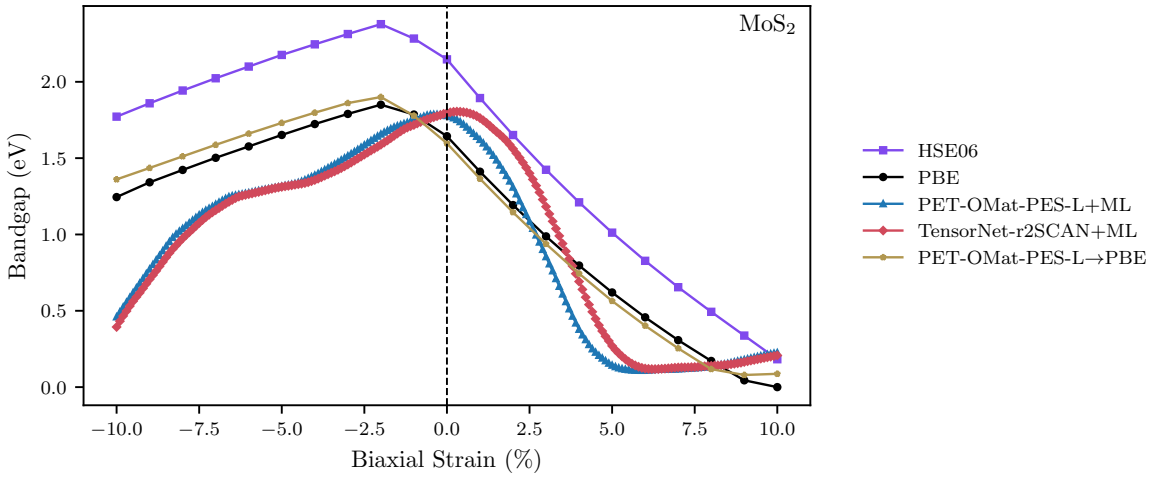


Figure 3: Bandgap evolution of monolayer MoS_2 under biaxial strain, comparing DFT and machine learning based approaches. Reference curves from HSE06 and PBE are shown alongside predictions from MLIP-based workflows. "+ML" refers to the PET-MAD-DOS bandgap model. Additionally, the PET-OMat-PES-L→PBE curve corresponds to PBE bandgaps evaluated on geometries relaxed using the PET-OMat-PES-L MLIP.

The PBE and HSE06 bandgaps follow an identical trend with strain, albeit with a slight offset such that the PBE functional more significantly underestimates bandgaps. Generally speaking, the bandgap decreases as the strain on the TMD increases. The blue and red curves depict the MLIP to PET-MAD-DOS bandgap prediction workflow, with extremely similar results. These entirely machine learning workflows produce the same qualitative trends as the DFT results, albeit with poor quantitative agreement. For example, the DFT results indicate a zero-bandgap near 10% biaxial strain, while the MLIP to PET-MAD-DOS workflows predict this transition at around 5% strain. This indicates that the rates of bandgap tunability agree quite poorly with DFT, and the PET-MAD-DOS model does not perform well on non-equilibrium structures. This may be expected since the MAD training set only includes equilibrium structures. While the bandgap agreement is poor, the model is able to create a realistic and continuous interpolation, correctly modeling the trends across the 0.1% strain grid. This once again emphasizes that among the workflows tested, the exact choice of MLIP has only a small impact on the predicted properties. Another issue is that a small "kink" is produced in the compressive strain regime using the PET-MAD-DOS model: this feature is not present in the DFT predictions. While it is not plotted for brevity, using PBE-relaxed structures in combination with the PET-MAD-DOS model for bandgap predictions yields results

identical to the blue curve, reinforcing the fundamental limitations of the PET-MAD-DOS model.

A more realistic workflow may include attaining MLIP-relaxed structures, then using a static DFT calculation to compute the bandgap or some other properties of interest. This is precisely the gold curve, using the PET-OMat-PES-L model for structural relaxations and the PBE functional for a final single-point calculation. This machine learning augmented workflow almost exactly reproduces the fully-PBE method, reinforcing MLIPs as promising for TMD structural relaxations. Abridged results for the other four TMDs are presented in Fig. 4 in Appendix A.2. The MLIP to PET-MAD-DOS workflow similarly captures the correct trends across the other TMDs, often with poor quantitative agreement. The model similarly introduces unexpected features like kinks, inconsistently under- or over-estimating the bandgaps. Remarkably, the workflow quantitatively captures the tensile strain trends of $MoSe_2$, $MoTe_2$, and WS_2 quite well compared to the baseline PBE results.

4 Limitations and Future Work

This study considers structural relaxations and bandgap predictions across five TMD materials– a fundamental limitation is that this analysis only applies to these five materials, and the results of machine learning workflows for more obscure TMDs may not be entirely consistent. This study is also limited to monolayer TMDs. While group-VI TMDs have historically seen the highest research interest,[1] it is also important to extend computational studies to TMDs with transition metals beyond the group-VI family, for example with Zr, Hf, Ti, V, Nb, Ta, Re, Pt, Pd, etc. These materials and their heterostructures may exhibit novel electronic properties and quantum phenomena useful for future technologies. Another limitation is that this study does not consider the exact speedups of MLIPs and machine learning augmented workflows compared to DFT. This would provide useful insights to better determine when machine learning provides significant speedups and where it may be more justified to use regular DFT.

While this work provides a quite comprehensive analysis of TMD monolayers across DFT and MLIP methods, further computational work is still needed in developing improved workflows. Given the high research interest in TMDs and other two-dimensional materials like graphene, black phosphorus (BP), and hexagonal boron nitride (h-BN), there is ample utility in tuning an exchange-correlation functional specifically for two-dimensional materials. Computationally, this could be as simple as tuning the empirical weighting parameters of a hybrid functional, or as complicated as weighting various exchange-correlation functionals together to reproduce experimental results across a wide range of compounds. Regardless, this would significantly aid in modeling the optical and electronic properties of current and future two-dimensional materials. Another plausible development could include fine-tuning a large, foundational MLIP specifically for use with two-dimensional materials. This could help in modeling complex heterostructures with large supercells while maintaining accurate predictions. It is also clear that machine learning workflows with enhanced capabilities at property prediction are needed, for example for bandgap predictions beyond the PBE level of theory.

This work primarily considers TMDs with ideal structures. However, using both DFT and MLIP-based methods it is important to continue accounting for phenomena and effects that are physically observed: for example, accounting for wrinkles in monolayers, dopants, defects, substrate effects, and in general, heterostructures, to better match experimental research. Significant research has already been performed on alloying and strain, and it is important to view these effects not only in isolation, but also in relation to one another.

5 Conclusion

This study evaluates various MLIPs and machine learning based workflows for structural relaxations and strained bandgap modeling in monolayer MoS_2, $MoSe_2$, $MoTe_2$, WS_2, and WSe_2. Across TMDs, all of the MLIPs tested attain similar structural relaxations with average RMSDs below 0.1Å compared to experimental unit-cells. In general, DFT and MLIP methods tend to underbind atoms compared to experiment. The r^{2}SCAN trained MLIPs achieve *slightly* better results for TMDs than MLIPs trained on PBE data, although the improvement is marginal. The TensorNet-MatPES-r^{2}SCAN model achieves the most accurate results and is used to viably relax a WS_2/WSe_2 lateral heterostructure, inducing a non-uniform strain distribution. This study demonstrates that for TMDs, an MLIP's level of theory is more important compared to its overall parameter count for its accuracy compared to experiment.

The PET-MAD-DOS bandgap model is used in conjunction with MLIP-relaxed structures to qualitatively predict correct bandgap trends under strain. However, the quantitative accuracy is quite poor; a modified workflow is introduced, relaxing structures using MLIPs and performing DFT single-point calculations to almost identically reproduce DFT bandgaps under strain. The PET-MAD-DOS model falls short due to its limited training dataset and its PBEsol level of theory.

Overall, MLIPs are extremely viable for speeding up structural relaxations for TMD modeling applications. While MLIPs are limited in terms of the properties they predict, they may be used in conjunction with DFT static calculations to accurately model specific material properties. This workflow is optimal for applications requiring large unit-cells or a significant number of geometric relaxations, such as in TMD lateral heterostructures or in TMD strain engineering.

A Supplemental Data

A.1 Structural Comparison

This appendix provides supplemental data to the summary table provided in the main text (Table 1). Raw measurements are provided across TMD materials and each of the MLIP methods, in addition to average errors per method.

Table 4: Lattice constants (in Å) for monolayer TMDs predicted by various MLIPs compared to experimental values from Samadi *et al.*[1] The final row reports the mean absolute percentage error (MAE) relative to experimental values.

Material	**Exp.**	**PBE**	**PET-OMat -PES-L**	**uma-s-1p2**	**uma-m-1p1**	**TensorNet -PBE**	**TensorNet -r2SCAN**
MoS_2	3.16	3.187	3.184	3.189	3.183	3.171	3.175
$MoSe_2$	3.29	3.320	3.320	3.324	3.319	3.315	3.313
$MoTe_2$	3.52	3.554	3.562	3.565	3.555	3.518	3.556
WS_2	3.15	3.186	3.182	3.196	3.187	3.174	3.174
WSe_2	3.28	3.322	3.318	3.328	3.321	3.324	3.302
MAE (%)	–	1.031%	1.008%	1.227%	1.009%	0.629%	0.729%

Table 5: Monolayer thickness (chalcogen-chalcogen distances) (in Å) for monolayer TMDs predicted by various MLIPs compared to experimental values from Samadi *et al.*[1] The final row reports the mean absolute percentage error (MAE) relative to experimental values.

Material	**Exp.**	**PBE**	**PET-OMat -PES-L**	**uma-s-1p2**	**uma-m-1p1**	**TensorNet -PBE**	**TensorNet -r2SCAN**
MoS_2	3.124	3.122	3.103	3.116	3.121	3.149	3.092
$MoSe_2$	3.173	3.342	3.306	3.336	3.339	3.355	3.298
$MoTe_2$	3.198	3.618	3.578	3.613	3.614	3.653	3.580
WS_2	3.124	3.143	3.120	3.137	3.143	3.161	3.107
WSe_2	3.173	3.362	3.322	3.355	3.361	3.365	3.311
MAE (%)	–	4.984%	3.986%	4.794%	4.927%	5.592%	3.725%

Table 6: Transition metal to chalcogen bond distances (d_{M-X}) (in Å) for monolayer TMDs predicted by various MLIPs compared to experimental values from Samadi *et al.*[1] The final row reports the mean absolute percentage error (MAE) relative to experimental values.

Material	**Exp.**	**PBE**	**PET-OMat -PES-L**	**uma-s-1p2**	**uma-m-1p1**	**TensorNet -PBE**	**TensorNet -r2SCAN**
MoS_2	2.402	2.413	2.406	2.412	2.411	2.415	2.398
$MoSe_2$	2.475	2.543	2.532	2.543	2.541	2.545	2.526
$MoTe_2$	2.586	2.736	2.726	2.739	2.735	2.731	2.724
WS_2	2.397	2.419	2.411	2.422	2.420	2.420	2.403
WSe_2	2.471	2.550	2.536	2.551	2.550	2.552	2.525
MAE (%)	–	2.626%	2.212%	2.660%	2.591%	2.649%	1.926%

Table 7: Root mean square deviations (RMSD) (in Å) for monolayer TMDs predicted by various MLIPs compared to experimental structures reported by Samadi *et al.*[1] The final row reports the mean RMSD value for each method.

Material	**PBE**	**PET-OMat -PES-L**	**uma-s-1p2**	**uma-m-1p1**	**TensorNet -PBE**	**TensorNet -r2SCAN**
MoS_2	0.012	0.012	0.009	0.007	0.011	0.014
$MoSe_2$	0.069	0.055	0.067	0.068	0.075	0.051
$MoTe_2$	0.172	0.156	0.170	0.170	0.186	0.156
WS_2	0.015	0.018	0.017	0.015	0.017	0.011
WSe_2	0.078	0.062	0.075	0.077	0.079	0.056
Mean	0.069	0.060	0.068	0.068	0.074	0.058

A.2 Strained Bandgap Results

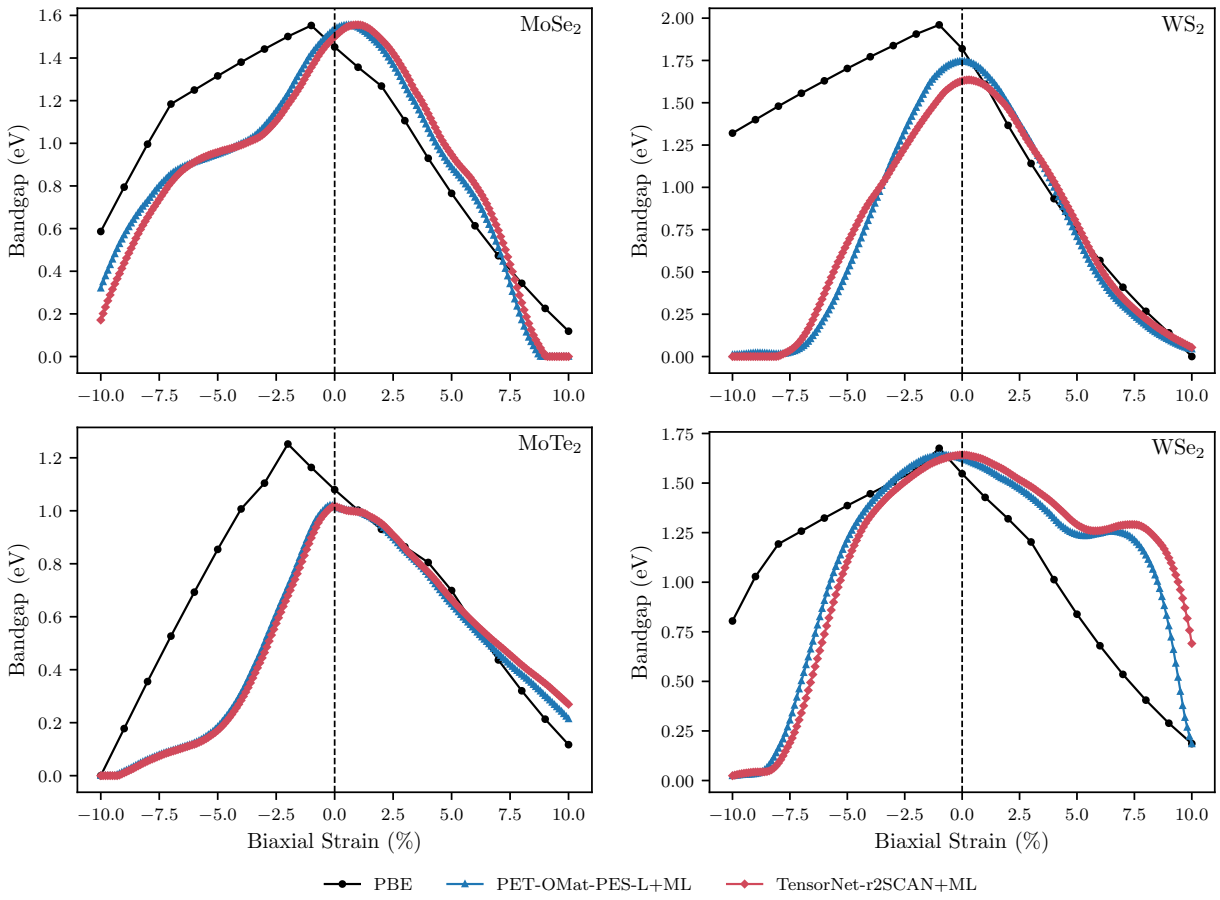


Figure 4: Bandgap evolution of various monolayer TMDs under biaxial strain, comparing DFT and MLIP approaches. PBE reference curves are shown alongside predictions from MLIP-based workflows. "+ML" refers to the PET-MAD-DOS bandgap model. Note the differing y-axis scales.

Works Cited

(1) Samadi, M.; Sarikhani, N.; Zirak, M.; Zhang, H.; Zhang, H.-L.; Moshfegh, A. Z. *Nanoscale Horizons* **2018**, *3*, 90–204.

(2) Bertolazzi, S.; Brivio, J.; Kis, A. *ACS Nano* **2011**, *5*, Publisher: American Chemical Society, 9703–9709.

(3) Roldán, R.; Castellanos-Gomez, A.; Cappelluti, E.; Guinea, F. *Journal of Physics: Condensed Matter* **2015**, *27*, 313201.

(4) Johari, P.; Shenoy, V. B. *ACS Nano* **2012**, *6*, Publisher: American Chemical Society, 5449–5456.

(5) Bendavid, L. I.; Zhong, Y.; Che, Z.; Konuk, Y. *Journal of Applied Physics* **2022**, *132*, 225303.

(6) Gong, C.; Zhang, H.; Wang, W.; Colombo, L.; Wallace, R. M.; Cho, K. *Applied Physics Letters* **2013**, *103*, 053513.

(7) Wei, W.; Dai, Y.; Sun, Q.; Yin, N.; Han, S.; Huang, B.; Jacob, T. *Phys. Chem. Chem. Phys.* **2015**, *17*, 29380–29386.

(8) Wei, W.; Dai, Y.; Huang, B. *Phys. Chem. Chem. Phys.* **2016**, *18*, 15632–15638.

(9) Mortazavi, B.; Zhuang, X.; Rabczuk, T.; Shapeev, A. V. *Mater. Horiz.* **2023**, *10*, 1956–1968.

(10) Batatia, I. et al. *The Journal of Chemical Physics* **2025**, *163*, 184110.

(11) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. a. *APL Materials* **2013**, *1*, 011002.

(12) Kresse, G.; Hafner, J. *Journal of Physics: Condensed Matter* **1994**, *6*, 8245.

(13) Kresse, G.; Furthmüller, J. *Computational Materials Science* **1996**, *6*, 15–50.

(14) Kresse, G.; Furthmüller, J. *Physical Review B* **1996**, *54*, 11169.

(15) Kresse, G.; Joubert, D. *Physical Review B* **1999**, *59*, 1758.

(16) Wang, V.; Xu, N.; Liu, J.-C.; Tang, G.; Geng, W.-T. *Computer Physics Communications* **2021**, *267*, 108033.

(17) Wang, V.; Tang, G.; Liu, Y.-C.; Wang, R.-T.; Mizuseki, H.; Kawazoe, Y.; Nara, J.; Geng, W. T. *The Journal of Physical Chemistry Letters* **2022**, *13*, PMID: 36480578, 11581–11594.

(18) Wood, B. M. et al. UMA: A Family of Universal Models for Atoms, 2025.

(19) Barroso-Luque, L.; Shuaibi, M.; Fu, X.; Wood, B. M.; Dzamba, M.; Gao, M.; Rizvi, A.; Zitnick, C. L.; Ulissi, Z. W. Open Materials 2024 (OMat24) Inorganic Materials Dataset and Models, 2024.

(20) Kaplan, A. D.; Liu, R.; Qi, J.; Ko, T. W.; Deng, B.; Riebesell, J.; Ceder, G.; Persson, K. A.; Ong, S. P. A Foundational Potential Energy Surface Dataset for Materials, 2025.

(21) Bigi, F.; Pegolo, P.; Mazitov, A.; Ceriotti, M. Pushing the limits of unconstrained machine-learned interatomic potentials, 2026.

(22) How, W. B.; Febrer, P.; Chong, S.; Mazitov, A.; Bigi, F.; Kellner, M.; Pozdnyakov, S.; Ceriotti, M. A universal machine learning model for the electronic density of states, 2025.

(23) Mazitov, A.; Chorna, S.; Fraux, G.; Bercx, M.; Pizzi, G.; De, S.; Ceriotti, M. *Scientific Data* **2025**, *12*, 1857.

(24) Xie, S.; Tu, L.; Han, Y.; Huang, L.; Kang, K.; Lao, K. U.; Poddar, P.; Park, C.; Muller, D. A.; DiStasio, R. A.; Park, J. *Science* **2018**, *359*, 1131–1136.