# Stillinger-Weber Potential for $MoS_2$: Parameterization and Sensitivity Analysis


Mingjian Wen, Sharmila Shirodkar, Petr Plecháč, Efthimios Kaxiras, Ryan Elliott, and Ellad Tadmor

University of Minnesota — Driven to Discover℠ — Aerospace Engineering & Mechanics

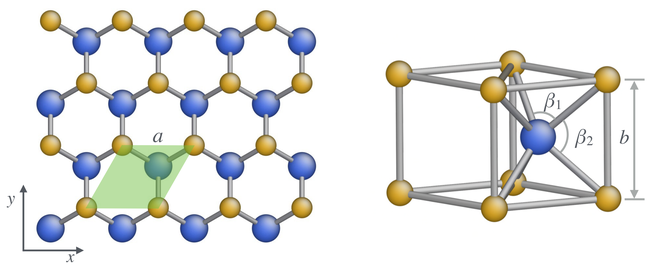


Science to Shape the Future of the Army — Army Research Office — 1951


## Objective:

- Parameterize a Stillinger-Weber interatomic potential for $MoS_2$ that is able to predict basic material properties, such as cohesive energy and lattice constants, with good transferability to properties outside the training and testing sets.
- Analyze the sensitivity of the potential to its fitting parameters. If some parameters are not identifiable, perform model reduction to simplify the model.

## Stillinger-Weber Potential for $MoS_2$:




$MoS_2$ structure

Equilibrium structural parameters predicted by SIESTA:

- $a = 3.20$ Å
- $b = 3.19$ Å
- $\beta_1 = 81.92°$
- $\beta_2 = 81.61°$

The Stillinger-Weber (SW) potential is widely used to describe the interaction in tetrahedral semiconductors, such as silicon (Stillinger and Weber, *PRB*, 1985), and compounds composed of the major II-VI elements Zn, Cd, Hg, S, Se, and Te (Zhou *et al*, *PRB*, 2013). It has also been adapted for 2D materials, such as monolayer $MoS_2$ and monolayer black phosphorus (Jiang *et al*, *Nanotechnology*, 2015). The total potential energy of a system consisting of $N$ atoms given by the SW potential is:

$$\mathcal{V}(\boldsymbol{r};\boldsymbol{\theta}) = \sum_{i=1}^{N}\sum_{j>i}^{N}\phi_2(r_{ij}) + \sum_{i=1}^{N}\sum_{j\neq i}^{N}\sum_{\substack{k>j\\k\neq i}}^{N}\phi_3(r_{ij},r_{ik},\beta_{jik}),$$

where the 2-body interactions are given by

$$\phi_2(r_{ij}) = A_{IJ}\left[B_{IJ}\left(\frac{r_{ij}}{\sigma_{IJ}}\right)^{-p} - \left(\frac{r_{ij}}{\sigma_{IJ}}\right)^{-q}\right]\exp\left(\frac{\sigma_{IJ}}{r_{ij}-r_{IJ}^{\mathrm{cut}}}\right),$$

and the 3-body bond angle term is

$$\phi_3(r_{ij},r_{ik},\beta_{jik}) = \lambda_{JIK}\left[\cos\beta_{jik} - \cos\beta_{jik}^0\right]^2\exp\left(\frac{\gamma_{IJ}}{r_{ij}-r_{IJ}^{\mathrm{cut}}} + \frac{\gamma_{IK}}{r_{ik}-r_{IK}^{\mathrm{cut}}}\right),$$

in which $r_{ij} = \|\boldsymbol{r}_i - \boldsymbol{r}_j\|$ is the bond length, and $\beta_{jik}$ is the angle between bonds $r_{ij}$ and $r_{ik}$. Only the bond angles of type S-Mo-S and Mo-S-Mo (the middle atom is the vertex) are considered. Both $\phi_2$ and $\phi_3$ vanish for $r > r^{\mathrm{cut}}$.

Equilibrium lattice constants and bond angles are pre-built into the potential by requiring:

$$(\partial\phi_2/\partial r)|_{r=a} = 0, \quad \text{and} \quad (\partial\phi_3/\partial\beta)|_{\beta=\beta^0} = 0.$$

## Force matching:

The potential parameters are optimized using a force-matching method in which the potential forces are matched as closely as possible with forces obtained from ab initio molecular dynamics (AIMD) calculations using the SIESTA code. This takes the form of a least-squares problem to minimize the following cost function:

$$C(\boldsymbol{\theta}) = \sum_{m=1}^{M}\frac{1}{2}w_m\|\boldsymbol{f}(\boldsymbol{\theta};\boldsymbol{r}_m) - \boldsymbol{f}_m^0\|^2,$$

where $M$ is the total number of configurations in the training set, $w_m$ is the weight for configuration $m$ (set to 1 in this work, since all forces are equally important), $\boldsymbol{f}_m^0$ is the reference forces from the AIMD training set, and

$$\boldsymbol{f}(\boldsymbol{\theta};\boldsymbol{r}_m) = (-\partial\mathcal{V}/\partial\boldsymbol{r})|_{\boldsymbol{r}_m}$$

is the forces on atoms (with coordinates $\boldsymbol{r}_m$) computed from the potential. The vector $\boldsymbol{\theta}$ represents the potential parameters.

## Fitting results and potential predictions:

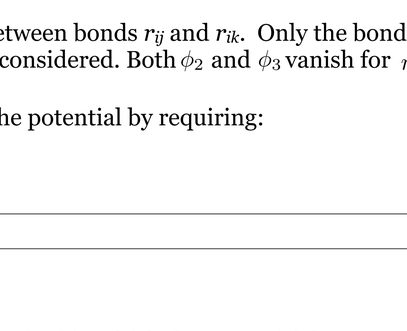
Fitted parameters in 2-body term

| Parameter | Mo-Mo | Mo-S | S-S |
|---|---|---|---|
| $A$ (eV) | 3.9781 | 11.3797 | 1.1907 |
| $B$ | 0.4446 | 0.5267 | 0.9015 |
| $\sigma$ (Å) | 2.8529 | 2.1751 | 2.8413 |
| $r^{\mathrm{cut}}$ (Å) | 5.5466 | 4.0269 | 4.5195 |
| $p = 5$ | $q = 0$ | | |

Fitted parameters in 3-body term

| $\gamma$(Å) | $\lambda_{\text{S-Mo-S}}$(eV) | $\lambda_{\text{Mo-S-Mo}}$(eV) | $r^{\mathrm{cut3}}_{\text{S-S}}$(Å) | $r^{\mathrm{cut3}}_{\text{Mo-Mo}}$(Å) |
|---|---|---|---|---|
| 1.3566 | 7.4767 | 8.1595 | 3.8609 | 4.3744 |

Energy versus lattice constant




The predictions of the new SW-TT750 potential are in better agreement with first principles (SIESTA) results for the energy versus lattice constant than other potentials available for $MoS_2$.

Two methods are used to compute linear thermal expansion coefficient (LTEC) using molecular dynamics. The direct method computes the LTEC from its definition:

$$\alpha_L = \frac{1}{\mathbb{E}[a]}\left.\frac{\partial\mathbb{E}[a]}{\partial T}\right|_{p=0},$$

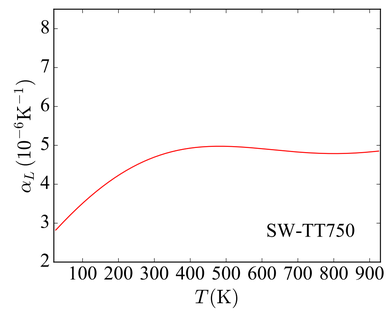
and the fluctuation method relates the LTEC to the covariance of Hamiltonian $\mathcal{H}$ and volume $V$ of the system by

$$\alpha_L = \frac{1}{2k_{\mathrm{B}}T^2\mathbb{E}[V]}\left(\mathbb{E}[\mathcal{H}V] - \mathbb{E}[\mathcal{H}]\mathbb{E}[V]\right),$$

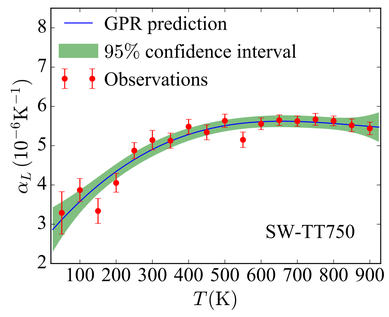
where $\mathbb{E}[\cdot]$ is the expectation, $k_{\mathrm{B}}$ is Boltzmann's constant, and $T$ is the temperature.

## Potential predictions (cont.):

Direct method




Fluctuation method




Curve fit is performed using Gaussian Process Regression (GPR).

Predictions of material properties

To test the transferability (accuracy away from the training set) of the fitted SW potential, we applied it to compute the lattice constants and stiffness of $MoS_2$, as well as their temperature dependence. The results of the new potential (SW-TT750) are in good agreement with first principles results (green).

| Method | $a$ (Å) | $b$ (Å) | $E_c$ (eV) | $C_{11}$ (N/m) | $C_{12}$ (N/m) |
|---|---|---|---|---|---|
| SW-TT750 | 3.19 | 3.19 | 15.28 | 119.2 | 41.0 |
| SW-Jiang2013 | 3.09 | 3.18 | 12.76 | 140.8 | 52.7 |
| SW-Jiang2015 | 3.11 | 3.12 | 3.72 | 105.0 | 28.7 |
| REBO | 3.17 | 3.24 | 21.48 | 154.4 | 45.8 |
| SIESTA (GGA) | 3.20 | 3.19 | 15.90 | - | - |
| VASP (GGA)[a] | 3.19 | 3.13 | 15.21 | - | - |
| VASP (GGA)[b] | - | - | - | 132.7 | 33.0 |
| VASP (GGA)[c] | - | - | - | 130.0 | 40.0 |

[a] Ding *et al*, *Physica B*, 2011 [b] Cakir *et al*, *APL*, 2014 [c] Copper *et al*, *PRB*, 2013

## Information-based sensitivity analysis:

The Fisher information matrix (FIM) provides a measure of the information change when the potential parameters are perturbed. For Langevin dynamics with a Boltzmann distribution, the FIM can be derived as

$$\boldsymbol{F}(\boldsymbol{\theta}) = \frac{1}{2k_{\mathrm{B}}T\eta}\frac{1}{M}\sum_{m=1}^{M}\left(\frac{\partial\boldsymbol{f}(\boldsymbol{\theta};\boldsymbol{r}_m)}{\partial\boldsymbol{\theta}}\right)^{\mathrm{T}}\left(\frac{\partial\boldsymbol{f}(\boldsymbol{\theta};\boldsymbol{r}_m)}{\partial\boldsymbol{\theta}}\right),$$

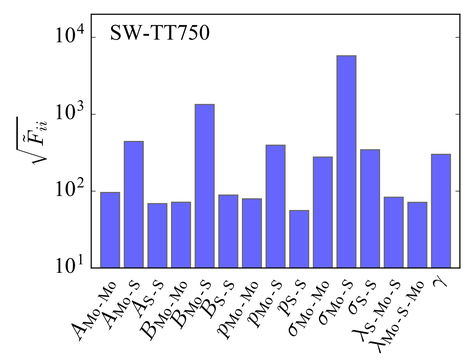
where $\eta$ is the damping coefficient in the equations of motion.

The inverse of the FIM provides a lower bound on the covariance of any unbiased estimator $\hat{\boldsymbol{\theta}}$ for the parameters (Cramér-Rao inequality):

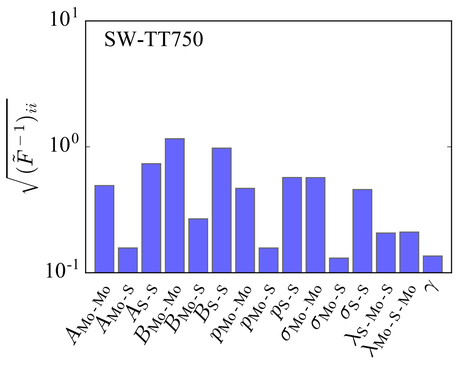
$$\mathrm{Cov}_{\boldsymbol{\theta}}[\hat{\boldsymbol{\theta}}] \geq \boldsymbol{F}^{-1}(\boldsymbol{\theta}).$$

The diagonal elements of the FIM and its inverse provide information on model sensitivity and uncertainty:

$\sqrt{F_{ii}}\rightarrow$ Sensitivity of model predictions to the fitting parameters




$\sqrt{[\boldsymbol{F}^{-1}]_{ii}}\rightarrow$ Uncertainty in parameters (lower bound of s.d.)




## Bounds for observables:

The FIM provides an upper bound for the change in the model prediction of any observable $t$ due to parameter perturbation $\Delta\boldsymbol{\theta}$ (a sharper version of Pinsker's inequality, Dupuis *et al*, *J. Uncertainty Quantif.*, 2016):

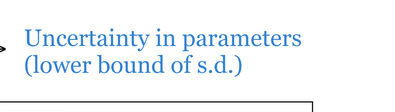
$$|\mathbb{E}_{\boldsymbol{\theta}+\Delta\boldsymbol{\theta}}[t] - \mathbb{E}_{\boldsymbol{\theta}}[t]| \leq \mathrm{Std}_{\boldsymbol{\theta}}[t]\sqrt{\Delta\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{F}\Delta\boldsymbol{\theta}} \qquad (*)$$

$\mathbb{E}_{\boldsymbol{\theta}}[t]$: expectation of observable $t$ using parameter $\boldsymbol{\theta}$

$\mathrm{Std}_{\boldsymbol{\theta}}[t]$: standard deviation of observable $t$ using parameter $\boldsymbol{\theta}$

As an example, consider the expected thickness of the $MoS_2$ sheet:

$$t = \mathbb{E}\left[\frac{1}{N}\left(\sum_{i=1}^{N}z_i^{\mathrm{top}} - \sum_{i=1}^{N}z_i^{\mathrm{bot}}\right)\right]$$




$N$: number of sulphur atoms in one layer $\quad z_i^{\mathrm{top}}$: $z$ coordinate of atom $i$ in top layer

Take the parameter $\sigma_{\text{Mo-S}}$ for example:

left of (*): 0.02011 < right of (*): 0.03029

It is seen that the bound (*) is satisfied, and the bound appears to be tight.

## Future work:

- Incorporateinformation-based sensitivity analysis into a general framework for interatomic potential fitting that is being developed, and use this framework to fit parametric potential models for 2D materials and heterostructures.
- Develop machine learning based nonparametric potential models for 2D materials and heterostructures, such as Neural Network and GPR models.