Supplementary Information for: Origin of the Machine Learning Forces Field Errors Across Metal Elements

Xingze Geng[1,2], Wentao Zhang[3], Lin-Wang Wang[3*] & Xiangying Meng [2*]

[1] College of Sciences, Northeastern University, Shenyang 110819, China

[2] Institute of Materials Intelligent Technology, Liaoning Academy of Materials, Shenyang 110004, China

[3] State Key Laboratory of Optoelectronic Materials and Devices, Institute of Semiconductors, Chinese Academy of Sciences, Beijing 100083, China

* Correspondence: lwwang@semi.ac.cn and x_y_meng@mail.neu.edu.cn

1. Metal-43 Dataset

Table 1 presents the detailed computational information of the Metal-43 dataset, including the initial structure IDs obtained from the Material Project (MP) database, space group, the PAW pseudopotentials used, and the temperatures at which the AIMD simulations were performed.

Supplementary Table 1: Detailed computational information of the Metal-43 dataset, including the initial structure ID from the Material Project database, spatial groups, PAW pseudopotentials used, and AIMD simulations at 300 K, 0.5 Tc, Tc, and 1.2 Tc temperatures. An empty temperature column indicates that no simulation was conducted at the corresponding temperature.

Element	mp-id	Space Group	300K	0.5Tc	Tc	1.2Tc	POTCAR
Li	mp-51	Fm-3m	300		454	548	Li_sv
Be	mp-87	P6_3/mmc	300	780	1560	1872	Be_sv
Na	mp-127	Im-3m	300		370	444	Na_pv
Mg	mp-153	P6_3/mmc	300	462	923	1108	Mg_pv
Al	mp-134	Fm-3m	300	467	933	1120	Al
K	mp-58	Im-3m	300			403	K_pv
Ca	mp-45	Fm-3m	300	558	1115	1338	Ca_sv
Sc	mp-67	P6_3/mmc	300	907	1814	2176	Sc_sv
Ti	mp-72	p6-mmm	300	970	1941	2297	Ti_pv
V	mp-146	Im-3m	300	1092	2183	2620	V_pv
Cr	mp-90	Im-3m	300	1090	2180	2616	Cr_pv

Mn	mp-35	I4-3m	300	760	1519	1823	Mn_pv
Fe	mp-13	Im-3m	300	906	1811	2173	Fe
Co	mp-54	P6_3/mmc	300	884	1768	2122	Co
Ni	mp-23	Fm-3m	300	864	1728	2138	Ni
Cu	mp-30	Fm-3m	300	678	1358	1630	Cu_pv
Zn	mp-79	P6_3/mmc	300		693	832	Zn
Ga	mp-142	Cmce	300			364	Ga
Rb		Im-3m	300			374	Rb_pv
Sr	mp-76	Fm-3m	300	525	1050	1260	Sr_sv
Y	mp-112	P6_3/mmc	300	900	1799	2159	Y_sv
Zr	mp-131	P6_3/mmc	300	1065	2128	2554	Zr_sv
Nb	mp-75	Im-3m	300	1375	2750	3300	Nb_sv
Mo	mp-129	Im-3m	300	1448	2896	3475	Mo_pv
Tc	mp-113	P6_3/mmc	300	1215	2430	2916	Tc_pv
Ru	mp-33	P6_3/mmc	300	1304	2607	3218	Ru_pv
Rh	mp-74	Fm-3m	300	1119	2237	2684	Rh_pv
Pd	mp-2	Fm-3m	300	914	1828	2194	Pd
Ag	mp-124	Fm-3m	300	618	1235	1482	Ag
Cd	mp-94	P6_3/mmc	300		594	713	Cd
In	mp-85	Fm-3m	300		430	516	In
Cs	mp-1	Im-3m	300			362	Cs_sv
Ba	mp-122	Im-3m	300	500	1000	1200	Ba_sv
Hf	mp-103	P6_3/mmc	300	1253	2506	3007	Hf_pv
Ta	mp-50	Im-3m	300	1645	3290	3948	Ta_pv
W	MP-91	Im-3m	300	1848	3695	4434	W_sv
Re	mp-8	P6_3/mmc	300	1730	3459	4151	Re
Os	mp-49	P6_3/mmc	300	1653	3306	3967	Os_pv
Ir	mp-101	Fm-3m	300	1370	2739	3287	It
Pt	mp-126	Fm-3m	300	1020	2041	2449	Pt
Au	mp-81	Fm-3m	300	669	1337	1606	Au
Hg	mp-10861	P6/mmm	300				Hg
Tl	mp-82	P6_3/mmc	300		577	692	Tl

2. Hyperparameters of NEP and DeePMD models

Unless otherwise specified, the MLFF models for the elements in this work adopt a uniform hyperparameters setting. Tables 2 and 3 summarize the main hyperparameters of the NEP and DeePMD models, including the cutoff radius, neural network architecture, and the number of training iterations.

Supplementary Table 2: DeePMD model hyperparameters

Parameter	**Value**
neighbor	100
cutoff	6.0
dtype	float64
R_max	6.0
R_min	0.5
smooth_fun	poly2_r
embedding network size	25,50,100
fitting_network size	200,200,200
activate_function	tanh
optimizer	Adam
start_energy_weight	100.0
limit_energy_weight	10.0
start_force_weight	100.0
limit_force_weight	10.0
epoch	50
start_lr	0.01
limit_lr	10^{-7}

Supplementary Table 3: NEP model hyperparameters

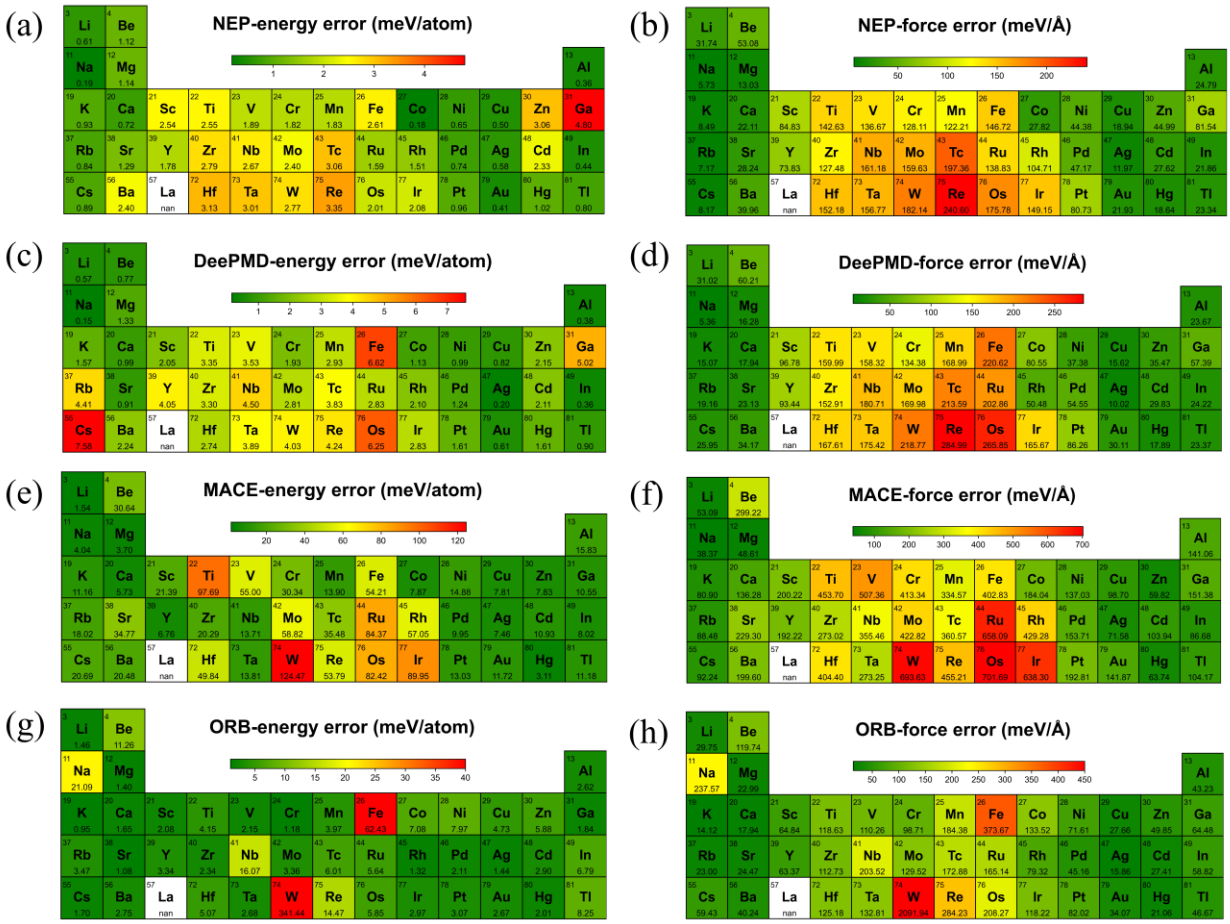
Parameter	**Value**
cutoff	6,4
n_max	4, 4

basis_size	8,8
l_max	4,2,0
neuron	50
lambda_e	1.0
lambda_f	1.0
population	50
generation	100000

3. Errors of different MLFF models

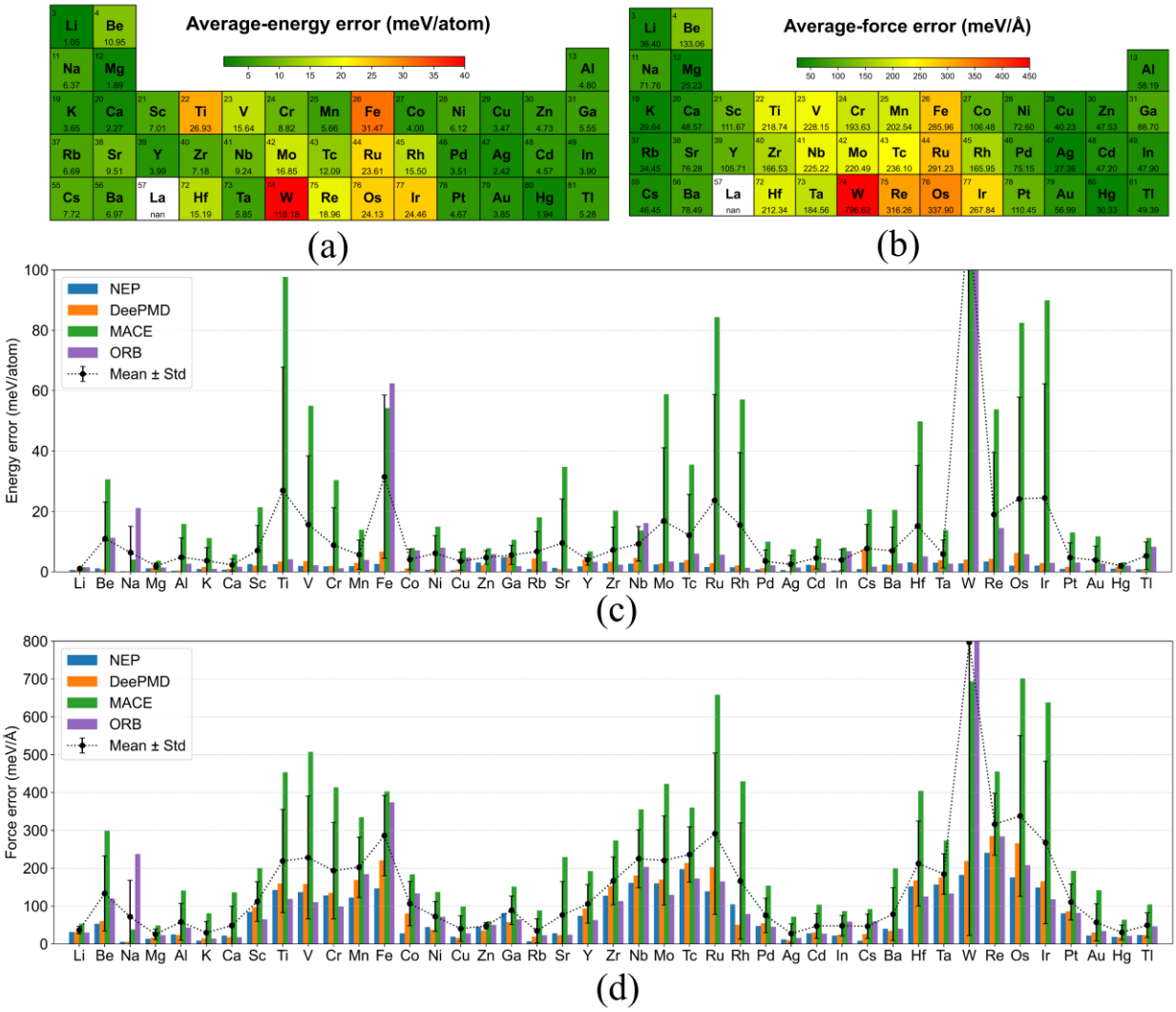
To further demonstrate that these errors originate from the intrinsic properties of elements rather than specific models or architectures, we also conducted tests on two General-Purpose large models. These models designed to cover a broader range of elements, featuring enhanced generalization capability and model size with parameter counts reaching the million scale. Representative examples of this category include the Higher Order Equivariant Message Passing Neural Networks (MACE) [1,2]and ORB[3].

Unlike the smaller models primarily used in this study (NEP and DeePMD), both MACE and ORB are pre-trained general-purpose models. Therefore, we directly employed the publicly released models for testing. Specifically, the MACE model used in this study is the large model trained on the Materials Project Relaxation Trajectory (MPTrj) dataset[4], while ORB used its latest version V3 model. Meanwhile, it is worth noting that due to differences in the DFT parameter settings between the datasets used for training the General-Purpose models and the dataset constructed in this study, there is an energy reference offset. Therefore, a constant energy shift was applied to the predicted energies during testing. Specifically, for each element, the shift was defined as the average difference between the mode-predicted energies and the reference DFT energies for the corresponding test system. Given that practical applications typically focus on relative differences rather than absolute energy, such a correction is both reasonable and commonly adopted in model evaluation.



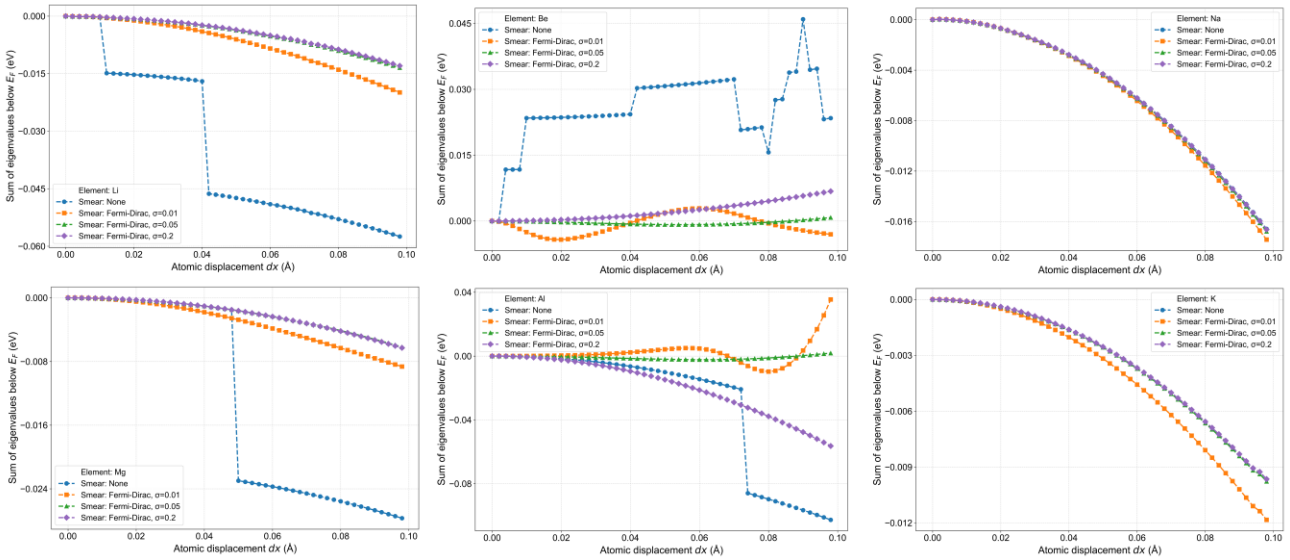
Supplementary Figure 1: Energy and force errors of four machine learning force fields (NEP, DeePMD, MACE, and Orb) across 43 metallic elements, presented in the form of the periodic table.

Figure 1 presents the energy and force errors of the four machine learning force fields, NEP, DeePMD, mace, and ORB, in the form of a periodic table of elements respectively. Figure. 2 (c) and (b) further compares the energy and force errors of the four MLFF models across different metallic systems. Among them, the MACE model exhibits slightly weaker overall performance, whereas the ORB model has approached the accuracy of small models (NEP and DeePMD) that undergo separate fitting optimization. Nevertheless, it can be seen that, regardless of the model, the error distribution trends across different metals are highly consistent, indicating that the distribution patteren of errors on the periodic table of elements has a certain universality and reflects the influence of the intrinsic properties of elements on fitting difficulty.

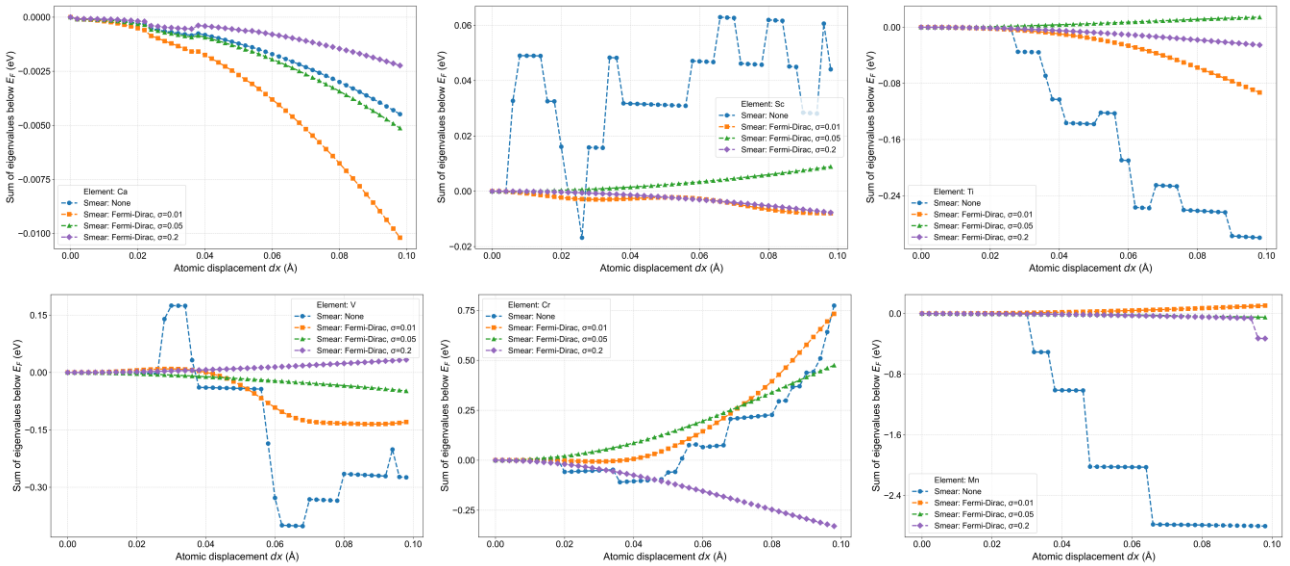


Supplementary Figure 2: (**a, b**) Average energy and force errors of NEP, DeePMD, MACE, and ORB models across 43 metallic elements, presented in the form of the periodic table. (**c, d**) Histogram of energy and force error distributions for four models across different elemental systems, with mean values and standard deviations indicated.

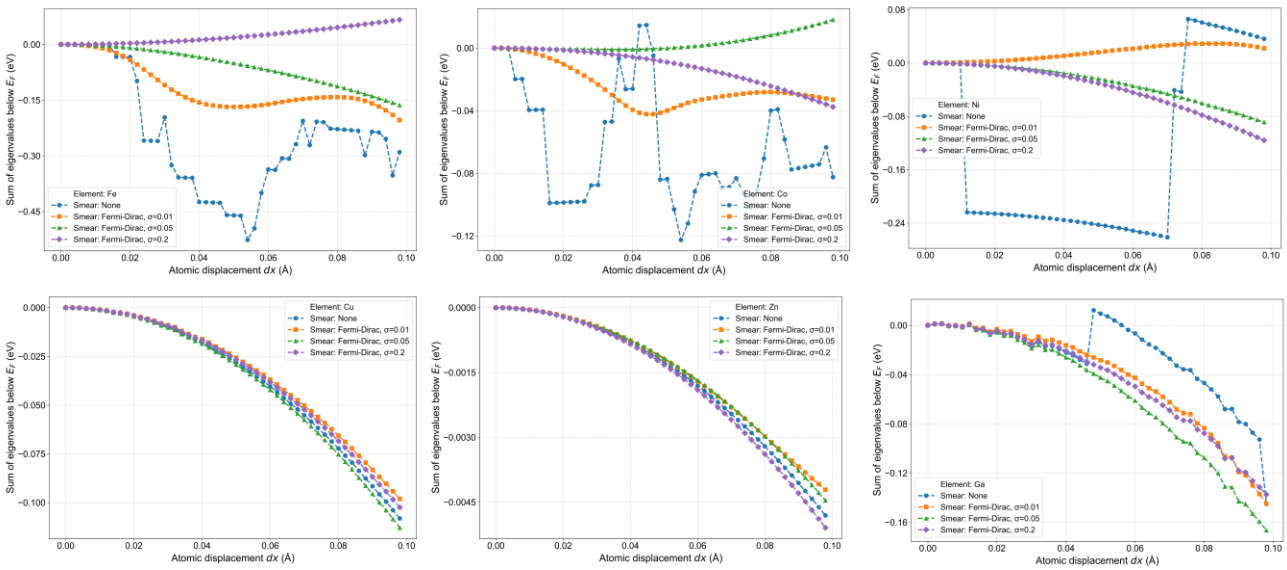
4. Variation of the Sum of Eigenvalues below the Fermi level under Different Condition



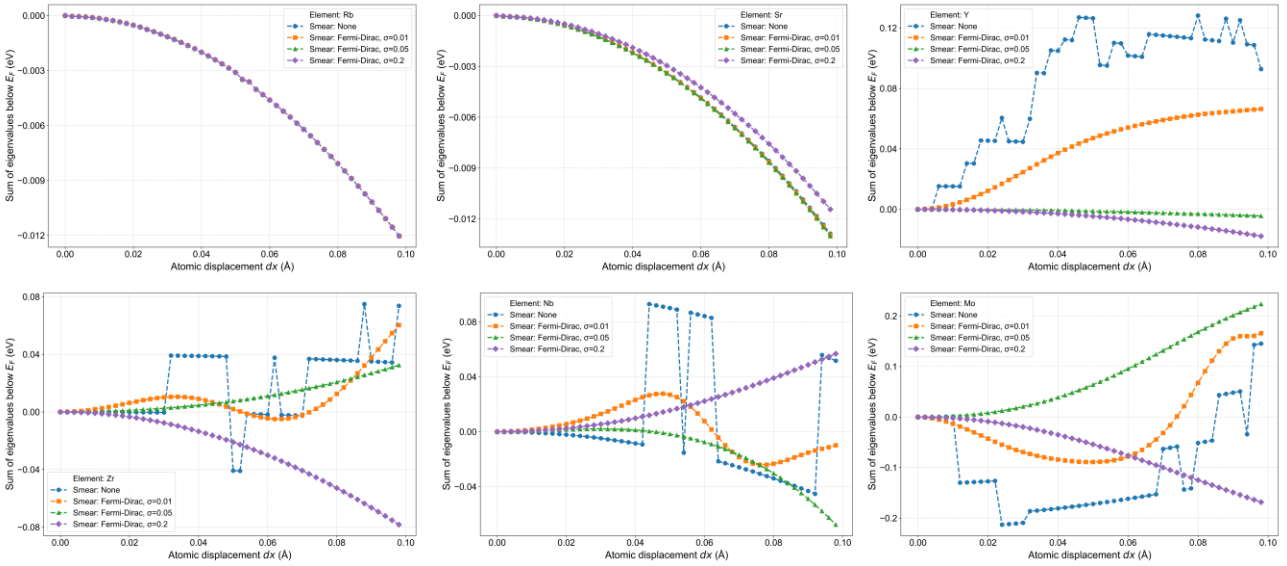
Supplementary Figure 3: Variation of the sum of eigenvalues below the Fermi level under atomic displacement perturbations for elements Li, Be, Na, Mg, Al, and K. Different curves represent various smearing methods applied.



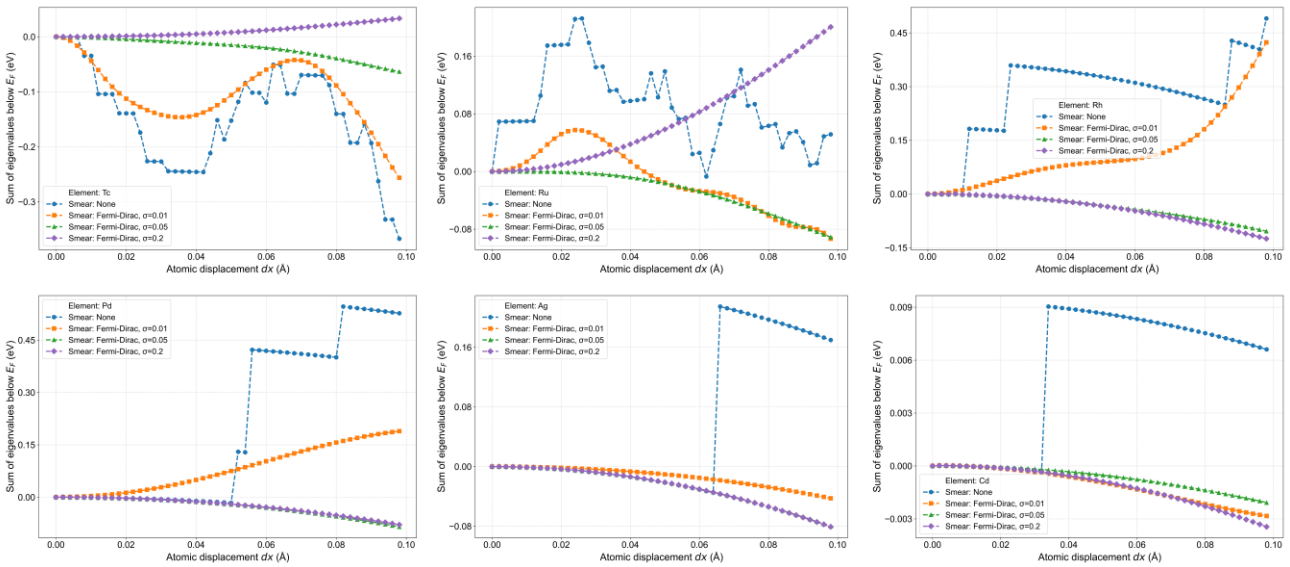
Supplementary Figure 4: Variation of the sum of eigenvalues below the Fermi level under atomic displacement perturbations for elements Ca, Sc, Ti, V, Cr, and Mn. Different curves represent various smearing methods applied.



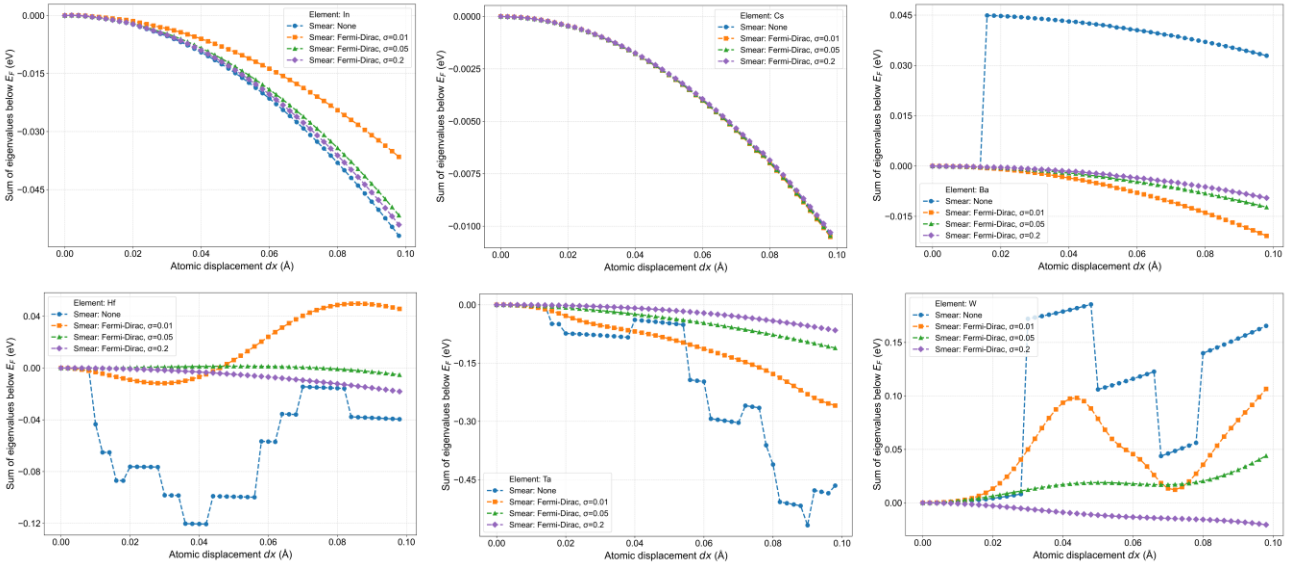
Supplementary Figure 5: Variation of the sum of eigenvalues below the Fermi level under atomic displacement perturbations for elements Fe, Co, Ni, Cu, Zn and Ga. Different curves represent various smearing methods applied.



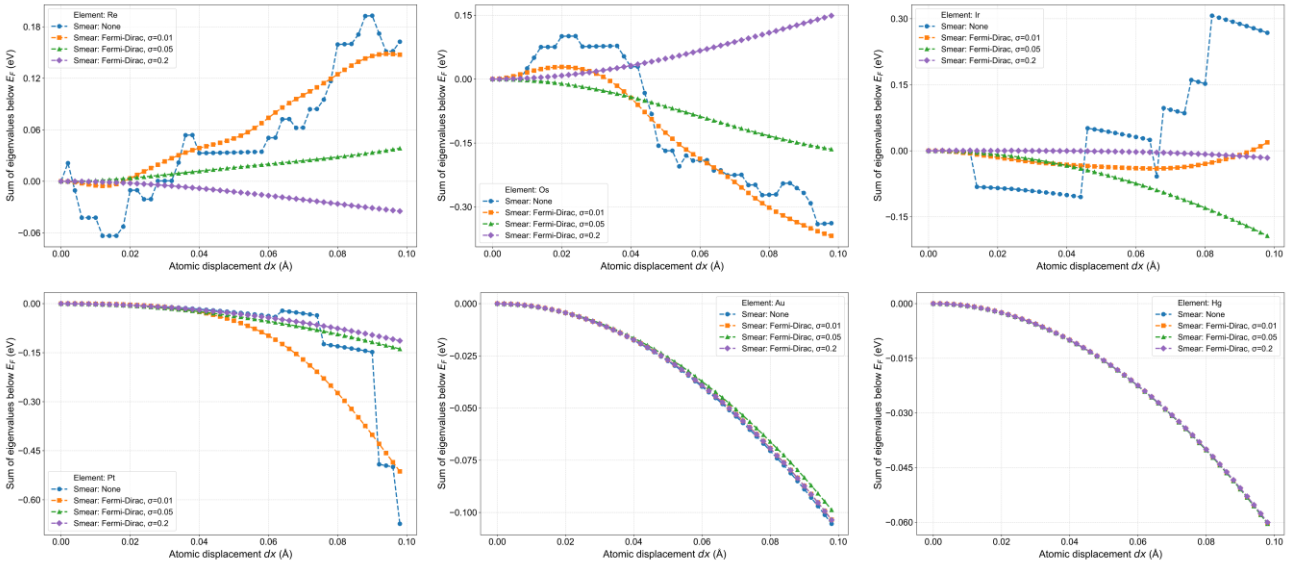
Supplementary Figure 6: Variation of the sum of eigenvalues below the Fermi level under atomic displacement perturbations for elements Rb, Sr, Y, Zr, Nb, and Mo. Different curves represent various smearing methods applied.



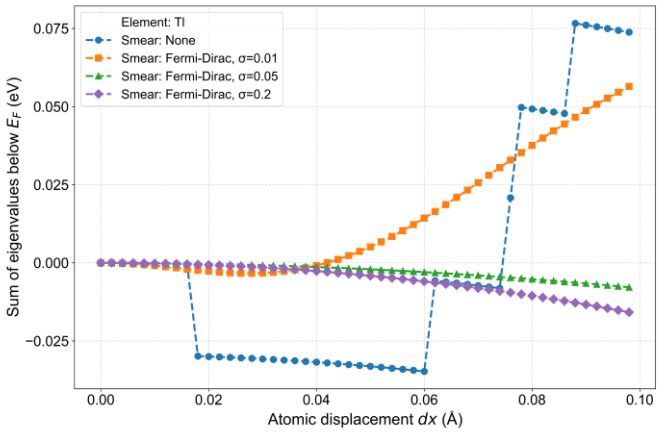
Supplementary Figure 7: Variation of the sum of eigenvalues below the Fermi level under atomic displacement perturbations for elements Tc, Ru, Rh, Pd, Ag, and Cd. Different curves represent various smearing methods applied.



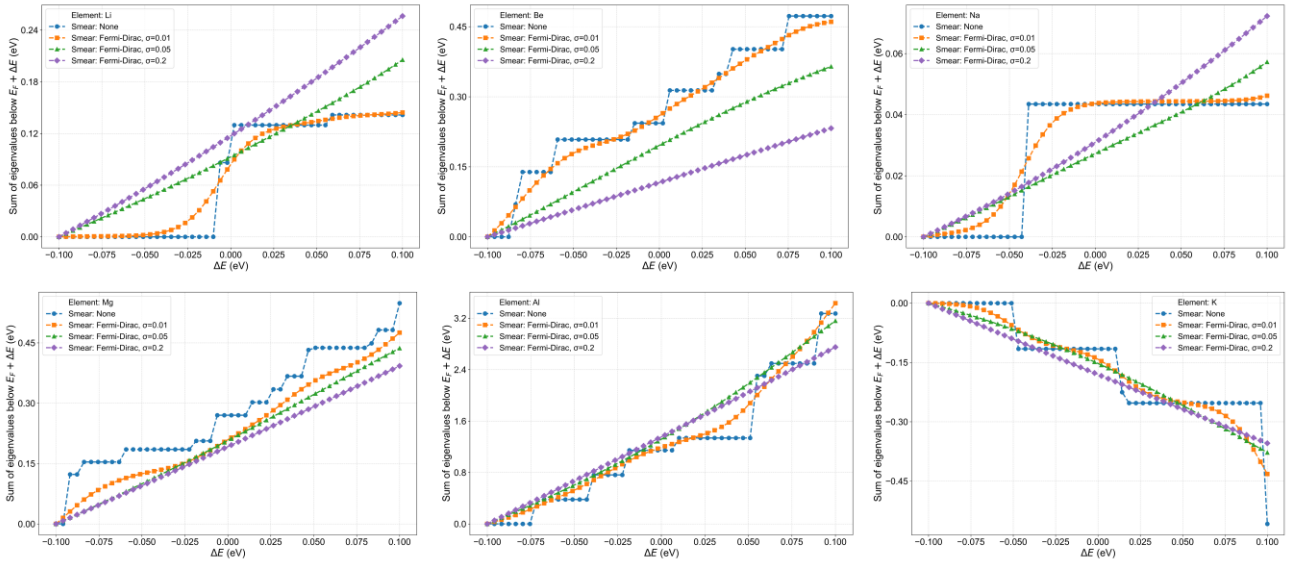
Supplementary Figure 8: Variation of the sum of eigenvalues below the Fermi level under atomic displacement perturbations for elements In, Cs, Ba, Hf, Ta, and W. Different curves represent various smearing methods applied.



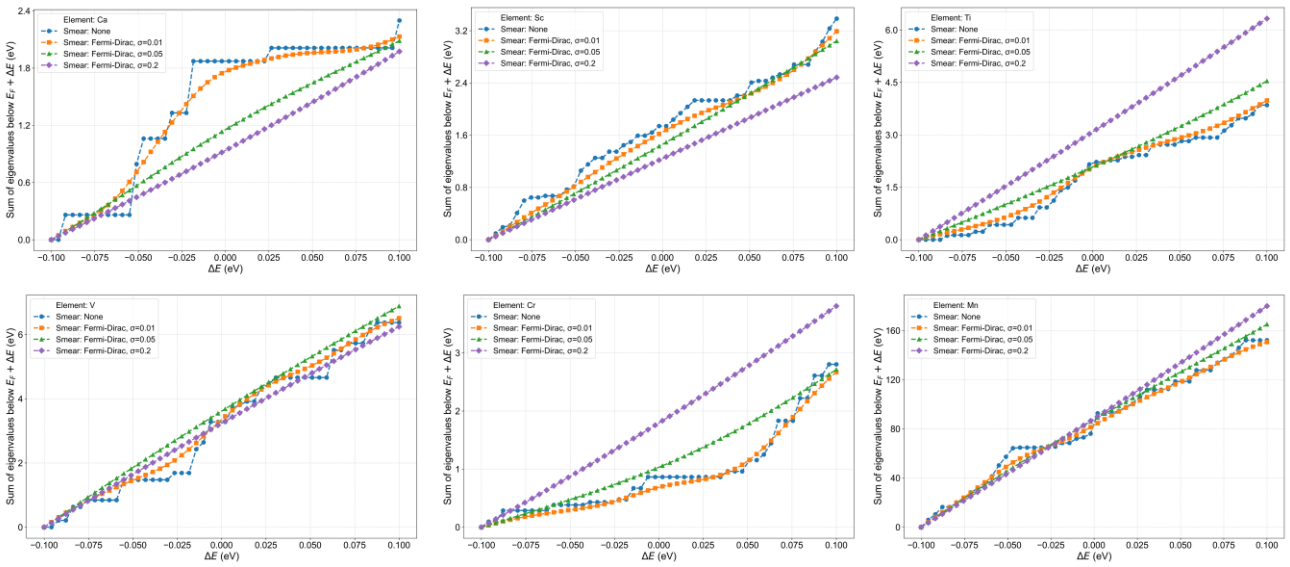
Supplementary Figure 9: Variation of the sum of eigenvalues below the Fermi level under atomic displacement perturbations for elements Re, Os, Ir, Pt, Au, and Hg. Different curves represent various smearing methods applied.



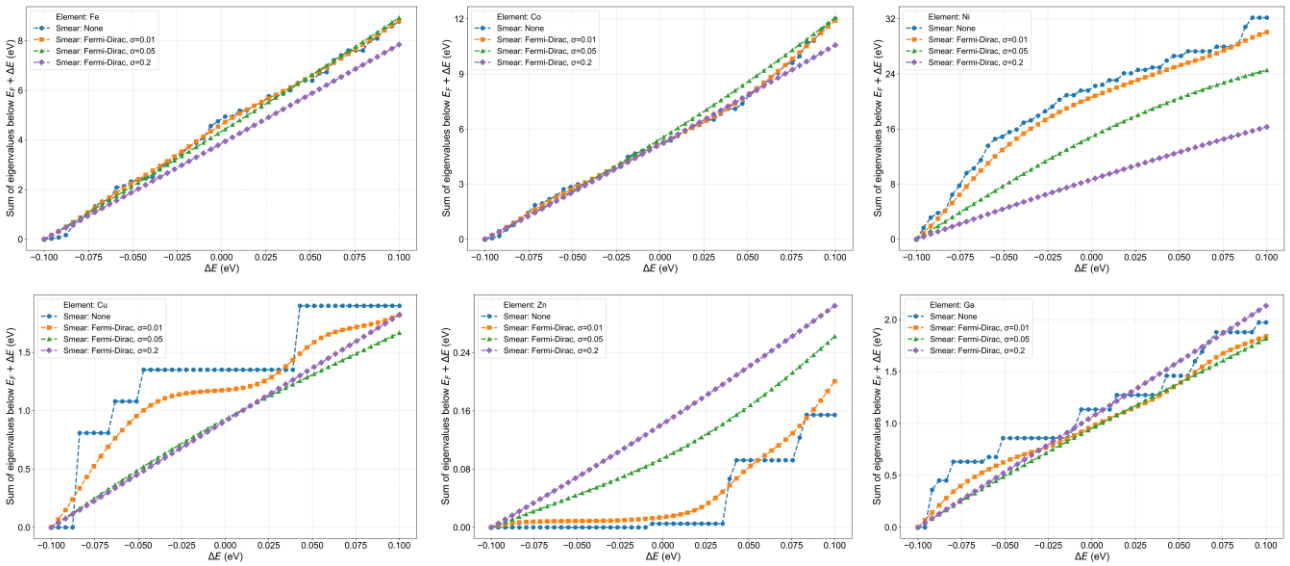
Supplementary Figure 10: Variation of the sum of eigenvalues below the Fermi level under atomic displacement perturbations for elements Tl. Different curves represent various smearing methods applied.



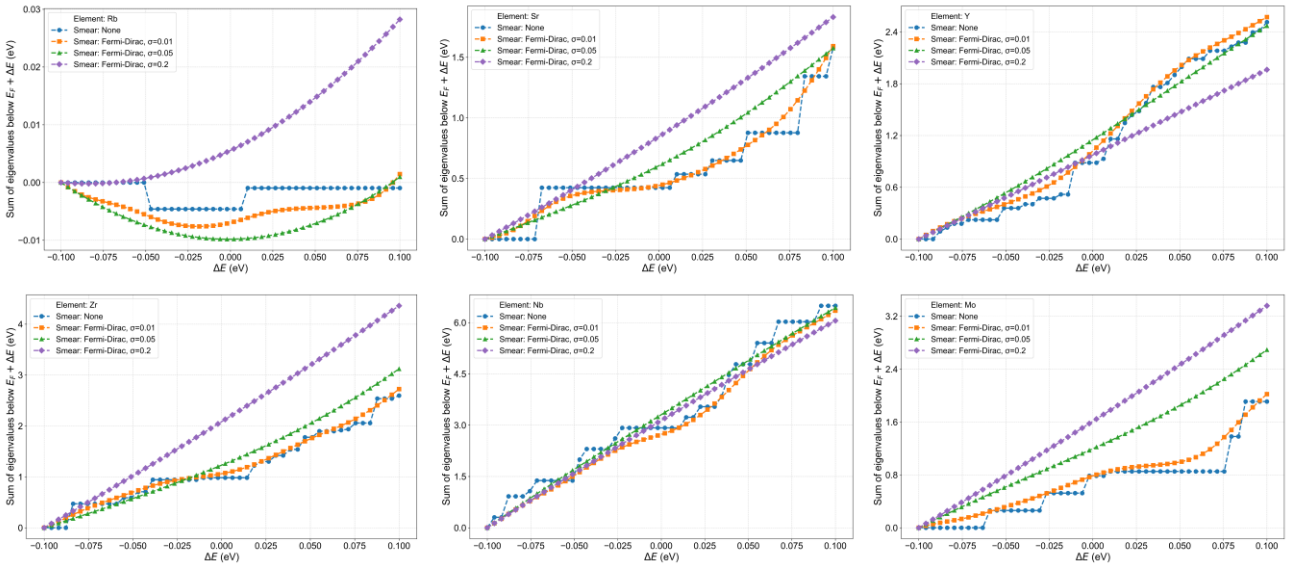
Supplementary Figure 11: Variation of the sum of eigenvalues below the Fermi level under Fermi level shifts for elements Li, Be, Na, Mg, Al, and K. Different curves represent various smearing methods applied.



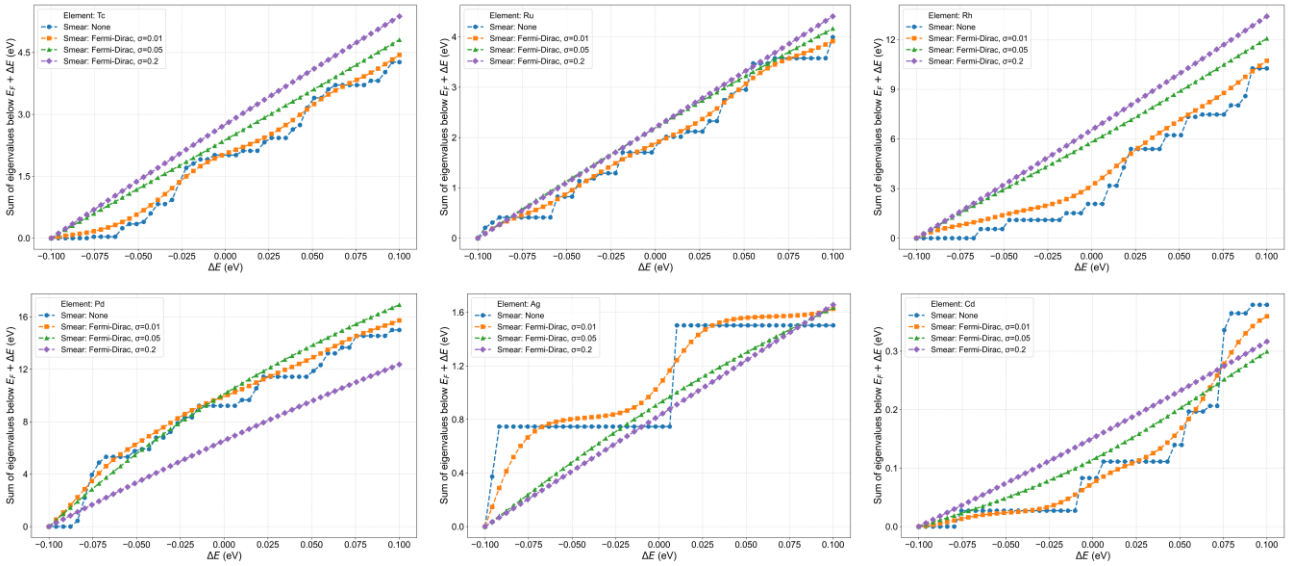
Supplementary Figure 12: Variation of the sum of eigenvalues below the Fermi level under Fermi level shifts for elements Ca, Sc, Ti, V, Cr, and Mn. Different curves represent various smearing methods applied.



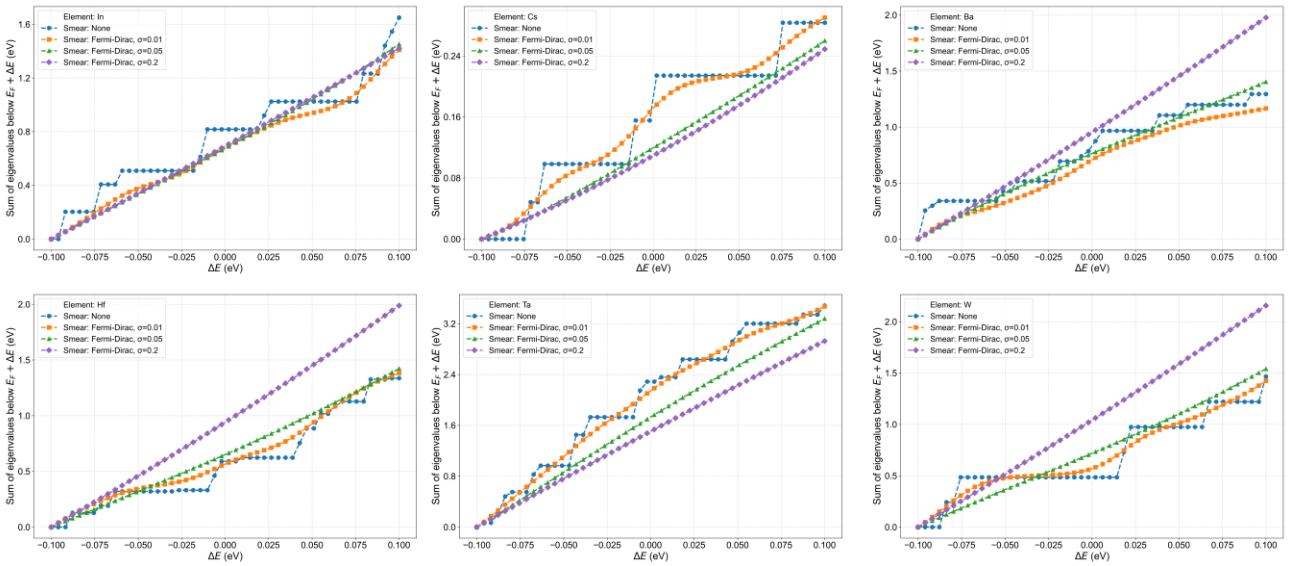
Supplementary Figure 13: Variation of the sum of eigenvalues below the Fermi level under Fermi level shifts for elements Fe, Co, Ni, Cu, Zn, and Ga. Different curves represent various smearing methods applied.



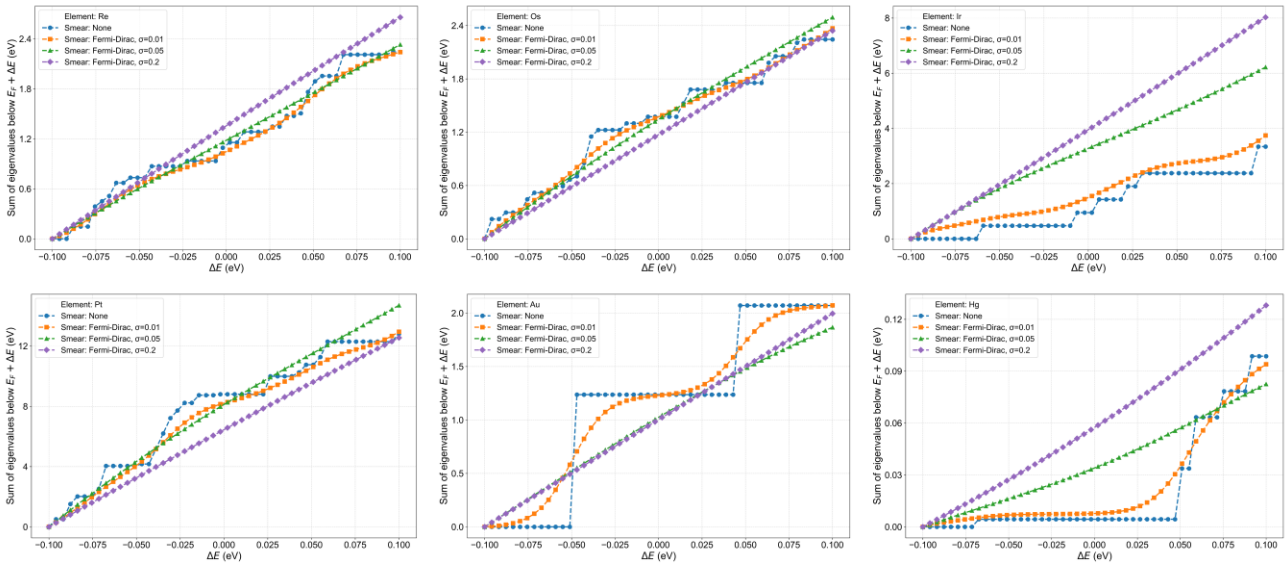
Supplementary Figure 14: Variation of the sum of eigenvalues below the Fermi level under Fermi level shifts for elements Rb, Sr, Y, Zr, Nb, and Mo. Different curves represent various smearing methods applied.



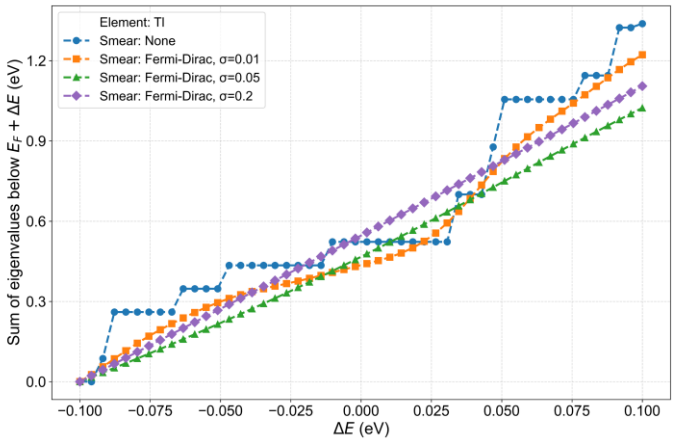
Supplementary Figure 15: Variation of the sum of eigenvalues below the Fermi level under Fermi level shifts for elements Tc, Ru, Rh, Pd, Ag, and Cd. Different curves represent various smearing methods applied.



Supplementary Figure 16: Variation of the sum of eigenvalues below the Fermi level under Fermi level shifts for elements In, Cs, Ba, Hf, Ta, and W. Different curves represent various smearing methods applied.



Supplementary Figure 17: Variation of the sum of eigenvalues below the Fermi level under Fermi level shifts for elements Re, Os, Ir, Pt, Au, and Hg. Different curves represent various smearing methods applied.



Supplementary Figure 18: Variation of the sum of eigenvalues below the Fermi level under Fermi level shifts for elements Tl. Different curves represent various smearing methods applied.

Reference

1. Batatia, I., Kovács, D. P., Simm, G. N. C., Ortner, C. & Csányi, G. MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields. Preprint at https://doi.org/10.48550/arXiv.2206.07697 (2023).

2. Batatia, I. The design space of E(3)-equivariant atom-centred interatomic potentials.

3. Neumann, M. *et al.* Orb: A Fast, Scalable Neural Network Potential. Preprint at https://doi.org/10.48550/arXiv.2410.22570 (2024).

4. Jain, A. *et al.* Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).