Supplementary Information for: “Representing Born effective charges with equivariant graph convolutional neural networks”

Alex Kutana[1], Koji Shimizu[2], Satoshi Watanabe[3], Ryoji Asahi[1]✉

[1]Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Japan
[2]National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan
[3]The University of Tokyo, Tokyo, Japan
✉email: ryoji.asahi@chem.material.nagoya-u.ac.jp

Bulk parent structures

The following bulk structures from the Materials Project database[1] were used for dataset generation: Pnma $CaTiO_3$ (mp-4019), Pnma Li_3PO_4 (mp-2878), Fm-3m cubic ZrO_2 (mp-1565), $I4_1$/amd tetragonal ZrO_2 (mp-2574, mp-754403), $P2_1$/c monoclinic ZrO_2 (mp-2858).

Model size and memory allocation

Model BM1 has 510081 parameters and 16323 buffers, with model size being 2105616 bytes in single precision and 4211232 bytes in double precision. Memory size allocated by CUDA to store the model, as given by `'torch.cuda.memory_allocated'`, is 2118656 bytes in single precision, and 4216832 bytes in double precision. The size of the model save file on disk is 2562091 bytes. Model BM2 has 139905 parameters and 9027 buffers, with model size being 595728 bytes in single precision and 1191456 bytes in double precision. Memory size allocated by CUDA to store the model is 610304 bytes in single precision, and 1200640 bytes in double precision. The size of the model save file on disk is 1051947 bytes.

Acoustic sum rule for the Born charges

The acoustic sum rule/charge neutrality condition for the Born charges states[2,3]:

$$\sum_i Z^*_{i,\alpha\beta} = 0$$

i.e., the sum over all atoms in the unit cell should be zero for each component of the Born charge tensor. We tested how well the outputs of the BM1 and BM2 models satisfy this condition. The mean absolute error is defined as

$$MAE = \frac{1}{9N_{struct}} \sum_{s=1}^{N_{struct}} \sum_{\alpha,\beta=1}^{3} \left| \sum_{i \in S_s} Z^*_{i,\alpha\beta} \right|$$

where N_{struct} is the total number of structures in the dataset, and S_s labels the structure with index s. The MAE of 0.154 and 0.173 were obtained for the BM1 and BM2 models, respectively for the combined dataset of perovskites, Li_3PO_4, and ZrO_2. A subsequent correction is applied to the model outputs to ensure that the sum rule is satisfied exactly.

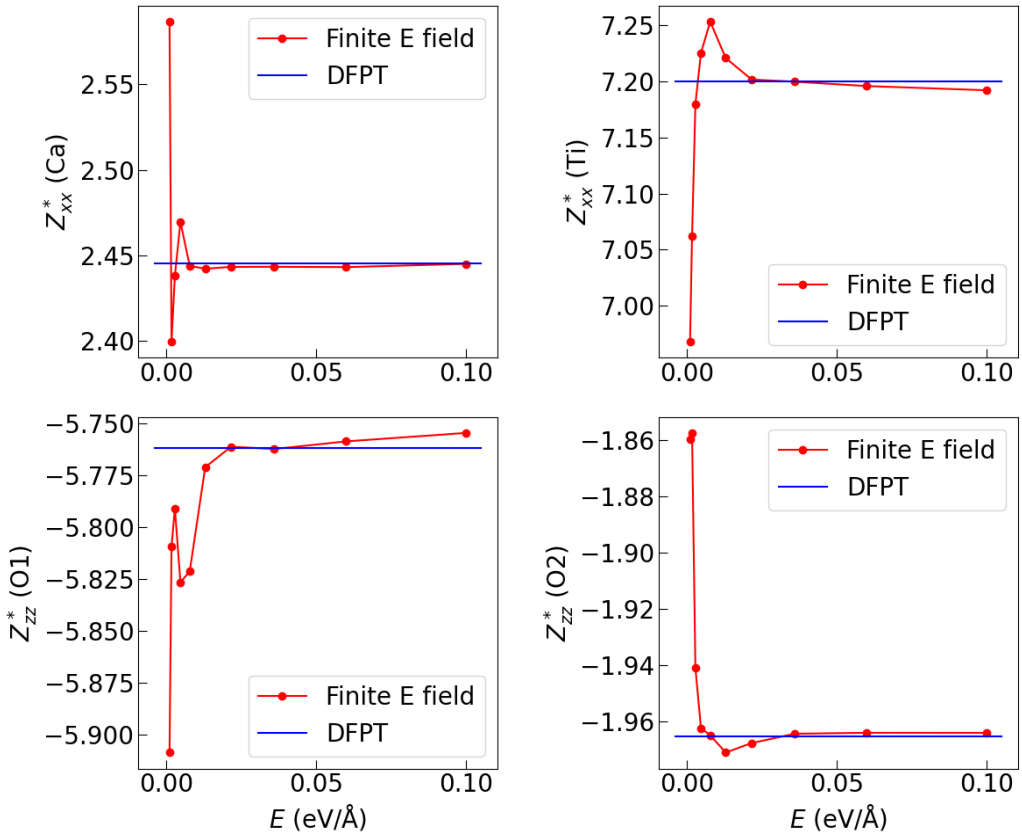

Figure S1. Comparison of the values of Born effective charges in $CaTiO_3$ obtained using density functional perturbation theory and with finite electric field method.

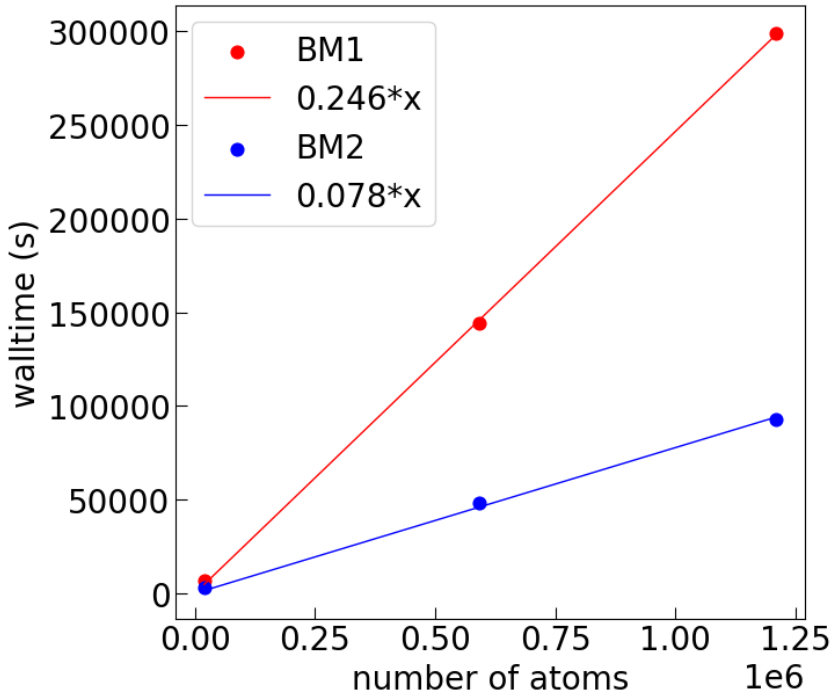

Figure S2. Training times for Equivar base model 1 (BM1) and base model 2 (BM2) on a single NVIDIA RTX A6000 GPU for 500 epochs. The numbers of parameters are 510,081 and 139,905 for BM1 and BM2, respectively. Training time scales approximately linearly with both the number of atoms in the dataset and number of model parameters, as 0.493 ms/atom/epoch for BM1, and 0.156 ms/atom/epoch for BM2.

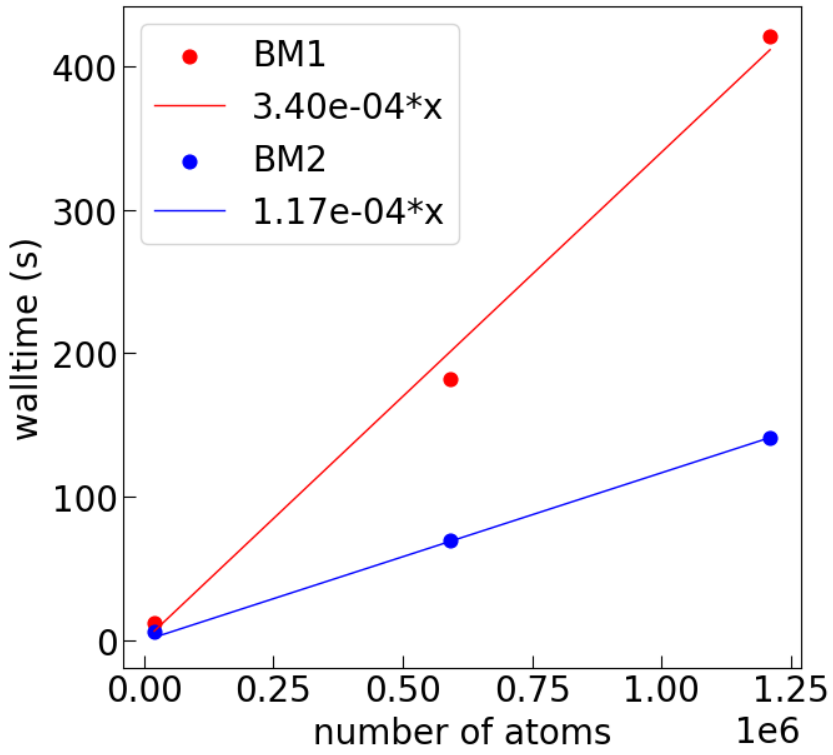

Figure S3. Inference times for Equivar base model 1 (BM1) and base model 2 (BM2) on a single NVIDIA RTX A6000 GPU. The numbers of parameters are 510,081 and 139,905 for BM1 and BM2, respectively. Inference time scales approximately linearly with both the number of atoms in the dataset and number of model parameters, as 0.340 ms/atom for BM1, and 0.117 ms/atom for BM2.

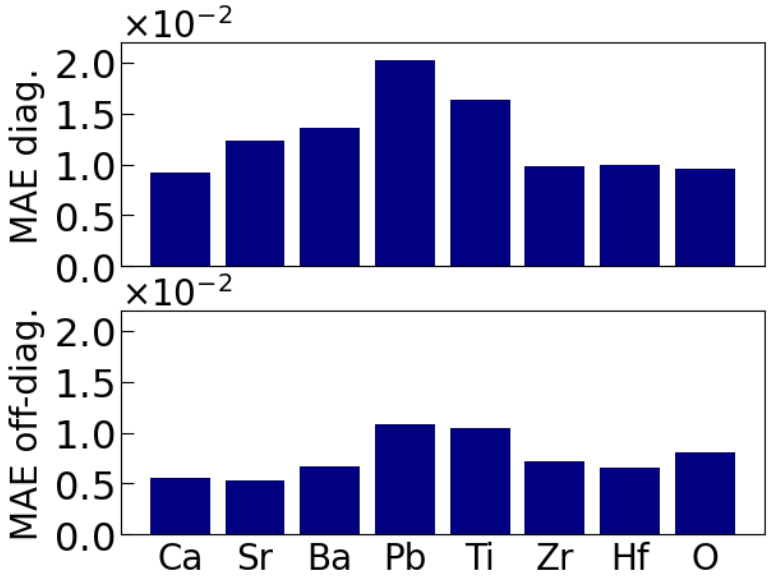

Figure S4. Five-fold cross-validation mean absolute error (MAE) of the Equivar model for diagonal and off-diagonal tensor components of atomic Born effective charges and individual atomic species in substituted perovskites.

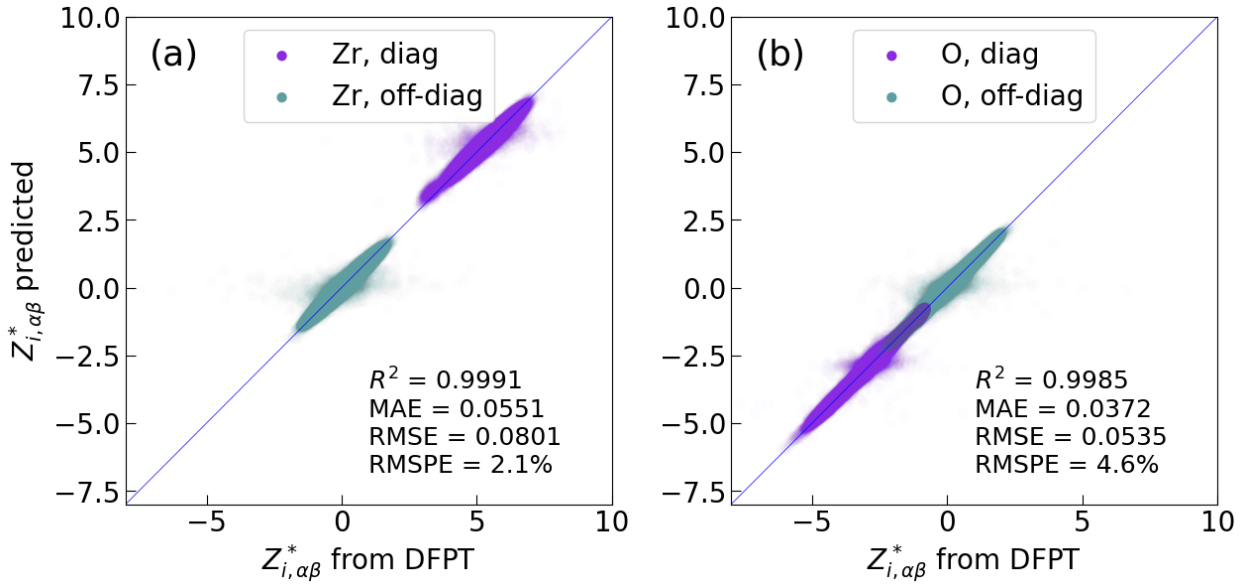

Figure S5. Five-fold cross-validation results for the model with 3 interaction layers and a cutoff radius of 5 Å, trained with tensors of Born effective charges of ZrO_2. Other hyperparameters are the same as those of BM1. Panels show results for different atomic species: (a) Zr, (b) O. The model performance is slightly worse than that for the model with 6 interaction layers and a cutoff radius of 3 Å, shown in figure 5.

Table S1. Benchmark of the model for Born effective charge tensor using a dataset of 224 perovskites while changing hyperparameters. '`model_size`' designates the number of trainable model parameters. The model was trained for 250 epochs, and best validation performance was recorded.

embedding_dim	**num_radial**	**int_num_feat**	**edge_sh_lmax**	**num_interactions**	**train_loss**	**validation_loss**	**model_size**
8	32	32	2	6	0.00798	0.01271	136737
16	32	32	2	6	0.0071	0.01186	137793
32	32	32	2	6	0.00655	0.0117	139905
64	32	32	2	6	0.00662	0.01136	144129
128	32	32	2	6	0.0071	0.01164	152577
32	8	32	2	6	0.0096	0.01288	126081
32	16	32	2	6	0.00825	0.0119	130689
32	32	32	2	6	0.00683	0.01104	139905
32	64	32	2	6	0.00574	0.0121	158337
32	128	32	2	6	0.00515	0.01489	195201
32	32	8	2	6	0.0138	0.01685	15489
32	32	16	2	6	0.01031	0.01417	42369
32	**32**	**32**	**2**	**6**	**0.00715**	**0.01196**	**139905**[(b)]

32	**32**	**64**	**2**	**6**	**0.00523**	**0.01092**	**510081**[a]
32	32	128	2	6	0.00331	0.00975	1950849
32	32	32	2	6	0.00685	0.01177	139905
32	32	32	4	6	0.00761	0.01211	139905
32	32	32	2	4	0.00898	0.01308	94721
32	32	32	2	6	0.00764	0.01181	139905
32	32	32	2	8	0.00613	0.01111	185089
32	32	32	2	10	0.00562	0.01085	230273
32	32	32	2	12	0.00503	0.00997	275457
32	32	32	2	14	0.00522	0.01045	320641
32	32	32	2	16	0.00459	0.01012	365825

(a) BM1 model
(b) BM2 model

References

1. Jain, A. *et al.* Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).
2. Pick, R. M., Cohen, M. H. & Martin, R. M. Microscopic Theory of Force Constants in the Adiabatic Approximation. *Phys Rev B* **1**, 910–920 (1970).
3. Gonze, X. & Lee, C. Dynamical matrices, Born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory. *Phys Rev B* **55**, 10355–10368 (1997).