Supplementary Information for:

Grain Boundary Sliding and Distortion on a Nanosecond Timescale Induce Trap States in $CsPbBr_3$: Ab Initio Investigation with Machine Learning Force Field

Dongyu Liu[a], Yifan Wu[b], Andrey S. Vasenko*[a,c] and Oleg V. Prezhdo*[b,d]

[a]*HSE University, 101000 Moscow, Russia*

[b]*Department of Chemistry, University of Southern California, Los Angeles CA 90089, USA*

[c]*I.E. Tamm Department of Theoretical Physics, P.N. Lebedev Physical Institute, Russian Academy of Sciences, 119991 Moscow, Russia*

[d]*Department of Physics & Astronomy, University of Southern California, Los Angeles CA 90089, USA*

**Corresponding author E-mail: avasenko@hse.ru (A. S. V.), prezhdo@usc.edu (O. V. P.)*

Computational Details

The cubic $CsPbBr_3$ phase was used to build the Σ5 (120) grain boundary (GB) model. Although the most stable configuration of $CsPbBr_3$ at room temperature is the tetragonal phase, the system still possesses the perovskite structure in the cubic phase, and differs from the tetragonal phase only via small local distortions. Therefore, using the cubic configuration as the initial structure is unlikely to affect the results significantly, but it notably reduces the computational cost due to the higher symmetry and smaller unit cell size.

For ab initio molecular dynamics (AIMD) simulations, a 10 ps trajectory was generated with a 2 fs timestep at 300 K, to explore the system changes during the GB sliding process, and the relevant results were shown in the manuscript. However, this trajectory is insufficient for training the machine learning (ML) force field (FF) because the GB can undergo complex and slow atomic rearrangements, and generate many configurations due to thermal fluctuations. To address this issue, we supplemented the 300 K trajectory with other AIMD calculations. We performed AIMD simulations with longer trajectories and at higher temperatures (600 K and 1000 K), and used the data for ML FF training. Further, after the initial training, long ML FF trajectories produced several partially collapsed structures. To address this problem, we used the failed ML FF simulations as the initial structures for additional AIMD simulations and used these AIMD trajectories for ML FF training as well. At the end, we collected ~30 000 configurations to compose the training set.

The ML FF was developed using the se_e2_a version DeepPot-SE descriptor to construct the deep neural network (DNN) potential. We used a cutoff radius of 9 Å for neighbor searching and a 0.5 Å smooth cutoff. The maximum number of neighbors within the cutoff radius was set to 128. The dimensions of the embedding and fitting layers were 25 × 50 × 100 and 240 × 240 × 240, respectively. The DNN was optimized using the Adam stochastic gradient descent method. The accuracy of the trained ML FF was checked by comparing the potential energies from the DFT calculations and ML predictions as shown in Fig. S1. The error was 6.09 meV/atom, and no large errors were seen for any particular configuration. The 1 nm ML FF trajectory used for the analysis was stable, as judged by conservation of the total energy, fluctuations in the potential and kinetic energies, visual inspection of the trajectory, and proper behavior of the electronic properties, such as the fundamental energy gap, Fig. 4a, b.

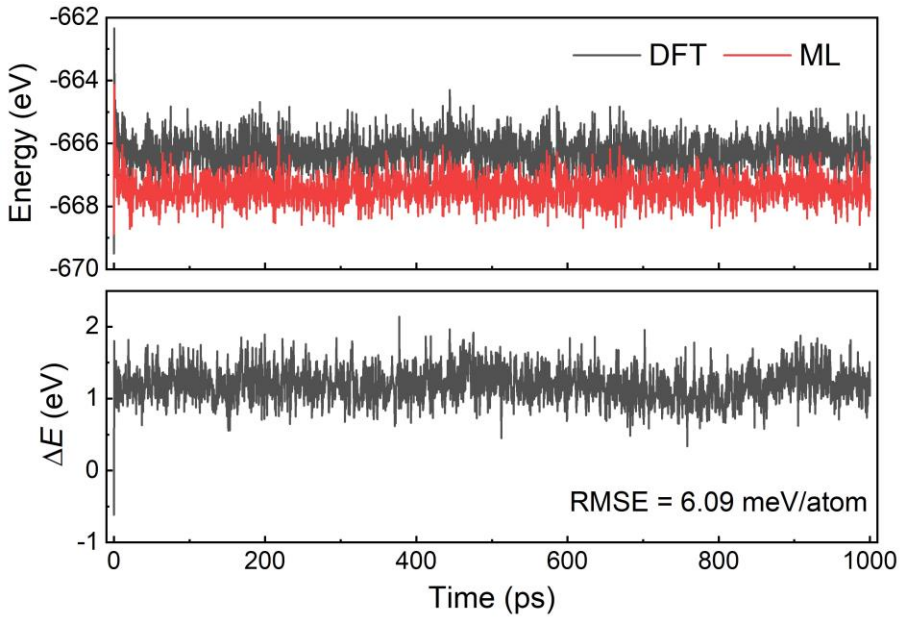

Figure S1. Potential energy evolutions of the GB system in a 1 ns trajectory with the timestep of 0.5 ps derived from DFT and ML calculations and their differences. The root-mean-square error (RMSE) is calculated to check the accuracy of the ML FF.

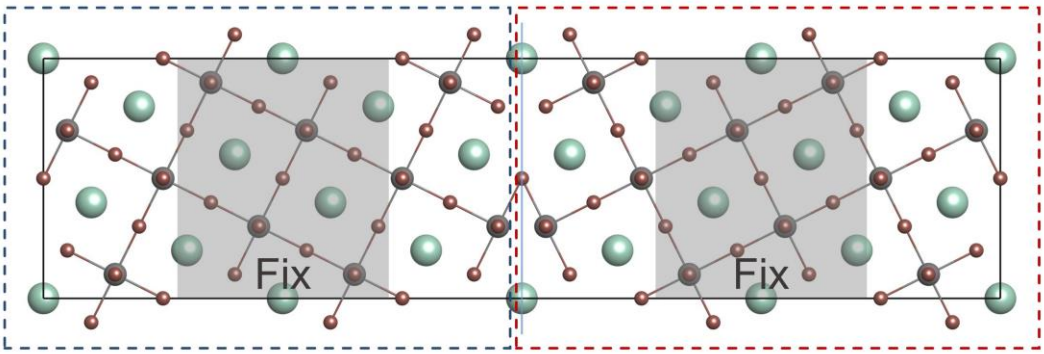

Figure S2. Scheme of the model used in calculating the potential energy changes during the GB sliding. The dashed boxes indicate two parts of the GB model. One of the boxes is moved entirely along various directions over different distances. We uniformly select 20 sliding distances along each direction within one periodic unit cell to interpolate the energy diagram. The atoms in the gray rectangles are fixed during the geometry optimization after sling to mimic the bulk restrictions. Color code: green for Cs, grey for Pb, and brown for Br.

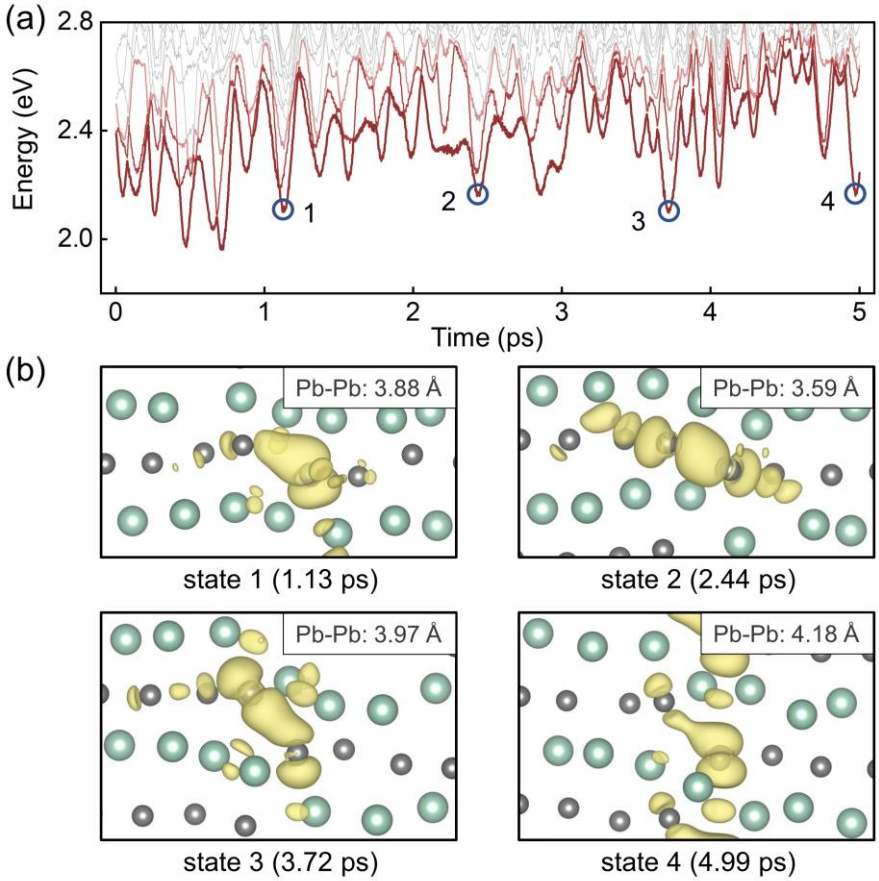

Figure S3. (a) Evolution of the conduction band energy levels in the 5 ps AIMD trajectory. Some low-energy LUMO levels are labeled with blue circles. **(b)** The corresponding charge density plots and the Pb−Pb distances.

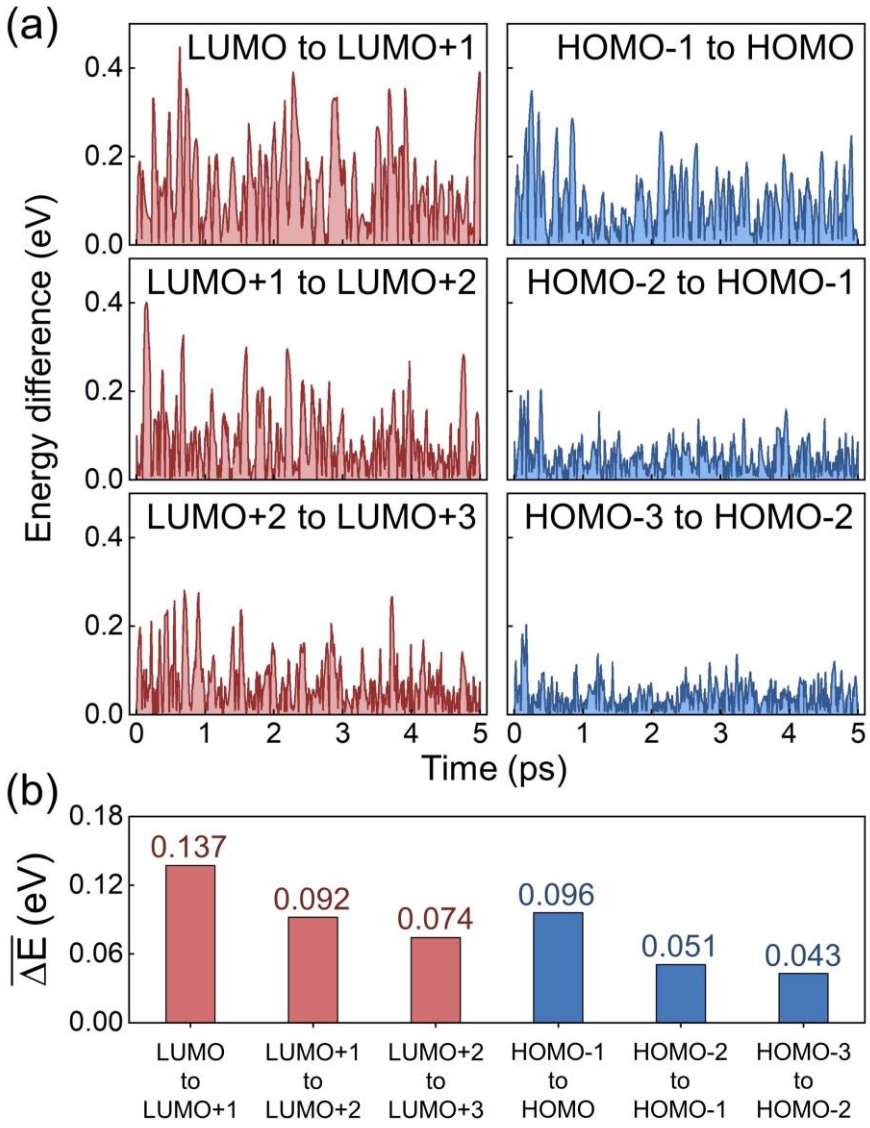

Figure S4. (a) Evolution of the energy level differences in the 5 ps AIMD trajectory. **(b)** The average value of these energy level differences.

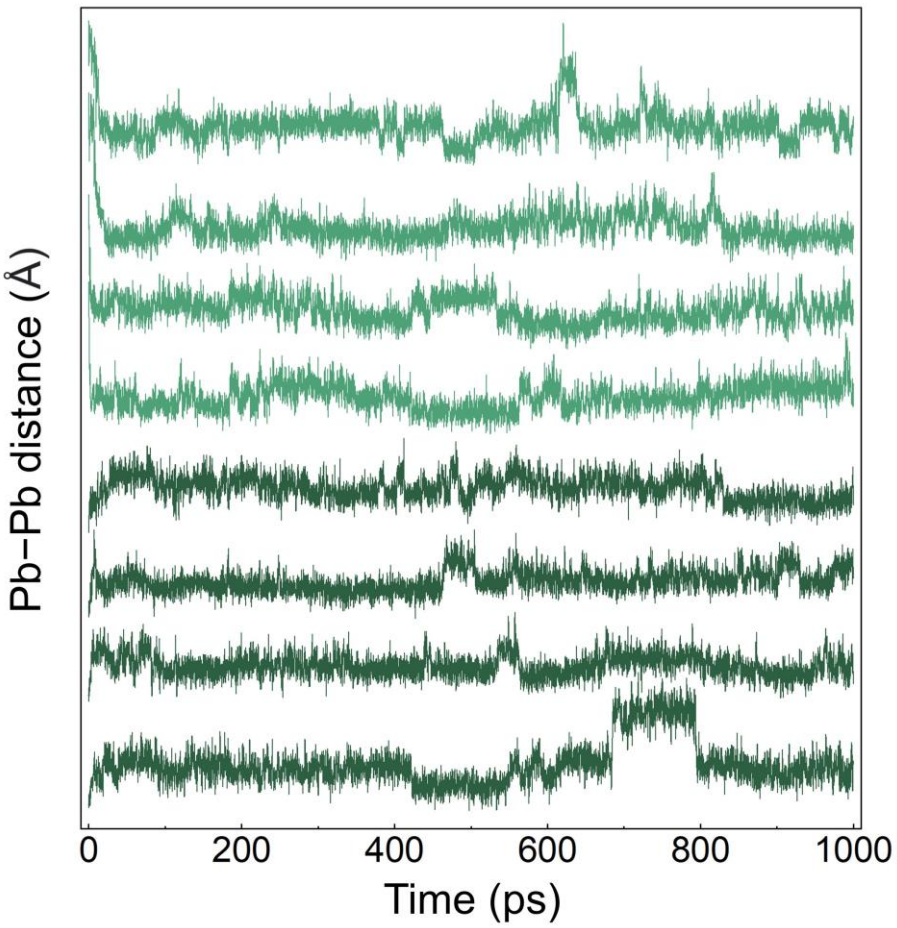

Figure S5. Evolutions of the Pb−Pb distances in the 1 ns MLMD trajectory.

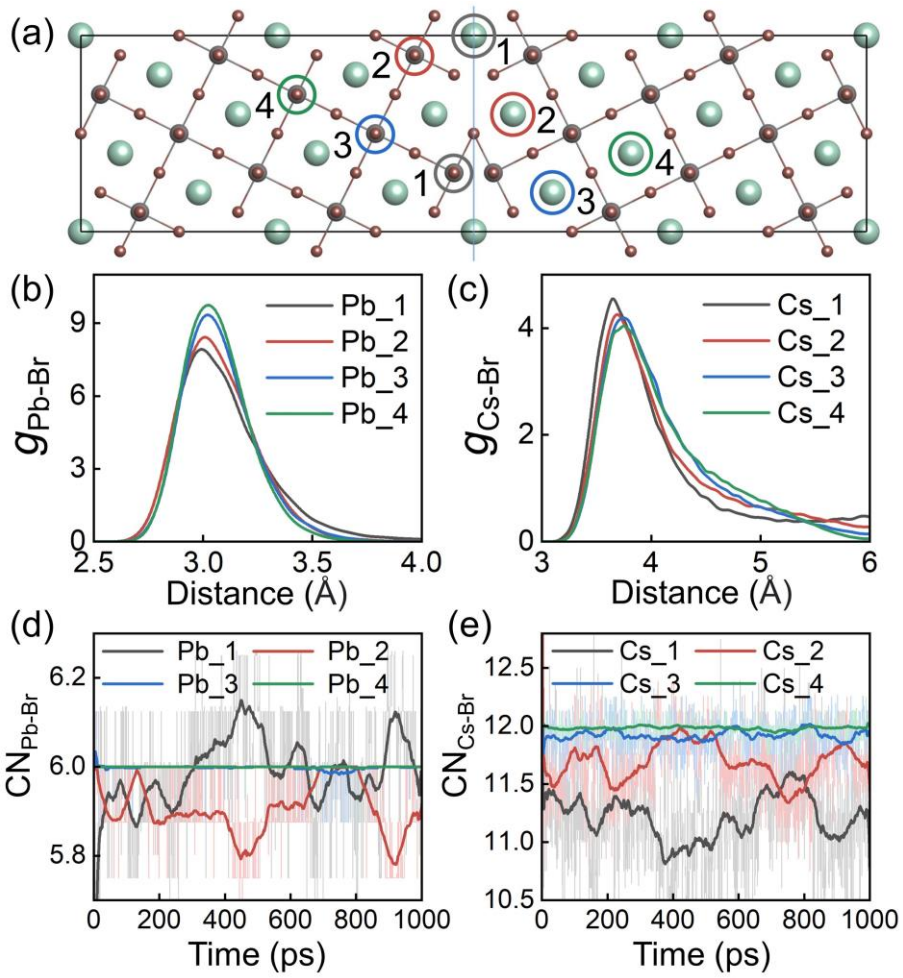

Figure S6. (a) Pb and Cs atoms in the GB and bulk regions (from 1 to 4). Each labelled atom has four equivalent sites, stemming from both sides of the two GBs, due to the system symmetry. All of these atoms are taken into account. Color code: green for Cs, grey for Pb, and brown for Br. **(b, c)** Canonical radial distribution function of the Pb-Br (g_{Pb-Br}) and Cs-Br (g_{Cs-Br}) pairs along the 1 ns MLMD trajectory. **(d, e)** Evolutions of the coordination number (CN) of Pb with Br (CN_{Pb-Br}) and Cs with Br (CN_{Cs-Br}). The 4 Å and 6 Å cutoff are used for the Pb-Br and Cs-Br pairs, respectively. The shallow lines are the raw data and the dark lines are the moving averages over 50 ps. The CNs of the GB atoms 1 and 2 (grey and red lines) fluctuate significantly, while the CNs of the bulk atoms 3 and 4 (blue and green lines) fluctuate little. The fluctuations of the CNs of the GB atoms are synchronously amplified in the GB distortion stage after around 350 ps, Fig. 3b.