Supplementary Information on Data-Efficient Machine Learning for Molecular Crystal Structure Prediction

Simon Wengert,[1] Gábor Csányi,[2] Karsten Reuter,[1,3] and Johannes T. Margraf[1,a]

[1] *Chair of Theoretical Chemistry, Technische Universität München, 85747 Garching, Germany*

[2] *Engineering Laboratory, University of Cambridge, Cambridge CB2 1PZ, United Kingdom*

[3] *Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany*

[a] Electronic mail: johannes.margraf@ch.tum.de

A. ML Fitting Workflow

The overall Δ-ML correction presented in the main document consists of two separate contributions for intra- and intermolecular interactions. Figure 1 provides an overview for the generation process of each of the two models, while the following section complement the description of the individual parts.

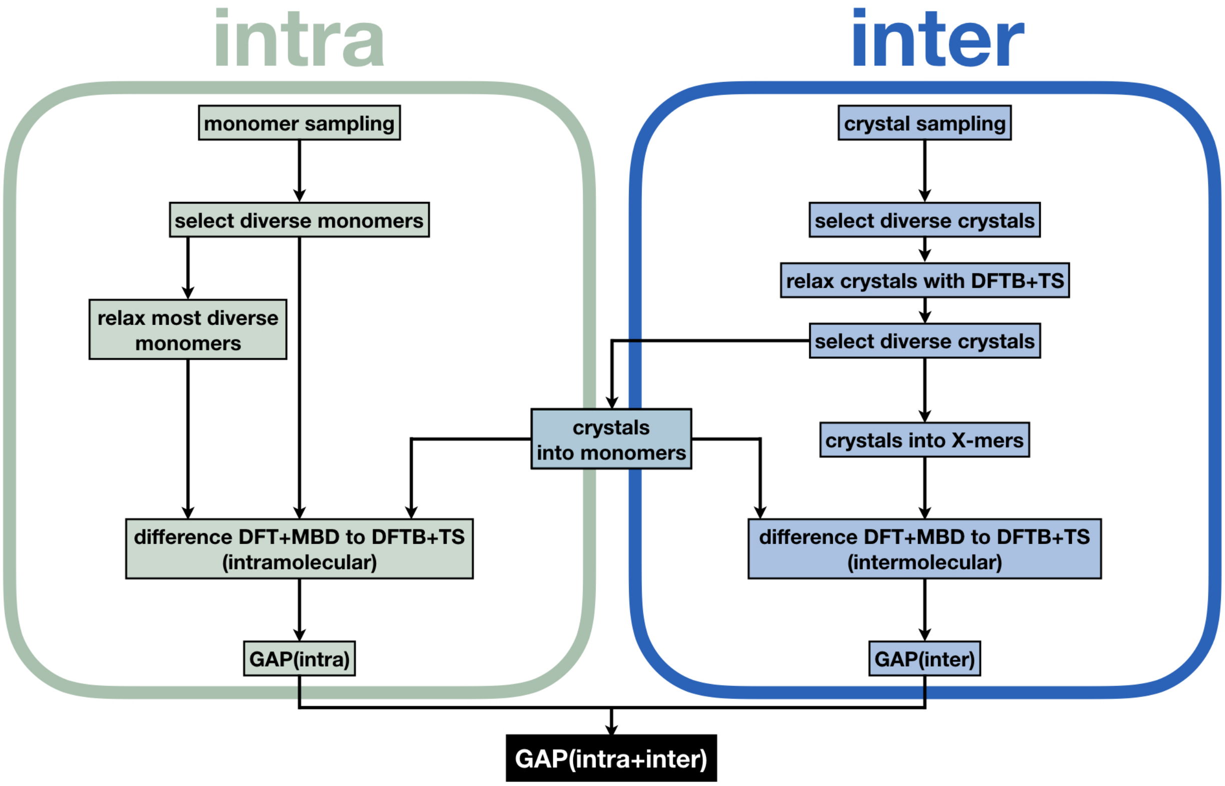

FIG. 1: Diagram to illustrate the workflow for generating the intra- and intermolecular machine-learning corrections.

B. Training Data Generation

Intramolecular Δ-ML: The training configurations for the intramolecular model are obtained from three distinct sources, namely local minima on the potential energy surface, configurations of the molecule in crystal environments and out-of-equilibrium geometries (see Figure 1, left).

For the latter, a long molecular dynamics (MD) simulation is performed at 500 K, using a classical force field (gaff2[1] or Dreiding[2]). From this trajectory, a diverse subset is selected via FPS. As a rule of thumb, we use maximally fifty structures per degree of freedom dependent on the rigidity of the molecule. Additionally, the fifty most dissimilar conformations from the MD are locally relaxed with both the DFTB+TS baseline and the DFT+MBD target level of theory. This step doubles as a conformer search, as well as providing information about the differences between the target and baseline methods at the bottom of the potential well. If conformers of the molecule are known *a priori*, they are also included.

In the main document, we already described the third source for intramolecular training data, namely condensed phases. Configurations are extracted from crystal structures relaxed at the low-cost DFTB+TS baseline level. These configurations provide information about the intramolecular strain the molecule undergoes upon crystalization. Note that these relaxed crystals are also used to generate the configurations for the intermolecular correction, so that they are available at no additional cost.

The energy and force differences between DFT+MBD target and DFTB+TS baseline are then used to generate the model for the $\Delta E^{\mathrm{GAP(intra)}}$ correction (see Equation 4 in the main document). Thus, each geometry enters the training with elements for the intramolecular correction and its derivative (in terms of force differences).

Intermolecular Δ-ML: An overview of the applied process for obtaining geometries for the intermolecular Δ-ML model has been described in the main document and is illustrated by the right-hand side of Figure 1. We will now provide a more detailed insight and separately discuss the individual steps together with their function. For this purpose, the selection process is illustrated with kernel principal component analysis (kPCA) in the upper part of Figure 2, for the example of oxalic acid.[3]

Overall, this yields a set of highly diverse structures, which are representative of the interactions expected in actual crystal polymorphs. This can be seen in the lower part of Figure 2, again for oxalic acid. The sampled crystal structures cover a wide range of unit cell shapes and volumes. Furthermore, the relaxed structures also display varied monomer configurations and sample the typical H-bond distances between 1.6 and 2.0 Å.[4] In particular, it is notable that some crystals exhibit molecules that are strongly distorted with respect to gas-phase conformations. Most prominently, the partial C-C double bond in the oxalic acid molecule is twisted in some cases, to allow for H-bond formation. This highlights the necessity for a CSP model that allows a flexible description of the monomers. While the

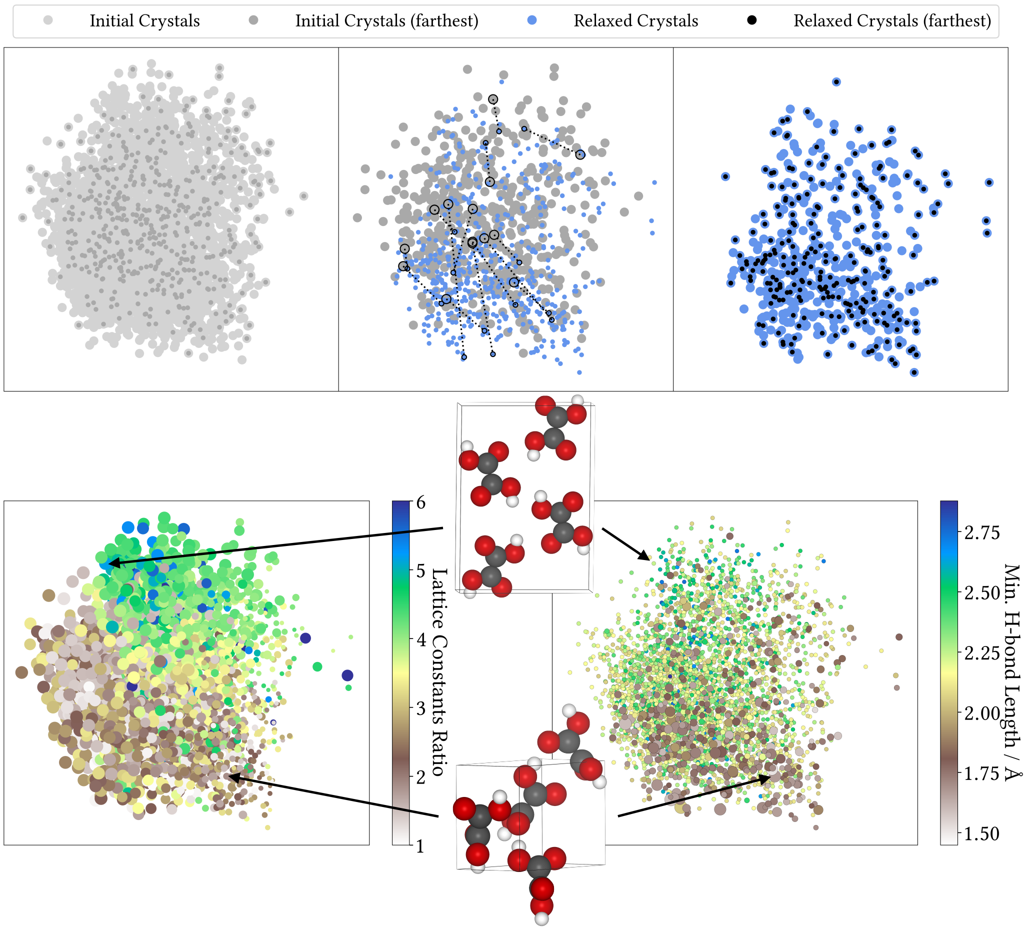

FIG. 2: Visualization of the crystal structure selection process with kernel principal component analysis. The top frames illustrate the FPS of crystals from an initial pool (left), DFTB+TS relaxations (center, including some exemplary paths from initial to relaxed geometries) and the final FPS selection from these crystals (right). The bottom frames illustrate the distribution of lattice parameters (left) and hydrogen bond lengths (right). The size of the symbols illustrates the unit cell volume (left, in the limits of 283 Å to 331 Å) and the distortion of the molecule, relative to the tTt gas-phase conformer (right, in the limits of 0 to 0.8 Å RMSD).

experimentally known polymorphs of oxalic acid both display a flat conformation[5,6] this is not known *a priori*.

Based on the above defined set of crystals, the X-mers are defined by cutting spherical clusters out of the crystal, so that all molecules touching the sphere are assigned to the X-mer. This procedure is repeated with multiple radii (typically 4-6 values between 2 and

5 Å) and central atoms to obtain a range of X-mers for each crystal. The precise values are system dependent and selected so that most X-mers contain 4-5 molecules. These X-mers form the final training set for the intermolecular correction and each cluster enters with the differences between DFTB+TS baseline and DFT+MBD target energies according to Equation 5.

Note that it is not necessary to compute DFT+MBD and DFTB+TS energies for each molecule of every X-mer explicitly, but only for the unit cell molecules of each crystal structure. Furthermore, distributing the corresponding calculations to available computational resources is trivial, most notably since there are no special demand–for instance in terms of memory or walltime restrictions—for these kind of systems. Overall, such a model can therefore be trained without access to high-performance computing resources.

C. Detailed Description of the ML Methods

We will briefly discuss the most important hyperparameters of SOAP/GAP models and how they are selected in this work, and subsequently provide a detailed listing of all the remaining settings. For details about the underlying concepts we refer to the original literature.[7–9]

SOAP: In SOAP, the atomic environment χ_a (within a given cutoff r_{cut}) around a reference atom a is described via the neighborhood density function $\rho_a(\boldsymbol{r})$ which, in case of multi-element systems, is evaluated separately for each species Z as

$$\rho_a^{\mathrm{Z}}(\boldsymbol{r}) = \sum_{i \in \chi_a^{\mathrm{Z}}} \exp\left(-\frac{(\boldsymbol{r} - r_i)^2}{2\sigma^2}\right) \cdot f_{\mathrm{cut}}(\boldsymbol{r}). \tag{1}$$

Thus, each atom i within χ_a is represented as a Gaussian with the parameter σ defining its width. The function f_{cut} ensures a smooth transition to zero at r_{cut} within a cutoff region of width d. It is defined as

$$f_{\mathrm{cut}}(\boldsymbol{r}) = \begin{cases} 1 & \text{for } \boldsymbol{r} \leq r_{\mathrm{cut}} - d \\ \left[\cos\left(\pi \frac{\boldsymbol{r} - r_{\mathrm{cut}} + d}{d} + 1\right)\right]/2 & \text{for } r_{\mathrm{cut}} - d < \boldsymbol{r} < r_{\mathrm{cut}} \\ 0 & \text{for } \boldsymbol{r} > r_{\mathrm{cut}}. \end{cases} \tag{2}$$

As described elsewhere,[7] the SOAP kernel is then evaluated by expanding the neighborhood density with spherical harmonics and forming a normalized polynomial kernel with the vector of expansion coefficients.

In principle the above hyperparameters can be optimized for each system, but we use fixed defaults, which we have found to be accurate for all systems studied herein. Specifically, for the intramolecular correction the Gaussian width σ is set to 0.3 Å and the cutoff is set to 3.0 Å, with a transition width d 0.5 Å. For the intermolecular correction with DFT+MBD reference all length-related hyperparameters are increased by 20 %. Thus, the atomic environment χ_a is now defined by a 3.6 Å cutoff, the transition region has a width of 0.6 Å and the atomic positions within the cutoff are broadened by a Gaussian with $\sigma = 0.36$ Å. Similarly, for corrections with MP2 reference the hyperparemeters have been increased by 50 % such that the cutoff is 4.5 Å, d is 0.75 Å and σ is 0.45 Å.

GAP: The GAP approach uses Gaussian Process Regression (GPR) to learn an interatomic potential from quantum mechanical data. The total energy is expanded as a sum of atomic contributions ϵ_a, ensuring size-extensivity and linear scaling of computational effort. The atomic contributions are in turn expressed as a weighted linear combination of kernel functions:

$$\epsilon_a = \sum_{\chi_b} \alpha_b k(\chi_a, \chi_b), \tag{3}$$

where the sum runs over atomic environments χ_b in the training set and $k(\chi_a, \chi_b)$ denotes the SOAP Kernel between two environments. The regression coefficients α are obtained by inverting the covariance matrix $\mathbf{C}$, with the covariance between training points n and n' defined as:

$$C_{nn'} = \delta^2 k(\chi_n, \chi_{n'}) + \sigma_T^2 \tag{4}$$

Here σ_T denotes the uncertainty of the target values (i.e. force or energy) that is to be fitted and δ denotes the standard deviation of the Gaussian process (related to the expected magnitude of the potential). The latter is set to the standard deviation of the target values per atom.

Meanwhile, the uncertainties in the training data (σ_E and σ_F for energies and forces, respectively) require more careful consideration. For the intramolecular correction, σ_E is a global parameter, optimized to minimize the error on a validation set. For this purpose, 1/3 of the training data is set aside and the root mean square error (RMSE) between actual and predicted energies is minimized. Here, σ_E is constrained to be smaller than 0.1δ. In contrast, values for σ_F are selected to be proportional to the forces acting on each atom. This is to ensure an approximately constant relative accuracy of predicted forces. Specifically, for each atom σ_F is set to 10 % of the corresponding DFT+MBD force norm. However, the values have been restricted, since forces vary by several orders of magnitude. To this end we define a lower bound for σ_F as 5 % of the mean force norm of a room temperature MD (gaff2[1] or Dreiding[2]) and an upper bound as 15 times the lower bound. Values outside this range have been set to the corresponding boundary value. Finally, σ_E for the intermolecular correction is set to 15 % of δ.

While the above heuristics are admittedly arbitrary, we have found that they offer good performance for a wide range of systems. In our view, this is a significant advantage, as it allows applying the framework routinely to different CSP tasks. In principle, a more system-specific optimization of parameters is also possible, of course.

Intramolecular Δ-ML: The hyperparameters that come along with the combination GAP+SOAP have been set to the following values for the generation of intramolecular models. Beside the length-related parameters from above we used l_max 8 and n_max 8 for the descriptor. Furthermore, we used covariance_type dot_product, a zeta value of 2.0 and CUR_POINTS as our sparse_method for selecting a total of 2,000 sparse points (n_sparse). Finally, the e0_method has been set to isolated and the atoms with the corresponding DFTB+TS to DFT+MBD differences have been added to the fit.

Intermolecular Δ-ML: Compared to the intramolecular fit the value for e0_method has been changed to average. Apart from that l_max, n_max, covariance_type, zeta, sparse_method and n_sparse remain the same leaving us with no additional hyperparameter to be optimized.

D. Force Accuracy for Intramolecular Δ-ML

For monomer configurations an analysis equivalent to the energetic consideration (see Figure 2, top, in the main document) has been performed for the second kind of target values, the force components, which is illustrated in Figure 3. It show the MAEs compared to the high-level target method DFT+MBD for the individual test systems. On the baseline (DFTB+TS) level of theory, this error can be as large as 630 $^{\mathrm{meV}}/_{\text{Å}}$ (for oxalic acid). After applying the Δ-ML correction, however, MAEs at the DFTB+TS+GAP level are reduced by at least one order of magnitude. For water the improvement is especially pronounced as the DFTB+TS baseline error for test monomers of above 370 $^{\mathrm{meV}}/_{\mathrm{Å}}$ vanishes almost completely after applying the Δ-ML model with a remaining error below 0.8 $^{\mathrm{meV}}/_{\mathrm{Å}}$. In addition to the finding for energetics, the good agreement between training and test errors on forces further confirms that the models are not overfitted.

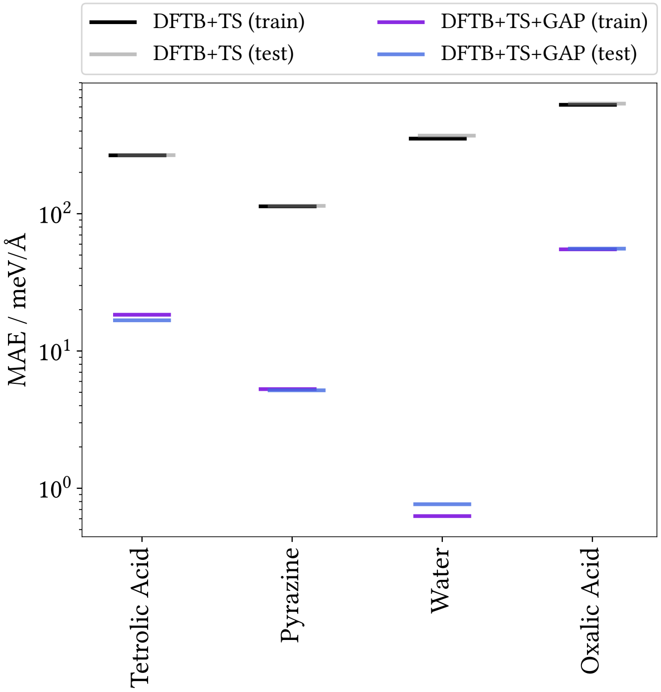

FIG. 3: Mean absolute error of force components obtained with the DFTB+TS baseline and Δ-ML corrected DFTB+TS+GAP against the DFT+MBD target for monomer conformations from training and test crystals.

E. Accuracy of Intermolecular Δ-ML for Crystals

In Section 3.2 of the main document we discussed the improvements achieved by applying the intermolecular correction on training targets (i.e. X-mers). Similarly, we have also been analysing the effect on the intermolecular description in periodic systems by applying the correction on the respective crystal structures. The resulting MAEs per molecule are shown in Figure 4. Note that intermolecular energies of crystals are not entering the fit of the Δ-ML model and, thus, even the ”training” crystals can be viewed as validation.

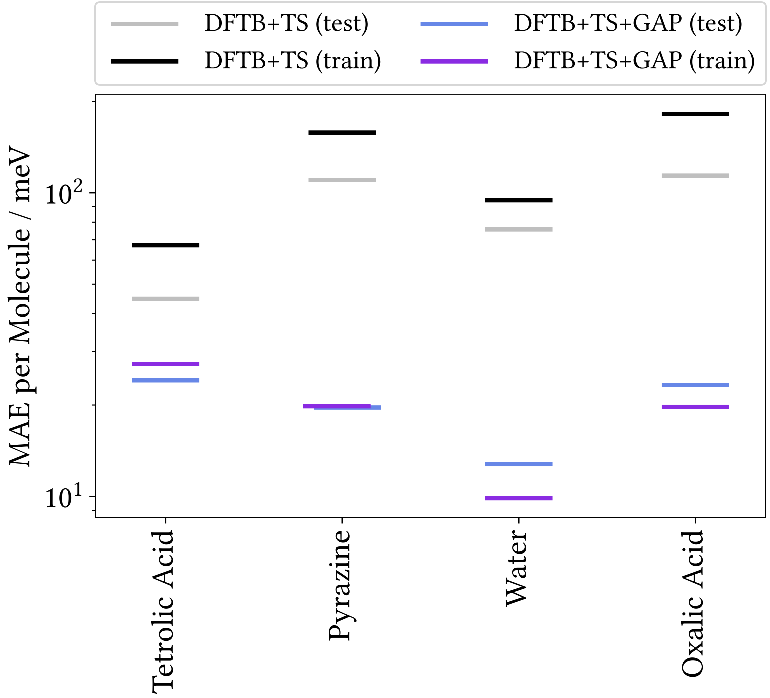

FIG. 4: Mean absolute error of intermolecular energies per molecule obtained with the DFTB+TS baseline and the Δ-ML corrected DFTB+TS+GAP against the DFT+MBD target method for training and test crystals.

In comparison to X-mers (see Figure 2, center, in the main document), the MAEs per molecule on the baseline (DFTB+TS) level are between 2.2 and 3.3 times larger for crystal structures. Despite the differences that arise when going from finite to periodic systems, the Δ-ML model achieves significant improvements with MAEs between 10 and 27 meV per molecule for test and training crystals which illustrates a successful translation to these kind of systems.

The fact that the remaining errors are on the same order as obtained for lattice energies (compare Figure 2, bottom, in the main document) is not surprising. In fact, the MAEs of Figure 4 can be viewed as lattice energy errors when neglecting intramolecular contributions. Keeping in mind that intramolecular errors are in most cases below a meV when applying the corresponding correction (see Figure 2, top, in the main document), thus, explains the resemblance between Δ-ML corrected results in Figure 4 and the corresponding plot for the lattice energies. As a results, we also obtain the excellent agreement between test and training crystals which substantiates a good generalization of the Δ-ML model, even for periodic systems. Moreover, we also recover DFTB+TS baseline MAEs for training crystals being consistently above the corresponding test crystals which, again, confirms the selection of a particular challenging and diverse training set.

F. Decomposition of Lattice Energies into Intra- and Intermolecular Contributions

In Section 3.3 of the main document we have been discussing different kinds of error the Δ-ML correction is able to cure when applied to the DFTB+TS baseline. These errors can be traced back to their origin in terms of intra- and intermolecular deviations from the DFT+MBD target method. For this purpose, we disassemble the lattice energy into the two contributions according to:

$$\begin{aligned} E_{crys}^{latt} &= \left[\sum_{i}^{N} E_{mono,i}^{intra} + E_{crys}^{inter} \right] /N - E_{mol} \\ &= \left[\frac{1}{N} \sum_{i}^{N} E_{mono,i}^{intra} - E_{mol} \right] + \frac{1}{N} E_{crys}^{inter} \\ &= E_{crys}^{latt,intra} + E_{crys}^{latt,inter}. \end{aligned} \tag{5}$$

Figure 5 illustrates the correlation of these contributions with the DFT+MBD target method for the test crystals before and after applying the Δ-ML correction. The left column clearly identifies intermolecular deviations as the origin of systematic errors for pyrazine and water lattice energies, while the effect of intramolecular deviations is less pronounced. Oxalic acid, on the other hand, exhibits additional intramolecular deviations which cause the offset in lattice energies described in the main document.

Overall, applying the Δ-ML corrections leads to a significantly decreased scattering in correlation for both the intra- and intermolecular contribution as can be seen from the right-hand side of Figuren 5.

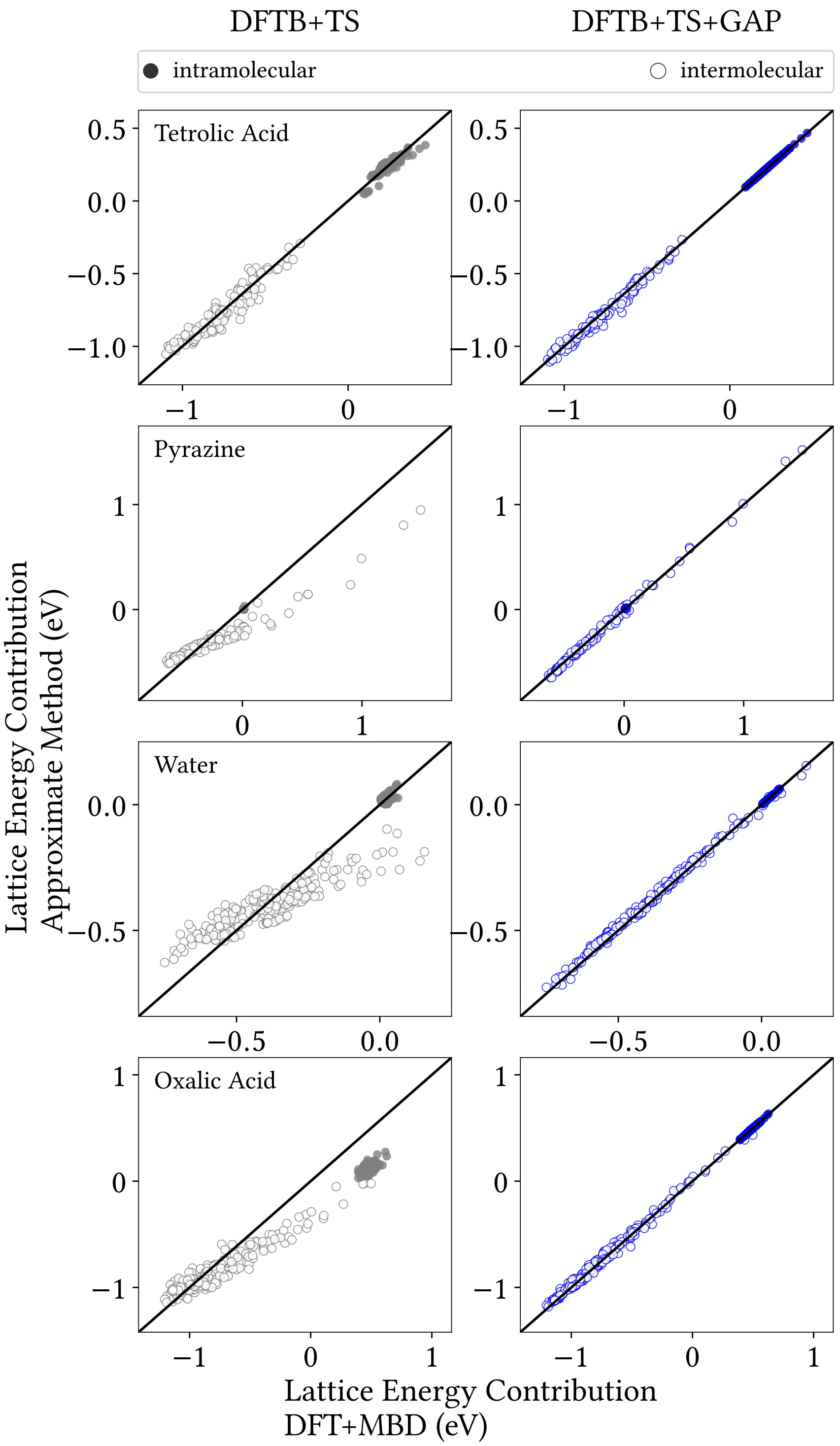

FIG. 5: Correlation plot for intra- and intermolecular lattice energy contributions of test crystals for the DFTB+TS baseline (left) and the Δ-ML corrected DFTB+TS+GAP method (right).

G. Details on the Crystal Structure Prediction Showcase

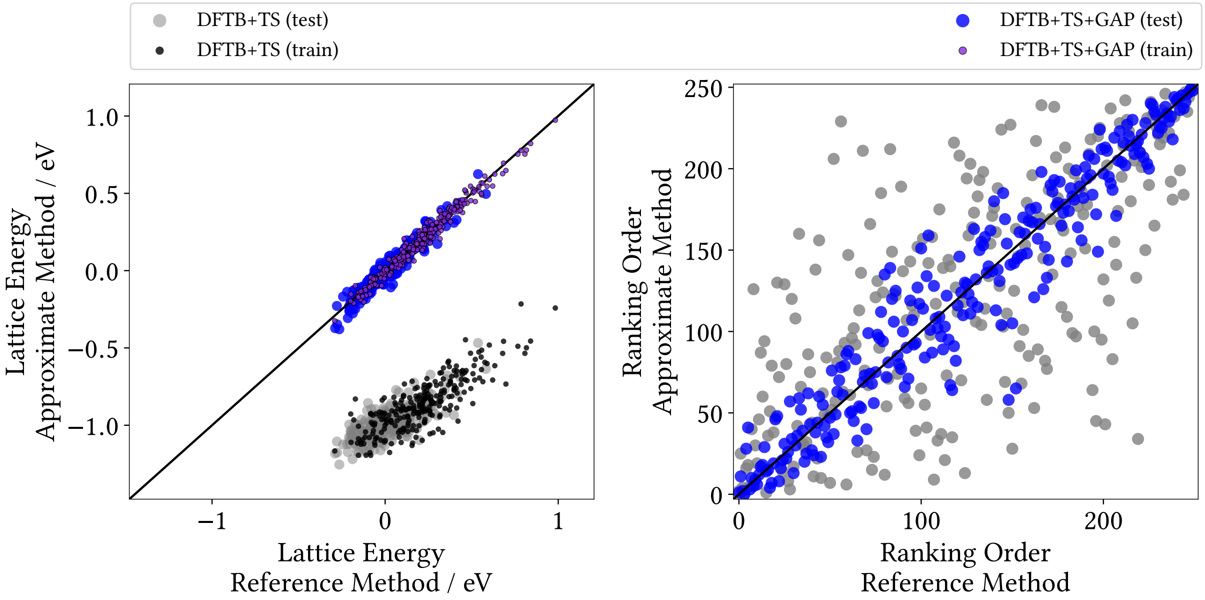

FIG. 6: Correlation plot for lattice energies of XXII crystals entering the training and test crystals obtained with the DFTB+TS baseline and the Δ-ML corrected DFTB+TS+GAP method (left) and ranking order of test crystals (right), both with respect to the DFT+MBD target values. Note that DFTB+TS(test) values are identical to Figure 5 (without the experimental XXII values) in the main document.

Lattice Energies of training and test crystals: For target XXII, we have been analysing the accuracy of the baseline (DFTB+TS) and Δ-ML method against the DFT+MBD target method for lattice energies of crystals entering the training and test crystals (similar to Section 3.3 in the main document). Applying the Δ-ML model decreases the MAE from above 1 eV (1.077 eV training and 1.005 eV test crystals) to only 28 meV for training and 34 meV for test crystals. The huge discrepancy largely originates from correcting for the offset in lattice energies obtained with the baseline method as can be seen from Figure 6. Figure 7 provides a more detailed insight by separate consideration of intra- and intermolecular contributions and shows that this is partly explained by intramolecular deviations. However, also intermolecular deviations contribute to the offset. Moreover, the right-hand side shows an excellent correction that removes the offset and leads to an overall improved correlation with the DFT+MBD target method. As a result, the coefficient

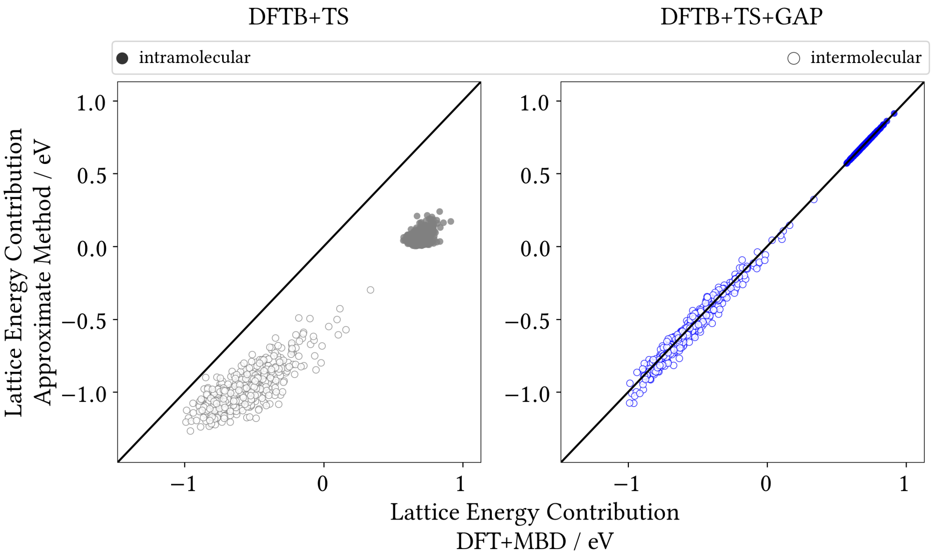

FIG. 7: Correlation plot for intra- and intermolecular lattice energy contributions of XXII test crystals for DFTB+TS (left) and DFTB+TS+GAP (right).

of determination improves from 0.478 for the DFTB+TS baseline to 0.924 for the Δ-ML corrected DFTB+TS+GAP method.

Relaxed Geometries of the Experimental Structure XXII: In Figure 8 we compare relaxed experimental XXII crystal structures (in terms of the often used overlay of 15-mers), obtained individually with the DFTB+TS baseline, the Δ-ML corrected DFTB+TS+GAP and the DFT+MBD target method. The DFTB+TS baseline guides the experimental geometry into an area that differs from the DFT+MBD minimum. The overlay reveals that the same geometrical differences discussed in the main document for monomers are also found in the condensed phase. In particular, the preference of DFTB+TS for the flat conformation and more acute C-S-N angles of the five-membered rings can be seen. As mentioned in the main document, the crystal structure predicted with the Δ-ML corrected DFTB+TS+GAP and the DFT+MBD target are in excellent agreement. More generally, the generation of sound structures via DFTB+TS+GAP relaxations can also be seen from Figure 5 in the main document as the DFT+MBD evaluation of each trial crystal, as well, yields a negative lattice energy.

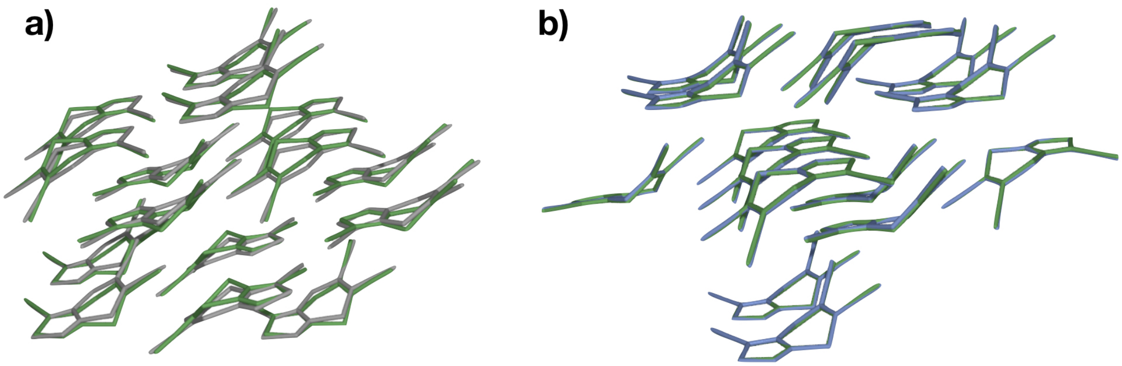

FIG. 8: Overlay of best-matching 15-mers from a) DFT+MBD- and DFTB+TS-relaxed (RMSD 0.317 Å) and b) DFT+MBD- and DFTB+TS+GAP-relaxed (RMSD 0.087 Å) experimental XXII crystals. (DFT+MBD: green, DFTB+TS: gray, DFTB+TS+GAP: blue)

Timing: In the following we will provide details about the timings for GAP training and crystal relaxations that have been outlined in the main document.

For the intramolecular part we used a total number of 3,000 monomer configurations to obtain robust values for average timings of DFT+MBD and DFTB+TS evaluations, as well as the GAP predictions. The total timing for generating DFT+MBD training data additionally includes relaxations of 52 monomer configurations.

Similarly, for the intermolecular part a total number of 41,655 X-mers has been used to obtain the corresponding timings, except for the GAP predictions where a subset of 10,000 X-mers has been used. For MP2 evaluations a subset of 31,682 X-mers has been used.

Average timings for crystal relaxations are computed from 20 geometries for DFT+MBD, 92 geometries for DFTB+TS+GAP and 51 geometries for DFTB+TS. In case of DFT+MBD and DFTB+TS the internal relaxation algorithm of the respective code (FHI-aims[10] and DFTB+[11]) has been used as it allows for a speed up by making use of information from previous steps. For DFTB+TS+GAP relaxations we used the BFGS[12–15] algorithm implemented in ASE[16]. In order to obtain comparable values, we calculated the average time required for a single relaxation step for each of the three methods and multiplied it with the average number of relaxation steps required to optimize rigid crystals (with DFTB+TS) obtained from genarris[17].

The timings required to generate a model trained on periodic data were computed from the average timing of 500 DFT+MBD single-point calculations.

H. Δ-ML Models Trained on Periodic Reference Data

In Section 3.3 of the main document we compare the X-mer based intermolecular models with alternative models that are exclusively trained on periodic reference calculations. The corresponding MAEs for test and training crystals of all molecules considered in this work (including target XXII) are shown in Figure 9. As discussed in the main document these models show slightly improved lattice energies for most systems, except for target XXII where the MAE is slightly higher. The figure also shows that Δ-ML models exclusively trained on crystals exhibit a considerably larger gap between training and test MAEs indicating that such models are more prone to overfitting.

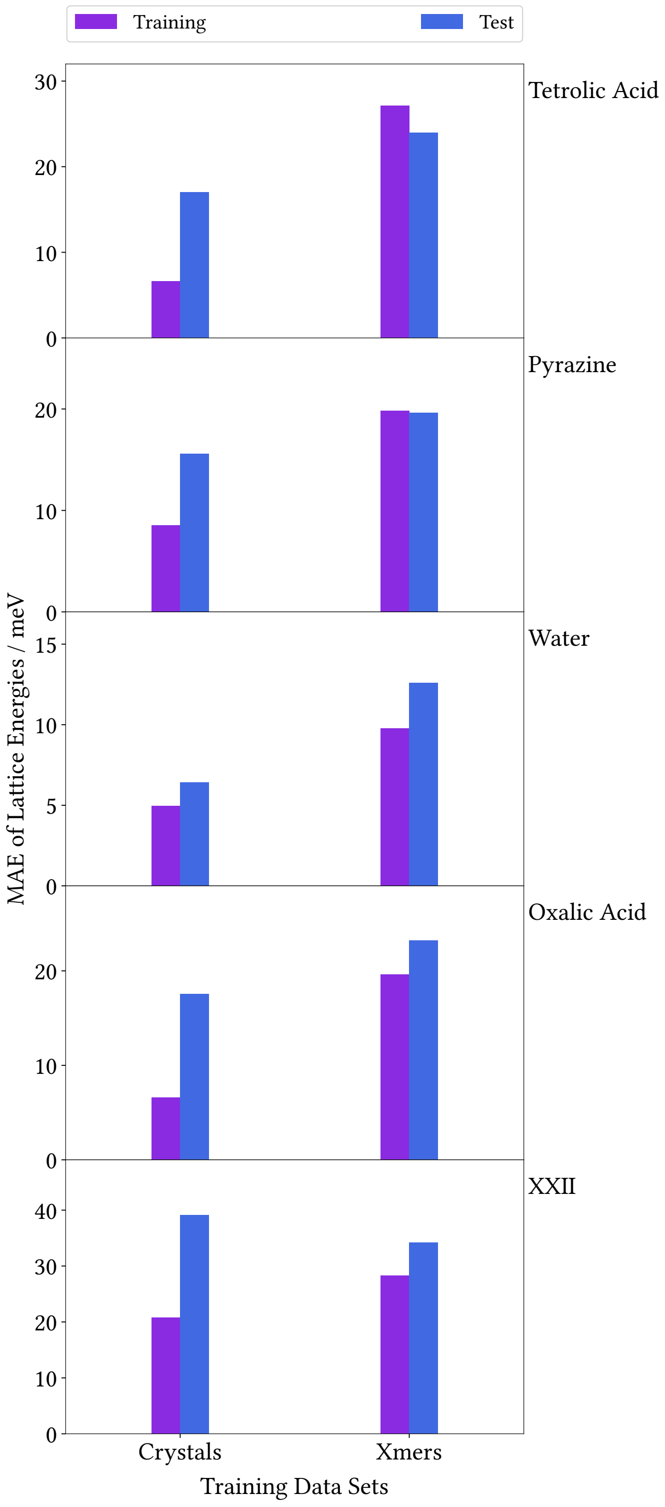

FIG. 9: Mean absolute error of lattice energies obtained with Δ-ML models trained exclusively on the underlying crystals or its X-mers against the DFT+MBD target method for training and test crystals.

I. Crystal Structure Prediction Beyond Density Functional Theory

In Section 3.5 of the main document we modified the XXII model of Section 3.4 by replacing its intermolecular part with a model trained on SCS-MP2 reference data. Figure 10 provides detailed insight regarding the predictive quality of the intermolecular energies of the training set, as well as 10,295 test X-mers. The corresponding results obtained with the DFT+MBD intermolecular GAP model of Section 3.4 are also shown, allowing for a direct comparison. Training and test MAEs obtained with the model trained on SCS-MP2 reference data are 3 meV and 7 meV and, thus, slightly lower than the ones obtained with the DFT+MBD reference (7 meV for training and 10 meV for test set). The former benefits from a better correlation with the baseline model compared to the DFT+MBD reference case.

In analogy to Section 3.4 the modified model has been used to relax a set of 251 XXII trial crystals (including the experimental structure). The subsequent energetic ordering of the resulting geometries is illustrated in Figure 11 and reveals that the model correctly identifies the experimental geometry to be the most stable one.

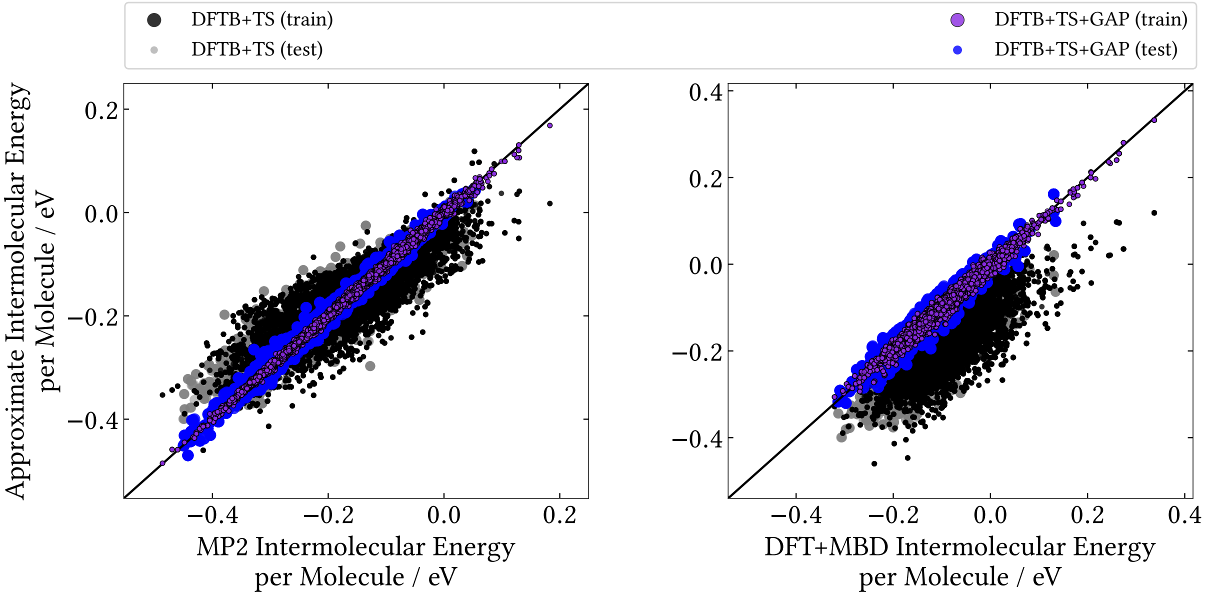

FIG. 10: Correlation plot for intermolecular energies per molecule of XXII training and test X-mers obtained with the DFTB+TS baseline and the Δ-ML corrected DFTB+TS+GAP methods. Results regarding the model trained on SCS-MP2 reference data are shown left, while the right frame shows the results for the model trained on DFT+MBD reference data.

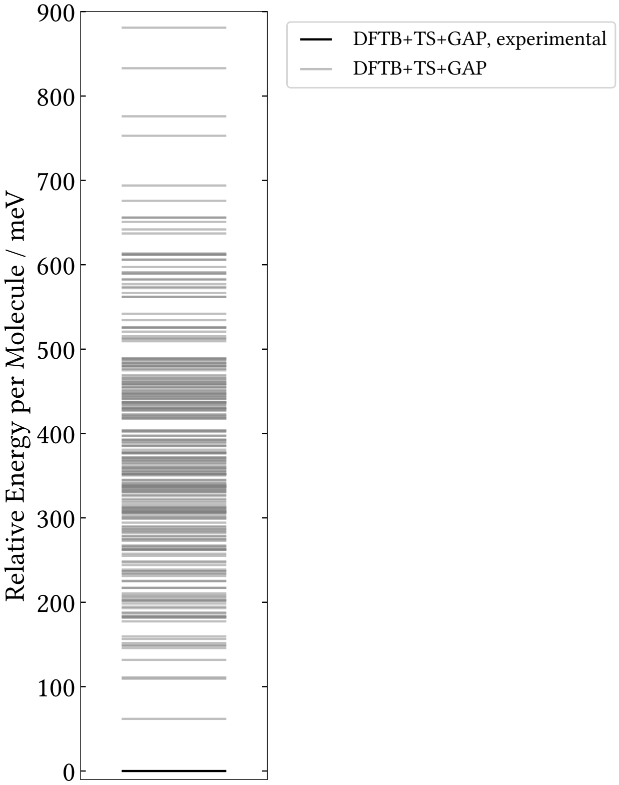

FIG. 11: Relative energies per molecule of XXII crystals relaxed with the Δ-ML corrected DFTB+TS+GAP method with SCS-MP2 as reference for intermolecular interactions.

J. k-Grid Convergence for Validation

The presented workflow for the Δ-ML model generation does not use periodic calculations, and is therefore not dependent on the convergence of any k-grid. For validation, however, periodic calculations are required, which do have a k-grid dependency. In Figure 12 we therefore test whether the convergence criterion of 1.5 $^{\mathrm{meV}}/_{\mathrm{atom}}$ is sufficiently accurate for a sound validation.

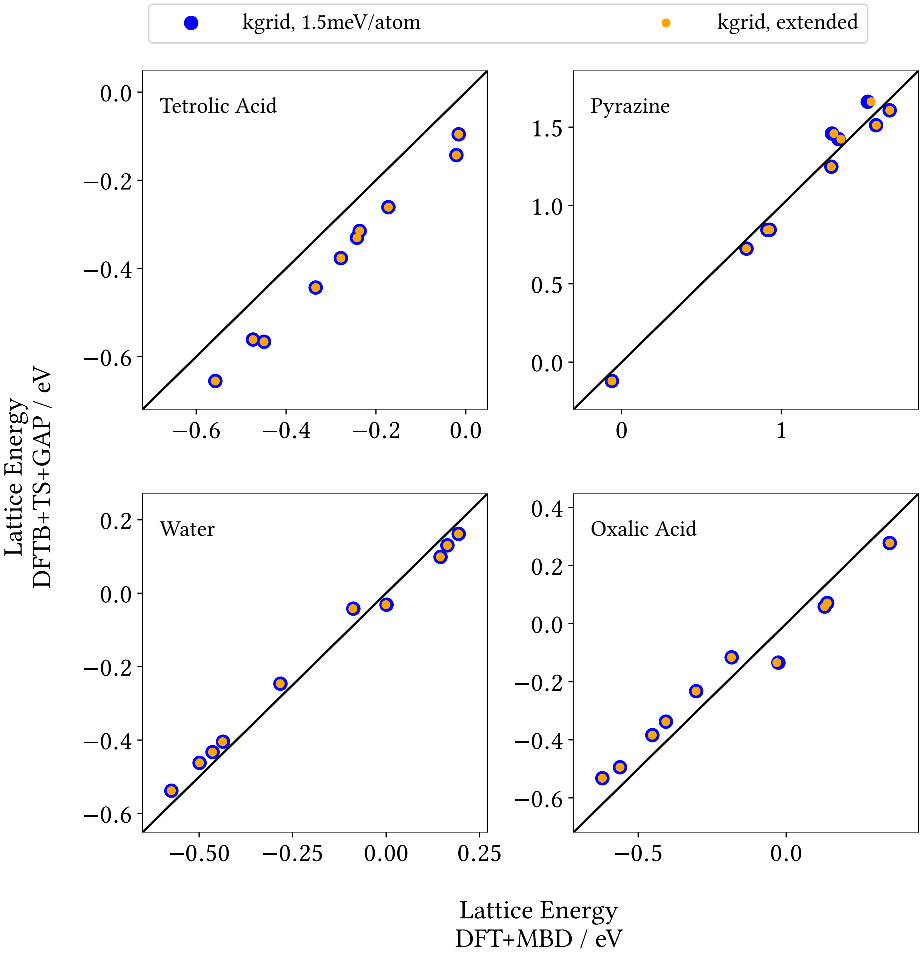

FIG. 12: Correlation plot of DFTB+TS+GAP vs DFT+MBD lattice energies with k-grids converged to 1.5 $^{\mathrm{meV}}/_{\mathrm{atom}}$ and a larger grid with two additional k-points in each direction.

For this purpose, we selected ten crystal structures with a particular large deviation between the Δ-ML corrected DFTB+TS+GAP and the target DFT+MBD lattice energies, for each of the four test systems. For these cases, we reevaluated the lattice energies with increased k-grids. Our findings show that our choice of k-grids is robust and increasing the grid has virtually no effect on the predicted lattice energies.

REFERENCES

[1] D. A. Case, R. Betz, D. Cerutti, I. T. Cheatham, T. Darden, R. Duke, T. Giese, H. Gohlke, A. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. Merz, G. Monard, H. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, D. Roe, A. Roitberg, C. Sagui, C. Simmeling, W. Botello-Smith, J. Swails, R. Walker, J. Wang, R. Wolf, X. Wu, L. Xiao and P. Kollman, 2016.

[2] S. L. Mayo, B. D. Olafson and W. A. Goddard, J. Phys. Chem., 1990, **94**, 8897–8909.

[3] B. Cheng, R.-R. Griffiths, S. Wengert, C. Kunkel, T. Stenczel, B. Zhu, V. L. Deringer, N. Bernstein, J. T. Margraf, K. Reuter and G. Csányi, Acc. Chem. Res., 2020.

[4] S. A. Blair and A. J. Thakkar, Chem. Phys. Lett., 2010, **495**, 198 – 202.

[5] V. R. Thalladi, M. Nüsse and R. Boese, J. Am. Chem. Soc., 2000, **122**, 9227–9236.

[6] J. L. Derissen and P. H. Smith, Acta Cryst. B, 1974, **30**, 2240–2242.

[7] A. P. Bartók, R. Kondor and G. Csányi, Phys. Rev. B, 2013, **87**, 184115.

[8] A. P. Bartók, M. C. Payne, R. Kondor and G. Csányi, Phys. Rev. Lett., 2010, **104**, 136403.

[9] A. P. Bartók and G. Csányi, Int. J. Quantum Chem., 2015, **115**, 1051–1057.

[10] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter and M. Scheffler, Comput. Phys. Commun., 2009, **180**, 2175–2196.

[11] B. Aradi, B. Hourahine and T. Frauenheim, J. Phys. Chem. A, 2007, **111**, 5678–5684.

[12] C. G. Broyden, IMA J. Appl. Math., 1970, **6**, 76–90.

[13] R. Fletcher, Comput. J., 1970, **13**, 317–322.

[14] D. Goldfarb, Math. Comput., 1970, **24**, 23–23.

[15] D. F. Shanno, Math. Comput., 1970, **24**, 647–647.

[16] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng and

K. W. Jacobsen, J. Phys. Condens. Matter, 2017, **29**, 273002.

[17] X. Li, F. S. Curtis, T. Rose, C. Schober, A. Vazquez-Mayagoitia, K. Reuter, H. Oberhofer and N. Marom, J. Chem. Phys., 2018, **148**, 241701.