# Density Functional Tight-Binding Model for Lithium-Silicon Alloys.

Ma. Belén Oviedo,[†,‡] Francisco Fernandez,[¶] Manuel Otero,[¶] Ezequiel P. M. Leiva,[†,‡] and S. Alexis Paz[*,†,‡]

†*Universidad Nacional de Córdoba. Facultad de Ciencias Químicas. Departamento de Química Teórica y Computacional. Córdoba, Argentina.*

‡*INFIQC, CONICET, (X5000HUA), Córdoba, Argentina.*

¶*IFEG, CONICET, (X5000HUA), Córdoba, Argentina.*

E-mail: apaz@unc.edu.ar

**Abstract**

The predictive power of molecular dynamic simulations is mainly restricted by the time scale and model accuracy. Many systems of current relevance are of such complexity that requires addressing both issues simultaneously. This is the case of silicon electrodes in Li-ion batteries, where different $Li_xSi$ alloys are formed during charge/discharge cycles. While first-principles treatments for this system are seriously limited by the computational cost of exploring its large conformational space, classical force-fields are not transferable enough to represent it accurately. Density Functional Tight-Binding (DFTB) is an intermediate complexity approach capable of capturing the electronic nature of different environments with a relatively low computational cost. In this work, we present a new set of DFTB parameters suited to model amorphous $Li_xSi$ alloys. The latter is the usual finding upon cycling the Si electrodes in the presence of Li-ions. The model parameters are constructed with a particular emphasis on their transferability for the entire $Li_xSi$ composition range. This is achieved by

introducing a new optimization procedure that weights stoichiometries differently to improve the prediction of their formation energies. The resulting model is shown to be robust for predicting crystal and amorphous structures for the different compositions, giving excellent agreement with DFT calculations and outperforming state-of-the-art ReaxFF potentials.

# 1 Introduction

Nowadays, the relevance of lithium batteries for renewable energies and multiple other technological applications is common knowledge and does not need further introduction, other than stating that their mentors were honoured with the Nobel Prize in Chemistry 2019. Several reviews and perspectives on the the subject can be found in literature elsewhere.[1–8] Current efforts focus on developing better materials to be used in the different components of the battery like anodes, electrolytes, separators, and cathodes. The challenge requires working on the edge where theoretical and experimental researchers from multiple disciplines come together. In this scenario, computer simulations must be the nexus between the different scientific approaches. Although several multiscale models and advanced computational methods have been developed in the last decades, further work is required to expand their implementations in systems like lithium batteries.[9]

One of the top promising materials considered for lithium battery anodes is silicon, mainly due to its large theoretical capacity (3579 mAh/g) and natural abundance.[10,11] The main drawback of this electrode is the large volume expansion it suffers during charge, which provokes fractures and loss of electrical contact with the corresponding hindrance in performance and useful life. Different strategies have been proposed to cope with this phenomenon through the use of thin films,[12] nanostructures,[13,14] carbon composites[15,16] or special binders.[17,18] Many of these ideas are based on speculative analysis that addresses the improvement of the material performance to the nature of the structures prepared, although a realistic picture based on an atomistic point of view is still missing.

Computer simulations can shorten the path to the industrial application of silicon anode. However, the discipline must face the intrinsic issues of the complexity of the system. Simulations must be able to follow the drastic changes that occur in the material during charge/discharge cycles. The silicon electrode passes through several alloys with different physicochemical properties. After electrode cycling, the Li-Si structures formed are found to be amorphous, making it difficult to collect structural information from experimental X-ray measurements. Two main concerns must be addressed: to count with an atomistic model transferable to the entire loading range (i.e. $Li_xSi$ / $x \in [0, 3.75]$) and to achieve a meaningful sampling of that model. During the last decade, progress has been made in both directions. Fan *et al.* have developed a ReaxFF parameterization using different $Li_xSi$ crystal structures.[19] Using this model, we proposed an enhanced sampling method that allowed us to predict several amorphous structures that probably exists in cycled $Li_xSi$ electrode.[20] However, improved sampling requires a highly transferable model capable of accurately representing unknown regions of configurational space, allowing true discovery of electrode structures and avoiding false positives.

When the silicon electrode is repeatedly cycled amorphous structures are produced. In order to develop a model transferable to these *a prior* unknown structures, we have parameterized a Density Functional Tight-Binding (DFTB) model for $Li_xSi$ which is the main goal of the present paper. Although DFTB implies a stronger level of theory than classical force fields like ReaxFF, it is faster than conventional Density Functional Theory (DFT) calculations, enabling to study the dynamics of thousands of atoms. We also present a new parameterization approach that allows us to optimize the model by reproducing other observables beyond absolute energies. We use this idea to focus the model on reproducing the relative formation energies of the different $Li_xSi$ structures, a key feature to achieve a robust and transferable model for the entire composition range. The present new model shows excellent performances in the prediction of formation energies against DFT for several crystalline and amorphous structures.

# 2 Computational Methods

## 2.1 DFTB method

The density functional tight-binding (DFTB) formalism applied in this work has been extensively described in the literature, for a complete formulation we address to references 21–24. DFTB is based on a second-order expansion of the density functional theory (DFT) energy with respect to a reference electron density fluctuation.[25] The resulting DFTB energy can be written as

$$E_{\text{DFTB}} = \sum_{i}^{\text{occ}} \langle \Psi_i | \hat{H}^0 | \Psi_i \rangle + \frac{1}{2} \sum_{AB} \gamma_{AB} \Delta q_A \Delta q_B + E_{\text{rep}}^{AB} \tag{1}$$

where the $\Psi_i$ are the single-particle Kohn-Sham (KS) orbitals. Within the linear combination of atomic orbitals formalism, $\Psi_i$ is expanded into a set of pseudoatomic valence (Slater-type) orbitals $\phi_\nu$:

$$\Psi_i(\boldsymbol{r}) = \sum_{\nu} c_{\nu i} \phi_\nu(\boldsymbol{r} - \boldsymbol{R}_A) \tag{2}$$

which are determined by solving the KS secular equation

$$\sum_{\mu} c_{\mu i} \left( H_{\nu\mu}^0 - \epsilon_i S_{\nu\mu} \right) = 0, \quad \forall \nu, i \tag{3}$$

where $S_{\nu\mu} = \langle \phi_\nu | \phi_\mu \rangle$ and $\epsilon_\nu$ are the overlap matrix and the eigenvalues in the isolated atom, respectively. $H_{\nu\mu}^0$ is the KS effective Hamiltonian generated with the reference electron density, $\rho^0$, and is defined as:

$$H_{\nu\mu}^0 = \begin{cases} \epsilon_\mu & \text{if } \nu = \mu \\ \langle \phi_\nu | -\frac{1}{2}\nabla^2 + v_{\text{eff}} \left[ \rho_A^0 + \rho_B^0 \right] | \phi_\nu \rangle & \text{if } \mu \in A,\ \nu \in B \text{ and } A \neq B \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where $\rho_A^0$ is the reference density of the neutral atom $A$ and $v_{\text{eff}}$ is the effective Kohn-Sham potential. In particular, the elements of the Hamiltonian matrix depend only on atoms $A$ and $B$, therefore only the two-center elements of the Hamiltonian and overlap matrices are explicitly calculated as a function of the distance and the orientation, using the Slater-Koster transformation rules.[26]

One of the crucial parts of the use of DFTB method is to calculate the basis functions $\phi$, and the atomic densities $\rho^0$. The pseudoatomic orbitals and density are obtained by solving the modified atomic Kohn-Sham (KS) equations in which a confining potential, $V_{\text{conf}}$, is added:

$$\left[\hat{T} + V_{\text{eff}} + V_{\text{conf}}\right] \phi_\mu = \epsilon_\mu \phi_\mu \tag{5}$$

A common practice in the DFTB community is to choose a confinement potential to be a parabolic or quadratic function. In this work, we have used the methodology implemented in the Hotcent code[27,28] where a power law function with the following form is used:

$$V_{\text{conf}}(r) = \left(\frac{r}{r_0}\right)^\sigma \tag{6}$$

$r_0$ and $\sigma$ are real numbers, which can be chosen differently for atomic orbitals $\phi$ and densities $\rho^0$.

The second term in equation 1 is the energy due to charge fluctuations and is parameterized analytically as a function of the orbital charges and $\gamma_{AB}$, which is a function of the interatomic separation and Hubbard parameter, $U$. $\Delta q_A = q_A - q_A^0$ is the self-consistent induced Mulliken charge on atom $A$; for details see reference 21. The remaining contribution to the total DFTB energy in equation 1 is $E_{\text{rep}}$, which corresponds to the distance-dependent diatomic repulsive potential and contains the core electron effects, ion–ion repulsion terms, as well as some exchange–correlation effects. This term can be considered as a practical equivalent to an xc-functional in DFT as it approximates the many-body correlation interactions with simple functions. The total repulsive energy of a system in DFTB

is a sum of contributions of repulsive potentials $V_{\text{rep}}(r)$ from each atom pair:

$$E_{\text{rep}} = \sum_{I<J} V_{\text{rep}}(R_{IJ}) \tag{7}$$

where $I$ and $J$ run over the atom indices in the system, and $R_{IJ}$ is the distance between pair of atoms. $V_{\text{rep}}$ is generally considered to be an empirical function that is determined by fitting data from a higher level electronic structure calculations such as DFT. In this work, the adjustment of this potential was carried out using the TANGO code[29,30] where the repulsive potential is approximated by the following function:

$$V_{\text{rep}}(R) = \begin{cases} e^{-a_1 R + a_2} + a_3 & 0 \leq R < R_{\text{min}} \\ \sum_{i=2}^{m} c_i \left(R_{\text{cut}} - R\right)^i & R_{\text{min}} \leq R < R_{\text{cut}} \\ 0 & R_{\text{cut}} \leq R \end{cases} \tag{8}$$

The values of $R_{\text{min}}$ and $R_{\text{cut}}$ used for this work are shown in table S1 of the supporting information. The $a_i$ parameters were fitted using the Levenberg-Marquardt algorithm to reproduce the DFT energies for each crystal system. The degree $m$ of the polynomial was chosen to be $8$ and the $c_i$ coefficients were optimized through a least-squares fit.

## 2.2 DFT calculations

DFT calculations of crystalline structures were performed using the GPAW simulation package[31,32] of the Atomic Simulation Environment.[33] The GPAW package is a real space grid algorithm based on the projector-augmented wavefunction method[34] that uses the frozen core approximation. Coordinates of Li, $Li_{15}Si_4$, $Li_{13}Si_4$, $Li_7Si_3$, $Li_{12}Si_7$, LiSi and Si were downloaded from The Materials Project[35] (mp codes 135, 569849, 672287, 1201871, 1314, 795 and 149) corresponding to BCC Li, x $\approx$ 3.75 , 3.25 , 2.33 , 1.71 , 1 and Si diamond, respectively. DFT calculations were performed using the PBE (Perdew-Burke-

Ernzerhof) exchange-correlation functional with a kinetic energy cutoff of 600 eV and the Brillouin zone integration was done with Monkhorst-Pack grids with a density of 2.5 k points per Å$^{-1}$.

Amorphous $Li_xSi$ structures were also studied by DFT calculations following the procedure proposed by Chevrier and Dahn.[36,37] We used a repeated cell scheme with 12 silicon atoms and $N$ lithium atoms per unit cell, with $N \in [0, 45]$ covering the hole range of $x \in [0, 3.75]$. Each $Li_{N+1}Si_{12}$ structure was obtained by adding a lithium atom in the greatest void of the $Li_NSi_{12}$ cell and performing a geometrical optimization of all atomic positions and cell volume. In this case, the calculations were performed with the QUANTUM ESPRESSO software,[38,39] using the PBE xc-functional with a kinetic energy cutoff of 1090 eV and the Brillouin zone integration was done with Monkhorst-Pack grids with a density of 7 k points per Å$^{-1}$. The atomic positions and cell volume were optimized employing the BFGS algorithm until the force was less than 0.08 eV/Å for each structure.

## 2.3 DFTB parameterization

In this work, all DFTB calculations were carried out using DFTB+,[40] and we follow the training method described in the work of van den Bossche *et al.*[28,30] to obtain two sets of DFTB parameters, named A and B, that differ from each other in the fitting of the band energy term. In group A, Li and Si band structures were fitted separately, while in group B the $Li_7Si_3$ band structure was fitted. The choice of this structure for the adjustment of the band structure was based on the value of its formation energy, which is the lowest of the 5 main alloys ($Li_{15}Si_4$, $Li_{13}Si_4$, $Li_{12}Si_7$, $Li_7Si_3$ and LiSi) as reported in The Materials Project,[35] and reproduced in our calculations. The parameterization of the pseudoatomic orbitals and electronic density consists in optimizing the values of $r_0$ and $s$ in equation 6 in order to fit the band structure to the reference DFT band structure. Table 1 shows the obtained confinement parameters.

On the other hand, the training dataset required for fitting the pair repulsion term

Table 1: Confinement potential parameters $r_0$ (Bohr) and $\sigma$, for atomic orbitals $\phi$ and densities $\rho^0$ of Li and Si.

| | set A | | | | set B | | | |
|---|---|---|---|---|---|---|---|---|
| Element | $r_0.\phi$ | $\sigma.\phi$ | $r_0.\rho^0$ | $\sigma.\rho^0$ | $r_0.\phi$ | $\sigma.\phi$ | $r_0.\rho^0$ | $\sigma.\rho^0$ |
| Li | 4.899 | 1.889 | 7.233 | 1.986 | 4.843 | 1.927 | 7.210 | 1.999 |
| Si | 4.558 | 6.909 | 6.318 | 2.188 | 3.556 | 2.382 | 6.292 | 1.891 |

was created using the fundamental crystalline structures of the different stoichiometries considered in this work, $S = \{\text{Li}, \text{Li}_{15}\text{Si}_4, \text{Li}_{13}\text{Si}_4, \text{Li}_{12}\text{Si}_7, \text{Li}_7\text{Si}_3, \text{LiSi}, \text{Si}\}$. These fundamental structures were isotropically compressed/expanded using a scaling factor that varied from 0.4 to 1.45 every 0.05 units, generating 22 structures per stoichiometry. The energy of each structure was computed using DFT, and those with more than 10 eV above the corresponding stochiometry minimum were excluded. This procedure gives a total of 108 structures in the training dataset. We denote with $N_s$ the number of structures associated with stoichiometry $s \in S$ in the dataset. Also, we denote with $\boldsymbol{r}_i^s$ the $i$-th structure of stoichiometry $s$ and with $\check{\boldsymbol{r}}^s$ the structure corresponding to the DFT energy minimum for $s$. Consistently, we will use the symbol " $\check{}$ " to denote the argument of the minimum for other functions.

We proceed to obtain a set of DFTB parameters, $\check{\boldsymbol{p}} = (\{\check{c}_i\}, \{\check{a}_i\})$ (see equation 8), that minimize the residual

$$\text{Res}_E(\boldsymbol{p}) = \sum_{s \in S} \sum_{i=1}^{N_s} \omega_i^s [E_{\text{DFT}}(\boldsymbol{r}_i^s) - E_{\text{DFTB}}(\boldsymbol{r}_i^s; \boldsymbol{p}) - C]^2 \tag{9}$$

where $C$ is a constant that shift the DFTB energy in order to correct systematic overbinding or underbinding tendencies,[28,30] $E_{\text{DFTB}}(\boldsymbol{r}_i^s; \boldsymbol{p})$ is computed using DFTB with a specific parameter set $\boldsymbol{p}$ and $\omega_i^s$ allowing to control the relative weight of each $\boldsymbol{r}_i^s$ structure in the fitting procedure. Note that this equation is just particular case of the one presented in reference,[28] with only two minor changes: we explicitly group the structures in subsets (*i.e.* stoichiometries) and we exclude the force terms. The first change is found to be convenient for the discussion below and the second one is for simplicity since all the structures

in the training dataset are in mechanical equilibrium.

In order to minimize equation 9, it is necessary to choose the $\omega_i^s$ relative weights. In reference 28, the authors suggest a Boltzmann-like distribution

$$\tilde{\omega}_i^s = \exp\left(-\frac{E_{\text{DFT}}(\boldsymbol{r}_i^s) - E_s}{b_i^s}\right) \tag{10}$$

where $b_i^s$ is taken to be proportional to the number of atoms $n_i^s$ in each structure and $E_s$ is an energy reference. As suggested in the documentation of the TANGO code we can fix $b_i^s = 0.1 n_i^s$ eV and a suitable choice for $E_s$ would be the lowest energy of stoichiometry $s$:

$$E_s = E_{\text{DFT}}(\check{\boldsymbol{r}}_s) \quad \text{and} \quad E_{\text{DFT}}(\check{\boldsymbol{r}}_s) \leq E_{\text{DFT}}(\boldsymbol{r}_i^s) \quad \forall i \in [1, N_s] \tag{11}$$

The underlying motivation for this equation is to increase the accuracy of the resulting DFTB model in predicting low-energy structures, to the detriment of the higher ones which are less probable to find during a canonical simulation. Note however that this choice does not weight the different stoichiometries used.

In general, by choosing the weights of the structures in the training dataset in equation 9, we can tune the scope of application for the resulting DFTB parameterization. In other words, for each set of weights there is a different set of DFTB parameters $\check{\boldsymbol{p}}$ that minimize equation 9. Inspired by this idea, we choose the weights to be

$$\omega_i^s = \xi_s \tilde{\omega}_i^s = \xi_s \exp\left(-\frac{E_{\text{DFT}}(\boldsymbol{r}_i^s) - E_{\text{DFT}}(\check{\boldsymbol{r}}_s)}{b_i^s}\right) \tag{12}$$

retaining the focus in low-energy structures for each stoichiometry but including a new set of coefficients $\boldsymbol{\xi} = (\xi_{\text{Li}}, \cdots, \xi_{\text{Si}})$ that control the relative weights between different stoichiometries. With this definition, the optimal DFTB parameters for equation 9 can be considered a function $\check{\boldsymbol{p}}(\boldsymbol{\xi})$. Next, we will introduce a second optimization procedure to obtain $\boldsymbol{\xi}$ coefficients.

The final goal of our work is to parameterize a DFTB model that allows the simulation of silicon anode lithiation. Note that this is a very complex process that involves different chemical environments, with a wide range of $Li_xSi$ compositions. However, it is important that the new parameterization maintains its accuracy for the widest possible range of concentrations, in order to avoid the need to change the model "on-the-fly" during a simulation. This could be particularly disturbing for approaches like grand-canonical MD simulations, where the relative composition is constantly changing. Thus, we require the parameterization to be as transferable as possible between different stoichiometries. In this sense, to predict relative formation energies, $F(\check{\boldsymbol{r}}_s)$:

$$F(\check{\boldsymbol{r}}_s) = \frac{E(\check{\boldsymbol{r}}_s) - (xE(\check{\boldsymbol{r}}_{\mathrm{Li}}) + E(\check{\boldsymbol{r}}_{\mathrm{Si}}))}{x+1} \tag{13}$$

with $x$ as in $Li_xSi$, is the main goal of our parameterization. Therefore, we choose the $\check{\xi}$ coefficients to minimize the residual:

$$\mathrm{Res}_F(\boldsymbol{\xi}) = \sum_{s\in S}[F_{\mathrm{DFT}}(\check{\boldsymbol{r}}^s) - F_{\mathrm{DFTB}}(\check{\boldsymbol{r}}^s; \check{\boldsymbol{p}}(\boldsymbol{\xi}))]^2 \tag{14}$$

Minimizing this residue leads to a set of DFTB parameters $\check{\boldsymbol{p}}(\check{\boldsymbol{\xi}})$ which gives a minimal energy residue for a training dataset where stoichiometries are weighted to also give a minimal residue in their formation energy.

Figure 1 shows a flow chart with the main steps of the fitting procedure described in the text above. The minimization of equation 9 is performed using the TANGO code.[29] In order to minimize equation 14, we developed a driver program called Milonga that executes several TANGO instances, one for each $\mathrm{Res}_F(\boldsymbol{\xi})$ evaluation required by the minimization algorithm.

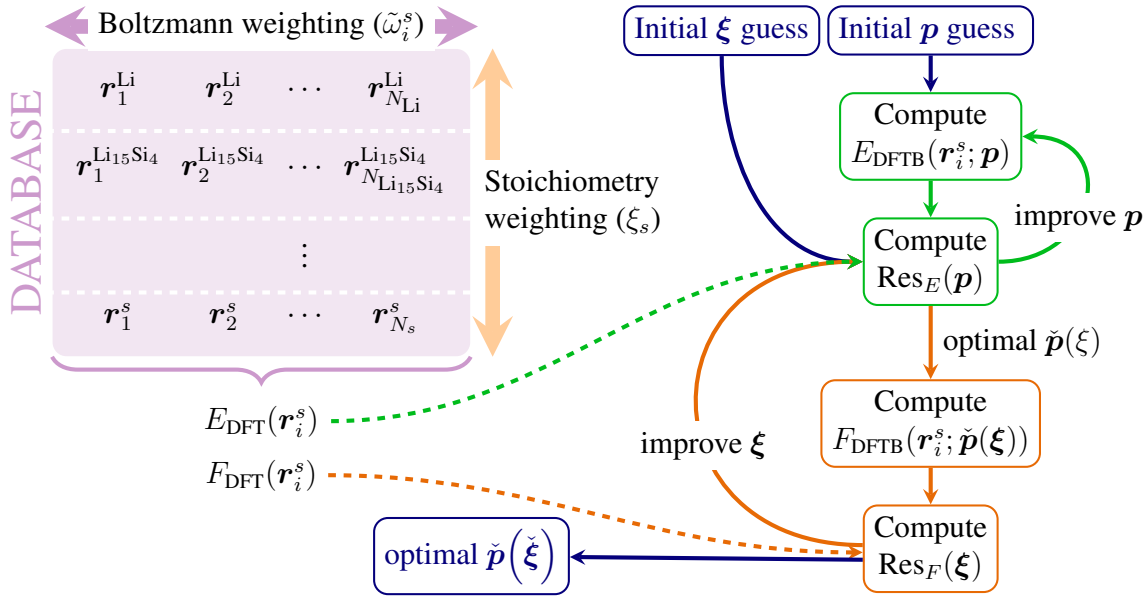


Figure 1: Flow chart for the fitting procedure used in this work. Two nested optimization procedures are performed: the minimization of $\mathrm{Res}_E$ using TANGO[29] code (highlighted in green) and the minimization of $\mathrm{Res}_F$ using a new driver code named Milonga (highlighted in orange). Note that each improvement of $\xi$ weights requires a full $\mathrm{Res}_E$ minimization to obtain the optimal DFTB associated parameters $\check{p}(\xi)$.

## Results and discussion

As indicated in the previous section, there are two groups of parameters to be determined, the electronic ones (pseudoatomic orbitals and electron density) and the repulsive potentials for each pair of chemical elements. These parameters are subject to optimization in order to reproduce certain desired properties, such as the electronic band structure, and atomization energies, among others. Following the work of van den Bossche *et al.*[28] only the 2s valence electron is considered for Li, meanwhile, the 3s and 3p valence electrons were considered for Si. Figure 2 shows a comparison of DFTB computed band structures using the set of parameters $B$ with DFT computed band structures for four systems: Li BCC, $Li_7Si_3$, an isolated Li atom in Si diamond ($LiSi_{15}$), and Si diamond. The same comparison of DFTB with parameter set $A$ is shown in figure S1 of Supporting Information. Overall, for Li and Si bulk systems, DFTB calculations mimic closely the DFT computed band structures near the Fermi energy level (0 eV). These agreements were not expected since the

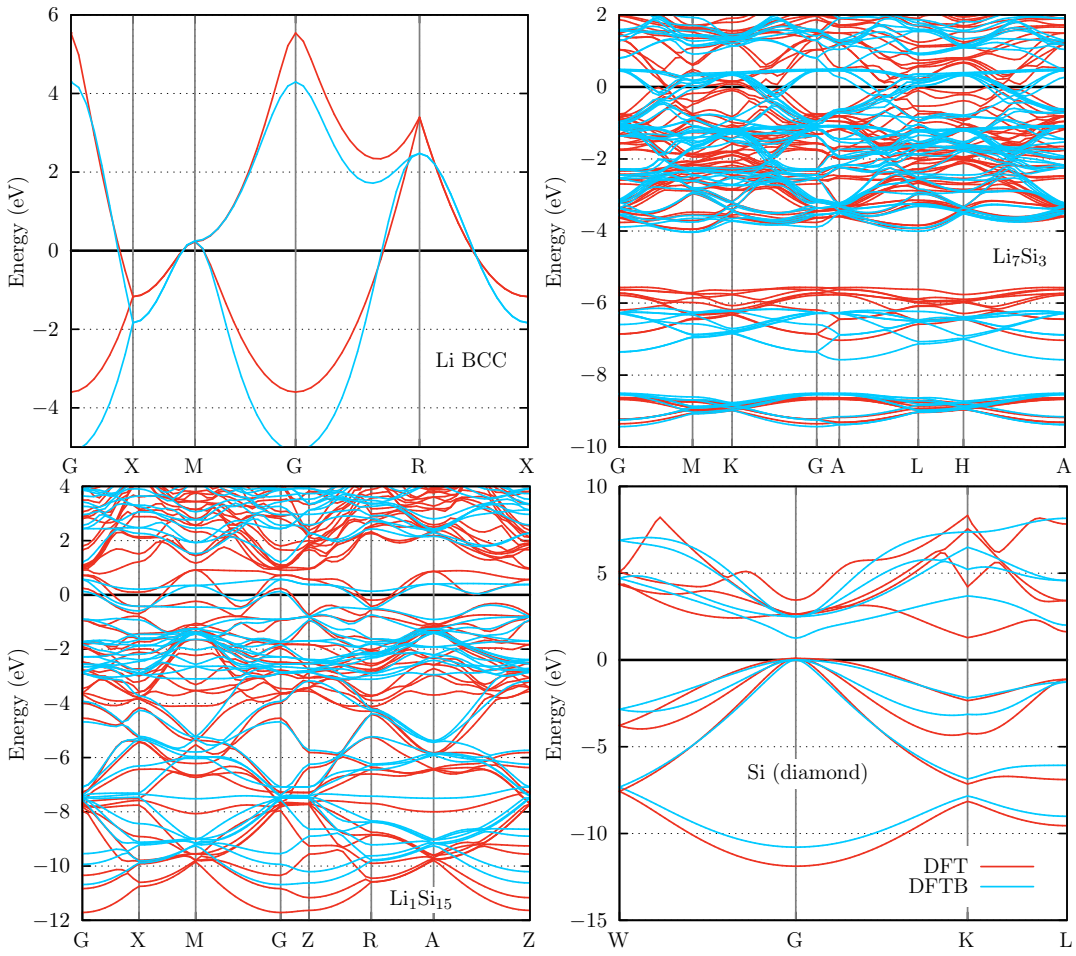


Figure 2: Band structures computed by DFTB using the set of parameters B, in comparison with the band structures computed by DFT/PBE for Li body-centered cubic (BCC), $Li_7Si_3$, an isolated Li in Si diamond ($Li_1Si_{15}$), and Si diamond. All the electronic bands are shifted to the respective Fermi levels (0 eV).

parameters were not fitted to bulk Li or bulk Si (as was for set $A$) instead, they are fitted to Li and Si in the $Li_7Si_3$ crystal structure. These results are the first test of the accuracy and transferability of the parameter set $B$. However, the DFTB computed electronic bands far away from the Fermi level differ from the DFT predicted values, something which is expected given the minimal basis used in DFTB. Figure 2 shows a pronounced mismatch in the conduction band at the $\Gamma$ point (G in the figure) in the Si structure. This difference can be attributed to the present simplified basis used in the DFTB model. For the two LiSi systems, we see a qualitative agreement of the band structures far from the Fermi level and a small degree of compression when approaching the Fermi level. Given that in general, the repulsive term is capable of correcting small imperfections in the electronic part quite effectively, we decided to continue with the parameterization using these confinement parameters. This criterion is supported also by the good results that we present below.

Table 2: Optimal weights, $\check{\xi}_s$, obtained for the different band energy parameterization.

| s | set A | set B |
|---|---|---|
| Li | $0.23\,10^{-2}$ | $0.49$ |
| $Li_{15}Si_4$ | $0.15$ | $0.28\,10^{-21}$ |
| $Li_{13}Si_4$ | $0.21$ | $0.17\,10^{-1}$ |
| $Li_7Si_3$ | $0.21$ | $0.11\,10^{-1}$ |
| $Li_{12}Si_7$ | $0.23$ | $0.11\,10^{-2}$ |
| LiSi | $0.21$ | $0.35\,10^{-3}$ |
| Si | $0.83\,10^{-7}$ | $0.49$ |

We proceed to parameterize the pair repulsion of the DFTB model by using each of the two energy band terms obtained as starting point. This gives us two different sets of DFTB parameters, say $A$ and $B$, obtained by using the weight optimization procedure described in the method section. The $\check{\xi}_s$ coefficients obtained for each set, A and B, are reported in table 2. It is interesting that for set $A$, the algorithm reduces the relative weight of pure Li and Si crystals, and increases the weights of the different alloys. On the contrary, for set $B$, the optimal weights are lower for the alloys and higher for the pure crystals. We rationalize

this behavior by noticing that the band energy term of set $A$ is built using pure element structures while the same term for set $B$ uses an alloy. It seems reasonable that the pair repulsion term, which attempts to compensate for the energy residue given by all the other energy contributions, will be less important for those structures used for fitting the band energies and more important for the others. In this sense, optimal $\check{\xi}_s$ coefficients seem to be able to sense this situation and try to compensate for it by focusing the parameterization on the structures that need it most.

Another interesting observation that arises from table 2 is the low weight resulting for $Li_{15}Si_4$, in comparison with the other alloys of each set. Furthermore, in the case of set $B$ this weight is practically zero, indicating that the inclusion of these structures in the training dataset does not contribute to improving the overall prediction of the relative formation energies. This leads us to assume that $Li_{15}Si_4$ structure has a particular nature that differs from all the other alloys and interferes somehow with the fitting of the model parameters. For instance, a plausible explanation may be that the chemical environment of this structure is quite different from the rest of the alloys so the prediction can be globally improved by focusing the model on all the other alloys but not $Li_{15}Si_4$. In this sense, the weight optimization procedure proposed in this work may be seen as a tool to detect peculiar information introduced in a training dataset. This should be addressed in further work.

Figure 3 shows the DFT energies for the 108 structures in the training dataset, together with the corresponding DFTB predictions with the parameters sets A and B. In general, we found a good agreement for both parameter sets, although there are slightly larger differences in the compression curves for the pure element with the parameter set A. This is consistent with the low weight obtained for pure elements in this set (see table 2). Therefore, this greater variation in the energy for the pure elements is mainly a consequence of our weight optimization procedure. Here we recall that our priority is to achieve transferability over a wide range of mole fractions by focusing our model on reproducing the

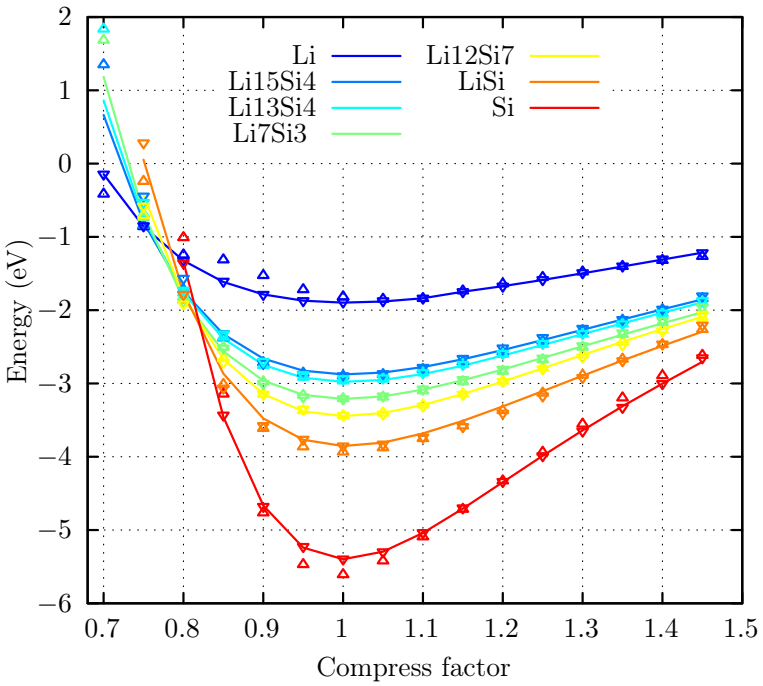


Figure 3: Energy profiles for the isotropic compression/expansion of the $Li_xSi_y$ crystal structures, computed using DFT (lines) and DFTB with the parameter sets A ($\triangle$) and B ($\triangledown$).

relative energies rather than their absolute values.

We turn now to inspect the energy of formation as defined in equation 13, although we present it as a function of the mole fraction of Si ($\Theta$ in $Li_{1-\Theta}Si_{\Theta}$) instead of $x = (1-\Theta)/\Theta$ because this is more usual in literature. The resulting values are shown in Figure 4 for DFT and the two DFTB parameters sets. We also include for comparison the formation energies obtained with the ReaxFF model parameterized in reference 19, which is a state-of-the-art reactive potential for these alloys. It is possible to observe a good agreement for the two DFTB models, where the parameter set B is shown to give the best of all approximations. The major differences between DFT and our model are for $Li_{13}Si_4$ ($\Theta \approx 0.24$) and $Li_{15}Si_4$ ($\Theta \approx 0.21$) alloys, the latter being the worst. Strikingly, this is the only structure that gets low weights during the optimization procedure. As we mentioned before, this may indicate that $Li_{15}Si_4$ structure has some peculiarity that makes it difficult to fit with the other alloys in the training dataset. Nevertheless, there is an improvement with respect to the

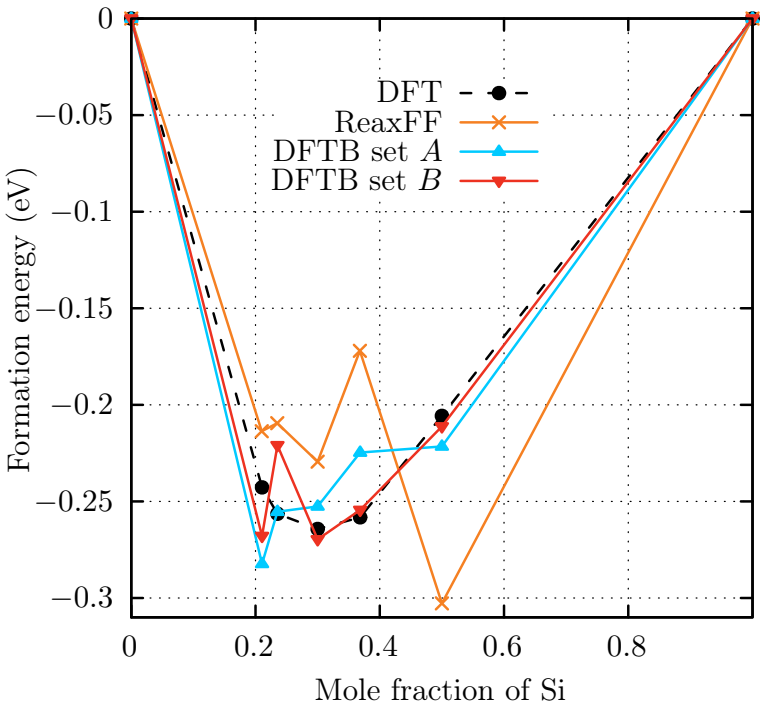


Figure 4: Formation energy as defined in equation 13 computed using DFT, ReaxFF[19] and DFTB using the two set of parameters optimized in this work.

ReaxFF potential for this system, and the differences obtained in the formation energies, in this case, are smaller.

In order to analyze the benefit of the weighted optimization procedure, we have computed a set of DFTB parameters using fixed equal weights for each alloy. In other words, we just skip the optimization of the $\xi$ coefficients. Let us call these sets of parameters $A_0$ and $B_0$, which correspond to a band energy term fitted from the pure elements or the $Li_7Si_3$ global minimum respectively. Figure 5 shows the formation energy residues for both sets and compares them with the obtained with $A$ and $B$ sets with optimized $\xi$ weights. Although some particular alloys might increase their formation energy residues after the weight optimization process, we can confirm a general trend toward relevant improvement of the prediction.

We proceed to test the prediction power of our DFTB models by computing the formation energy for different amorphous structures and compare these results with the ReaxFF and DFT methods. These structures were prepared as described in section 2.2 and the

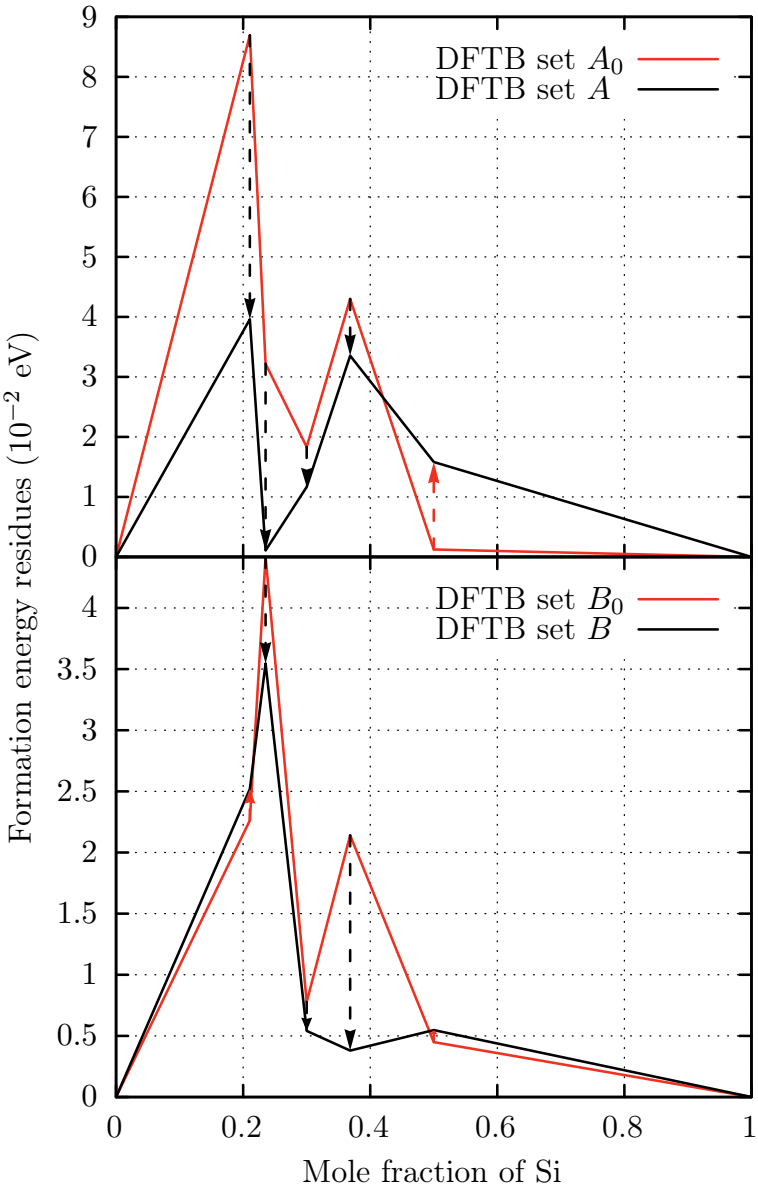


Figure 5: Formation energy residues for crystal structures obtained using DFTB with the new parameter sets before the optimization of $\xi$ weights ($A^0$ and $B^0$) and after it ($A$ and $B$). DFT energies are used as a reference. The gain (or loss) of accuracy is indicated by black (or red) arrows.

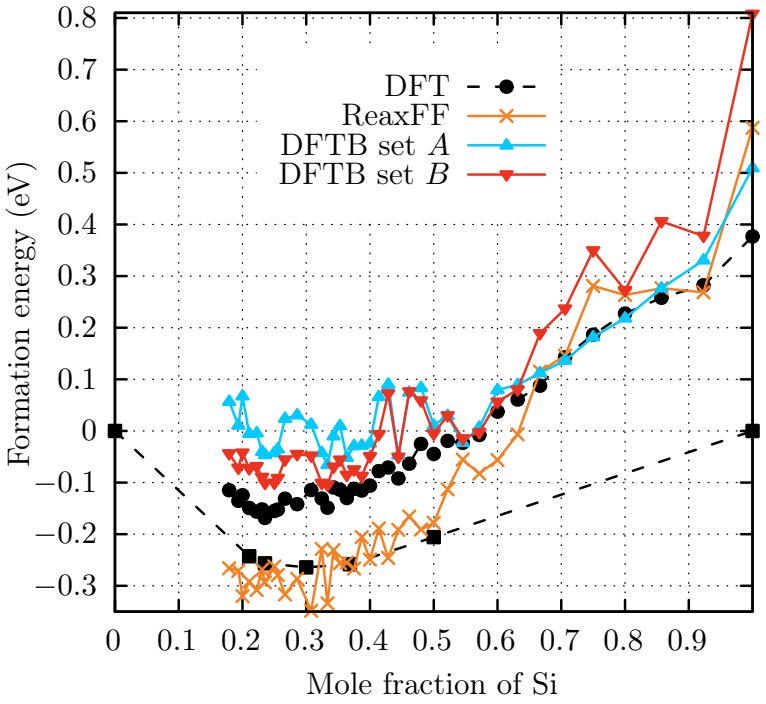


Figure 6: Formation energy of amorphous Li-Si alloys of different compositions. Crystalline structures calculated with DFT are included in black squares for comparison.

results are shown in figure 6. In general, it is possible to see that there is a good agreement between DFT and both DFTB parameterizations. Furthermore, a careful inspection allows concluding that DFTB is able to mimic the trend of the DFT data, following the occurrence of different maxima and minima. On the other hand, the formation energies obtained with ReaxFF show a significant deviation from DFT data at high lithium concentration ($\Theta < 0.5$), where amorphous structures are erroneously predicted to be even more stable than crystals. Another interesting observation is that DFTB predictions with parameter set $A$ are significantly better for alloys at low lithium concentration, while the opposite is true for set $B$. Note that the high precision for each set occurs for those alloys with similar concentrations to the structures used for fitting the band energy term, which are $Li_7Si_3$ ($\Theta = 0.3$) for set $B$ and pure Li and Si elements for set $A$.

Finally, we repeated the calculations for amorphous structures with DFTB using the parameter sets $A_0$ and $B_0$. A comparison between the resulting formation-energy residues with the corresponding quantities for $A$ and $B$ sets can be seen in Figure 7. From this data,

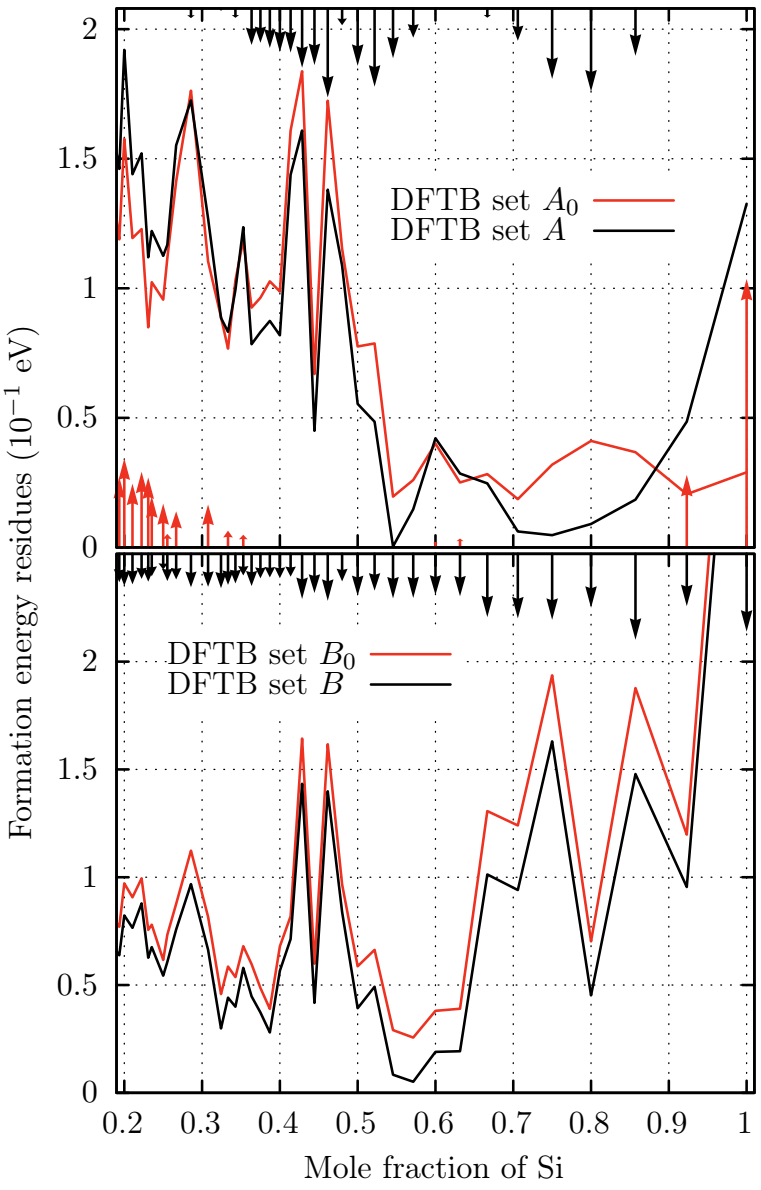


Figure 7: Formation energy residues obtained using DFTB with the new parameter sets before the optimization of $\xi$ weights ($A^0$ and $B^0$) and after it ($A$ and $B$). DFT energies are used as a reference. The gain (or loss) of accuracy is indicated by black (or red) arrows.

we can conclude that the weight optimization procedure helps to increase the accuracy of the formation energy. This is especially true for set $B$, which gives lower residues than set $B_0$ for all concentrations. In the case of set $A$, the weight optimization procedure allows a better precision for $0.35 < \Theta < 0.9$ but not for the extreme $\Theta$ values. We remark that the optimized weights for this set focus the pair-repulsion parameterization in these intermediate alloys, while the weights for the pure elements are considerably smaller.

The largest discrepancies between the formation energies obtained by DFT and the new DFTB models correspond to pure Si amorphous structures. Therefore, we have performed additional tests for this case. The radial distribution functions (RDF) are shown in Figure 8 together with experimental data. To obtain the amorphous structures we started with crystalline silicon using a super-cell with 64 atoms. Then, simulated annealing was performed in the NVT ensemble using the Nosé-Hoover thermostat, with a linear heating from room temperature to 3000 K for 100 ps, a melting at that temperature for 750 ps and an exponential quenching of 600 ps, always with a time step of 1 fs. An equilibration at 300 K for 100 ps was then run to obtain the production frames. We can highlight that the DFTB model of set B shows excellent agreement with the experimental data of reference 41, making this parameterization the most adequate for further simulation studies.

# Conclusions

We have developed a new DFTB parameterization scheme aiming to optimize other observables beyond absolute energies. We used this scheme to construct a DFTB model with an optimal prediction of formation energies for several crystallines and amorphous structures in a wide range of compositions. In the case of pure amorphous silicon, which exhibits the largest associated error in the prediction of formation energies, the experimental radial distribution function was perfectly reproduced by the model. The excellent performance shown by this model in comparison with DFT, in addition to being more than 1000

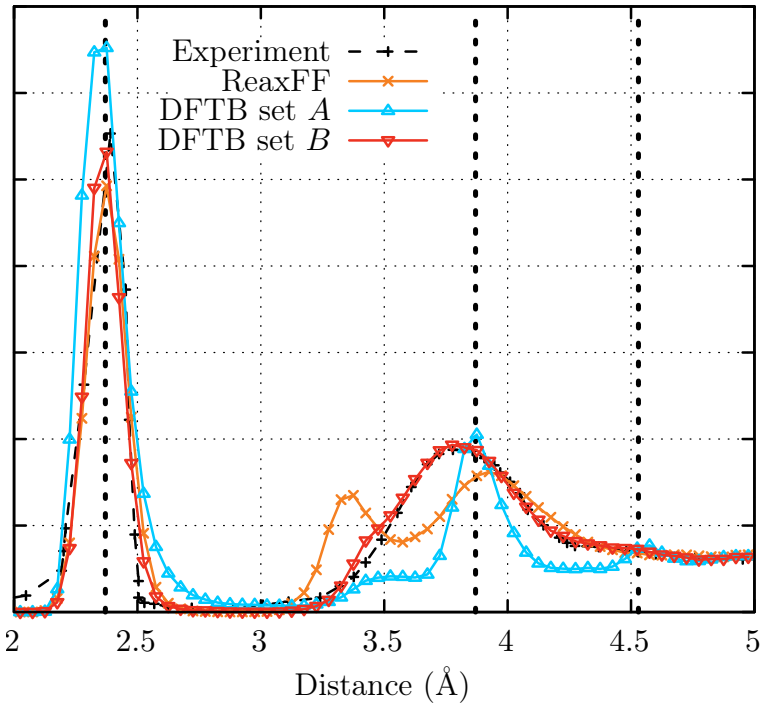


Figure 8: Radial distribution function (RDF) of amorphous silicon for the different parameterizations of the DFTB models developed here, set A and set B, as defined in the text. The results are compared with experimental values from reference 41, and with the ReaxFF.[19] The vertical dashed black lines show where the peaks of crystalline silicon would be at 0 K. An excellent agreement is found with parameterization B.

times faster, allows us to consider it as the potential of choice for future simulations of $Li_xSi$ alloys.

## Acknowledgement

We especially acknowledge Prof. Jorge Oviedo for his useful discussions and mathematical inputs. We acknowledge financial support from the Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT-FONCyT Grant PICT-2017-0621, PICT-2017-0795 and 2020-SERIEA-02139). F. Fernandez wishes to thank CONICET for his fellowship. E.P.M. Leiva acknowledges grants PIP CONICET 1220200101189CO, PUE/2017 CONICET, FONCYT 2020-SERIEA-03689 and SECyT of the Universidad Nacional de Córdoba. This work used computational resources from CCAD, Universidad Nacional de Córdoba (https://ccad.unc.edu.ar/), which are part of SNCAD, MinCyT of República Argentina.

## Supporting Information Available

Values of $R_{\text{min}}$, $R_{\text{cut}}$ and energy bands for DFTB parameter set A.

## References

(1) Zhang, L.; Zhu, C.; Yu, S.; Ge, D.; Zhou, H. Status and Challenges Facing Representative Anode Materials for Rechargeable Lithium Batteries. *Journal of Energy Chemistry* **2022**, *66*, 260–294.

(2) Cheng, X.-B.; Liu, H.; Yuan, H.; Peng, H.-J.; Tang, C.; Huang, J.-Q.; Zhang, Q. A Perspective on Sustainable Energy Materials for Lithium Batteries. *SusMat* **2021**, *1*, 38–50.

(3) Ali, H.; Khan, H. A.; Pecht, M. G. Circular Economy of Li Batteries: Technologies and Trends. *Journal of Energy Storage* **2021**, *40*, 102690.

(4) dos Reis, G.; Strange, C.; Yadav, M.; Li, S. Lithium-Ion Battery Data and Where to Find It. *Energy and AI* **2021**, *5*, 100081.

(5) Kulova, T. L.; Fateev, V. N.; Seregina, E. A.; Grigoriev, A. S. A Brief Review of Post-Lithium-Ion Batteries. **2020**, *15*, 7242–7259.

(6) Walter, M.; Kovalenko, M. V.; Kravchyk, K. V. Challenges and Benefits of Post-Lithium-Ion Batteries. *New Journal of Chemistry* **2020**, *44*, 1677–1683.

(7) Chu, S.; Cui, Y.; Liu, N. The Path towards Sustainable Energy. *Nature Materials* **2016**, *16*, 16–22.

(8) Nitta, N.; Wu, F.; Lee, J. T.; Yushin, G. Li-Ion Battery Materials: Present and Future. *Materials Today* **2015**, *18*, 252–264.

(9) Leiva, E. P. M. Modeling of Lithium-Ion Batteries Is Becoming Viral: Where to Go? *Journal of Solid State Electrochemistry* **2020**, *24*, 2117–2120.

(10) Feng, K.; Li, M.; Liu, W.; Kashkooli, A. G.; Xiao, X.; Cai, M.; Chen, Z. Silicon-Based Anodes for Lithium-Ion Batteries: From Fundamentals to Practical Applications. *Small* **2018**, *14*, 1702737.

(11) McDowell, M. T.; Lee, S. W.; Nix, W. D.; Cui, Y. 25th Anniversary Article: Understanding the Lithiation of Silicon and Other Alloying Anodes for Lithium-Ion Batteries. *Advanced Materials* **2013**, *25*, 4966–4985.

(12) Salah, M.; Murphy, P.; Hall, C.; Francis, C.; Kerr, R.; Fabretto, M. Pure Silicon Thin-Film Anodes for Lithium-Ion Batteries: A Review. *Journal of Power Sources* **2019**, *414*, 48–67.

(13) Zhao, X.; Lehto, V.-P. Challenges and Prospects of Nanosized Silicon Anodes in Lithium-Ion Batteries. *Nanotechnology* **2020**, *32*, 042002.

(14) Wu, H.; Cui, Y. Designing Nanostructured Si Anodes for High Energy Lithium Ion Batteries. *Nano Today* **2012**, *7*, 414–429.

(15) Li, P.; Kim, H.; Myung, S.-T.; Sun, Y.-K. Diverting Exploration of Silicon Anode into Practical Way: A Review Focused on Silicon-Graphite Composite for Lithium Ion Batteries. *Energy Storage Materials* **2021**, *35*, 550–576.

(16) Wu, J.; Cao, Y.; Zhao, H.; Mao, J.; Guo, Z. The Critical Role of Carbon in Marrying Silicon and Graphite Anodes for High-energy Lithium-ion Batteries. *Carbon Energy* **2019**, *1*, 57–76.

(17) Chen, H.; Wu, Z.; Su, Z.; Chen, S.; Yan, C.; Al-Mamun, M.; Tang, Y.; Zhang, S. A Mechanically Robust Self-Healing Binder for Silicon Anode in Lithium Ion Batteries. *Nano Energy* **2021**, *81*, 105654.

(18) Zhang, X.; Wang, D.; Qiu, X.; Ma, Y.; Kong, D.; Müllen, K.; Li, X.; Zhi, L. Stable High-Capacity and High-Rate Silicon-Based Lithium Battery Anodes upon Two-Dimensional Covalent Encapsulation. *Nat Commun* **2020**, *11*, 3826.

(19) Fan, F.; Huang, S.; Yang, H.; Raju, M.; Datta, D.; Shenoy, V. B.; van Duin, A. C. T.; Zhang, S.; Zhu, T. Mechanical Properties of Amorphous Li $_x$ Si Alloys: A Reactive Force Field Study. *Modelling and Simulation in Materials Science and Engineering* **2013**, *21*, 074002.

(20) Fernandez, F.; Paz, S. A.; Otero, M.; Barraco, D.; Leiva, E. P. M. Characterization of Amorphous LixSi Structures from ReaxFF via Accelerated Exploration of Local Minima. *Phys. Chem. Chem. Phys.* **2021**, *23*, 16776–16784.

(21) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. Self-Consistent-Charge Density-Functional Tight-Binding Method for Simulations of Complex Materials Properties. *Phys. Rev. B* **1998**, *58*, 7260–7268.

(22) Frauenheim, T.; Seifert, G.; Elsterner, M.; Hajnal, Z.; Jungnickel, G.; Porezag, D.; Suhai, S.; Scholz, R. A Self-Consistent Charge Density-Functional Based Tight-Binding Method for Predictive Materials Simulations in Physics, Chemistry and Biology. *phys. stat. sol. (b)* **2000**, *217*, 41–62.

(23) Seifert, G. Tight-Binding Density Functional Theory: An Approximate Kohn-Sham DFT Scheme. *J. Phys. Chem. A* **2007**, *111*, 5609–5613.

(24) Gaus, M.; Cui, Q.; Elstner, M. DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB). *J. Chem. Theory Comput.* **2011**, *7*, 931–948.

(25) Foulkes, W. M. C.; Haydock, R. Tight-Binding Models and Density-Functional Theory. *Phys. Rev. B* **1989**, *39*, 12520–12536.

(26) Slater, J. C.; Koster, G. F. Simplified LCAO Method for the Periodic Potential Problem. *Phys. Rev.* **1954**, *94*, 1498–1524.

(27) Van den Bossche, M. Hotcent. 2019.

(28) Van den Bossche, M. DFTB-Assisted Global Structure Optimization of 13- and 55-Atom Late Transition Metal Clusters. *J. Phys. Chem. A* **2019**, *123*, 3038–3045.

(29) Van den Bossche, M. TANGO: Tight-binding Approximation eNhanced Global Optimization. 2018.

(30) Van den Bossche, M.; Grönbeck, H.; Hammer, B. Tight-Binding Approximation-Enhanced Global Optimization. *J. Chem. Theory Comput.* **2018**, *14*, 2797–2807.

(31) Enkovaara, J.; Rostgaard, C.; Mortensen, J. J.; Chen, J.; Dułak, M.; Ferrighi, L.; Gavnholt, J.; Glinsvad, C.; Haikola, V.; Hansen, H. A.; Kristoffersen, H. H.; Kuisma, M.; Larsen, A. H.; Lehtovaara, L.; Ljungberg, M.; Lopez-Acevedo, O.; Moses, P. G.; Ojanen, J.; Olsen, T.; Petzold, V.; Romero, N. A.; Stausholm-Møller, J.; Strange, M.; Tritsaris, G. A.; Vanin, M.; Walter, M.; Hammer, B.; Häkkinen, H.; Madsen, G. K. H.; Nieminen, R. M.; Nørskov, J. K.; Puska, M.; Rantala, T. T.; Schiøtz, J.; Thygesen, K. S.; Jacobsen, K. W. Electronic Structure Calculations with GPAW: A Real-Space Implementation of the Projector Augmented-Wave Method. *J. Phys.: Condens. Matter* **2010**, *22*, 253202.

(32) Mortensen, J. J.; Hansen, L. B.; Jacobsen, K. W. Real-Space Grid Implementation of the Projector Augmented Wave Method. *Phys. Rev. B* **2005**, *71*, 035109.

(33) Larsen, A. H.; Mortensen, J. J.; Blomqvist, J.; Castelli, I. E.; Christensen, R.; Dułak, M.; Friis, J.; Groves, M. N.; Hammer, B.; Hargus, C.; Hermes, E. D.; Jennings, P. C.; Jensen, P. B.; Kermode, J.; Kitchin, J. R.; Kolsbjerg, E. L.; Kubal, J.; Kaasbjerg, K.; Lysgaard, S.; Maronsson, J. B.; Maxson, T.; Olsen, T.; Pastewka, L.; Peterson, A.; Rostgaard, C.; Schiøtz, J.; Schütt, O.; Strange, M.; Thygesen, K. S.; Vegge, T.; Vilhelmsen, L.; Walter, M.; Zeng, Z.; Jacobsen, K. W. The Atomic Simulation Environment—a Python Library for Working with Atoms. *J. Phys.: Condens. Matter* **2017**, *29*, 273002.

(34) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50*, 17953–17979.

(35) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A. Commentary: The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation. *APL Materials* **2013**, *1*, 011002.

(36) Chevrier, V. L.; Dahn, J. R. First Principles Model of Amorphous Silicon Lithiation. *J. Electrochem. Soc.* **2009**, *156*, A454, Publisher: IOP Publishing.

(37) Chevrier, V. L.; Dahn, J. R. First Principles Studies of Disordered Lithiated Silicon. *J. Electrochem. Soc.* **2010**, *157*, A392, Publisher: IOP Publishing.

(38) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; Corso, A. D.; Gironcoli, S. d.; Fabris, S.; Fratesi, G.; Gebauer, R.; Gerstmann, U.; Gougoussis, C.; Kokalj, A.; Lazzeri, M.; Martin-Samos, L.; Marzari, N.; Mauri, F.; Mazzarello, R.; Paolini, S.; Pasquarello, A.; Paulatto, L.; Sbraccia, C.; Scandolo, S.; Sclauzero, G.; Seitsonen, A. P.; Smogunov, A.; Umari, P.; Wentzcovitch, R. M. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys.: Condens. Matter* **2009**, *21*, 395502.

(39) Giannozzi, P.; Andreussi, O.; Brumme, T.; Bunau, O.; Nardelli, M. B.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Cococcioni, M.; Colonna, N.; Carnimeo, I.; Corso, A. D.; Gironcoli, S. d.; Delugas, P.; DiStasio, R. A.; Ferretti, A.; Floris, A.; Fratesi, G.; Fugallo, G.; Gebauer, R.; Gerstmann, U.; Giustino, F.; Gorni, T.; Jia, J.; Kawamura, M.; Ko, H.-Y.; Kokalj, A.; Küçükbenli, E.; Lazzeri, M.; Marsili, M.; Marzari, N.; Mauri, F.; Nguyen, N. L.; Nguyen, H.-V.; Otero-de-la Roza, A.; Paulatto, L.; Poncé, S.; Rocca, D.; Sabatini, R.; Santra, B.; Schlipf, M.; Seitsonen, A. P.; Smogunov, A.; Timrov, I.; Thonhauser, T.; Umari, P.; Vast, N.; Wu, X.; Baroni, S. Advanced capabilities for materials modelling with Quantum ESPRESSO. *J. Phys.: Condens. Matter* **2017**, *29*, 465901, Publisher: IOP Publishing.

(40) Hourahine, B.; Aradi, B.; Blum, V.; Bonafé, F.; Buccheri, A.; Camacho, C.; Cevallos, C.; Deshaye, M. Y.; Dumitrică, T.; Dominguez, A.; Ehlert, S.; Elstner, M.; van der Heide, T.; Hermann, J.; Irle, S.; Kranz, J. J.; Köhler, C.; Kowalczyk, T.; Kubař, T.; Lee, I. S.; Lutsker, V.; Maurer, R. J.; Min, S. K.; Mitchell, I.; Negre, C.; Niehaus, T. A.; Niklasson, A. M. N.; Page, A. J.; Pecchia, A.; Penazzi, G.; Persson, M. P.; Řezáč, J.; Sánchez, C. G.; Sternberg, M.; Stöhr, M.; Stuckenberg, F.; Tkatchenko, A.; Yu, V. W.-z.;

Frauenheim, T. DFTB+, a software package for efficient approximate density functional theory based atomistic simulations. *J. Chem. Phys.* **2020**, *152*, 124101.

(41) Laaziri, K.; Kycia, S.; Roorda, S.; Chicoine, M.; Robertson, J. L.; Wang, J.; Moss, S. C. High Resolution Radial Distribution Function of Pure Amorphous Silicon. *Phys. Rev. Lett.* **1999**, *82*, 3460–3463.