# Efficient Atomic-Resolution Uncertainty Estimation for Neural Network Potentials Using a Replica Ensemble

Wonseok Jeong, Dongsun Yoo, Kyuhyun Lee, Jisu Jung, and Seungwu Han*

ACCESS |   Metrics & More   |   Article Recommendations   |   SI Supporting Information

**ABSTRACT:** Neural network potentials (NNPs) are gaining much attention as they enable fast molecular dynamics (MD) simulations for a wide range of systems while maintaining the accuracy of density functional theory calculations. Since NNP is constructed by machine learning on training data, its prediction uncertainty increases drastically as atomic environments deviate from training points. Therefore, it is essential to monitor the uncertainty level during MD simulations to judge the soundness of the results. In this work, we propose an uncertainty estimator based on the replica ensemble in which NNPs are trained over atomic energies of a reference NNP that drives MD simulations. The replica ensemble is trained quickly, and its standard deviation provides atomic-resolution uncertainties. We apply this method to a highly reactive silicidation process of Si(001) overlaid with Ni thin films and confirm that the replica ensemble can spatially and temporally trace simulation errors at atomic resolution, which in turn guides the augmentation of the training set. The refined NNP completes a 3.6 ns simulation without any noticeable problems. By suggesting an efficient and atomic-resolution uncertainty indicator, this work will contribute to achieving reliable MD simulations by NNPs.
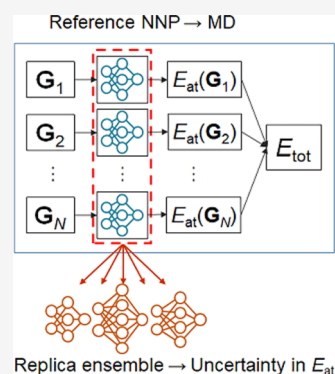
By replacing a knowledge-based formulation with data-driven parametrization in potential development, machine-learning (ML)-based interatomic potentials are extending the application range of the classical molecular dynamics (MD) simulation to multicomponent systems with complex bonding natures.[1] For example, ML potentials have been successfully applied to simulating phase-change materials,[2] catalysts,[3] nanoclusters,[4] and solid−liquid interfaces,[5] all of which are beyond the current scope of conventional classical or first-principles MD simulations. The ML potentials are trained over first-principles total energies and their derivatives (forces and stresses) of $10^2$−$10^4$ structures that are sampled from MD trajectories or crystalline structures with deformations or defects.

Machine learning is essentially an interpolative algorithm, so its prediction uncertainty grows rapidly as input features deviate from the training domain. In terms of ML potentials, the accuracy of the energy prediction degrades unacceptably if local atomic configurations are substantially different from those in the training set. If this happens during MD simulations, computational results may not be fully meaningful even if the simulation terminates without any drastic failures such as diverging energies. Therefore, it is critical to monitor uncertainty levels during MD simulations when utilizing ML potentials. Note that such problems are less acute with traditional classical potentials because principle-based, hard-coded functions can safeguard atomic configurations against unintended structures.

To date, several types of ML potentials have been applied to material research, in particular, the Gaussian-approximation potential (GAP),[6] the high-dimensional neural network potential (NNP),[7] and the kernel-ridge-regression (KRR)-based potential.[8] While GAP can automatically estimate the prediction uncertainty using posterior predictive variances,[9] no such formula exists with the NNP- or KRR-based potential. In the machine-learning community, the prediction uncertainty of neural network models is often assessed by employing a model ensemble that is formed by varying the training data or network structure.[10] The output variances within the ensemble are used as an indicator of uncertainty. The ensemble method was also applied to NNP.[11−13] In ref 11, an ensemble of NNPs was generated by bootstrapping training data, and it was found that the variance among ensemble NNPs increased when the simulation encountered configurations outside the training data. In refs 14−17, multiple NNPs with different network dimensions were generated with the same training data, and output discrepancies among NNPs were used as a reliability indicator or in selecting training points for adaptive learning. However, these ensemble methods require training multiple NNPs over first-principles total energies and atomic forces, which will impede the potential development in the case of big training data of structures. In addition, while the ensemble
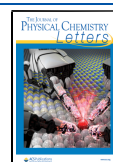
half-spread of total energies can signal instances when the simulation enters environments that were not learned, it is difficult to localize atoms with high uncertainties out of $10^3$–$10^5$ atoms in the simulation box, making it hard to refine NNPs by supplementing the training set. We note that the ensemble half-spreads of atomic energies in ref [11] could identify problematic configurations at the atomic scale. However, atomic energy mapping under the given total energies is not unique,[18] introducing a certain degree of randomness into atomic energies for each training of NNP in the ensemble. This should obscure the atomic-scale resolution of the reliability indicator.

In this Letter, we present an efficient and atomic-resolution uncertainty indicator for NNP that is based on a "replica" ensemble. The NNPs in the replica ensemble differ in the network structure and initial weights, and they are trained directly over atomic energies of the reference NNP that drives MD simulations. The training time for replica NNPs is much shorter than that for the reference NNP, and the standard deviation within the ensemble plays as the atomic-resolution uncertainty indicator during MD simulations. We apply this method in simulating silicidation processes of Ni contacts in semiconductor devices and demonstrate that the method can reveal spots with high uncertainty at the atomic scale. By analyzing the problematic region, we can improve the NNP with a repairing data set, thereby achieving a reliable simulation for an extended simulation time. The rest of the Letter is organized as follows. We first explain the notion of replica NNP and how it can be utilized during MD simulations. Next, we provide details of the model system and NNP training for the silicidation simulation. Then, a full cycle of NNP development for the Ni–Si interface reaction is demonstrated. Finally, we summarize and conclude.

In NNP, the total energy of a system ($E_{tot}$) is expressed as the sum of atomic energies ($E_{at}$) that depend on the symmetry function ($\mathbf{G}$) reflecting local environments[19]

$$E_{tot} = \sum_i E_{at}(\mathbf{G}_i)$$

(1)

where $i$ is the atom index and the atomic energy is given by a feed-forward neural network.[7] (For simplicity, we assume a unary system, but the discussions easily generalize to multicomponent systems.) We first obtain a reference NNP in a standard way by training it over first-principles results based on density functional theory (DFT). The loss function is defined with respect to DFT total energies and atomic forces.[20] This reference NNP is used in evaluating potential energy surfaces for MD simulations. Similar to the literature,[11−13] we employ an ensemble approach to gauge uncertainties in $E_{tot}$ and $E_{at}$. The main difference in the present approach is that NNPs in the ensemble (called replica NNPs hereafter) directly learn atomic energy $\{E_{at}(\mathbf{G}); \mathbf{G} \in \text{training set}\}$ output by the trained reference NNP (Figure 1a). To calibrate numeric scales in uncertainty, we control the initial range of randomized NN weights and also diversify the network size of replica NNPs. The standard deviation in the atomic energy ($\sigma_{at}(\mathbf{G})$) is calculated among replica NNPs and used as the prediction uncertainty at $\mathbf{G}$. During the MD simulation, $E_{at}$ of each atom is calculated by the reference NNP, and it is associated with the uncertainty value of $\sigma_{at}$ from replica NNPs (Figure 1b).

There are two benefits to employing such double-tier NNP: first, by training only over atomic energies, excluding
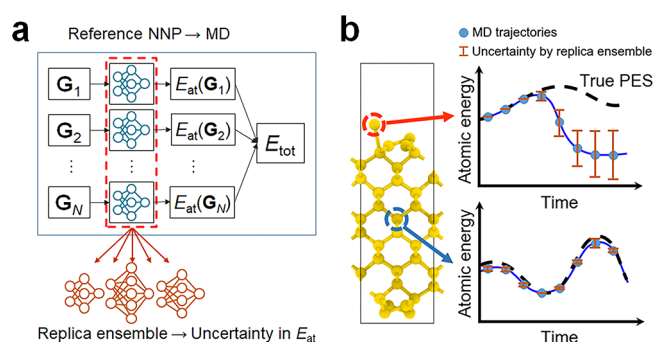


**Figure 1.** (a) Schematic description of reference NNP and the replica NNP ensemble. Reference NNP drives MD simulations, and the replica NNP ensemble measures the uncertainty level of atomic configurations by the standard deviation ($\sigma_{at}$) in atomic energies. (b) Schematic description of how $\sigma_{at}$ can be used as the uncertainty indicator of atomistic simulations with the example of a Si surface model. When the MD trajectory of a certain atom deviates from the true potential energy surface (PES) (e.g., a surface atom in the red dashed circle), the uncertainty in the replica ensemble becomes larger. In contrast, the uncertainty of atoms within normal MD trajectories (e.g., a bulk atom in the blue dashed circle) remains small.

derivatives such as forces and stresses, the computational cost of training replica NNPs becomes negligible compared to that for training the reference NNP over the whole set of properties. Second, replica NNPs by construction are designed to produce the same atomic energies for local configurations in the training set, but their inferences disagree for chemical environments outside the training set. This leads to atomic resolution in the uncertainty estimation by replica NNPs. As mentioned previously, in the ensemble NNPs that are trained over total energies,[11] such atomic resolution is not always guaranteed because of freedom in partitioning total energies into atomic contributions.

As an example of applying the replica ensemble as an uncertainty indicator, we choose the Ni-silicidation process, which is one of the key processes in the fabrication of semiconductor devices. The metal-silicidation process reduces the contact resistance at gate-source/drain contacts by lowering Schottky barriers. The process itself is rather simple; the metal layer is deposited on the Si substrate at room temperatures and annealed at 350−450 °C, which yields high-quality silicides with a controlled thickness.[21] Recently, the metal-silicidation process has received renewed attention in sub-10-nm device architectures because the mechanism of silicide formation changes when the deposited Ni thickness is below 4 nm.[22,23] Furthermore, additive elements such as Pt and Co are known to control the silicide phase and redistribute dopants in the Si channel. However, a microscopic understanding at the atomic scale is incomplete for the silicidation process, which may hinder further optimization of the process in highly scaled devices.

In ref [24], the formation of Ni silicide on the Si(001) substrate was studied by first-principles MD simulations, focusing on the role of Pt as a stabilizer of the NiSi phase. However, the sheer computational costs of DFT calculations severely limited the system size and simulation time. Furthermore, in order to accelerate interface reactions, the simulations were executed at an unrealistically high temperature of 2000 K, which may affect the simulation results significantly in view of the temperature dependence of the process.[22,25] These limitations of DFT calculations could be

**Table 1. Summary of Reference Structures and Root-Mean-Square Errors (RMSEs) for the Validation Set[a]**

| structure type | number of training points | temp (K) | RMSE$_{energy}$ (meV/atom) | RMSE$_{force}$ (eV/Å) |
|---|---|---|---|---|
| Ni | 3704 | 0, 500−1500 | 3.6 | 0.16 |
| Ni(001) | 8000 | 1000 | 2.1 | 0.16 |
| Si | 7408 | 0, 500−1500 | 3.0 | 0.16 |
| Si(001) | 15 360 | 1000 | 4.4 | 0.38 |
| $\delta$-Ni$_2$Si | 18 240 | 0, 1000−2000 | 3.9 | 0.28 |
| $l$-Ni$_2$Si | 24 000 | 2000 | 5.1 | 0.45 |
| $a$-Ni$_2$Si | 9600 | 800 | 3.8 | 0.28 |
| Ni$_2$Si ($l \rightarrow a$) | 14 400 | 2000−300 | 3.1 | 0.28 |
| NiSi | 12 160 | 0, 1000−2000 | 7.1 | 0.45 |
| $l$-NiSi | 16 000 | 2000 | 7.0 | 0.41 |
| $a$-NiSi | 6400 | 800 | 7.4 | 0.43 |
| NiSi ($l \rightarrow a$) | 9600 | 2000−300 | 6.8 | 0.41 |
| $\alpha$-NiSi$_2$ | 18 240 | 0, 1000−2000 | 4.8 | 0.33 |
| $l$-NiSi$_2$ | 24 000 | 2000 | 5.2 | 0.30 |
| $a$-NiSi$_2$ | 9,600 | 800 | 8.3 | 0.50 |
| NiSi$_2$ ($l \rightarrow a$) | 14 400 | 2000−300 | 5.5 | 0.49 |
| interface (1) | 108 410 | 300, 1000, 1300, 1500 | 4.8 | 0.40 |
| interface (2) | 92 160 | 1000, 1300 | 2.5 | 0.30 |
| interface (3) | 46 400 | 1000 | 3.0 | 0.28 |
| total | 458 082 | | 5.1 | 0.34 |

[a]The first column describes the structure type of training data. The second column is the number of training points (the number of atoms) corresponding to the structure type. The third and fourth columns show RMSEs for the energy and force, respectively. $l$ and $a$ indicate liquid and amorphous phases, respectively.

mitigated by employing NNP. However, the development of NNP aiming at the silicidation process is challenging because the system concurrently involves salient covalent (Si) and metallic (Ni) bonds, along with complicated mixed bonding (Ni$_x$Si$_y$). Furthermore, the bonding nature constantly changes during the reactive interdiffusion of Ni and Si atoms, posing challenges in selecting a training set that can properly sample the configurational space explored during MD simulations. The complexity of the training set also necessitates monitoring MD simulations using an uncertainty indicator. Below, a full cycle of NNP development for the Ni−Si interface reaction will be discussed.

Table 1 summarizes the training set for the reference NNP. Besides static distortions of crystals, the training set includes DFT MD snapshots of crystals, liquids, amorphous structures, and surface slabs. (Details of DFT calculations are provided in the Supporting Information.) Crystal structures are annealed at relatively high temperatures of up to 2000 K because the highly exothermic reaction can significantly increase local temperatures at the interface. In addition, we consider MD simulations of Ni−Si interfaces at various temperatures (300−1500 K) in order to train NNP over interface reactions directly. The training set contains 4944 structures with 198 694 and 259 388 training points for Ni and Si, respectively. Full details and atomistic models of the training data can be found in Table S1 and Figures S1−S3 of the Supporting Information.

In training NNPs and performing MD simulations, we use the SIMPLE-NN package (https://github.com/MDIL-SNU/SIMPLE-NN)[20,26] and the LAMMPS package.[27] (A more detailed description of the training procedure and training RMSE can be found in the Supporting Information.) After training, we compare NNP and DFT on key properties such as the equation of states for crystals, density and energy of amorphous structures, and radial distribution function of liquid and amorphous structures. The results are compiled in the Supporting Information, and all of the properties agree well

between NNP and DFT. As a more direct test, we also perform Ni silicidation for a relatively small Ni−Si interface structure consisting of 60 Ni atoms and 120 Si atoms (top-left figure in Figure 2). This is the same structure as interface (2) in Table 1, but the test simulation starts with different initial velocities and lasts for 20 ps, much longer than 5 ps in the training set.
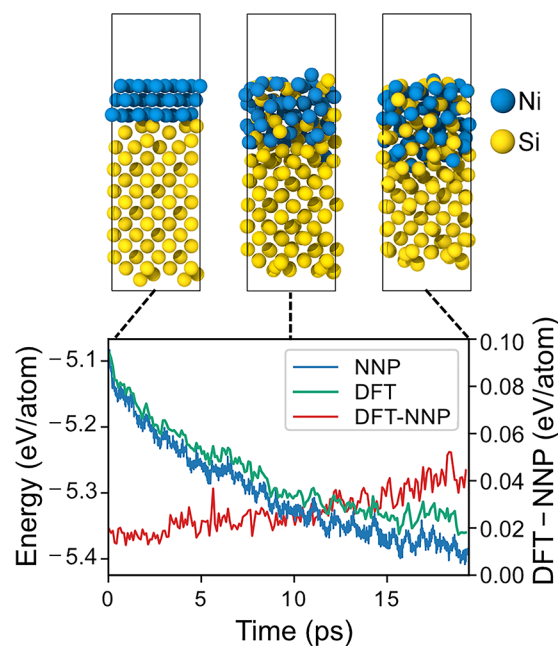


**Figure 2.** Comparison of the DFT and NNP total energies per atom along MD trajectories for a small Ni−Si interface structure with 60 Ni atoms and 120 Si atoms. The blue line represents the energy from 1000 K NNP-MD with a time step of 1 fs. The green line represents the DFT energy of the NNP-MD trajectory sampled with a 50 fs interval. The red line with the right-hand $y$ scale represents differences between DFT and NNP energies.

The blue line in Figure 2 shows total energies by NNP with respect to the simulation time, and the top figures are structures at selected instances. For comparison, DFT energies are calculated every 50 fs along the NNP-MD trajectory (the green line). Overall agreement between NNP and DFT is good, and fine features in the total energy are well reproduced by NNP. However, the energy difference linearly increases with time (red line). This is attributed to a small discrepancy in the reaction energy (Ni + Si → NiSi) between DFT and NNP.

We train a replica ensemble which consists of five NNPs that are independently trained over atomic energies of the reference NNP at $\mathbf{G}$ values in the training set. To obtain output variances that are large enough for uncertainty quantification, NNPs in the ensemble should be decorrelated. This can be achieved by diversifying model parameters and training conditions such as initial weights, training/validation set selection from a data set (for example, bootstrap aggregating), network design such as the layer size and depth, and training hyperparameters. Among these methods, we find that varying the weight parameters and network size achieve a sufficient range of prediction variation. In detail, the initial weights are chosen from a normal Gaussian distribution with a zero mean and a standard deviation of $\sigma_w$. When $\sigma_w$ is too small, the replica NNPs tend to converge and energy variances of the ensemble become much smaller than actual errors relative to DFT results. On the other hand, if $\sigma_w$ is too large, then NNPs easily fall into local minima, degrading training. To strike a balance, we choose $\sigma_w = 3.0$; when $\sigma_w$ exceeds this value, some replica NNPs produce RMSEs larger than 100 meV, which is a signature of poor training quality. In addition to tuning $\sigma_w$, we also diversify the network structure (70−30−30−1, 70−50−50−1, 70−70−70−1, 70−100−100−1, and 70−120−120−1) and select a random 70% of atomic energies by the reference NNP as a training set, independently for each replica NNP. In training replica NNPs, we use the same hyperparameters as for the reference NNP. Figure S7 in the Supporting Information shows the learning curves of the replica NNPs, and it is seen that resulting RMSEs range from 30 to 70 meV (RMSE of the ensemble average is 16 meV). Notably, the standard deviation among replica NNPs ($\sigma_{at}(\mathbf{G})$) is 50 meV for $\mathbf{G}$ in the training set, confirming that the prediction uncertainty is low for the learned local configurations. Note that RMSEs of replica NNPs are an order of magnitude larger than RMSEs for typical NNPs, which are usually a few meV/atom. This is because direct learning of the atomic energy is far more constrained than learning total energies that are sums of atomic energies. In addition, $\sigma_w$ is larger than for typical training, causing the training result to fall into a local minimum.

For a large-scale simulation of the silicidation process using NNP, we construct a Ni/Si(001) interface slab of ~8.5 nm thickness as shown in Figure 3, which includes 3008 Si and 2080 Ni atoms. The Ni layer is 2 nm thick in the simulation, while Ni films with a thickness of 1−10 nm were deposited in experiments. The lattice parameters in the $xy$ plane are fixed to those of crystalline Si at theoretical equilibrium. In Figure 3, the bottom two layers are fixed and the top three layers are heated to 1000 K under the NVT condition controlled by the Nosé−Hoover thermostat.[28] The rest of the atoms are simulated under the NVE condition. This allows for simulating the transport of heat energies that are generated during interfacial reactions and dissipated into the bottom layer through thermal conduction.
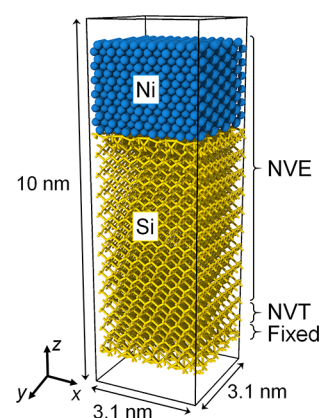


**Figure 3.** Model structure of an 8.5-nm-thick Ni/Si(001) interface slab. The two bottom layers are fixed during the MD simulations. The next three layers are heated to 1000 K using the NVT ensemble. Other parts are simulated using the NVE ensemble.

In the experiment, Ni is deposited at room temperature. When we carry out the simulation at this temperature, only a few Ni and Si atoms react at the interface. Therefore, we start the high-temperature annealing right after the initial relaxation of the structure. Furthermore, at experimental annealing temperatures of 350−450 °C, the reaction almost stops after four Ni layers (~0.7 nm) are silicided. In order to accelerate the dynamics and observe meaningful reactions within the simulation time, we increase the temperature to 726 °C or 1000 K.

Figure 4 displays characteristic instances during the MD simulation that is carried out for up to 1 ns. In a very early
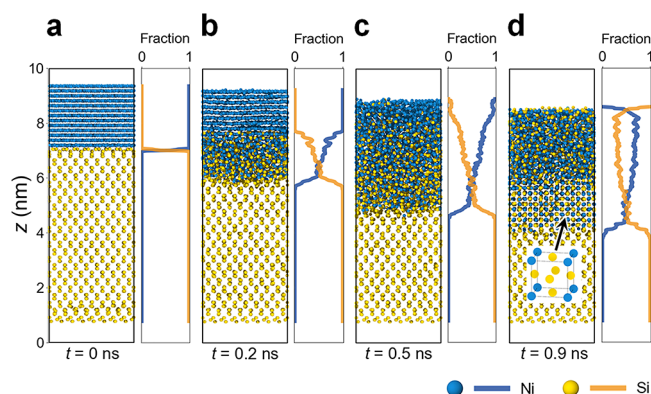


**Figure 4.** Snapshots and concentration profiles of Ni and Si with respect to $z$ from the large-scale NNP MD trajectory. (a) 0, (b) 0.2, (c) 0.5, and (d) 0.9 ns. The inset in (d) indicates that the L1$_2$ phase formed at the interface.

stage ($t < 0.2$ ns), vigorous interfacial reactions are observed, which originate from the highly exothermic reaction of Ni + Si → NiSi (1.0 eV per formula unit). As a result, amorphous Ni$_x$Si$_y$ layers grow up to a thickness of ~2 nm (Figure 4b). For $0.2 < t < 0.5$ ns, the concentration gradient of Ni decreases along the $z$ axis and so does the driving force for reactive diffusion. Consequently, the silicidation process gradually slows. In Figure 4c, the concentration profile within the Ni$_x$Si$_y$ layer is approximately linear with the atomic fraction of Si varying continuously from 0.55 to 0.3, implying the absence of phase separation along the $z$ direction. At around 0.6 ns, the L1$_2$ phase of NiSi nucleates and grows from the interface as

shown in Figure 4d. Once the crystalline phase with a thickness of ∼2 nm forms at the interface, the diffusion is virtually stopped and no further interdiffusion is observed until 1 ns. As a result, the system arrives at a steady-state distribution comprising three distinct layers ($a$-$Ni_2Si$/$L1_2$-NiSi/Si). (The NNP MD movie is provided in the Supporting Information.)

Next, we check the soundness of the simulation using the replica ensemble. Figure 5a shows atom-by-atom coloring of
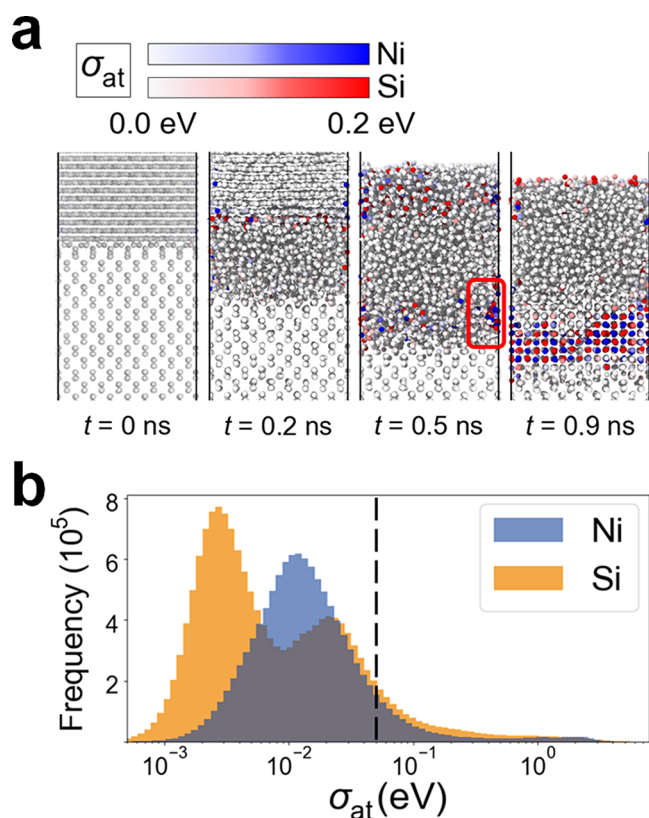


**Figure 5.** (a) Snapshots from the Ni-silicidation MD trajectory at 0, 0.2, 0.5, and 0.9 ns with the color scale by $\sigma_{at}$. (Parts of the bottom Si layers are removed.) (b) Histogram of $\sigma_{at}$ collected over the 1 ns simulation.

$\sigma_{at}$ for instances selected in Figure 4. The frequency histogram of $\sigma_{at}$ for the whole simulation is shown in Figure 5b, and it is seen that $\sigma_{at}$ is mostly less than 50 meV, which is the average $\sigma_{at}$ for training points (methods discussion) and an uncertainty threshold. As such, MD simulations appear to be reasonable for most configurations as confirmed by dominantly light colors in Figure 5a. Nevertheless, the uncertainty indicator also detects concerning values of $\sigma_{at} \gtrsim 0.1$ eV for a number of regions near Ni/NiSi and NiSi/Si interfaces as indicated by blue (Ni) and red (Si) atoms. Most notably, $\sigma_{at}$ is largest for the $L1_2$ phase formed at the interface. In fact, the $L1_2$ phase was not considered in the training set, and DFT calculations show that the $L1_2$-NiSi phase is unstable against the most stable MnP-type NiSi by 0.21 eV/atom. However, the present NNP estimates that $L1_2$ is more stable than the MnP type by 0.08 eV/atom. This means that the nucleation of the $L1_2$ phase is an artifact triggered by a subtle deviation from the training set.

The atomic resolution of $\sigma_{at}$ allows for localizing simulation defects temporally as well as spatially. At ∼0.5 ns, a cluster of atoms near the NiSi/Si interface begins to develop high

uncertainties (a round box in Figure 5a) and later nucleates into the $L1_2$ phase. In addition, some atoms near the NiSi/Ni interface and Si atoms at the Ni surface exhibit notable uncertainties. These errors should follow from deficiencies of the interface models in the training set. By close inspection, we find that the $Ni_xSi_y$ layer in the training set is rather Ni-rich compared to the boxed region in Figure 5a. This is because the simulation time is too short to permit sufficient interdiffusion. Furthermore, there are only three Ni layers in interface (1) and (2) models, which is too thin to sample Ni/NiSi interfaces properly.

Inspired by the above analysis, we augment the training set with additional structures from DFT calculations. We first add MD trajectories of a model with five Ni layers (similar to interface (3)) at a higher temperature of 1300 K and a longer simulation time of 40 ps, thereby sampling interfacial compositions similar to NNP-MD simulations. Furthermore, we add annealing trajectories at 1000−2000 K for crystals such as $\gamma$-$Ni_{31}Si_{12}$, $\beta_1$-$Ni_3Si$, $\beta_2$-$Ni_3Si$, $\varepsilon$-$Ni_3Si_2$, $\theta$-$Ni_2Si$, and NiSi ($\varepsilon$-FeSi type) to calibrate the formation energies of $Ni_xSi_y$ mixed layers with fine stoichiometric variations. We also consider a surface model of amorphous NiSi because Si atoms with high uncertainties are noticeable in the surface region ($t = 0.9$ ns in Figure 5a). These supplementary structures add 258 030 points to the final training set, which is then used to train a refined NNP (r-NNP). The RMSEs of r-NNP are 5.62 meV/atom (energy) and 0.23 eV/Å (force) for the validation set. As shown in Figures S8−S11 and Table S4, the computational results on test structures of crystals, liquids, and amorphous phases are similar between r-NNP and the original NNP. However, Figure S11 shows that the interface reaction is described more precisely by r-NNP in comparison with Figure 2, implying that r-NNP might be more accurate for interfacial reactions.

Figure 6a−d shows the sequence of the silicidation process simulated by r-NNP up to 0.9 ns (total simulation time 3.6 ns). The model and simulation conditions are the same as in the previous simulation, and the prediction uncertainties by $\sigma_{at}$ in Figure 6e,f are evaluated similarly to Figure 5. In contrast with Figure 5b, almost all of the atoms exhibit $\sigma_{at} < 0.1$ eV throughout the whole simulation except for brief fluctuations above the threshold, indicating that the supplementary training structures enabled reliable and stable simulations. Overall, the simulation proceeds similarly to Figure 5 up to ∼0.5 ns, namely, vigorous initial reactions until 0.2 ns followed by sluggish interdiffusion. The amorphous $Ni_xSi_y$ layer is thicker than in the previous simulation in Figure 4d by ∼1.5 nm. Unlike the previous simulation, the incubation of any crystal phase is absent with r-NNP, which contributes to the formation of thick interfacial layers (the dashed box in Figure 6c) rather than the abrupt interface in Figure 4c. After 0.6 ns, the thickness of the $Ni_xSi_y$ layer remains nearly constant, and atoms diffuse mainly within the amorphous layer. Further simulation up to 3.6 ns changes only the stoichiometry within the amorphous layer close to 1:1 (not shown). In experiments, epitaxial $NiSi_{2-\delta}$ ($\delta = 0.5$−0.6) phases grow at high annealing temperatures.[29] Energy differences between crystal and amorphous phases are similar between r-NNP and DFT, so amorphous NiSi will eventually crystallize in the simulation. However, the incubation time for crystallization is much longer than the feasible simulation time,[2] so we could not observe the formation of crystalline phases within 3.6 ns. (The r-NNP MD movie is provided in the Supporting Information.)
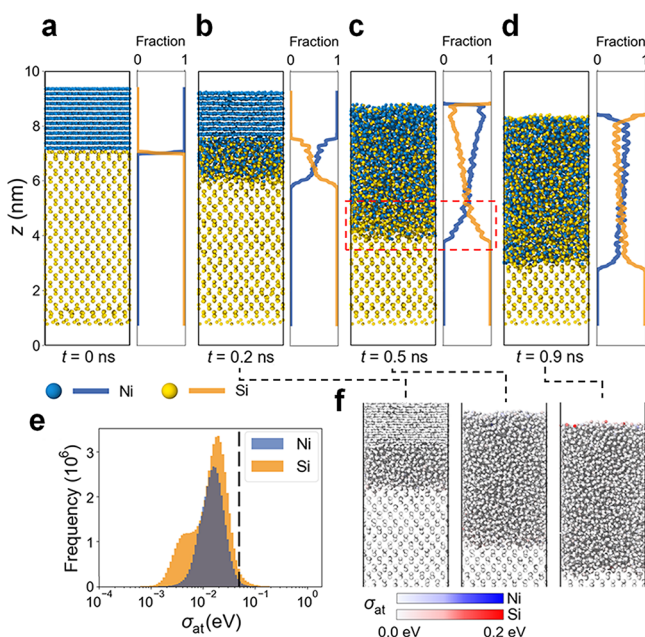
**Figure 6.** Snapshots and concentration profiles of Ni and Si with respect to $z$ from the large-scale r-NNP MD trajectory at (a) 0, (b) 0.2, (c) 0.5, and (d) 0.9 ns. (e) Histogram of $\sigma_{at}$ of the 3.6 ns r-NNP simulation trajectory atoms. (f) Snapshots from the large-scale Ni silicidation r-NNP MD trajectory at 0.2, 0.5, and 0.9 ns. The color intensity is scaled by $\sigma_{at}$. (Part of the bottom Si layers are removed.)

For further analysis of the electronic property, we theoretically estimate the Schottky barrier height (SBH, $\Phi_b$) for the as-obtained $Ni_xSi_y$/Si interface. The $Ni_xSi_y$/Si interface structure is obtained from the r-NNP MD simulation of the model structure after 3.6 ns of 1000 K MD. Before the DFT calculation, atomic positions are relaxed with r-NNP at 0 K. Then, the interface region of the whole structure is sliced to reduce the number of atoms to below 1000. Dangling bonds of the bulk-terminated Si surface are passivated with H atoms. As a result, the interface structure consists of seven layers of Si and a 6.2-Å-thick $Ni_xSi_y$ layer (Figure S12a in the Supporting Information). To address the finite-size effect, the crystalline Si is also computed independently. The SBHs for n- and p-type Si are obtained as follows

$$\Phi_{b,n} = E_C - E_F - (\langle V \rangle_{Si}^{bulk} - \langle V \rangle_{Si}^{interface}) \quad (2)$$

$$\Phi_{b,p} = E_F - E_V + (\langle V \rangle_{Si}^{bulk} - \langle V \rangle_{Si}^{interface}) \quad (3)$$

where $E_C$ and $E_V$ denote the conduction and valence band edges obtained from the bulk Si structure and $E_F$ indicates the Fermi energy obtained from the $Ni_xSi_y$/Si interface structure. In eqs 2 and 3, $\langle V \rangle_{Si}^{interface}$ refers to the average electrostatic potential of the bulk Si part in the $Ni_xSi_y$/Si interface system, while $\langle V \rangle_{Si}^{bulk}$ is the corresponding value for the crystalline Si.[30] In addition, to compensate for the band gap underestimation in semilocal functionals, the band edge energies are corrected with the hybrid functional method (HSE06).[31] As a result, the interface model gives SBH values of 0.83 and 0.21 eV for $\Phi_{b,n}$ and $\Phi_{b,p}$, respectively, which is in reasonable agreement with experimental values of 0.7 and 0.35 eV for NiSi/Si interfaces.[32,33]

In summary, we proposed the replica ensemble that enables efficient and atomic-resolution uncertainty estimation in energy prediction by NNPs. The replica NNPs are trained

directly over atomic energies of the reference NNP that calculates potential energy surface during MD simulations. By excluding energy derivatives from the loss function, the construction of the replica ensemble takes much less time than for conventional NNPs. The usefulness of the suggested indicator was demonstrated by simulating the Ni silicidation process. The replica ensemble was able to trace highly uncertain atomic configurations at interfaces during the MD simulation, which later evolved into the formation of the unphysical $L1_2$ phase. The uncertainty analysis revealed structures that were deficient in the training set. With the augmented training set, a refined NNP was developed and reliable Ni silicidation simulation was carried out until 3.6 ns. We have implemented the replica-ensemble method into the SIMPLE-NN package, so it can be used by practitioners without significant effort. In conclusion, by providing an efficient atomic-resolution uncertainty indicator, this work will contribute to achieve reliable MD simulations by NNPs.

## ■ ASSOCIATED CONTENT

**Ⓢ Supporting Information**

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.0c01614.

> Details of the DFT calculation; details of NNP training; NNP validation with material properties; and r-NNP validation with material properties (PDF)

> MD movie of Ni silicidation using NNP with uncertainty estimation (MP4)

> MD movie of Ni silicidation using r-NNP with uncertainty estimation (MP4)

## ■ AUTHOR INFORMATION

**Corresponding Author**

**Seungwu Han** − *Department of Materials Science and Engineering, Seoul National University, Seoul 08826, Korea;* ⓘ orcid.org/0000-0003-3958-0922; Email: hansw@snu.ac.kr

**Authors**

**Wonseok Jeong** − *Department of Materials Science and Engineering, Seoul National University, Seoul 08826, Korea;* ⓘ orcid.org/0000-0001-8894-1857

**Dongsun Yoo** − *Department of Materials Science and Engineering, Seoul National University, Seoul 08826, Korea;* ⓘ orcid.org/0000-0002-4889-8396

**Kyuhyun Lee** − *Department of Materials Science and Engineering, Seoul National University, Seoul 08826, Korea*

**Jisu Jung** − *Department of Materials Science and Engineering, Seoul National University, Seoul 08826, Korea*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jpclett.0c01614

**Notes**

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

## ■ REFERENCES

(1) Schleder, G. R.; Padilha, A. C. M.; Reily Rocha, A.; Dalpian, G. M.; Fazzio, A. Ab initio Simulations and Materials Chemistry in the Age of Big Data. *J. Chem. Inf. Model.* **2020**, *60*, 452−459.

(2) Lee, D.; Lee, K.; Yoo, D.; Jeong, W.; Han, S. Crystallization of Amorphous GeTe Simulated by Neural Network Potential Addressing Medium-range Order. *Comput. Mater. Sci.* **2020**, *181*, 109725.

(3) Ulissi, Z. W.; Tang, M. T.; Xiao, J.; Liu, X.; Torell, T. L.; et al. Machine-learning Methods Enable Exhaustive Searches for Active Bimetallic Facets and Reveal Active Site Motifs for $CO_2$ Reduction. *ACS Catal.* **2017**, *7*, 6600−6608.

(4) Sun, G.; Sautet, P. Metastable Structures in Cluster Catalysis from First-principles: Structural Ensemble in Reaction Conditions and Metastability Triggered Reactivity. *J. Am. Chem. Soc.* **2018**, *140*, 2812−2820.

(5) Quaranta, V.; Hellström, M.; Behler, J. Proton-transfer Mechanisms at the Water−ZnO Interface: The Role of Presolvation. *J. Phys. Chem. Lett.* **2017**, *8*, 1476−1483.

(6) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. *Phys. Rev. Lett.* **2010**, *104*, 136403.

(7) Behler, J.; Parrinello, M. Generalized Neural-network Representation of Hgh-dimensional Potential-energy Surfaces. *Phys. Rev. Lett.* **2007**, *98*, 146401.

(8) Huan, T. D.; Batra, R.; Chapman, J.; Krishnan, S.; Chen, L.; Ramprasad, R. A Universal Strategy for the Creation of Machine Learning-Based Atomistic Force Fields. *npj Comput. Mater.* **2017**, *3*, 1−8.

(9) Vandermause, J.; Torrisi, S. B.; Batzner, S.; Xie, Y.; Sun, L.; Kolpak, A. M.; Kozinsky, B. On-the-fly Active Learning of Interpretable Bayesian Force Fields for Atomistic Rare Events. *npj Comput. Mater.* **2020**, *6*, 1−11.

(10) Dietterich, T. G. Ensemble Methods in Machine Learning. *Lect. Notes Comput. Sci.* **2000**, *1857*, 1−15.

(11) Peterson, A. A.; Christensen, R.; Khorshidi, A. Addressing Uncertainty in Atomistic Machine Learning. *Phys. Chem. Chem. Phys.* **2017**, *19*, 10978−10985.

(12) Gastegger, M.; Behler, J.; Marquetand, P. Machine Learning Molecular Dynamics for the Simulation of Infrared Spectra. *Chem. Sci.* **2017**, *8*, 6924−6935.

(13) Smith, J. S.; Nebgen, B.; Lubbers, N.; Isayev, O.; Roitberg, A. E. Less is More: Sampling Chemical Space with Active Learning. *J. Chem. Phys.* **2018**, *148*, 241733.

(14) Behler, J. Representing Potential Energy Surfaces by High-dimensional Neural Network Potentials. *J. Phys.: Condens. Matter* **2014**, *26*, 183001.

(15) Li, W.; Ando, Y. Dependence of a Cooling Rate on Structural and Vibrational Properties of Amorphous Silicon: A Neural Network Potential-based Molecular Dynamics study. *J. Chem. Phys.* **2019**, *151*, 114101.

(16) Shen, L.; Yang, W. Molecular Dynamics Simulations with Quantum Mechanics/Molecular Mechanics and Adaptive Neural Networks. *J. Chem. Theory Comput.* **2018**, *14*, 1442−1455.

(17) Schran, C.; Behler, J.; Marx, D. Automated Fitting of Neural Network Potentials at Coupled Cluster Accuracy: Protonated Water Clusters as Testing ground. *J. Chem. Theory Comput.* **2020**, *16*, 88−99.

(18) Yoo, D.; Lee, K.; Jeong, W.; Lee, D.; Watanabe, S.; Han, S. Atomic Energy Mapping of Neural Network Potential. *Phys. Rev. Mater.* **2019**, *3*, No. 093802.

(19) Behler, J. Atom-centered Symmetry Functions for Constructing High-dimensional Neural Network Potentials. *J. Chem. Phys.* **2011**, *134*, No. 074106.

(20) Lee, K.; Yoo, D.; Jeong, W.; Han, S. SIMPLE-NN: An Efficient Package for Training and Executing Neural-network Interatomic Potentials. *Comput. Phys. Commun.* **2019**, *242*, 95−103.

(21) Lavoie, C.; d'Heurle, F. M.; Detavernier, C.; Cabral, C. Towards Implementation of a Nickel Silicide Process for CMOS Technologies. *Microelectron. Eng.* **2003**, *70*, 144−157.

(22) Zhang, Z.; Zhang, S.; Yang, B.; Zhu, Y.; Rossnagel, S. M.; Gaudet, S.; et al. Morphological Stability and Specific Resistivity of Sub-10 nm Silicide Films of $Ni_{1−x}Pt_x$ on Si Substrate. *Appl. Phys. Lett.* **2010**, *96*, No. 071915.

(23) Geenen, F. A.; van Stiphout, K.; Nanakoudis, A.; Bals, S.; Vantomme, A.; Jordan-Sweet, J.; et al. Controlling the Formation and Stability of Ultra-thin Nickel Silicides - An Alloying Strategy for Preventing Agglomeration. *J. Appl. Phys.* **2018**, *123*, No. 075303.

(24) Christensen, M.; Eyert, V.; Freeman, C.; Wimmer, E.; Jain, A.; Blatchford, J.; et al. Formation of Nickel-Platinum Silicides on a Silicon Substrate: Structure, Phase Stability, and Diffusion from Ab Initio Computations. *J. Appl. Phys.* **2013**, *114*, No. 033533.

(25) Foggiato, J.; Yoo, W. S.; Ouaknine, M.; Murakami, T.; Fukada, T. Optimizing the Formation of Nickel Silicide. *Mater. Sci. Eng., B* **2004**, *114*, 56−60.

(26) Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; et al.TensorFlow: Large-scale Machine Learning on Heterogeneous Systems. http://tensorflow.org/, 2015.

(27) Plimpton, S. Fast Parallel Algorithms for Short-range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1−19.

(28) Nosé, S. A Unified Formulation of the Constant Temperature Molecular Dynamics Methods. *J. Chem. Phys.* **1984**, *81*, 511−519.

(29) Luo, J.; Qiu, Z.; Zha, C.; Zhang, Z.; Wu, D.; Lu, J.; et al. Surface-energy Triggered Phase Formation and Epitaxy in Nanometer-thick Silicide Films. *Appl. Phys. Lett.* **2010**, *96*, No. 031911.

(30) Shan, B.; Cho, K. Ab Initio Study of Schottky Barriers at Metal-nanotube Contacts. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2004**, *70*, 233405.

(31) Park, S.; Lee, B.; Jeon, S. H.; Han, S. Hybrid Functional Study on Structural and Electronic Properties of Oxides. *Curr. Appl. Phys.* **2011**, *11*, S337−S340.

(32) Coe, D. J.; Rhoderick, E. H. Silicide Formation in Ni-Si Schottky Barrier Diodes. *J. Phys. D: Appl. Phys.* **1976**, *9*, 965−972.

(33) Alptekin, E.; Ozturk, M. C. Tuning of the Nickel Silicide Schottky Barrier Height on p-Type Silicon by Indium Implantation. *IEEE Electron Device Lett.* **2009**, *30*, 1272−1274.