Mechanism of local lattice distortion effects on vacancy migration barriers in fcc alloys

Zhucong Xi, Mingfei Zhang, Louis G. Hector, Jr., Amit Misra, and Liang Qi

Phys. Rev. Materials **6**, 073601 — Published 13 July 2022

DOI: 10.1103/PhysRevMaterials.6.073601

Mechanism of Local Lattice Distortion Effects on Vacancy Migration Barriers in FCC Alloys

Zhucong Xi,* Mingfei Zhang,* Louis G. Hector Jr.,† Amit Misra,* and Liang Qi‡

(Dated: June 21, 2022)

Accurate prediction of vacancy migration energy barriers, ΔE_a, in multi-component alloys is extremely challenging yet critical for the development of diffusional transformation kinetics needed to model alloy behavior in many technological applications. In this paper, results from ΔE_a and the energy driving force ΔE of many (>1000) vacancy migration events calculated using density functional theory and nudged elastic band method show large changes ($\sim$ 1 eV) of ΔE_a in different local chemical environments of the model face-centered cubic (FCC) Al-Mg-Zn alloys. Due to local lattice distortion effects induced by solute atoms (such as Mg) with different sizes than the matrix element (Al), the changes of ΔE_a for one type of migrating atoms originate primarily from fluctuations of $\Delta e_a \equiv \Delta E_a - \frac{1}{2}\Delta E$ (instead of $\frac{1}{2}\Delta E$ according to the widely used Kinetic Ising model). To understand these fluctuations, a quartic function of the reaction coordinate is shown to accurately describe the energy landscape of the minimum energy path (MEP) for each vacancy migration event studied in this paper. Analyses of the quartic function show that Δe_a can be approximated with $\Delta e_a \approx \alpha k_f D^2$, where $\alpha \sim 0.022$ is a constant value of all types of migrating atoms in the Al lattice. Here D is the distance of a migrating atom between two adjacent equilibrium positions and k_f is the average vibration spring constant of this atom at these two equilibrium positions. k_f and D quantitatively describe the lattice distortion effects on the curvatures and locations of the MEP at its initial and final states in different local chemical environments. We also used the local lattice occupations as inputs to train surrogate models to predict the coefficients of the quartic function, which accurately and efficiently output both ΔE_a and ΔE as the necessary inputs for the mesoscale studies of diffusional transformation in Al-Mg-Zn alloys.

I. INTRODUCTION

Diffusion kinetics in metallic alloys and associated material mechanisms (e.g., aging), which control properties such as strength and ductility, are critically dependent upon vacancy-mediated migration of matrix atoms and substitutional solutes[1, 2]. A migrating species in an alloy encounters complex and varying local chemical environments, especially in multicomponent alloys, which in turn change the energy barrier ΔE_a of a vacancy migration event between two adjacent lattice sites[3–7]. Accurate descriptions of such local chemical effects on ΔE_a are necessary to construct the kinetic master equations in mesoscale methods, such as kinetic Monte Carlo (kMC)[8–11], phase-field crystal (PFC)[12, 13], and diffusive molecular dynamics (DMD) simulations[14], to study diffusion and precipitation. However, potentially large variations of local chemical environments present significant challenges which have yet to be overcome.

A typical strategy to predict ΔE_a in different local chemical environments is the Kinetic Ising Model detailed by two vacancy migration events with the same migrating atom in Fig. 1 (a) and (b)[10] (Path 1 and 2). Path 1 occurs in a dilute local environment with zero energetic driving force ΔE and its ΔE_a is easily obtainable using first-principles calculations[15, 16]. Path 2 represents a general vacancy migration case with non-zero ΔE. Fig. 1 (b) describes the energy landscape of the minimum energy path (MEP) for each of these two events based on two assumptions. First, the MEP curves are approximated as linear functions of the reaction coordinate with almost the same slope ($\theta_1 = \theta_2$) away from both the initial and final states; second, the only changes from Path 1 to 2 are that the MEP curves near the final state shift rigidly along the energy coordinate (Y-axis) by ΔE of Path 2. Therefore, it is easy to demonstrate that ΔE_a of Path 2 is equal to one-half of its ΔE plus ΔE_a of Path 1. In practice, ΔE of Path 2 can be predicted by the bond counting model[17] or cluster expansion (CE) methods[18, 19], which use the local lattice occupations as inputs with parameters fitted based on first-principles calculations. This strategy to predict ΔE_a as a linear function of ΔE for a general vacancy migration event was used to model many metallic alloys[9–11, 17, 20, 21].

The above two assumptions on MEPs can be incorrect. Fig. 1 (c) illustrates the detailed MEP plot of Path 2 without the two assumptions: the distance between the initial and final states along the reaction coordinate, defined as D_{MEP} in Fig. 1 (c), can vary due to the lattice distortion induced by changes of local chemical compositions; in addition, the MEP curves near the initial states can have different shapes (such as local curvatures) compared with those at the final states. These variations change the position of the transition state along the reaction coordinate and its energy. Thus, a robust model of ΔE_a should provide accurate descriptions of MEPs and

* Department of Materials Science and Engineering, University of Michigan, Ann Arbor, Michigan, 48109, USA
† GM Global Technical Center, General Motors Company, Warren, Michigan, 48092, USA
‡ Department of Materials Science and Engineering, University of Michigan, Ann Arbor, Michigan, 48109, USA; qiliang@umich.edu

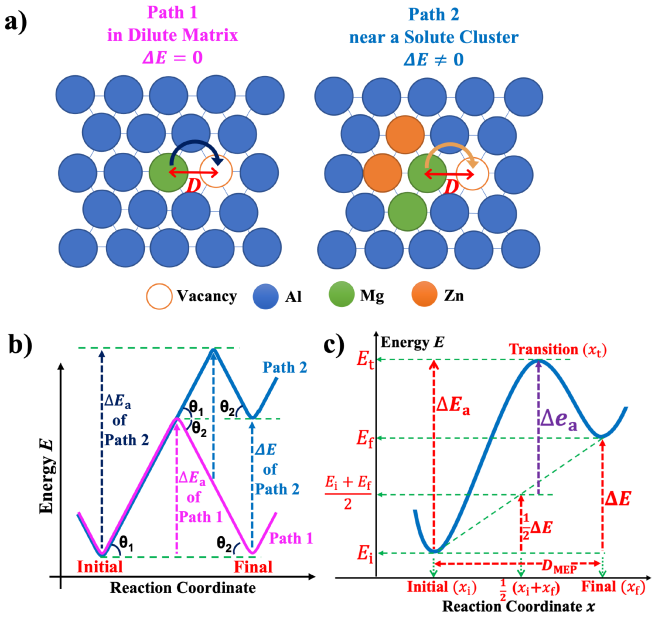

FIG. 1. Models of energy barrier ΔE_a and driving force ΔE of vacancy migrations in Al alloys. (a): Two vacancy (open circle) migration events between adjacent lattice sites in Al alloys. Colored solid circles indicate the chemical elements on occupied sites. (b): Simplified energy landscape plots of the minimum energy paths (MEPs) for two events in (a) based on the two assumptions described in the text. (c): An energy landscape plot of the MEP for Path 2 in (b) without the two assumptions.

the corresponding transition states[8, 22–24]. One strategy is to investigate Δe_a defined as the transition-state energy (E_t) relative to the average of the initial-state (E_i) and final-state (E_f) energies:

$$\Delta e_a \equiv E_t - \frac{1}{2}(E_i + E_f) = \Delta E_a - \frac{1}{2}\Delta E \qquad (1)$$

Δe_a is a variable and a function of local lattice occupations. This function of Δe_a can be fit using a local cluster expansion method[19, 25, 26]. ΔE_a is then obtained by the summation of Δe_a and $\frac{1}{2}\Delta E$. Note that the Kinetic Ising Model is recovered if Δe_a is a fixed value as ΔE_a in Path 1 of Fig. 1 (a). This method requires sufficient samples of transition states to construct the training data set for fitting Δe_a. The quantitative understanding of the mechanisms that determine Δe_a and ΔE_a can benefit the selections of the representative vacancy migration cases for fitting and verifying the functions of Δe_a and ΔE_a in different local chemical environments, which are critical for the investigations of diffusion kinetics in multiple precipitation stages of advanced alloys.

To clarify the mechanisms that determine ΔE_a and Δe_a, we applied high-throughput first-principles calculations to study vacancy migrations in model Al-Mg-Zn systems. As the 7XXX series of aerospace grade Al alloys, they achieve high strengths (∼700 MPa) after appropriate heat treatments[27], but applications outside of aerospace are limited since solute clustering during natural aging limits formability[28–31]. This issue exists in several types of Al alloys, and it can be mitigated if vacancy-mediated diffusion can be understood and manipulated [2, 32–34] since this controls solute clustering.

We performed density functional theory (DFT) calculations for many (> 1000) model Al-Mg-Zn alloys. MEPs and ΔE_a of vacancy migrations were computed via DFT plus the climbing image nudged elastic band (CI-NEB) method[35, 36]. Details of DFT+CI-NEB methods are described in Sec. II A. Our results in Sec. III A show that large fluctuations (∼ 1 eV) of ΔE_a under different local chemical environments originate primarily from changes in Δe_a rather than the commonly assumed variations of ΔE, which are typically small (mostly ± ∼ 0.2 eV). A quartic function of the reaction coordinate (x in Fig. 1 (c)) is proposed in Sec. III B to accurately describe and analyze the MEPs of all investigated vacancy migration events. Analyses in Sec. III C reveal that Δe_a is linearly correlated to $k_f D^2$: D is the Cartesian distance of a migrating atom between two adjacent equilibrium positions illustrated by the double-headed arrows in Fig. 1 (a), and k_f is the average vibration spring constant of this atom at these two equilibrium positions. D and k_f are parameters that quantify the local lattice distortion effects on, respectively, the locations and shapes of the MEP at local energy minimum states. Specifically, D is correlated with $D_{\rm MEP}$ in Fig. 1 (c) and k_f is related to the second derivatives of the MEP curves at the local energy minimum states in Fig. 1 (c). Both D and k_f can be calculated relatively easily without accurate descriptions of MEPs obtained from the DFT + CI-NEB method. Details of the calculation methods for D and k_f are described in Sec. II B.

In Sec. III D, based on our DFT+CI-NEB calculations, surrogate models using local lattice occupations as inputs are proposed to predict the coefficients of the quartic function of the vacancy migration MEP in Al-Mg-Zn alloys. This leads to a new approach to accurately and efficiently predict the MEPs and the corresponding ΔE_a and ΔE as functions of local chemical compositions. With this new method to estimate ΔE_a and ΔE, more accurate mesoscale studies, such as kMC, can be conducted. Finally, discussion of the major developments in the paper and conclusions are provided in Sec. IV.

II. METHODS

A. Transition-State Calculations

To compute migration energy barriers and the minimum energy paths (MEP) of vacancy migrations, we performed high-throughput density functional theory (DFT) calculations for model Al-Mg binary alloys, Al-Zn binary alloys, and Al-Mg-Zn ternary alloys. The energies of the atomic configurations at the initial and the final states (E_i and E_f) were first calculated with the

Vienna Ab-Initio Simulation Package (VASP) [37, 38], with all-electron projector-augmented wave potentials (PAW) method with the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional[39, 40]. All calculations used $4 \times 4 \times 4$ supercells, constructed from the FCC Al unit cell, with 255 atoms and 1 vacancy. All supercells used to calculate the vacancy migration barriers can be divided into three categories. The configurations in the first category, as shown in Fig. 2 (a), are randomly generated solid solution structures with different local concentrations of solute atoms (Mg and Zn) around the vacancy site and the migrating atom (Al, Mg, or Zn). These structures simulate vacancy diffusion in the solid-solution state. For the configurations in the second category, as shown in Fig. 2 (b), either $2 \times 2 \times 2$ or $2 \times 2 \times 4$ ordered cluster structures are embedded in the $4 \times 4 \times 4$ pure Al matrix. The data from these configurations are designed to describe the vacancy moving inside the precipitates or along the boundary between the ordered precipitates and the solid-solution Al matrix. These ordered structures were chosen from proposed GuinierPreston (GP) zone precipitates [41] and ordered ($L1_0$, $L1_2$, $L1_0^*$, W2, CH, and Z1) intermetallic structures on an FCC lattice[42]. The third category, as shown in Fig. 2 (c), consists of supercells with a single solute atom (Mg or Zn) embedded in the lattice of neighboring sites (including 1^{st}, 2^{nd}, and 3^{rd} nearest neighbors) of the vacancy and the migrating atom in the $4 \times 4 \times 4$ pure Al matrix. These configurations address the effect of a single solute atom on the vacancy migration barrier.

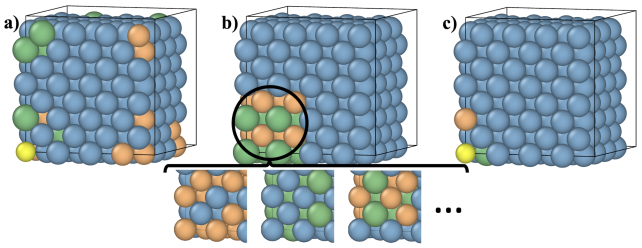

FIG. 2. Schematic diagrams of the model Al alloys supercells used to calculate vacancy migration barriers. Blue, orange, and green spheres represent Al, Zn, and Mg atoms, respectively. Yellow spheres represent vacancy sites. (a): A typical $4 \times 4 \times 4$ FCC supercell with random distributions of solute (Mg and Zn) atoms. (b): A typical $4 \times 4 \times 4$ FCC supercell with a $2 \times 2 \times 2$ ordered cluster structure embedded on Al matrix. (c): A typical $4 \times 4 \times 4$ FCC supercell with a single solute atom (Zn) embedded on a neighboring site around the vacancy and the migrating atom (Mg).

For all DFT calculations, the total energies for supercells of the initial and final states were converged to 10^{-6} eV/cell for the ionic relaxation loop and 10^{-7} eV for the electronic self-consistency loop using a plane-wave cutoff energy of 450.0 eV and MethfesselPaxton smearing of 0.4 eV. A $2 \times 2 \times 2$ k-point grid was applied for all supercells. Each grid was generated using the Monkhorst-Pack scheme [43]. See Supplemental Material at **??** for the results of K-Points convergence tests. The supercell sizes in all vacancy migration investigations were always fixed as four times that of the FCC lattice constant of the model 7XXX series Al alloy. Hence. a supercell of $4 \times 4 \times 4$ conventional FCC cells with 256 atoms (based upon the FCC unit cell) was used to calculate the lattice constant. The supercell consisted of 244 Al atoms, 7 Mg atoms, and 5 Zn atoms, which were within the range of compositions of 7075 Al alloys. Lattice occupations inside this supercell were optimized by the special quasi-random structures (SQS) method using the Alloy Theoretic Automated Toolkit (ATAT)[44]. The lattice constant of this SQS-optimized supercell was 4.046 Å after DFT relaxation of nuclear coordinates and the cell volume. This value is close to the lattice constant of a pure Al crystal at 0K (4.041 Å from DFT calculations with the same setups described in this section). See Supplemental Material at **??** for the effects of lattice constant variations on vacancy migration barriers.

For each vacancy migration event, the energy of the transition state (E_t) and the energy barrier ($\Delta E_a \equiv E_t - E_i$) were gathered by utilizing the climbing image nudged elastic band (CI-NEB) method after evaluating the energy difference (ΔE) by using E_f minus E_i. This was accomplished with VASP and the Transition States Tools (VTST) package [35, 36]. Five images between the relaxed initial and final images were set. The artificial spring constant was set to 5 eV/Å^2. The electronic self consistency-loop breaking criteria was set to 10^{-4} eV and the force convergence criteria for all models was set to be less than 0.05 eV/Å. The force-based quick-min optimizer provided by VTST was used for the CI-NEB calculations[45]. See Supplemental Material at **??** for the verifications of the transition state via phonon calculations[46, 47]. Finally, 2500 ΔE and ΔE_a pairs were obtained from 1250 CI-NEB calculations by considering both forward and backward vacancy migrations.

B. Calculations of Migration Distances and Vibration Spring Constants

As mentioned in Sec. I and discussed in Sec. III C, the migration distances D and vibration spring constants k_f of the migrating atoms are critical to describe the local lattice distortion effects on the vacancy migration MEPs and their ΔE_a. In this subsection, the detailed methods to calculate them and the related parameters are presented. First, the relative distance between the initial and final states along the MEP, D_{MEP}, as indicated in Fig. 1 (c), can be obtained from the outputs of the CI-NEB calculations. Here we set N as the number of intermediate images inserted between the initial and final states, and I_j represents the configuration of the j^{th} intermediate image in the CI-NEB calculations. Specifically, $I_0 = I_i$ and $I_{N+1} = I_f$ denote the initial and final configurations, respectively. $D_{\mathrm{RHD}}(I_a, I_b)$ is a function that returns the magnitude of the relative high-

dimensional distance between I_a and I_b[35, 36]:

$$D_{\mathrm{RHD}}(I_a, I_b) = \sqrt{\sum_{k=1}^{N_{\mathrm{atom}}} \left((\boldsymbol{x}_{b,k} - \boldsymbol{x}_{a,k})^{\mathrm{T}} (\boldsymbol{x}_{b,k} - \boldsymbol{x}_{a,k}) \right)} \tag{2}$$

Here, $\boldsymbol{x}_{j,k}$ is a three-dimensional vector representing the Cartesian positions of the k^{th} atom in the j^{th} image, and N_{atom} is the total number of atoms in each configuration. Since only 5 intermediate images were chosen between the relaxed initial and final images for all CI-NEB calculations in this study, D_{MEP} reduces to:

$$D_{\mathrm{MEP}} = \sum_{i=0}^{N=5} D_{\mathrm{RHD}}(I_j, I_{j+1}) \tag{3}$$

Alternatively, the migration distance D of a migrating atom between two adjacent equilibrium positions (its Cartesian positions in initial and final states) can be directly calculated as:

$$D = \sqrt{(\boldsymbol{x}_{\mathrm{f}} - \boldsymbol{x}_{\mathrm{i}})^{\mathrm{T}} (\boldsymbol{x}_{\mathrm{f}} - \boldsymbol{x}_{\mathrm{i}})} \tag{4}$$

Here $\boldsymbol{x}_{\mathrm{i}}$ and $\boldsymbol{x}_{\mathrm{f}}$ denote the Cartesian position of the migrating atom in the equilibrium initial and final states, respectively. Because most atoms are almost stationary during the vacancy migration process, there are strong correlations between D_{MEP} and D, so the value of D is utilized to quantify the lattice distortion effects on the MEP and the corresponding $\Delta e_{\mathrm{a}}/\Delta E_{\mathrm{a}}$ for each vacancy migration case in Sec. III C. See Supplemental Material at **??** for additional discussions of the correlations between D_{MEP} and D.

The vibration spring constants k_f of migrating atoms are calculated based on the Hessian matrix $\boldsymbol{H}$, which is the matrix of the second derivatives of the energy with respect to the atomic positions, obtained using the finite difference method implemented in VASP. $\boldsymbol{H}$ should be a $3N_{\mathrm{atom}}$ dimensional matrix if all N_{atom} atoms can be displaced in the supercell. In principle, k_f at the initial and final states can be acquired by finding the eigenvalues of $\boldsymbol{H}$, of which the corresponding eigenvectors describe the motions of atoms along the MEP of the vacancy migration. Using the Harmonic approximation, the energy landscape, V, of the MEP at the initial and final states can be expressed as $V = \frac{1}{2}k_f x^2$. Here x is the displacement along the MEP.

However, it is expensive to calculate $\boldsymbol{H}$ for all the investigated cases in this study if all 255 atoms in a supercell are displaced. Since most atoms are nearly stationary during the vacancy migration process, we can approximate the value of k_f by fixing the positions of atoms far away from the vacancy during the calculation of $\boldsymbol{H}$. In this study, only the migrating atom is displaced during the calculation of $\boldsymbol{H}$ for the initial and final states, but all other atoms are fixed. The calculated vibration spring constant values under this fixed-atom condition were obtained for both the initial and final states, and the average value was used in Sec. III C to estimate the lattice distortion effect on the MEP and the corresponding Δe_{a} and ΔE_{a} for each vacancy migration case. More accurate k_f can be obtained if more atoms in the supercells were displaced during the calculation of $\boldsymbol{H}$. See Supplemental Material at **??** for detailed discussions of k_f calculation.

III. RESULTS

A. DFT Calculations of ΔE_{a} and ΔE

Correlations of ΔE_a (Y-axis) and ΔE (X-axis) from our computational results are plotted in Fig. 3 (a) to (c) for different types of migrating atoms. In the Al-Zn binary system (Fig. 3 (a)), ΔE_a and ΔE data are scattered; however, the data still follows (approximately) linear relationships for both the migrating Al and Zn atoms, respectively. Simple linear regressions suggest that the slope of each fitted straight line is close to $\frac{1}{2}$, so Δe_{a} is approximately a constant according to Eq. (1) and the Kinetic Ising Model is still approximately valid for vacancy migrations in the Al-Zn binary system. However, as seen in Fig. 3 (b) and (c), the ΔE_a and ΔE data become significantly scattered, and a linear relationship does not apply when Mg is added as a solute element for all types of migrating atoms (Al, Zn and Mg) in both binary Al-Mg and ternary Ag-Mg-Zn systems. In these cases, ΔE values are still distributed in a similar range as those in the Al-Zn system in Fig. 3 (a), mostly from ~ -0.2 eV to ~ 0.2 eV. However, E_a values are scattered in much wider ranges from almost 0 eV up to ~ 1 eV. This indicates that the fluctuations of ΔE_{a} in different local chemical environments are mostly dependent upon changes in Δe_{a} rather than the small variations of ΔE according to Eq. (1). These deviations demonstrate that even simple vacancy migrations in alloys with close-packed lattices are complex; hence the assumptions behind Fig. 1 (b) are incorrect and the Kinetic Ising Model is not broadly applicable.

One problem in Fig. 1 (b) is the neglect of changes in MEP along the reaction coordinate axis. Here, we define the distance between the initial and final states along the reaction coordinate as D_{MEP} as indicated in Fig. 1 (c). The reaction coordinate x and D_{MEP} for all investigated cases result from DFT + CI-NEB calculations[35, 36]. Fig. 3 (d), (e), and (f) show the kernel density estimations [48], which indicate smoothed probability distributions, of D_{MEP} for different types of migrating atoms (Al, Mg and Zn) in all migration events. In the Al-Zn binary system (Fig. 3 (d)), D_{MEP} values are centered-distributed with negligible standard deviation $\sigma_{D_{\mathrm{MEP}}}$ for both migrating Al atoms ($\sigma_{D_{\mathrm{MEP}}}$ =0.035 Å) and migrating Zn atoms ($\sigma_{D_{\mathrm{MEP}}}$ =0.018 Å). These distributions indicate occupations of Zn atoms near vacancy sites induce small lattice distortions. However, in both Al-Mg systems (Fig. 3 (e)) and Al-Mg-Zn systems (Fig. 3 (f)), $\sigma_{D_{\mathrm{MEP}}}$ is much larger for all migrating Al, Mg, and Zn atoms ($\sigma_{D_{\mathrm{MEP}}}$ =$\sim$ 0.2 Å in all cases). These distribu-

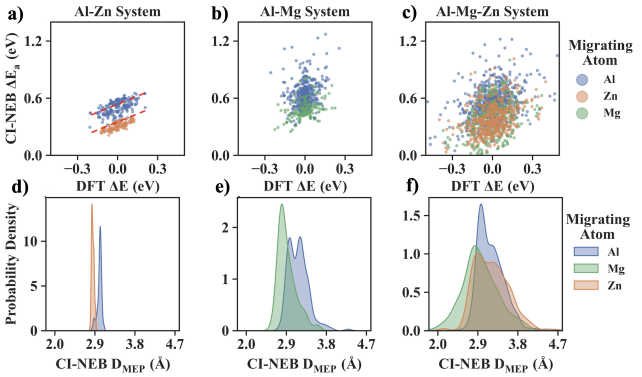

FIG. 3. Correlations between ΔE_a and ΔE for vacancy migration events in Al alloys. (a)-(c): Correlations between ΔE_a and ΔE for vacancy migration events from DFT + CI-NEB calculations. Migrating atoms are Al (blue dots), Zn (orange dots), and Mg (green dots) in binary Al-Zn (d), Al-Mg (e), and ternary Al-Mg-Zn (f) systems, respectively. The scattering distributions in (e) and (f) indicate the variations of ΔE_a do not only depend on the variations of ΔE as suggested in Fig. 1 (b). (g)-(i): Kernel density estimations [48] of D_{MEP}, the distance between the initial and final states defined in Fig. 1 (c), are plotted for migrating Al (blue), Zn (orange), and Mg (green) atoms in Al-Zn (g), Al-Mg (h), and Al-Mg-Zn (i) systems, respectively. Large variations of D_{MEP} in (h) and (i) suggest strong lattice distortion effects in different local chemical environments. The color coding applies to subsequent figures.

tions indicate occupations of Mg atoms near vacancy sites induce relatively large lattice distortions. These lattice distortions are understandable because of the atomic size differences and large fluctuations of local Mg/Zn concentrations for all investigated supercells. The size of Mg atoms is much larger than those of Zn and Al atoms (the radii of Mg and Zn and Al atoms are 1.50, 1.35, and 1.25 Å[49], respectively), so the lattice distortion effects due to Mg atoms in the Al matrix are much stronger than those due to Zn atoms. The large fluctuations of local Mg/Zn concentrations originate from the multiple types of supercells used in our calculations as shown in Fig. 2, which correspond to different precipitation stages of Al alloys. A key question is how to construct accurate MEP models illustrated in Fig. 1 (c) to accommodate the lattice distortion effects if we want to understand the physical mechanisms behind Δe_a and ΔE_a variations.

B. Quartic Functions of the MEP

An accurate and quantitative model to describe the MEP in Fig. 1 (c) has to satisfy several physical conditions, including zero first derivative at initial (x_i), transition (x_t) and final (x_f) states. Thus, we propose that the energy landscape of a general vacancy migration MEP, as a function of the reaction coordinate x with a single local energy maximum, is described by a simple quartic function, $E_{\mathrm{MEP}}(x)$:

$$E_{\mathrm{MEP}}(x) = ax^4 + bx^3 + cx^2 \tag{5}$$

Here, the coefficients (a, b, and c) are assumed to depend on the local lattice occupations near a vacancy/adjacent migrating atom pair. Values of x_i, x_t, and x_f are determined by the zero-first-derivative requirements mentioned above. The first derivative of Eq. (5) is $E'_{\mathrm{MEP}}(x) = 4ax^3 + 3bx^2 + 2cx$, which roots $x_0 = 0$, $x_1 = \frac{-3b-\sqrt{9b^2-32ac}}{8a}$, and $x_2 = \frac{-3b+\sqrt{9b^2-32ac}}{8a}$. When $a > 0$ and $c < 0$, Eq. (5) has two local minima and one local maximum, which corresponds to the shape of the energy landscape along the MEP in Fig. 1 (c). As plotted in Fig. 4 (a), we can shift the energy landscape to make the transition state at the origin point by denoting the position of the transition state at $x_t = 0$ and denoting its energy on the MEP $E_{\mathrm{MEP}}(x_t = 0) = 0$. Then we make the positions of the initial state and final state at two local minima as $x_i = x_1$ and $x_f = x_2$. If Eq. (5) is accurate enough to describe the MEP for each migration event and its coefficients (a, b, and c) can be predicated, ΔE_a, ΔE, and D_{MEP} of the corresponding migration event can be predicted as suggested in Fig. 4 (a) (then its Δe_a is from Eq. (1)).

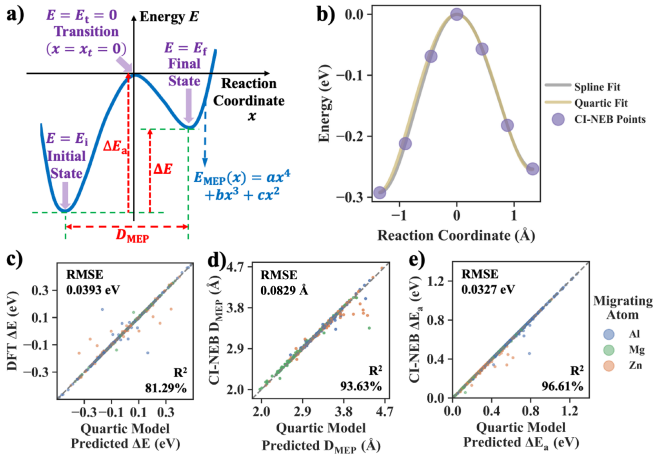

FIG. 4. The quartic function in Eq. (5) is used to fit vacancy migration MEPs from DFT + CI-NEB calculations. (a) Schematic plot of $E_{\mathrm{MEP}}(x)$ of Eq. (5) showing ΔE, D_{MEP}, and ΔE_a. (b) A specific example where Eq. (5) is used to fit an MEP curve from a DFT + CI-NEB calculation. (c)-(e): Comparisons of ΔE (c), D_{MEP} (d), and ΔE_a (e) from DFT + CI-NEB calculations and those predicted from the fitted Eq. (5) for all Al (blue), Zn (orange), and Mg (green) migrating atoms in all investigated supercells. The root-mean-square error (RMSE) is denoted at the upper left, and the number at the bottom-right corner shows the coefficient of determination R^2 (close to 100% means high accuracy). Small RMSE and large R^2 values in (c)-(e) demonstrate that Eq. (5) is accurate to describe vacancy migration MEPs. The same RMSE and R^2 symbols are used in Fig. 6 and Fig. 8.

Thus, by assuming Eq. (5) is accurate to describe all the MEPs from our DFT+CI-NEB calculations, we ap-

plied a least-squares fitting method with a weight matrix to fit the coefficients a, b, and c for each migration event. Since CI-NEB methods use a series of images along the reaction path to calculate MEPs, we can not only collect the energetics of the initial, final, and transition states, but also those of other intermediate images, which are at certain coordinates along MEPs. The following conditions are included in the quartic equation fitting: the energies of initial and final states predicted by the quartic function equal those from DFT calculations, $E_{\mathrm{MEP}}(x_{\mathrm{i}}) = E_{\mathrm{i}}$ and $E_{\mathrm{MEP}}(x_{\mathrm{f}}) = E_{\mathrm{f}}$; the energies of the other intermediate images equal to those from the DFT+CI-NEB calculations, $E_{\mathrm{MEP}}(x_j) = E_j$; the first derivatives at initial and final states zero, $E'_{\mathrm{MEP}}(x_{\mathrm{i}}) = 0$ and $E'_{\mathrm{MEP}}(x_{\mathrm{f}}) = 0$. Thus, the following equation can be obtained:

$$\begin{pmatrix} x_{\mathrm{i}}^4 & x_{\mathrm{i}}^3 & x_{\mathrm{i}}^2 \\ \vdots & \vdots & \vdots \\ x_j^4 & x_j^3 & x_j^2 \\ \vdots & \vdots & \vdots \\ x_{\mathrm{f}}^4 & x_{\mathrm{f}}^3 & x_{\mathrm{f}}^2 \\ 4x_{\mathrm{i}}^3 & 3x_{\mathrm{i}}^2 & 2x_{\mathrm{i}} \\ 4x_{\mathrm{f}}^3 & 3x_{\mathrm{f}}^2 & 2x_{\mathrm{f}} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} E_{\mathrm{i}} \\ \vdots \\ E_j \\ \vdots \\ E_{\mathrm{f}} \\ 0 \\ 0 \end{pmatrix} \tag{6}$$

Here x_j is the location of the j^{th} intermediate image along the reaction coordinate and E_j is its energy relative to the transition state (since $x_{\mathrm{t}} = 0$ and $E_{\mathrm{MEP}}(x_{\mathrm{t}}) = 0$ according to Eq. (5)). All values of x_j and E_j are directly from DFT + CI-NEB calculations. We denote the left matrix as $\boldsymbol{X}$, the quartic coefficients vector as $\boldsymbol{\beta}$, and the right vector as $\boldsymbol{y}$ for Eq. (6), which can be re-written as $\boldsymbol{X\beta} = \boldsymbol{y}$. To find the best description of each MEP, the weighted linear regression is applied, which is a generalization of ordinary least squares:

$$\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{X}\right)\hat{\boldsymbol{\beta}} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{y} \tag{7}$$

Here, $\boldsymbol{W}$ is a diagonal matrix, with each of its elements representing a weighting coefficient used for each data point. The estimated quartic coefficients vector is $\hat{\boldsymbol{\beta}} = \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{y}$. To emphasize the accuracy of the computed energy terms ΔE and ΔE_{a}, we increase the weight elements of the first condition mentioned above to large finite numbers and retain other weight elements equal to 1. Each MEP curve of all DFT-CI-NEB calculations was fitted by Eq. (7).

Fig. 4 (b) shows an MEP curve from the DFT + CI-NEB calculation is accurately described by both the standard spline fitting and our quartic fitting curve based on Eq. (5). Overall, Fig. 4 (c), (d), and (e) depict close matches between ΔE, D_{MEP}, and ΔE_{a} from direct DFT + CI-NEB calculations (Y-axis) and those from the quartic function $E_{\mathrm{MEP}}(x)$ with fitted coefficients (X-axis), respectively. Low values of the root-mean-square error (RMSE) (close to 0) and high values of the coefficient of determination R^2 (close to 100%) confirm that Eq. (5) is accurate and robust enough to describe the MEP of vacancy migrations in Al-Mg-Zn alloys while incorporating the requisite physics associated with the vacancy migration MEPs.

Fig. 5 shows the kernel density estimations of the fitted coefficients of Eq. (5) for all vacancy migration cases in different alloy systems (Fig. 5 (a)-(c) in Al-Zn binary systems, Fig. 5 (d)-(f) in Al-Mg binary systems, and Fig. 5 (g)-(i) in Al-Mg-Zn ternary systems). The results show the distributions of a and c vary significantly for different types of migrating atoms in three alloy systems. The wide ranges in a and c indicate the shapes of the MEPs in Fig. 1 (c) and Fig. 4 (a) can change significantly because both a and c determine the coordinates and curvatures of MEPs at local energy minimum states (as discussed in Sec. III C, the ratio of a to c is also important to determine the MEPs and the values of $\Delta E/D_{\mathrm{MEP}}/\Delta e_{\mathrm{a}}$, so these distributions of a and c can not be used to explain the differences between Al-Zn alloys and Al-Mg/Al-Mg-Zn alloys in Fig. 3). Alternatively, the distributions of b for all types of migrating atoms in all alloy systems are always in narrow ranges close to zero, which is consistent with the small variations of ΔE in Fig. 1 (d) ($\Delta E = 0$ if $b = 0$). This special feature of b provides us a relatively easy and accurate way to predict Δe_{a} based on lattice distortion effects as follows.

C. Estimations of Δe_{a} based on Lattice Distortion Effects

As indicated by Fig. 5, we can assume $b \approx 0$ giving

$$E_{\mathrm{MEP}}(x) \approx ax^4 + cx^2 \tag{8}$$

This is the same free energy formalism of second-order phase transitions in Landau theory[50]. Thus, $x_{\mathrm{i}} \approx -\sqrt{\frac{-c}{2a}}$, $x_{\mathrm{t}} = 0$, and $x_{\mathrm{f}} \approx \sqrt{\frac{-c}{2a}}$, respectively. Accordingly, $D_{\mathrm{MEP}} = x_{\mathrm{f}} - x_{\mathrm{i}} \approx \sqrt{\frac{-2c}{a}}$ and $\Delta e_{\mathrm{a}} \approx -ax_{\mathrm{i}}^4 - cx_{\mathrm{i}}^2 \approx \frac{c^2}{4a}$. We can also estimate $\Delta E \approx 2bx_{\mathrm{f}}^3 \approx \frac{\sqrt{2}}{2}(\frac{-c}{a})^{\frac{3}{2}}b$ based on Eq. (5). These approximate relations are confirmed in Fig. 6 (a)-(c) by the linear correlations between the DFT + CI-NEB results (ΔE, D_{MEP}, and Δe_{a} on the Y-axis) and their estimations based on Eq. (5) and Eq. (8) (the X-axis), respectively.

In addition, the second derivative of the MEP at the local-minimum states $E''_{\mathrm{MEP}}(x_{\mathrm{i}}) \approx E''_{\mathrm{MEP}}(x_{\mathrm{f}}) \approx -4c$ according to Eq. (8). Thus, we can get an approximate relation as $\Delta e_{\mathrm{a}} \approx \frac{E''_{\mathrm{MEP}}(x_{\mathrm{i}})D^2_{\mathrm{MEP}}}{32}$. Accurate E''_{MEP} and D_{MEP} are obtained from the MEP curves produced by the DFT + CI-NEB calculations. However, because the migrating atom moves a distance ($> \sim$2 Å in the Al lattice) much larger than the other atoms for a general migration event, the motion of the migrating atom is the most important factor for the reaction coordinate x. Therefore, we assume E''_{MEP} is proportional to the average vibration spring constant k_f of the migrating atom at

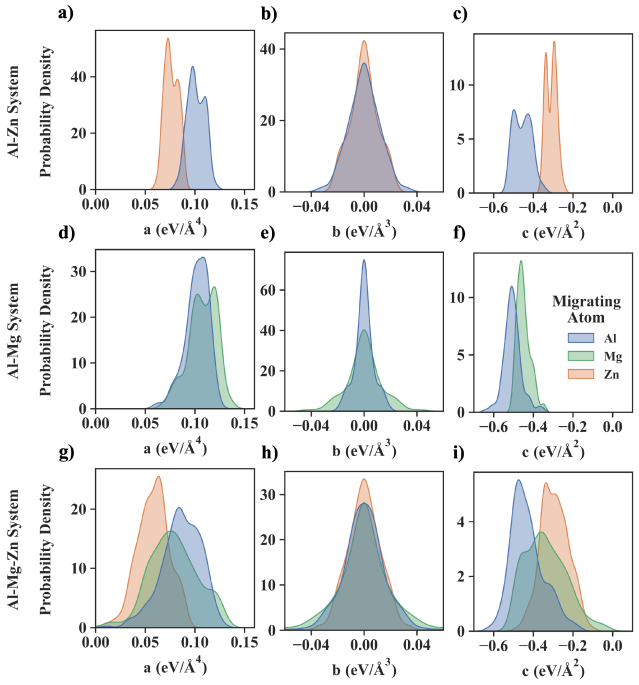

FIG. 5. Kernel density estimations of fitted coefficients (a, b, and c) of Eq. (5) for all Al (blue), Zn (orange), and Mg (green) migrating atoms in all investigated supercells. (a)-(c): Probability densities of a, b, and c in Al-Zn systems. (d)-(f): Probability densities of a, b, and c in Al-Mg systems. (g)-(i): Probability densities of a, b, and c in Al-Mg-Zn systems. The narrow probability densities variations of b in all investigated supercells is consistent with the small variations of ΔE in Fig. 3 (a)-(c).

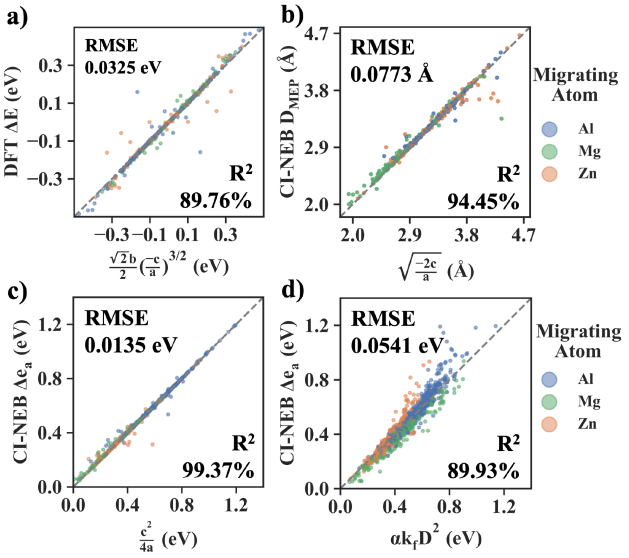

FIG. 6. Methods to estimate ΔE (a), D_{MEP} (b), and Δe_{a} (c) of vacancy migration MEPs. (a)-(c) Comparisons between DFT+CI-NEB calculated ΔE (a), D_{MEP} (b), and Δe_{a} (c) and those estimated based on Eq. (5) and Eq. (8). (d) Correlations between DFT+CI-NEB calculated Δe_{a} and $k_f D^2$ to verify Eq. (9).

the initial and final states, and we also assume D_{MEP} is proportional to the distance D of the same atom at these two states as illustrated in Fig. 1 (a). These assumptions give:

$$\Delta e_{\mathrm{a}} \approx \alpha k_f D^2 \tag{9}$$

where α is a unitless constant. Because the migrating atom moves a distance much larger than all other atoms, the variations of k_f and D can be used to approximate the local lattice distortion effects on the shape and locations of local-minimum states along a MEP. These two parameters are obtained from DFT calculations of $\boldsymbol{H}$ and the coordinate of the migrating atom in fully relaxed structures as described in Sec. II B. The validity of Eq. (9) is confirmed by comparing Δe_{a} from DFT + CI-NEB calculations (Y-axis) and $\alpha k_f D^2$ (X-axis) for all migrating atoms in all investigated supercells in Fig. 6 (d). This shows that Eq. (9) with the same α value (≈ 0.022 fitted by Fig. 6 (d)) works for all Al, Mg, Zn migrating atoms in these Al alloys. Eq. (9) therefore provides an efficient way to estimate Δe_{a} and ΔE_{a} using standard DFT calculations without the CI-NEB method.

To further verify the generality and accuracy of Eq. (9), we compute k_f and D of specific examples of vacancy migration with different types of migrating atoms in a dilute Al alloy in Table I. In these cases, except for the migrating atom, there is no solute atom in the supercell so that the initial and final states are equivalent since the MEP is symmetric on two sides of the transition state. Thus, the migration energetic driving force $\Delta E = 0$ and $\Delta e_{\mathrm{a}} \equiv \Delta E_{\mathrm{a}} - \frac{1}{2}\Delta E = \Delta E_{\mathrm{a}}$ according to Eq. (1). We also present these results of $\Delta E_{\mathrm{a}} = \Delta e_{\mathrm{a}}$ from DFT+CI-NEB calculations in Table I. These results show that, for all types of migrating atoms (Al, Mg, and Zn) in a dilute Al matrix, the ratio of Δe_{a} to $k_f D^2$ is almost a constant value close to the α value (0.022) fitted from all migration cases with different values of ΔE, thereby further supporting the generality and validity of Eq. (9).

D. Surrogate Models to Predict the MEP

Although Eq. (9) can be helpful to estimate Δe_{a} and ΔE_{a} without using the computationally expensive CI-NEB method for Al-Mg-Zn and potentially other multicomponent FCC alloys, it still requires DFT calculations that need considerable computational resources. To study diffusion and precipitation in mesoscale methods such as kMC simulations[8, 9], we still need to accurately and efficiently predict ΔE_{a} and ΔE in different local chemical environments. A practical approach is to construct surrogate models that can predict the coefficients (a, b and c) of Eq. (5) with respect to the local lattice occupations, then the properties of the MEP (ΔE_{a}, ΔE, and D_{MEP}) can be automatically obtained based

TABLE I. Results of vacancy migrations in a dilute Al matrix (at most one solute atom in a supercell) are listed as the migration barrier $\Delta E_a = \Delta e_a$, the average vibration spring constant k_f of the migrating atom at the initial and final states calculated under this fixed-atom condition described in Sec. II B, the high dimensional distance along the minimum energy path between the initial and final states D_{MEP} defined in Eq. (3), the Cartesian distance of the migrating atom between initial and final states D defined in Eq. (4), the values of $k_f D^2$, and the coefficient $\alpha \equiv \frac{\Delta e_a}{k_f D^2}$ of selected migrating atoms. As a reference, the value of α in Eq. (9) fitted from the whole database of vacancy migrations is 0.0220 as shown in Fig. 6 (d).

Migrating atom	$\Delta E_a = \Delta e_a$ (eV)	k_f (eV/Å^2)	D_{MEP} (Å)	D (Å)	$k_f D^2$ (eV)	$\alpha \equiv \frac{\Delta e_a}{k_f D^2}$
Al	0.58	3.60	3.00	2.75	27.28	0.0213
Mg	0.47	3.39	2.78	2.58	22.62	0.0207
Zn	0.34	2.02	2.80	2.69	14.66	0.0233

on Eq. (5) and Fig. 4 (a). The general strategy to construct these surrogate models and train them based on DFT+CI-NEB results is described as follows.

To train the surrogate models for coefficients of Eq. (5), 2000 training data points of ΔE and ΔE_a pairs (plus the corresponding supercell configurations of the initial and final states) were chosen randomly from the total 2500 data points generated from the DFT+CI-NEB calculations. The remaining 500 data points were utilized as testing data to evaluate the predictive accuracy of the surrogate models. All of the data were divided into three different groups based on the chemical type of the migrating atom (Al, Mg or Zn)[51, 52]. The input information was chosen to be the type of the migrating atom and the type of all atoms on the 1$^{\text{st}}$, 2$^{\text{nd}}$, and 3$^{\text{rd}}$ nearest-neighbor lattice sites relative to the vacancy site before and after the migration event, as shown Fig. 7 (a). This difference between the even-order-term coefficients of Eq. (5) (a and c) and the odd-order-term coefficient (b) suggests we should use different symmetry constraints to construct the inputs of the surrogate models for coefficients. Thus, the input information (only 1$^{\text{st}}$ nearest-neighbor lattice sites are shown) for surrogate models of a and c is constructed based on the symmetry operations of the mmm point group shown in Fig. 7 (b), and input for the surrogate model of b is constructed based on the symmetry operations of the $mm2$ point group shown in Fig. 7 (c).

In both Fig. 7 (b) and (c), the vacancy (black color) and the migration atom (yellow color) are aligned along the $\langle 110 \rangle$ direction (x-axis). For the mmm point group illustrated in Fig. 7 (b), there is a mirror symmetry plane perpendicular to the $\langle \bar{1}10 \rangle$ (y-axis), a mirror symmetry plane perpendicular to the $\langle 001 \rangle$ (z-axis), and a mirror symmetry plane perpendicular to the $\langle 110 \rangle$ direction (x-axis). For the $mm2$ point group illustrated in Fig. 7 (c), there is a mirror symmetry plane perpendicular to the $\langle \bar{1}10 \rangle$ (y-axis), a mirror plane perpendicular to the $\langle 001 \rangle$ (z-axis), a 2-fold rotation axis along the $\langle 110 \rangle$ direction (x-axis). Thus, the neighboring sites can be divided into different sets based on their symmetry relative to the vacancy and the migrating atom. As shown in Fig. 7 (b) and (c), the mmm point group sorts the 1$^{\text{st}}$ nearest neighbor sites into 4 sets; the $mm2$ point group sorts the 1$^{\text{st}}$ nearest neighbor sites into 7 sets; each set of lattice sites is plotted in the same color. The same strategy is applied to 2$^{\text{nd}}$, and 3$^{\text{rd}}$ nearest neighbors and 2-atom clusters (atoms at two lattice sites not apart than 3$^{\text{rd}}$-nearest-neighbor distance (4.955 Å)) as well. Atoms and clusters that are symmetrically equivalent should have the same contribution to the inputs of the surrogate models for coefficients of Eq. (5).

Besides the symmetry effect, the encoding strategy of the lattice occupations has a significant impact on the surrogate model. In this work, we applied the one-hot encoding method [53, 54] to construct feature vectors to describe types of single atoms and 2-atom clusters. The advantage of using the one-hot encoding for categorical data is that since it represents each type of the variable by a unique digit, there is no quantitative relationship between the values of variables. Hence, one-hot encoding without introducing any fictional ordinal relationship can be more accurate. The symmetry properties related to vacancy migrations in the FCC lattice illustrated in Fig. 7 are applied to construct these feature vectors. Because a and c of Eq. (5) are the coefficients of fourth-order and second-order terms, respectively, each should have the same values for the forward and backward migration processes in a vacancy migration case. However, b of Eq. (5) is the coefficient of a third-order term, so it should have the opposite values in forward and backward migration processes. Consequently, the feature vectors for the surrogate models of a and c are constructed based on the symmetry operations of the mmm point group as illustrated in Fig. 7 (b), and the feature vectors for the surrogate model of b are constructed based on the symmetry operations of the $mm2$ point group as illustrated in Fig. 7 (c). In both Fig. 7 (b) and (c), the symmetrically equivalent lattice sites are of the same color, so the contributions of chemical elements on these symmetrically equivalent sites to the feature vectors should be averaged.

Examples of feature vectors and the symmetry constraints on feature vectors are described as follows. We used a feature vector $\boldsymbol{v} \in \mathbb{R}^3$ to represent the chemical

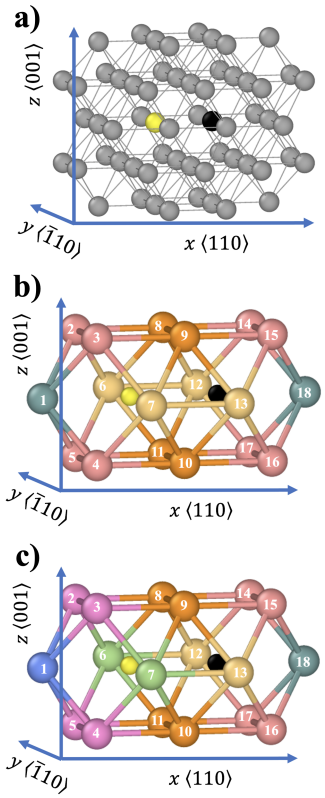

FIG. 7. Illustrations of local lattice sites related to the vacancy migration and their symmetry properties considered in the surrogate models of coefficients of Eq. (5). (a): The plot of the 1st, 2nd, and 3rd nearest neighboring sites of the vacancy (black) and the migrating atom (yellow) aligned along the ⟨110⟩ direction (x-axis). The vacancy and the migrating atom are plotted in the same way in (b) and (c). (b) and (c): Effects of *mmm* and *mm*2 point group symmetry applied on the 1st neighboring lattice sites, respectively. Atoms with the same color are at the symmetrically equivalent lattice sites, so their contributions to the inputs of the surrogate models are averaged together. Here the *mmm* point group shown in (b) has a mirror symmetry plane perpendicular to the ⟨1̄10⟩ (y-axis), a mirror symmetry plane perpendicular to ⟨001⟩ (z-axis), and a mirror symmetry plane perpendicular to ⟨110⟩ direction (x-axis). The *mm*2 point group shown in (c) has a mirror symmetry plane perpendicular to the ⟨1̄10⟩ (y-axis), a mirror symmetry plane perpendicular to ⟨001⟩ (z-axis), and 2-fold symmetry along ⟨110⟩ direction (x-axis).

type of a single atom:

$$\begin{aligned} \boldsymbol{v}_{\text{Al}} &= (1,\ 0,\ 0) \\ \boldsymbol{v}_{\text{Mg}} &= (0,\ 1,\ 0) \\ \boldsymbol{v}_{\text{Zn}} &= (0,\ 0,\ 1) \end{aligned} \tag{10}$$

For 2-atom clusters, if both of two lattice sites are from the same symmetry-equivalent sets (two sites with the same color in Fig. 7 (b) or (c)), such as the cluster of atom 8 and atom 9 shown in Fig. 7 (b) or (c), then their orientation and order relative to the vacancy site and the migrating atom can be neglected. Therefore, there are 6 combinations in total to put different types of chemical elements in these two sites. We used a feature vector $\boldsymbol{v} \in \mathbb{R}^6$ to represent each type:

$$\begin{aligned} \boldsymbol{v}_{\text{Al-Al}} &= (1,\ 0,\ 0,\ 0,\ 0,\ 0) \\ \boldsymbol{v}_{\text{Al-Mg}} &= (0,\ 1,\ 0,\ 0,\ 0,\ 0) \\ \boldsymbol{v}_{\text{Al-Zn}} &= (0,\ 0,\ 1,\ 0,\ 0,\ 0) \\ \boldsymbol{v}_{\text{Mg-Mg}} &= (0,\ 0,\ 0,\ 1,\ 0,\ 0) \\ \boldsymbol{v}_{\text{Mg-Zn}} &= (0,\ 0,\ 0,\ 0,\ 1,\ 0) \\ \boldsymbol{v}_{\text{Zn-Zn}} &= (0,\ 0,\ 0,\ 0,\ 0,\ 1) \end{aligned} \tag{11}$$

However, if two lattice sites are from different symmetry sets, for instance, the cluster of atom 3 and atom 9 shown in Fig. 7 (b) or (c), then their orientations and order can affect the vacancy migration energetics. Therefore, there are 9 combinations to put different types of chemical elements in these two sites. This required use of a feature vector $\boldsymbol{v} \in \mathbb{R}^9$ to represent each type:

$$\begin{aligned} \boldsymbol{v}_{\text{Al-Al}} &= (1,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0) \\ \boldsymbol{v}_{\text{Al-Mg}} &= (0,\ 1,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0) \\ \boldsymbol{v}_{\text{Al-Zn}} &= (0,\ 0,\ 1,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0) \\ \boldsymbol{v}_{\text{Mg-Al}} &= (0,\ 0,\ 0,\ 1,\ 0,\ 0,\ 0,\ 0,\ 0) \\ \boldsymbol{v}_{\text{Mg-Mg}} &= (0,\ 0,\ 0,\ 0,\ 1,\ 0,\ 0,\ 0,\ 0) \\ \boldsymbol{v}_{\text{Mg-Zn}} &= (0,\ 0,\ 0,\ 0,\ 0,\ 1,\ 0,\ 0,\ 0) \\ \boldsymbol{v}_{\text{Zn-Al}} &= (0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 1,\ 0,\ 0) \\ \boldsymbol{v}_{\text{Zn-Mg}} &= (0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 1,\ 0) \\ \boldsymbol{v}_{\text{Zn-Zn}} &= (0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 1) \end{aligned} \tag{12}$$

After using feature vectors to label single atoms and 2-atom clusters on the local lattice occupations near the vacancy and the migrating atom, we can average the one-hot feature encoding vectors from the clusters that share the same symmetry. A feature vector that represents the averaged information can be obtained. For example, if the 18 first-nearest-neighboring sites shown in Fig. 7 (b) have the following lattice occupations (σ_i, where i is the site index plotted in Fig. 7 (b)): σ_1 = Al, σ_2 = Mg, σ_3 = Al, σ_4 = Al, σ_5 = Zn, σ_6 = Mg, σ_7 = Al, σ_8 = Al, σ_9 = Zn, σ_{10} = Mg, σ_{11} = Zn, σ_{12} = Al, σ_{13} = Al, σ_{14} = Al, σ_{15} = Al, σ_{16} = Mg, σ_{17} = Zn and σ_{18} = Al, respectively, then four feature vectors can be obtained for the single-atom occupations in 4 sets of 1st nearest neighbor sites

by considering the mmm point group:

$$\begin{aligned}
\hat{\boldsymbol{v}}_1 &= \frac{1}{2}\boldsymbol{v}_{Al} + \frac{1}{2}\boldsymbol{v}_{Al} \\
&= \left(1,\ 0,\ 0\right) \\
\hat{\boldsymbol{v}}_2 &= \frac{1}{8}\boldsymbol{v}_{Mg} + \frac{1}{8}\boldsymbol{v}_{Al} + \frac{1}{8}\boldsymbol{v}_{Al} + \frac{1}{8}\boldsymbol{v}_{Zn} \\
&\quad + \frac{1}{8}\boldsymbol{v}_{Al} + \frac{1}{8}\boldsymbol{v}_{Al} + \frac{1}{8}\boldsymbol{v}_{Mg} + \frac{1}{8}\boldsymbol{v}_{Zn} \\
&= \left(0.5,\ 0.25,\ 0.25\right) \\
\hat{\boldsymbol{v}}_3 &= \frac{1}{4}\boldsymbol{v}_{Mg} + \frac{1}{4}\boldsymbol{v}_{Al} + \frac{1}{4}\boldsymbol{v}_{Al} + \frac{1}{4}\boldsymbol{v}_{Al} \\
&= \left(0.75,\ 0.25,\ 0\right) \\
\hat{\boldsymbol{v}}_4 &= \frac{1}{4}\boldsymbol{v}_{Al} + \frac{1}{4}\boldsymbol{v}_{Zn} + \frac{1}{4}\boldsymbol{v}_{Mg} + \frac{1}{4}\boldsymbol{v}_{Zn} \\
&= \left(0.25,\ 0.25,\ 0.5\right)
\end{aligned} \tag{13}$$

Here, each $\mathbb{R}^3$ feature vector of a single atom is multiplied by a weighting factor $\frac{1}{n_s}$, where n_s is the number of symmetry-equivalent sites in each of these 4 sets. Concatenating these feature vectors together, we can obtain a combined feature vector $\hat{\boldsymbol{v}} = \left(\hat{\boldsymbol{v}}_1,\ \hat{\boldsymbol{v}}_1,\ \hat{\boldsymbol{v}}_1,\ \hat{\boldsymbol{v}}_4\right) \in \mathbb{R}^{12}$. When we extended this method to 2-atoms clusters within 3$^{\text{rd}}$ nearest neighboring distance among all lattice sites shown in Fig. 7 (a), we obtained the combined feature vectors that describe the local environment of a vacancy migration event.

The dimensionalities of the combined feature vectors of lattice occupations in lattice sites of Fig. 7 (a) were 1401 based on the $mm2$ point group symmetry operations and 711 based on the mmm point group symmetry operations. These large dimensionalities were at the same scale as the size of our three training datasets (for three different elements of migrating atoms), which reflect a typical downside of one-hot encoding: it tends to create multicollinearity among individual variables because it creates multiple new variables. However, we can apply principal component analysis (PCA) to reduce the dimensionality of the feature vectors. Overall, using one-hot encoding and the PCA method together, we can eliminate potential quantitative relationships and multicollinearity between the individual variables at the same time, which significantly increases the accuracy and robustness of the surrogate model. See Supplemental Material at **??** and **??** for more details regarding the dimensionality and PCA methods.

After the dimensionality reduction, the ridge regression (linear least squares with the L_2 regularization) was applied to the training data. It can be described in the form of least squares as:

$$\hat{\boldsymbol{X}}\hat{\boldsymbol{\beta}}_{\text{ridge}} = \boldsymbol{y} \tag{14}$$

where, the estimated parameters $\hat{\boldsymbol{\beta}}_{\text{ridge}}$ minimizes the objective function:

$$\min_{\boldsymbol{\beta}} \left\{ \| \boldsymbol{y} - \hat{\boldsymbol{X}}\boldsymbol{\beta}\|_2^2 + \lambda|\boldsymbol{\beta}\|_2^2 \right\} \tag{15}$$

Here, $\hat{\boldsymbol{X}}$ is the dimension-reduced feature space. Each row in $\hat{\boldsymbol{X}}$ represents a dimension-reduced feature vector, and it has m rows in total, where m is the size of the training dataset. $\boldsymbol{y}$ is a vector that contains the results of the targeted coefficients a, b or c. Since there are two constraints ($a > 0$ and $c < 0$) to make sure that Eq. (5) represents the MEP in Fig. 1 (c), elements in $\boldsymbol{y}$ can be $\log(a)$, b or $\log(-c)$ for each data point. The scalar λ is a user-defined regularization parameter, which was set to 1 in our calculations. Based on Eq. (15), the estimated parameters vector is $\hat{\boldsymbol{\beta}}_{\text{ridge}} = \left(\boldsymbol{X}^{\text{T}}\boldsymbol{X} + \lambda\boldsymbol{I}\right)^{-1}\boldsymbol{X}^{\text{T}}\boldsymbol{y}$, where $\boldsymbol{I}$ is an identity matrix.

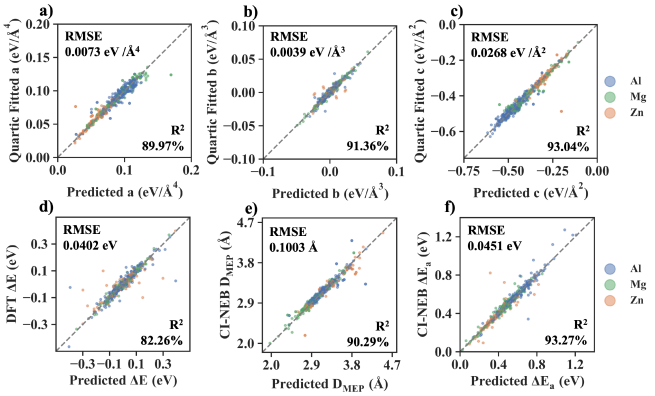

FIG. 8. Performances of surrogate models to predict vacancy migration energetics based on Eq. (5). (a)-(c): Comparisons between directly fitted results and predictions from our surrogate models for $E_{\text{MEP}}(x)$ coefficients (a, b, and c in Eq. (5)). (d)-(f): Comparisons between DFT+CI-NEB calculated results and predictions based on our surrogate-model-predicted $E_{\text{MEP}}(x)$ for ΔE (d), D_{MEP} (e), and ΔE_{a} (f).

After training the surrogate models to output the coefficients a, b and c based on the local lattice occupations, we can calculate the properties of the MEP (ΔE_{a}, ΔE, and D_{MEP}) from Eq. (5) as follows: $\Delta E = E_{\text{MEP}}(x_{\text{f}}) - E_{\text{MEP}}(x_{\text{i}}) = b^2(x_{\text{f}}^3 - x_{\text{i}}^3) = \frac{b\left(9b^2-32ac\right)^{3/2}}{256a^3}$, $D_{\text{MEP}} = x_{\text{f}} - x_{\text{i}} = \frac{\sqrt{9b^2-32ac}}{4a}$, and $\Delta E_{\text{a}} = -E_{\text{MEP}}(x_{\text{i}})$. See Supplemental Material at **??** for details of the training performance of surrogate models.

Fig. 8 (a)-(c) show how predictions of the coefficients of Eq. (5) from our surrogate models (X-axis) match with the coefficients of Eq. (5) directly fitted based on DFT+CI-NEB results (Y-axis) for 500 test cases chosen randomly from the total 2500 DFT+CI-NEB calculations. With the predicted coefficients, the values of ΔE, D_{MEP}, and ΔE_{a} can then be calculated based on Eq. (5) and Fig. 4 (a). Fig. 8 (d)-(f) compare these predicated values from surrogate models (X-axis) with ΔE, D_{MEP}, and ΔE_{a} directly from DFT+CI-NEB calculations (Y-axis). All plots in Fig. 8 indicate accurate matches between the surrogate model predictions and DFT+CI-NEB calculations, with low RMSE and high R^2 values (close or larger than 90%). Particularly, the

RMSE values of both ΔE and ΔE_a values are less than 0.04 eV, indicating our surrogate models can give precise descriptions of both the energetic driving force and energy barrier of vacancy migrations in complex local chemical environments. Using the one-hot encoding methods to describe the local lattice occupations as the inputs, these surrogate models can be easily implemented into kMC simulations for studies of early-stage precipitation kinetics in Al-Mg-Zn alloys.

IV. DISCUSSION AND CONCLUSIONS

Several previous studies support the generality of our studies of lattice distortion effects on vacancy migration barriers. For example, Eq. (9) is similar to the general linear correlation between ΔE_a and $a_l^3 B_0$ (a_l is the lattice constant and B_0 is the bulk modulus) for many pure metals with stable (such as Al, Ni, Cu, and Pt) or metastable (such as Fe and Ti) FCC structures[55, 56]. Both k_f and B_0 are related to second derivatives of the energy landscape at local-minimum states. As another example, correlations between site distortions and Li-ion migration barriers and correlations between Li-ion vibrational frequencies and Li-ion migration barriers were found separately in superionic conductors with antiperovskite structures (related to FCC lattice)[57]. Yet another example is that an equation similar to Eq. (9) was proposed to estimate the local free energy barriers in glass materials[58]. These results suggest Eq. (5) and Eq. (9) can be applied to atomic migrations in many other materials with FCC and similar crystal structures if each migration MEP only has one local energy maximum as plotted in Fig. 1 (c). Thus, not only are these equations (Eq. (5) and Eq. (9)) and the related surrogate models suitable for describing the energetics of vacancy migrations in multicomponent Al alloys, but they can also be applied in other multicomponent alloys such as high entropy alloys (HEAs) and the related concept of complex concentrated alloys (CCAs), where there can be strong lattice distortion effects on diffusion kinetics due to fluctuations in local chemical compositions[5–7, 59, 60]. However, the validity and the detailed forms of these equations in the ionic compounds need further investigations due to the different characteristics of interatomic bonds compared with the metallic alloys studied here.

The surrogate models to predict coefficients of Eq. (5) can be further improved from different aspects. First, only the feature vectors related to 2-atom clusters have been considered. We have confirmed that the accuracy of the surrogate models can increase if the feature vectors related to 3-atom clusters are considered (the R^2 values of the predictions of ΔE can be more than 90% in these cases). Second, high-order methods other than the linear ridge regression can be applied to train the surrogate models. However, since these surrogate models will be implemented into kMC simulations, these improvement strategies may increase the computational cost significantly and impede the ability of the kMC simulations to study the relatively long-time and large-scale diffusion and precipitation kinetics. Thus, the trade-off between accuracy and efficiency should be carefully considered for the construction of these surrogate models. These decisions can be made if kMC simulations are performed and compared with experimental validations, which will be the subject of future research.

In addition, physical mechanisms (including the symmetry properties discussed in Sec. III C and Sec. III D) will be applied to discover more efficient approaches to construct the DFT+CI-NEB data set to train the surrogate models. For example, Eq. (9) provides a criterion to select the representative data with appropriate distributions of Δe_a and ΔE_a as the training data set. Last but not least, the generality of our surrogate models based on Eq. (5) for different alloy compositions should also be verified. We have performed the DFT+CI-NEB calculations and analyses of quaternary Al-Mg-Zn-X alloy systems, where X is the alloying element possibly affecting the vacancy migration kinetics. Our preliminary results show that surrogate models based on Eq. (5) can also accurately describe the MEPs and the related $\Delta E_a/\Delta E$ in these quaternary alloy systems, which will be discussed in our future work.

In summary, the major conclusions of this study are

1. DFT+CI-NEB calculations provide energy barriers ΔE_a and driving forces ΔE of many (> 1000) vacancy migration events in different local chemical environments within the face-centered cubic (FCC) lattices of Al-Mg-Zn alloys.

2. The widely applied Kinetic Ising model [10], which states $\Delta E_a = \Delta e_a + \frac{1}{2}\Delta E$ and Δe_a is a constant value for one type of migrating atom in different local chemical environments inside a lattice, is not broadly applicable to FCC alloys, such as multicomponent Al alloys (Al-Mg system and Al-Mg-Zn system). This is because of the local lattice distortion effects resulting from changes in the chemical environment experienced by a migrating atom. Only Zn atoms near vacancy cites in Al lattices induce small lattice distortions due to the relatively small size difference between Al and Zn atoms[49]. Alternatively, large fluctuations (~ 1 eV) of ΔE_a in Al-Mg and Al-Mg-Zn alloys originate primarily from changes in $\Delta e_a = \Delta E_a - \frac{1}{2}\Delta E$ due to local lattice distortion effects because of the relatively large size of Mg atoms compared with Al and Zn atoms[49]. Here Δe_a can be regarded as the transition-state energy (E_t) relative to the average of the initial-state (E_i) and final-state (E_f) energies [25].

3. Based upon comparisons with DFT+CI-NEB results, a quartic function of the reaction coordinate x, $E_{MEP}(x) = ax^4 + bx^3 + cx^2$, accurately describes the energy landscape of the minimum energy path

(MEP) for each vacancy migration event in the FCC lattice, where $E_{\mathrm{MEP}}(x)$ of a vacancy migration event only has a single maximum at the transition state.

4. The quartic functions of the MEPs in Al-Mg-Zn alloys suggest that Δe_{a} of all types of migrating atoms in the FCC lattice of Al can be approximated with $\Delta e_{\mathrm{a}} \approx \alpha k_f D^2$, where $\alpha \sim 0.022$ is a constant value. Here D is the distance of a migrating atom between two adjacent equilibrium positions and k_f is the average vibration spring constant of this atom at these two equilibrium positions. This relation provide a a significant speedup in estimating ΔE_{a} without computational costly CI-NEB calculations since k_f is calculated rapidly by displacing only the migrating atom from its equilibrium positions.

5. Surrogate models using local lattice occupations as inputs were trained to predict the coefficients of the quartic function. The quartic function can then predict both ΔE_{a} and ΔE with the ab-initio accuracy but without additional DFT or CI-NEB calculations. The efficient and accurate predictions of ΔE_{a} and ΔE using these surrogate models will facilitate mesoscale studies, such as kinetic Monte Carlo simulations, of diffusional transformations that are critical for the processing and applications of Al-Mg-Zn-based and other FCC alloys, such as the solute clustering and early-stage precipitations during the natural aging of 7XXX series of Al alloys[28–31].

ACKNOWLEDGEMENT

This research is support by NSF-DMR-GOALI, Award Number: 1905421. The calculations were performed by using the Extreme Science and Engineering Discovery Environment (XSEDE) Stampede2 at the TACC through allocation TG-DMR190035.

[1] A. Borgenstam, L. Höglund, J. Ågren, and A. Engström, Dictra, a tool for simulation of diffusional transformations in alloys, Journal of phase equilibria **21**, 269 (2000).
[2] S. Pogatscher, H. Antrekowitsch, M. Werinos, F. Moszner, S. S. Gerstl, M. Francis, W. Curtin, J. F. Löffler, and P. J. Uggowitzer, Diffusion on demand to control precipitation aging: application to al-mg-si alloys, Physical Review Letters **112**, 225701 (2014).
[3] A. Van der Venand G. Ceder, First principles calculation of the interdiffusion coefficient in binary alloys, Physical review letters **94**, 045901 (2005).
[4] M. Mantina, Y. Wang, L. Chen, Z. Liu, and C. Wolverton, First principles impurity diffusion coefficients, Acta Materialia **57**, 4102 (2009).
[5] Y. N. Osetsky, L. K. Beland, and R. E. Stoller, Specific features of defect and mass transport in concentrated fcc alloys, ACTA MATERIALIA **115**, 364 (2016).
[6] S. Zhao, G. M. Stocks, and Y. Zhang, Defect energetics of concentrated solid-solution alloys from ab initio calculations: Ni0.5co0.5, ni0.5fe0.5, ni0.8fe0.2 and ni0.8cr0.2, PHYSICAL CHEMISTRY CHEMICAL PHYSICS **18**, 24043 (2016).
[7] S. L. Thomasand S. Patala, Vacancy diffusion in multi-principal element alloys: The role of chemical disorder in the ordered lattice, ACTA MATERIALIA **196**, 144 (2020).
[8] E. Clouet, M. Nastar, and C. Sigli, Nucleation of al 3 zr and al 3 sc in aluminum alloys: From kinetic monte carlo simulations to classical theory, Physical Review B **69**, 064109 (2004).
[9] G. Shaand A. Cerezo, Kinetic monte carlo simulation of clustering in an al–zn–mg–cu alloy (7050), Acta materialia **53**, 907 (2005).
[10] F. Soisson, C. Becquart, N. Castin, C. Domain, L. Malerba, and E. Vincent, Atomistic kinetic monte carlo studies of microchemical evolutions driven by diffusion processes under irradiation, Journal of Nuclear Materials **406**, 55 (2010).
[11] H. Miyoshi, H. Kimizuka, A. Ishii, and S. Ogata, Temperature-dependent nucleation kinetics of guinier-preston zones in al–cu alloys: An atomistic kinetic monte carlo and classical nucleation theory approach, Acta Materialia **179**, 262 (2019).
[12] K. Elder, M. Katakowski, M. Haataja, and M. Grant, Modeling elasticity in crystal growth, PHYSICAL REVIEW LETTERS **88**, 10.1103/PhysRevLett.88.245701 (2002).
[13] V. Fallah, B. Langelier, N. Ofori-Opoku, B. Raeisinia, N. Provatas, and S. Esmaeili, Cluster evolution mechanisms during aging in al-mg-si alloys, ACTA MATERIALIA **103**, 290 (2016).
[14] J. Li, S. Sarkar, W. T. Cox, T. J. Lenosky, E. Bitzek, and Y. Wang, Diffusive molecular dynamics and its application to nanoindentation and sintering, PHYSICAL REVIEW B **84**, 10.1103/PhysRevB.84.054103 (2011).
[15] L. Messina, M. Nastar, T. Garnier, C. Domain, and P. Olsson, Exact ab initio transport coefficients in bcc Fe $-$ x (x = Cr, Cu, Mn, Ni, p, Si) dilute alloys, Phys. Rev. B **90**, 104203 (2014).
[16] H. Wu, T. Mayeshiba, and D. Morgan, High-throughput ab-initio dilute solute diffusion database, Scientific data **3**, 1 (2016).
[17] T. Rautiainenand A. Sutton, Influence of the atomic diffusion mechanism on morphologies, kinetics, and the mechanisms of coarsening during phase separation, Physical Review B **59**, 13681 (1999).
[18] J. M. Sanchez, F. Ducastelle, and D. Gratias, Generalized cluster description of multicomponent systems, Physica A: Statistical Mechanics and its Applications **128**, 334 (1984).
[19] X. Zhangand M. H. Sluiter, Cluster expansions for thermodynamics and kinetics of multicomponent alloys, Journal of Phase Equilibria and Diffusion **37**, 44 (2016).

[20] E. Vincent, C. Becquart, C. Pareige, P. Pareige, and C. Domain, Precipitation of the fecu system: A critical review of atomic kinetic monte carlo simulations, Journal of Nuclear Materials **373**, 387 (2008).

[21] C. Pareige, M. Roussel, S. Novy, V. Kuksenko, P. Olsson, C. Domain, and P. Pareige, Kinetic study of phase transformation in a highly concentrated fe–cr alloy: Monte carlo simulation versus experiments, Acta Materialia **59**, 2404 (2011).

[22] F. Soissonand G. Martin, Monte carlo simulations of the decomposition of metastable solid solutions: Transient and steady-state nucleation kinetics, Physical Review B **62**, 203 (2000).

[23] F. Soissonand C.-C. Fu, Cu-precipitation kinetics in α-fe from atomistic simulations: Vacancy-trapping effects and cu-cluster mobility, Physical Review B **76**, 214102 (2007).

[24] C. Danielsand P. Bellon, Hybrid kinetic monte carlo algorithm for strongly trapping alloy systems, Computational Materials Science **173**, 109386 (2020).

[25] A. Van der Ven, G. Ceder, M. Asta, and P. Tepesch, First-principles theory of ionic diffusion with nondilute carriers, Physical Review B **64**, 184307 (2001).

[26] J. G. Goiri, S. K. Kolli, and A. Van der Ven, Role of short-and long-range ordering on diffusion in ni-al alloys, Physical Review Materials **3**, 093402 (2019).

[27] P. V. Liddicoat, X.-Z. Liao, Y. Zhao, Y. Zhu, M. Y. Murashkin, E. J. Lavernia, R. Z. Valiev, and S. P. Ringer, Nanostructural hierarchy increases the strength of aluminium alloys, Nature Communications **1**, 63 (2010).

[28] G. Shaand A. Cerezo, Early-stage precipitation in al–zn–mg–cu alloy (7050), Acta Materialia **52**, 4503 (2004).

[29] S. Liu, C. Li, S. Han, Y. Deng, and X. Zhang, Effect of natural aging on quench-induced inhomogeneity of microstructure and hardness in high strength 7055 aluminum alloy, Journal of Alloys and Compounds **625**, 34 (2015).

[30] W. Huo, L. Hou, Y. Zhang, and J. Zhang, Warm formability and post-forming microstructure/property of high-strength aa 7075-t6 al alloy, Materials Science and Engineering: A **675**, 44 (2016).

[31] A. Chatterjee, L. Qi, and A. Misra, In situ transmission electron microscopy investigation of nucleation of gp zones under natural aging in al-zn-mg alloy, Scripta Materialia **207**, 114319 (2022).

[32] C. Wolverton, Solutevacancy binding in aluminum, Acta Materialia **55**, 5867 (2007).

[33] H. Zuroband H. Seyedrezai, A model for the growth of solute clusters based on vacancy trapping, Scripta Materialia **61**, 141 (2009).

[34] M. Werinos, H. Antrekowitsch, T. Ebner, R. Prillhofer, W. Curtin, P. J. Uggowitzer, and S. Pogatscher, Design strategy for controlled natural aging in al–mg–si alloys, Acta Materialia **118**, 296 (2016).

[35] G. Henkelman, B. P. Uberuaga, and H. Jónsson, A climbing image nudged elastic band method for finding saddle points and minimum energy paths, The Journal of chemical physics **113**, 9901 (2000).

[36] G. Henkelmanand H. Jónsson, Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points, The Journal of chemical physics **113**, 9978 (2000).

[37] G. Kresseand J. Furthmüller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, Computational materials science **6**, 15 (1996).

[38] G. Kresseand J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, Physical review B **54**, 11169 (1996).

[39] P. BLOCHL, Projector augmented-wave method, PHYSICAL REVIEW B **50**, 17953 (1994).

[40] J. Perdew, K. Burke, and M. Ernzerhof, Generalized gradient approximation made simple, PHYSICAL REVIEW LETTERS **77**, 3865 (1996).

[41] L. Berg, J. Gjønnes, V. Hansen, X. Li, M. Knutson-Wedel, D. Schryvers, and L. Wallenberg, Gp-zones in al–zn–mg alloys and their role in artificial aging, Acta materialia **49**, 3443 (2001).

[42] I. Zhuravlev, S. Barabash, J. An, and K. Belashchenko, Phase stability, ordering tendencies, and magnetism in single-phase fcc au-fe nanoalloys, Physical Review B **96**, 134109 (2017).

[43] H. J. Monkhorstand J. D. Pack, Special points for brillouin-zone integrations, Physical review B **13**, 5188 (1976).

[44] A. Zunger, S.-H. Wei, L. Ferreira, and J. E. Bernard, Special quasirandom structures, Physical Review Letters **65**, 353 (1990).

[45] D. Sheppard, R. Terrell, and G. Henkelman, Optimization methods for finding minimum energy paths, The Journal of chemical physics **128**, 134106 (2008).

[46] E. Wigner, The transition state method, Transactions of the Faraday Society **34**, 29 (1938).

[47] C. Schmidtand J. Bocquet, Calculation of the diffusion parameters in an ordered ni3al-alloy for a relaxed lattice, MRS Online Proceedings Library (OPL) **527** (1998).

[48] E. Parzen, On estimation of a probability density function and mode, The annals of mathematical statistics **33**, 1065 (1962).

[49] J. C. Slater, Atomic radii in crystals, The Journal of Chemical Physics **41**, 3199 (1964).

[50] L. Landauand E. Lifshitz, *Statistical Physics* (Butterworth-Heinemann, 1980).

[51] T. Lindseyand B. Fultz, Microstructural dependence of vacancy diffusion in ordered alloys, Journal of applied physics **75**, 1467 (1994).

[52] T. Rautiainen, *Modelling microstructural evolution in binary alloys*, Ph.D. thesis, University of Oxford (1998).

[53] T. Xieand J. C. Grossman, Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties, Physical review letters **120**, 145301 (2018).

[54] Z. Chen, N. Andrejevic, T. Smidt, Z. Ding, Q. Xu, Y.-T. Chi, Q. T. Nguyen, A. Alatas, J. Kong, and M. Li, Direct prediction of phonon density of states with euclidean neural networks, Advanced Science **8**, 2004214 (2021).

[55] C. P. Flynn, Atomic migration in monatomic crystals, Phys. Rev. **171**, 682 (1968).

[56] T. Angsten, T. Mayeshiba, H. Wu, and D. Morgan, Elemental vacancy diffusion database from high- throughput first-principles calculations for fcc and hcp structures, New Journal of Physics **88**, 015018 (2014).

[57] R. Chen, Z. Xu, Y. Lin, B. Lv, S.-H. Bo, and H. Zhu, Influence of structural distortion and lattice dynamics on li-ion diffusion in li3ocl1xbrx superionic conductors, ACS Applied Energy Materials **4**, 2107 (2021).

[58] R. HALLand P. WOLYNES, The aperiodic crystal picture and free-energy barriers in glasses, JOURNAL OF

CHEMICAL PHYSICS **86**, 2943 (1987).

[59] K.-Y. Tsai, M.-H. Tsai, and J.-W. Yeh, Sluggish diffusion in co-cr-fe-mn-ni high-entropy alloys, Acta Materialia **61**, 4887 (2013).

[60] Y. Osetsky, A. V. Barashev, L. K. Béland, Z. Yao, K. Ferasat, and Y. Zhang, Tunable chemical complexity to control atomic diffusion in alloys, npj Computational Materials **6**, 1 (2020).