

# Nonequilibrium *ab initio* molecular dynamics determination of Ti monovacancy migration rates in B1 TiN

D. Gambino,<sup>1,\*</sup> D. G. Sangiovanni,<sup>2,1</sup> B. Alling,<sup>1,3</sup> and I. A. Abrikosov<sup>1,4</sup>

<sup>1</sup>*Department of Physics, Chemistry, and Biology (IFM) Linköping University, SE-58183 Linköping, Sweden*

<sup>2</sup>*ICAMS, Ruhr-Universität Bochum, D-44780 Bochum, Germany*

<sup>3</sup>*Max-Planck-Institut für Eisenforschung GmbH, D-402 37 Düsseldorf, Germany*

<sup>4</sup>*Materials Modeling and Development Laboratory, National University of Science and Technology "MISIS," 119049 Moscow, Russia*

(Received 8 March 2017; revised manuscript received 29 June 2017; published 18 September 2017)

We use the color diffusion (CD) algorithm in nonequilibrium (accelerated) *ab initio* molecular dynamics simulations to determine Ti monovacancy jump frequencies in NaCl-structure titanium nitride (TiN), at temperatures ranging from 2200 to 3000 K. Our results show that the CD method extended beyond the linear-fitting rate-versus-force regime [Sangiovanni *et al.*, *Phys. Rev. B* **93**, 094305 (2016)] can efficiently determine metal vacancy migration rates in TiN, despite the low mobilities of lattice defects in this type of ceramic compound. We propose a computational method based on gamma-distribution statistics, which provides unambiguous definition of nonequilibrium and equilibrium (extrapolated) vacancy jump rates with corresponding statistical uncertainties. The acceleration-factor achieved in our implementation of nonequilibrium molecular dynamics increases dramatically for decreasing temperatures from 500 for  $T$  close to the melting point  $T_m$ , up to 33 000 for  $T \approx 0.7 T_m$ .

DOI: [10.1103/PhysRevB.96.104306](https://doi.org/10.1103/PhysRevB.96.104306)

## I. INTRODUCTION

Transition-metal (TM) nitride refractory ceramics possess outstanding physical and mechanical properties, including extreme hardness [1–3], high toughness [4–6], thermal stability [7,8], chemical inertness [9,10], and good electrical conductivity [2,11], which renders them important for a large variety of applications ranging from wear-resistant protective coatings on tools employed in industrial machining [12,13] to diffusion barrier in electronic devices [14,15].

TM nitride binary systems can maintain the NaCl structure even for large deviations from stoichiometry [16,17], which is primarily regulated by the concentration of nitrogen and/or metal vacancies [18]. Kinetic control of (N/TM) ratios during crystal growth is exploited to tune the structural [19,20], electrical [21,22], optical [23,24], and mechanical properties [25–27] of these ceramics. Hence, clarification of the mechanisms governing vacancy formation and evaluation of vacancy diffusivities as a function of temperature may allow determining optimal synthesis conditions to deposit TM nitride thin films with desired properties.

To date, mass transport in TM nitride systems is poorly understood. Experimental measurements of atomic diffusivities generally employ isotopes. Although scattered over several orders of magnitude, values for nitrogen defect diffusion parameters in TM nitrides have been reported in various experimental works [8,28,29]. In contrast, experimental data for metal-element self-diffusivities are lacking for most of these ceramics. This motivates us to undertake computational investigation of these challenging problems. Moreover, even though nitrogen vacancies are the most common type of defect in binary TM nitrides, experiments indicate that migration of metal vacancies may be the mechanism which primarily controls spinodal decomposition of (Ti,Al)N pseudobinary

alloys at high (>1000 K) temperatures [30,31]. Titanium nitride (TiN), the prototype and most studied among all TM nitrides [32], crystallizes in the cubic B1 lattice structure over a wide range of temperatures (from 0 K up to its melting point,  $T_m \approx 3250$  K [33]) and a wide range of stoichiometries  $\text{TiN}_x$ , with  $0.6 < x < 1.2$  [16]. In this work, TiN is chosen as representative TM nitride system to investigate the diffusion properties of metal monovacancies in the dilute limit.

*Ab initio* density functional theory (DFT) calculations combined with the nudged elastic band (NEB) [34,35] or the string method [36,37] are typically employed to estimate minimum energy paths and migration energies at 0 K [38–43]. Kinetic properties are extrapolated to finite temperatures using transition state theory (TST) [44] by assuming essentially fully harmonic lattice vibrations, or employing quasiharmonic approximations [45,46]. These approaches, however, are not applicable to crystal phases which are unstable at 0 K (e.g., Group-VB TM nitrides [47,48]) and may yield inaccurate predictions when the role of anharmonic lattice vibrations becomes relevant [49–51]. Quantitatively reliable evaluation of kinetically controlled properties at given temperatures and pressures of interest necessarily requires the use of computer simulations reproducing atomic trajectories. *Ab initio* molecular dynamics (AIMD) inherently addresses the problems mentioned above by integrating Newton's equations of motion for each atom in a system, obtaining reliable interatomic forces using DFT [52]. AIMD allows for direct visualization of reaction pathways and provides corresponding kinetic rates. In addition, this technique has revealed nonintuitive system configurations and reaction pathways at finite temperatures [53–55]. Nevertheless, given that AIMD is highly computationally-intensive, it has been employed in very few theoretical studies of mass transport in TM nitrides [53–58], materials characterized by inherently low defect mobilities, at least within the metal sublattice [8].

The problem of evaluating vacancy migration frequencies via computer simulations has been undertaken decades ago

\*davide.gambino@liu.se

by Charles Bennett [59]. Subsequently, several theoretical methods have been developed with the aim of determining the rate of *rare* events within feasible computational times [60]. These approaches are often based on phase-space sampling (PSS) or selective acceleration of a process of interest. Some of these techniques give direct access to kinetic properties, while others as, e.g., the string method at 0 K [36] and finite temperatures [61,62], metadynamics [63], and canonical adiabatic free-energy sampling [64], which are useful for accurately determining minimum-energy paths and activation energies, need to be combined with TST or other theories [61] to extract reaction rates.

Transition path sampling (TPS) [65–67] and transition interface sampling (TIS) [68,69] are examples of techniques employing the PSS. Both can provide the occurrence-probability of reaction pathways connecting known initial and final states by following any possible trajectory. TPS and TIS are practically useful to identify transition-state configurations on complex energy landscapes and/or when various competitive reactions may take place [70]. Other methods based on the PSS are, e.g., forward flux sampling [71–73], partial path sampling [74], and (based on similar principles) milestoneing [75,76].

Techniques based on the acceleration of a process of interest through a fictitious bias can yield speed-up factors of up to  $\times 10^9$  [77], but may also be affected by nonphysical behaviors or require non-trivial definition of collective variables [77]. Among these, we find metadynamics-based approaches [78], hyperdynamics [79], a combination of the latter two methods [77], temperature-accelerated molecular dynamics [80], and the color diffusion (CD) algorithm [81,82], which is an extension of the tagged-particle method [83].

Supercomputer simulations can be parallelized to investigate the physical properties of relatively large systems. However, supercomputing capacity alone cannot circumvent the limits of time scales. For this, the parallel replica method [84] offers a strategy in which, by exploiting the properties of stochastic processes, efficient evaluation of kinetic rates of rare events can be accomplished without having to apply external biases.

Some of the techniques listed above, implemented in the framework of AIMD, provide considerable gain in computational efficiency in comparison with nonaccelerated *brute-force* AIMD simulations, while maintaining the same level of accuracy. The convenience of using one or the other acceleration method may vary depending on time and length scales involved, as well as on the complexity of the process under investigation [60]. In particular, we have recently proposed a development for the CD algorithm [85], which allows to retrieve monovacancy migration rates in crystalline solids at much lower computational efforts in comparison with the former (yet more general) application [81].

Nonequilibrium molecular dynamics schemes, including CD simulations, entail applying external bias constraints to assist diffusion of species of interest, while rapidly dissipating the excess energy introduced in the system by controlling (rescaling) the kinetic energies of the unbiased particles. Thus, knowing the dependence of accelerated reaction rates on the magnitude of the external constraint allows retrieving equilibrium properties at the limit for vanishing biases. In the original formulation of CD-accelerated dynamics [81,82,86],

the migration velocities of diffusing species of interest are boosted by mean of an external constant force field  $\vec{F}$ , using in most cases force intensities  $F$  for which the nonequilibrium jump frequency  $k_{\text{NE}}(F)$  dependence on  $F$  remains linear. Equilibrium diffusivities are retrieved to zero field by linear fitting of  $k_{\text{NE}}(F)$  versus  $F$  data. The method allows discovering diffusion pathways without any initial assumption.

The technique that we describe in Ref. [85] is instead specifically tailored to determining the kinetics of monoatomic mass-transport pathways occurring over a single energy barrier, where initial and final states are known *a priori*. Our implementation of CD in nonequilibrium AIMD (NE-AIMD) yields accurate results for vacancy diffusion in bcc Mo and, by exploiting force-field intensities well beyond the linear-response regime [81], provides speed-up factors which increase by several order of magnitudes with decreasing simulation temperatures [85]. The fact that our current method is focused on quantifying the kinetics of a relatively narrow class of diffusion processes does not preclude the possibility of extending its usage to a broader range of problems in practical applications.

The rapid increase in computational efficiency achieved via CD simulations for decreasing temperatures [85] comes to the cost of much broader uncertainty ranges on extrapolated equilibrium-rate values. Thus, it is necessary to provide an unambiguous and reliable definition of statistical error bars. In Ref. [85], we do not include any rigorous treatment of statistical uncertainties on accelerated  $k_{\text{NE}}(F)$  rates which, in turn, affect the predictions of lattice-defect jump frequencies  $k_{\text{NE}\rightarrow\text{E}}$ . Here, we propose a procedure based on gamma-distribution statistics which provides unambiguous definition of  $k_{\text{NE}}(F)$  and  $k_{\text{NE}\rightarrow\text{E}}$  average values with corresponding confidence ranges. Results of the present work show that our implementation of the CD formalism in AIMD (i) allows calculating metal-vacancy migration rates in TiN, a binary compound material characterized by high thermal stability, within reasonable computational times, (ii) provides systematic reductions on the uncertainty of equilibrium jump frequencies  $k_{\text{NE}\rightarrow\text{E}}$  upon increasing the number and/or accuracy of  $k_{\text{NE}}$  interpolation points, and (iii) enhances the efficiency of AIMD simulations by several orders of magnitude for evaluation of vacancy migration rates at temperatures well below melting.

The paper is organized as follows. Section II A describes the DFT computational methods employed for calculating metal vacancy formation and interaction energies, as well as 0 K vacancy diffusion pathways with corresponding migration energies in bulk TiN. Section II B briefly presents the CD algorithm and summarizes the scheme used in equilibrium and non-equilibrium AIMD simulations of vacancy diffusion. Sections II C and II D detail the statistical analysis and computational procedure employed for determination of non-equilibrium jump rates with corresponding uncertainties. Section III A presents DFT results of vacancy formation and interaction energies, whereas Sec. III B is dedicated to the results of non-equilibrium AIMD simulations. In Sec. III C, we discuss the gain in computational efficiency provided by our implementation of the CD algorithm. The conclusions of this work are presented in Sec. IV.

## II. COMPUTATIONAL DETAILS AND THEORETICAL METHODS

DFT calculations and DFT-based AIMD simulations are carried out using the Projector Augmented Wave (PAW) method [87] as implemented in the Vienna *ab initio* simulation package (VASP) and the Armiento-Mattsson (AM05) functional for the description of one-electron exchange-correlation potential and energy [88]. The accuracy of self-consistent calculations is  $10^{-5}$  eV/supercell. The energy cutoff used for the plane-wave basis set and the density of  $k$ -point grids employed for Brillouin-zone sampling depend on the property under analysis (see below). In static calculations, atomic positions and cell shape are optimized via conjugate-gradient energy minimization. Equilibrium volumes are obtained by least squares fitting of energy versus volume curves.

### A. Vacancy formation, interaction, and migration energies at temperature $T = 0$ K

DFT calculations of vacancy formation and interaction energies, as well as Ti vacancy migration, employ defect-free and defective TiN supercell structures consisting of  $4 \times 4 \times 4$   $B1$  conventional unit cells, for a total of 256 metal and 256 nitrogen sites. The use of an energy cutoff of 400 eV and  $3 \times 3 \times 3$  Monkhorst-Pack [90]  $k$ -point meshes for Brillouin zone integration ensures convergence of calculated energies within  $\sim 0.1$  meV/atom. As shown below (see Sec. III A), our supercells should be large enough to avoid vacancy self-interactions.

The formation energy  $E_{nV}^f$  of  $n$  Ti vacancies in  $B1$  TiN is calculated as

$$E_{nV}^f = E_{nV} + n \cdot E_{\text{hcp Ti}} - E_0, \quad (1)$$

where  $E_{nV}$  is the total energy of the system with  $n$  vacancies,  $E_{\text{hcp Ti}}$  is the energy of one Ti atom in hexagonal close-packed (hcp) titanium (ground-state structure of titanium), and  $E_0$  is the total energy of the defect-free TiN system. The interaction energy  $E_{2V}^i$  between two Ti vacancies in TiN is obtained as a function of the vacancy/vacancy separation distance  $d_V$  by using the expression

$$E_{2V}^i(d_V) = E_{2V}^f(d_V) - 2 \cdot E_{1V}^f. \quad (2)$$

Positive (negative) interaction energy values indicate repulsive (attractive) vacancy/vacancy interactions.

Recent classical and *ab initio* molecular dynamics simulations of  $N$  point-defect diffusion in bulk TiN demonstrated that  $N$  monovacancy migration is the primary mechanism for transport of  $N$  atoms [54]. Our present theoretical investigations focus on migration of Ti monovacancies at the dilute limit, as this is expected to be the most important reaction pathway for diffusion of metal atoms in overstoichiometric (TM/ $N$  ratio  $< 1$ ) bulk TiN. It should be noted, however, that previous 0 K *ab initio* calculations indicate that the formation of a Ti-N divacancy complex is energetically more favored with respect to a pair of isolated  $N$  and Ti monovacancies [91]. Our work is a necessary prestep for any study of more complex types of diffusion.

The minimum energy path with corresponding migration energy of a Ti lattice-atom moving to a neighboring metal

vacancy is evaluated at 0 K through NEB [34,35] calculations as implemented in VASP. We employ the default spring parameter and nine images between prerelaxed initial and final states. For these calculations, we use an energy cutoff of 500 eV.

### B. AIMD and NE-AIMD simulations

Equilibrium and nonequilibrium AIMD is carried out in the NVT canonical ensemble, integrating the equations of motion at 1-fs time steps, while sampling the Brillouin zone  $\Gamma$  point. Simulation supercells contain 215 atoms ( $3 \times 3 \times 3$  TiN conventional  $B1$  unit cells with one Ti vacancy). We notice that TiN is an electrical conductor [89]. This implies long correlation lengths in the electronic structure of the system. Thus, one cannot rule-out the possibility that increasing supercell sizes and/or  $k$ -point mesh densities would have some quantitative effect on our results. However, this should not qualitatively affect the conclusions obtained from AIMD and NE-AIMD simulations. At each investigated temperature  $T$ , the supercell lattice parameter  $a_0(T)$  is obtained by rescaling the 0 K DFT value [ $a_0(0 \text{ K}) = 4.214 \text{ \AA}$ ] accounting for the experimental linear thermal expansion coefficient  $\alpha_L \approx 9 \times 10^{-6} \text{ K}^{-1}$  [33].

Prior to counting migration events in nonequilibrium simulations, AIMD system equilibration is performed for  $\sim 3$  ps controlling the temperature  $T$  via the Nosé-Hoover thermostat. Thus, the system is immersed in a  $\langle 110 \rangle$ -oriented constant force-field, which is used to accelerate one Ti lattice atom (colored-atom) toward the neighboring vacancy site (see details in Ref. [85]). The force-field acts on the colored-atom with intensity  $F$  and on each other atom in the simulation box with opposite forces of intensities  $F/(N-1)$  (where  $N$  is the total number of atoms in the supercell), thus maintaining the system in mechanical equilibrium (Fig. 1). The energy increase due to external work is rapidly dissipated by rescaling the velocities of all atoms but the colored one at each time step.

Equilibrium jump rates  $k_{\text{NE} \rightarrow \text{E}}(T)$  are extrapolated at the limit  $F \rightarrow 0$  by fitting nonequilibrium jump rates  $k_{\text{NE}}(F, T)$  determined at different force field intensities  $F$  with the expression (see Supplemental Material in Ref. [85] for the derivation):

$$\ln[k_{\text{NE}}(F, T)] = \ln[k_{\text{NE} \rightarrow \text{E}}(T)] + \frac{x_{\text{TS0}}(T)}{k_B T} F - \alpha(T) \cdot F^2. \quad (3)$$

In Eq. (3),  $k_B$  is the Boltzmann's constant,  $x_{\text{TS0}}(T)$  is the equilibrium transition-state position at a temperature  $T$ , and  $\alpha(T)$  is a fitting parameter. Under the assumption that the effective potential energy landscape along the migration path remains approximately sinusoidal at any  $T$ ,  $x_{\text{TS0}}(T)$  is conveniently set to half of the distance between the initial and final states of the diffusion process:  $x_{\text{TS0}}(T) = a_0(T)/(2\sqrt{2})$ . Having fixed the parameter  $x_{\text{TS0}}(T)$  reduces the minimum number (from three to two) of  $k_{\text{NE}}(F, T)$  interpolation points needed to retrieve  $k_{\text{NE} \rightarrow \text{E}}(T)$  [see Eq. (3)].  $k_{\text{NE} \rightarrow \text{E}}(T)$  are extrapolated to zero force-field at temperatures  $T = 2200, 2400, 2600, 2800,$  and  $3000$  K from  $k_{\text{NE}}(F, T)$  values calculated by employing  $\langle 110 \rangle$ -oriented force-fields with the two  $\langle 100 \rangle$  force-components  $F^{(100)}$  of equal intensities 1.4, 1.6, 1.8, 2.0, and 2.2 eV/ $\text{\AA}$ .

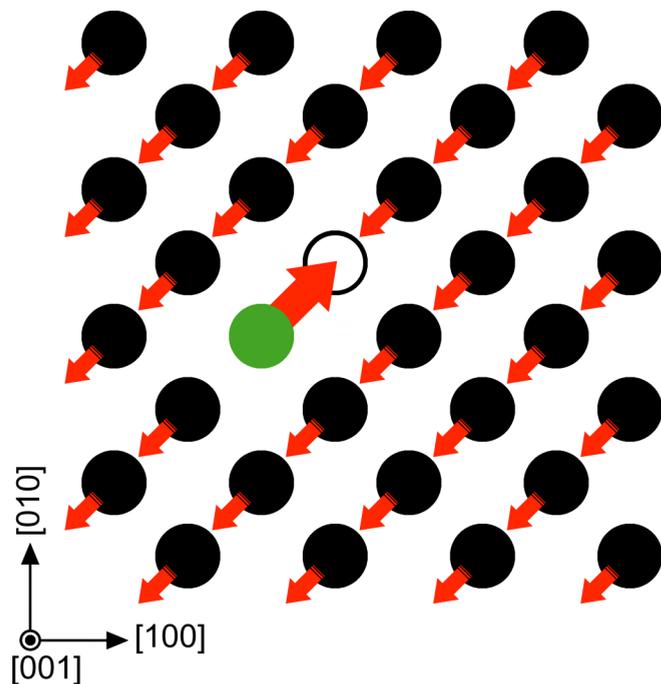


FIG. 1. Schematic illustration of the color-diffusion method applied to monovacancy migration in the metal fcc sublattice of  $B1$  TiN. The colored-atom (green circle) is accelerated toward the neighboring vacancy (empty circle) by a constant force of magnitude  $F$  (red arrow oriented along the  $[110]$  direction). Balancing forces (smaller red arrows) are applied to all other lattice atoms (black filled circles) to yield zero total force and zero torque on the system.

### C. Statistical analysis: Determination of $k_{NE}$ values and corresponding uncertainties

In Ref. [85], we presented a scheme to calculate nonequilibrium vacancy jump rates  $k_{NE}(F, T)$ . For each given  $F$ ,  $k_{NE}(F, T)$  was obtained as the ratio between the number of colored-atom jumps recorded over a set of independent sample-runs and the total simulation time, regardless of whether colored-atom jump was observed in all runs or not. Such scheme yields statistically meaningful average jump rates (as demonstrated in Appendix A), but does not provide clear definition of uncertainties on nonequilibrium migration frequencies, which ultimately determine the uncertainty  $\sigma_{NE \rightarrow E}$  on extrapolated equilibrium  $k_{NE \rightarrow E}$  values. Obtaining reliable  $\sigma_{NE \rightarrow E}$  error bars is of critical importance when studying reactions characterized by relatively large activation energies  $E_a/T$  (i.e., low  $T$  and/or high  $E_a$ ). For these cases, accelerated rates  $k_{NE}$  can be computed over feasible times only for  $F$  values close to  $F_{max}$ . This poses a question mark on the trustworthiness of extrapolated  $k_{NE \rightarrow E}$  values. In this paper, we introduce a method to define confidence ranges on  $k_{NE}$  and  $k_{NE \rightarrow E}$  via usage of gamma-distribution statistics.

In vacancy-mediated diffusion processes, one often considers atomic jumps as independent and uncorrelated events. The distribution of jump occurrence-times  $t$ , that is, time intervening between consecutive migration events, is described by an exponential probability density function  $\exp[-t]$  as numerically demonstrated in, for example, Ref. [92]. Here, we show that the distribution of colored-atom jump occurrence-

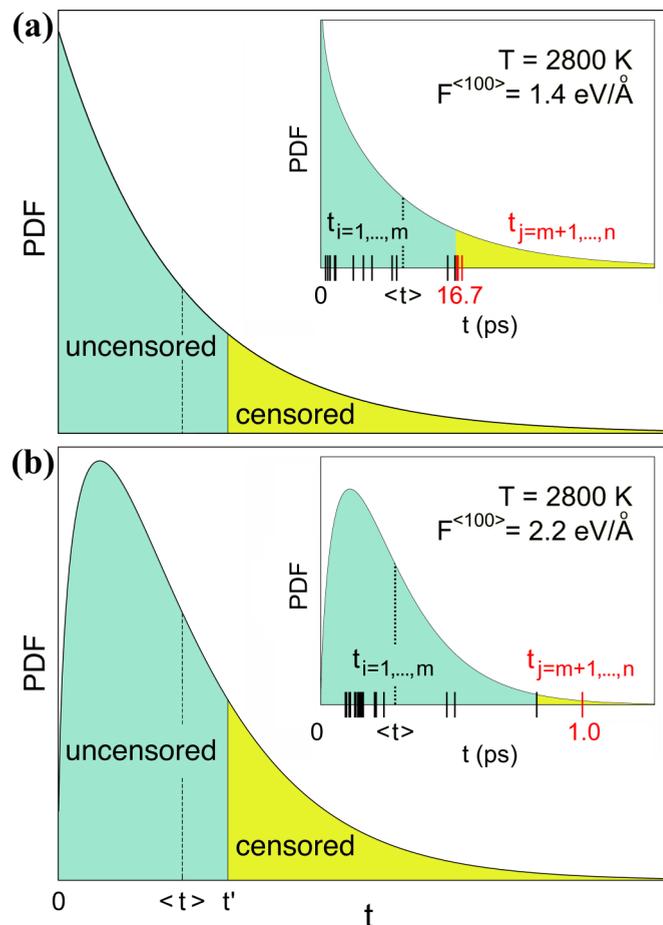


FIG. 2. Examples of gamma probability density functions with different shape ( $\lambda$ ) and scale ( $\theta$ ) parameters used to model the distribution of Ti vacancy jump occurrence-times  $t$  obtained from NE-AIMD simulations. Panel (a) shows an exponential distribution ( $\lambda = 1$ ) with  $\theta = 3$  time-units. The curve in panel (b) has shape and scale parameters  $\lambda = 1.5$  and  $\theta = 2$  time-units. Uncensored and censored data occupy different domains on the time scale  $t$  (between 0 and  $t'$  and from  $t'$  to  $+\infty$ , respectively). The value of  $t'$  is chosen such that the area delimiting uncensored data corresponds to 75% of the total probability. This ensures that the average occurrence-time ( $\langle t \rangle$ ) is within the range of uncensored data. The insets in (a) and (b) illustrate PDF obtained from NE-AIMD results at 2800 K for the smallest ( $1.4 \text{ eV/\AA}$ ) and largest ( $2.2 \text{ eV/\AA}$ ) force-field  $\langle 100 \rangle$ -components  $F^{(100)}$  employed in this work. Jump occurrence-times  $t_i$  (uncensored data) and the simulation times  $t_j$  of unsuccessful runs (censored data) are marked by black and red solid vertical ties on the time axis, respectively.

times obtained at force fields of modestly low intensities is well represented by exponential curves. However, for forces of intensities approaching the maximum value employed in the present simulations, the spread of  $t$  values, which clearly deviates from an exponential probability trend, closely resembles a gamma distribution, indicating that strong external biases increase the degree of correlation on migration times (Fig. 2).

The exponential distribution is a special case of the gamma distribution, a family of probability density functions (PDF) defined on domains  $t := (0, +\infty)$  [93,94] and characterized by

a shape parameter  $\lambda$  and a scale parameter  $\theta$  ( $\theta$  is expressed in units of time):

$$\text{PDF}_{\lambda,\theta}(t) = \frac{t^{\lambda-1} \cdot e^{-t/\theta}}{\theta^\lambda \cdot \Gamma(\lambda)}, \quad (4)$$

where  $\Gamma(\lambda)$  is the gamma function  $\Gamma(\lambda) = \int_0^{+\infty} x^{\lambda-1} \cdot e^{-x} dx$ . For  $\lambda = 1$ , an exponential PDF is recovered, and  $\theta$  remains the only fitting parameter. For  $\lambda > 1$ , the gamma distribution has 0 probability density at  $t = 0$ , exhibits a maximum for  $t > 0$  and vanishes at the limit for  $t \rightarrow +\infty$ . For  $\lambda \gg 0$ , the distribution becomes approximately symmetric, that is, similar to a standard distribution. The mean value of gamma distributions  $\langle t \rangle$  (average occurrence time of the process considered) is given by  $\langle t \rangle = \lambda \theta$ .

Let us assume that performing  $n$  independent simulations (or experiments), one records a process of interest in  $m$  ( $< n$ ) of the cases, while no event of interest is observed in the remaining  $[n-m]$  cases. The  $m$  successful observations (labeled as *uncensored* data, set of runs  $i$ ) occur with times  $t_i := \{t_1, t_2, \dots, t_m\}$ , whereas  $[n-m]$  unsuccessful tests (*censored* data, set of simulations  $j$ ) are terminated after times  $t_j := \{t_{m+1}, t_{m+2}, \dots, t_n\}$ . Average occurrence times  $\langle t \rangle$  and corresponding uncertainties  $\sigma_{\langle t \rangle}$  can be evaluated via fitting the distribution of  $t_i$  and  $t_j$  values, with a gamma PDF. Determination of  $\lambda$  and  $\theta$  parameters based on the  $n$  collected tests is performed with the Maximum Likelihood Estimation (MLE). A MLE entails solving a system of equations  $\partial[\ln(L)]/\partial\lambda = 0 \wedge \partial[\ln(L)]/\partial\theta = 0$  [93] to maximize the logarithm of the likelihood function

$$L(\lambda, \theta) = \underbrace{\prod_{i=1}^m [\text{PDF}_{\lambda,\theta}(t_i)]}_{\Pi_i} \cdot \underbrace{\prod_{j=m+1}^n \left[ 1 - \int_0^{t_j} \text{PDF}_{\lambda,\theta}(t) \cdot dt \right]}_{\Pi_j}. \quad (5)$$

In Eq. (5), the products  $\Pi_i$  and  $\Pi_j$  implicitly assign different statistical meaning to the outcome of  $i$  and  $j$  sets of simulations.

Imposing  $\lambda = 1$  in MLE, as done in Ref. [92], corresponds to evaluating  $k_{\text{NE}}$  via exponential distribution statistics. The mathematical demonstration contained in Appendix A shows that average occurrence times calculated as the mean of an exponential PDF are equivalent to the  $\langle t \rangle$  values obtained with the scheme employed in our previous work [85]. Letting, instead, both  $\theta$  and  $\lambda$  in Eq. (5) as free adjustable parameters provides higher degrees of freedom for optimizing the representation of NE-AIMD jump occurrence-time distributions at any  $F$ .

#### D. Evaluation of nonequilibrium jump rates and corresponding uncertainties

This subsection details the procedure used to calculate accelerated jump frequencies  $k_{\text{NE}}$  and related confidence ranges  $\sigma_{\text{NE}}$  via NE-AIMD.

For each force field intensity  $F$ , average accelerated jump rates

$$k_{\text{NE}}(F) \propto \langle t(F) \rangle^{-1} = [\lambda(F)\theta(F)]^{-1}, \quad (6)$$

with corresponding uncertainties

$$\sigma_{\text{NE}}(F) = k_{\text{NE}}[(\sigma_\lambda/\lambda)^2 + (\sigma_\theta/\theta)^2 + 2\sigma_{\lambda\theta}/(\lambda\theta)]^{1/2}, \quad (7)$$

where  $\sigma_{\lambda\theta}$  is a covariance, are evaluated from the outcome of  $n$  NE-AIMD runs, which satisfy

$$n \geq 15, \quad (8a)$$

$$m \geq 0.75n, \quad (8b)$$

$$t_j > t' \geq t_i \forall \{i \in [1, m] \wedge j \in [m+1, n]\}. \quad (8c)$$

We remind the reader that  $m$  is the number of runs terminating with the colored-atom jumping into the vacancy at times  $t_i$ , with  $i := \{1, \dots, m\}$ . The remaining  $(n-m)$  *unsuccessful* runs end at times  $t_j$ , with  $j := \{m+1, \dots, n\}$ . Imposing condition (8c) comports that censored and uncensored data collected by NE-AIMD simulations are represented on two distinguished  $t$  domains (see Fig. 2). Colored-atom jump occurrence-times  $t_i$  are represented by the gamma distribution curve within a  $(0, t']$  interval, while the simulation times in which no colored-atom jump has occurred are gathered in the  $(t', +\infty)$  curve tail. Overall, requiring that NE-AIMD runs satisfy condition (8a)–(8c) renders  $\langle t \rangle < t'$  and ensures accuracy of estimated  $k_{\text{NE}}(F)$  values.

For each simulation temperature  $T$ , 50 AIMD runs are initialized with atomic velocities randomly selected from the Boltzmann-Maxwell distribution, requiring null drift from the center of mass and zero torque acting on the system. After equilibration ( $\sim 3$  ps), the TiN+vacancy system is immersed in an external constant force field  $F$ . At each force field intensity, nonequilibrium simulations are performed for 15 (or more) configurations [condition (8a)] randomly chosen from the 50 equilibrated states. The simulations are restarted until colored-atom jump is recorded in (at least) 75% of the cases [condition (8b)]. In addition, each NE-AIMD run  $\tilde{j}$  in which colored-atom jump is not observed is continued until either (i) colored-atom migration occurs or (ii)  $t_j > t_i \forall i \in [1, m]$  [see condition (8c)].

$t_i$  and  $t_j$  data collected during  $n$  NE-AIMD simulations are fitted with the *censored*-gamma distribution by setting the lower boundary of the censored-data interval  $t'$  equal to the largest among all recorded  $t_i$  values (note that  $t'$  varies with  $F$ ). At all simulation temperatures, the curves resulting from MLE are typically close to exponential distributions for low  $F^{(100)}$  values (see, for example, the inset of Fig. 2(a), where  $\lambda = 0.894$  and  $\theta = 11.6$  ps), while deviate from exponential trends at high  $F^{(100)}$  (see the inset of Fig. 2(b), where  $\lambda = 1.69$  and  $\theta = 0.164$  ps). For this reason, we retain appropriate to leave both  $\lambda$  and  $\theta$  as free adjustable PDF parameters. Nevertheless, as described in Appendix B, constraining  $\lambda = 1$  (as done in Ref. [92]) returns essentially the same  $\langle t \rangle$  results. MLE are carried out with the software Rstudio [95].

Nonequilibrium jump rates can be calculated as  $k_{\text{NE}}(F) = 12\langle t(F) \rangle^{-1}$ , where the factor 12 is the number of nearest neighbors in a face-centered-cubic (fcc) lattice. This factor accounts for the fact that, in a real system, all nearest neighbors have equal probability of diffusing to the vacancy, whereas in our NE-AIMD runs, only colored atom jumps are statistically relevant. Equilibrium jump rates  $k_{\text{NE} \rightarrow \text{E}}$  are extrapolated from nonequilibrium  $k_{\text{NE}}(F)$  data [Eq. (3)]. Confidence ranges  $\sigma_{\text{NE} \rightarrow \text{E}}$  are evaluated by projection of  $\sigma_{\text{NE}}(F)$  uncertainties [53]. Briefly,  $\ln[\sigma_{\text{NE} \rightarrow \text{E}}]$  correspond to the widths of  $\ln[k_{\text{NE} \rightarrow \text{E}}]$  normal distributions obtained via Eq. (3) by fitting a large

TABLE I. DFT formation energies  $E^f$  calculated for one and two Ti vacancies placed on first, second, third, and fourth neighboring metal-sublattice shells with corresponding interaction energies  $E^i$ . Our results are compared with those of previous DFT investigations [91,97].

	$E^f$ (eV)	$E^i$ (eV)
One Ti vacancy	2.86 (2.76 [91])	–
Two Ti vacancies		
First neighbors	6.37	0.66 (0.68 [97])
Second neighbors	5.92	0.21 (0.31 [97])
Third neighbors	5.79	0.07
Fourth neighbors	5.78	0.06

number of accelerated rates  $\varepsilon_i$  versus  $F$  data sets, where  $\varepsilon_i$  values are picked in the range  $\ln[k_{\text{NE}}(F_i)/\sigma_{\text{NE}}(F_i)] \leq \varepsilon_i \leq \ln[k_{\text{NE}}(F_i) \times \sigma_{\text{NE}}(F_i)]$  with probabilities given by Gaussian distributions centered at  $\ln[k_{\text{NE}}(F_i)]$  and of standard deviations  $\ln[\sigma_{\text{NE}}(F_i)]$ . All the uncertainty intervals presented below for  $\ln[k]$  rates correspond to twice standard deviations, i.e., rate values are expressed as  $\ln[k] \pm 2 \cdot \ln[\sigma]$ .

Diffusion of atoms other than the colored one is extremely unlikely for systems, such as TiN, characterized by high thermal stability. In cases for which this occurs, the times  $t_q (q = m + 1, \dots, m + \varepsilon \leq n)$  at which an atom migrates into the vacancy are conveniently considered as censored data.

### III. RESULTS AND DISCUSSION

#### A. 0 K Ti vacancy formation and interaction energies in TiN

For compound materials such as TiN, metal vacancy formation energy values depend on the choice of the reference chemical potential  $\mu(\text{Ti})$ , energy of the reservoir of metal atoms with which the binary system exchanges particles at the equilibrium.  $\mu(\text{Ti})$  varies, in turn, as a function of the environmental conditions. By setting  $\mu(\text{Ti}) = E_{\text{hcp Ti}}$ , we obtain a 0 K DFT Ti vacancy ( $\text{Ti}_V$ ) formation energy  $E_V^f$  of 2.86 eV, in good agreement with another *ab initio* result, 2.76 eV, reported in the literature [91]. Comparison of  $E_V^f$  with  $E_{2V}^f$  or, more in general, with  $E_{nV}^f$  values allows predicting whether vacancy clustering is likely to occur.

Estimated Ti monovacancy ( $\text{Ti}_V$ ) formation energies  $E_V^f$  and  $\text{Ti}_V/\text{Ti}_V$  interaction energies  $E_{2V}^i$  in TiN are presented in Table I. DFT  $E_{2V}^i$  values, calculated with the two vacancies placed at first-, second-, third-, and fourth-neighbor metal-sublattice positions, are positive for any  $\text{Ti}_V/\text{Ti}_V$  separation distance  $d_V$  and decrease monotonically as  $\sim d_V^{-1}$  (Fig. 3). The fact that Ti vacancies repel each other indicates instability of  $\text{Ti}_V/\text{Ti}_V$  pairs in bulk TiN. For distances  $d_V$  corresponding to  $\text{Ti}_V/\text{Ti}_V$  third and fourth neighbor configurations, the interaction energy reaches a saturation value close to zero. This means that, for  $d_V \geq 5 \text{ \AA}$ , the two Ti vacancies can be effectively considered as isolated/uncorrelated.

Although lattice vibrations induce variations in point-defect formation free-energies [49,96], thus affecting the relative stability of vacancies vs. interstitials, as well as the magnitude of vacancy/vacancy interaction energies at finite temperatures, the observed strong short-range  $\text{Ti}_V/\text{Ti}_V$  repulsion energy

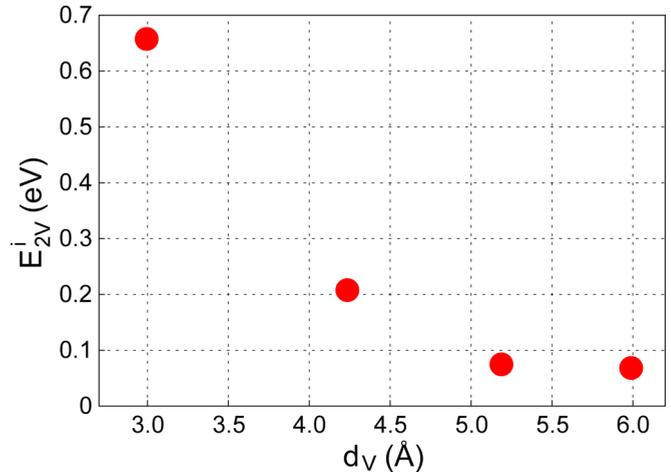


FIG. 3.  $\text{Ti}_V/\text{Ti}_V$  repulsion energies in TiN (Table I) plotted as a function of the vacancy/vacancy distance  $d_V$ .

(Fig. 3) suggests that  $\text{Ti}_V$  agglomeration in TiN is unlikely to occur at any temperature. In addition, considering that self-interstitial formation is thermodynamically much less favored than monovacancy formation in TM nitrides [97], it is reasonable to assume that  $\text{Ti}_V$  are the most common metal point-defect in bulk TiN at any temperature.

#### B. Ti vacancy equilibrium jump rates

DFT + NEB calculations carried out at  $T = 0 \text{ K}$  confirm that the energy landscape probed by a Ti lattice atom travelling toward a neighboring vacancy exhibits a sinusoidal-like curvature with a single transition state located half way between the initial and final configurations (Fig. 4). The NEB minimum energy path indicates that  $\text{Ti}_V$  migration in TiN is a straight jump along a  $\langle 110 \rangle$  direction. Assuming that the shape of

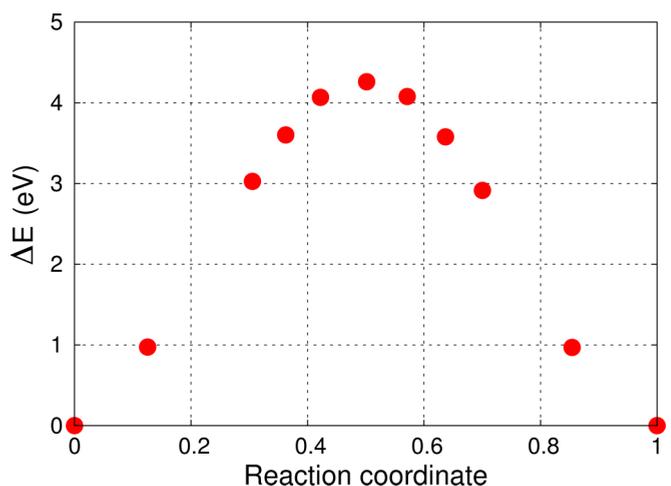


FIG. 4. DFT + NEB 0 K energy profile for Ti atom  $\langle 110 \rangle$  jump to a neighboring vacancy site. The reaction coordinate on the horizontal axis represents the fractional displacement of the Ti atom from its initial to final position. The 0 K activation energy along the minimum energy path is 4.26 eV.

the migration energy landscape is not qualitatively modified by lattice vibrations, the implementation of CD in AIMD simulations proposed in Ref. [85] can therefore be used to efficiently retrieve equilibrium jump frequencies as a function of temperature. Indeed, the CD algorithm [85] allows one to efficiently determine monovacancy equilibrium diffusivities  $k_{\text{NE} \rightarrow \text{E}}$  in crystalline solids via fitting nonequilibrium jump frequencies  $k_{\text{NE}}$  which increase exponentially with  $F$  [see Eq. (3) in Sec. II B]. Being often point-defect migration in crystalline solids a simple process involving a single saddle-point in energy along the reaction coordinate [98–101], Eq. (3) is based on the assumption that the potential energy landscape along the migration path has a sinusoidal-like shape (see supplemental material in Ref. [85]). Note, however, that the implementation of CD proposed in Ref. [85] may not be directly applicable to cases for which atomic diffusion is characterized by complex energy profiles and/or involves concerted migration of several atoms [102]. Results presented in Fig. 4 confirm that this is not the case for the diffusion process investigated here.

The DFT + NEB activation energy  $E_a^{\text{OK}}$  estimated at  $T = 0$  K for  $\text{Ti}_\text{V}$  diffusion, 4.26 eV, is larger than the *ab initio* value, 3.8 eV, obtained for N vacancy ( $\text{N}_\text{V}$ ) jump in TiN [40]. Using  $E_a = 4.26$  eV in the expression  $F_{\text{max}} \approx 0.75 \pi E_a / (2x_{\text{Tso}})$ , we estimate that the maximum force-field intensity  $F_{\text{max}}$  to be used in the present NE-AIMD simulations is  $\sim 3.3$  eV/Å (see details in the Supplemental Material of Ref. [85]).

NE-AIMD accelerated jump rates  $k_{\text{NE}}$  are computed at temperatures of 2200, 2400, 2600, 2800, and 3000 K (the TiN melting temperature is  $\sim 3250$  K) using (110)-oriented constant force fields at five different intensities: the  $\langle 100 \rangle$  force components range from 1.4 to 2.2 eV/Å in steps of 0.2 eV/Å. Fig. 5 shows that  $\ln[k_{\text{NE}}(F)]$  versus  $F$  data follow a parabolic trend. Irrespective of the temperature  $T$ , for  $F$  close to  $F_{\text{max}} (\approx 3.3$  eV/Å) all  $\ln[k_{\text{NE}}(T, F)]$  values are approximately equal to  $31.5 \ln(\text{s}^{-1})$ . These results are consistent with our theoretical description of CD accelerated-dynamics of vacancy migration [85].

The uncertainty on the extrapolated jump rate values  $\sigma_{\text{NE} \rightarrow \text{E}}$  is highly sensitive to the number of non-equilibrium jump rate interpolation points  $i_p$ . In Fig. 6, equilibrium jump rates  $k_{\text{NE} \rightarrow \text{E}}$  are represented as a function of  $i_p$ . Five different force-field intensities  $n_f$  are employed in this study. Thus, the number of extrapolated  $k_{\text{NE} \rightarrow \text{E}}$  values varies with  $i_p$  as  $n_f! / [i_p! (n_f - i_p)!]$ . As expected,  $\sigma_{\text{NE} \rightarrow \text{E}}$  (interpreted as scatter of  $k_{\text{NE} \rightarrow \text{E}}$  rates) progressively decreases for increasing  $i_p$ . This can be understood by noting that the frequency-intervals which include all  $k_{\text{NE} \rightarrow \text{E}}(T)$  values narrow down for larger  $i_p$  while maintaining approximately constant means. This confirms that  $\sigma_{\text{NE} \rightarrow \text{E}}$  can be systematically reduced upon increasing the number of interpolated points and/or the accuracy of each individual  $k_{\text{NE}}$  value.

The downward curvature  $\alpha(T)$  of  $\ln[k_{\text{NE}}(F, T)] = \ln[k_{\text{NE} \rightarrow \text{E}}(T)] + \frac{x_{\text{Tso}}(T)}{k_B T} F - \alpha(T) \cdot F^2$  parabolas increases very rapidly for decreasing temperatures (see Fig. 5). This implies that the extrapolated equilibrium rates  $k_{\text{NE} \rightarrow \text{E}}$  and their uncertainties  $\sigma_{\text{NE} \rightarrow \text{E}}$  become progressively more sensitive to the number of interpolation points and to the accuracy of  $k_{\text{NE}}$  values as  $T$  is reduced. Nevertheless, as demonstrated in Appendix B, both gamma and exponential distribution statistics

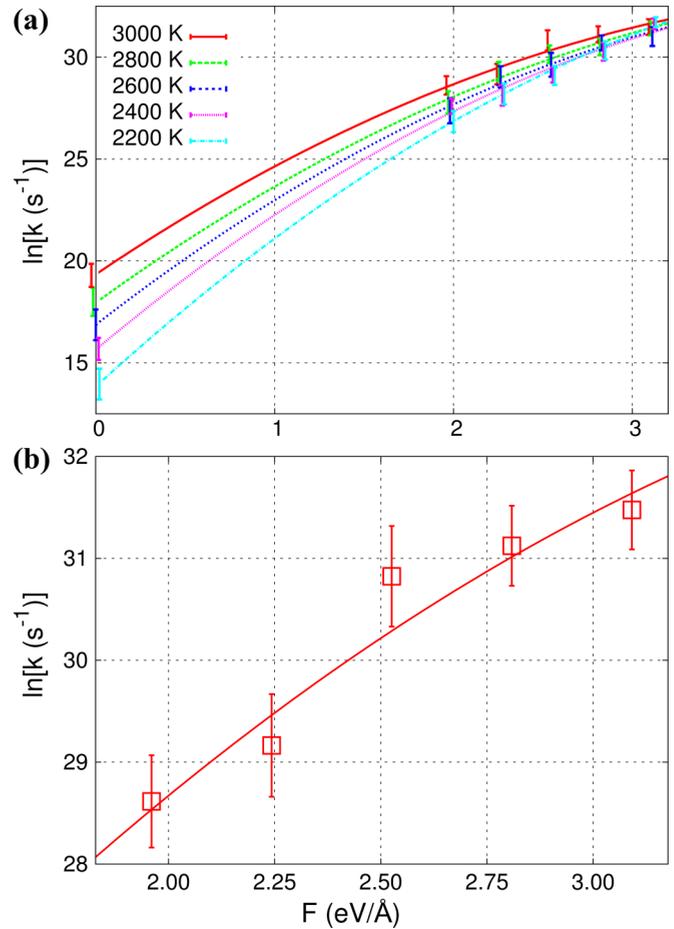


FIG. 5. (a) Accelerated vacancy jump rates  $k_{\text{NE}}$  plotted as a function of the force field intensity  $F$  at different temperatures and corresponding equilibrium  $k_{\text{NE} \rightarrow \text{E}}$  values extrapolated to zero force-intensity via Eq. (3). (b) Enlarged view of  $\ln[k_{\text{NE}}] \pm 2 \cdot \ln[\sigma_{\text{NE}}]$  values calculated at 3000 K. For calculation of  $k_{\text{NE}}$  and  $\sigma_{\text{NE}}$  values, MLE are performed by optimizing both  $\lambda$  and  $\theta$   $\Gamma$ -PDF parameters. To facilitate visualization, in panel (a), rates at the same force-field intensity are slightly laterally-shifted with respect to each other.

provide reliable estimations of vacancy jump frequencies (i.e.,  $k_{\text{NE} \rightarrow \text{E}} \pm \sigma_{\text{NE} \rightarrow \text{E}}$  results overlap with the uncertainty intervals of equilibrium nonaccelerated rates) even when only two  $k_{\text{NE}}(F)$  values (with  $F$  close to the limit  $F_{\text{max}}$ ) are used for extrapolation of equilibrium rates.

Equilibrium Ti vacancy jump rates  $k_{\text{NE} \rightarrow \text{E}}$  are plotted as a function of the inverse temperature  $1/T$  in Fig. 7. These are approximately one order of magnitude smaller than N vacancy migration rates, obtained from classical molecular dynamics simulations [54], at all investigated temperatures (see Fig. 7 and Table II). The  $k_{\text{NE} \rightarrow \text{E}}$  values shown in Fig. 7 are calculated from  $k_{\text{NE}}(F)$  results which, in turn, are obtained by maximizing  $\ln[L(\lambda, \theta)]$  [Eq. (5)] with respect to  $\lambda$  and  $\theta$ . Modeling the distribution of nonequilibrium jump occurrence times with an exponential PDF yields  $k_{\text{NE} \rightarrow \text{E}}$  values which are essentially equivalent to those shown in Fig. 7 (see Table II for comparison of exponential-versus-gamma-distribution results). Ti vacancy jump activation energies  $E_a$  and attempt frequencies  $A$  obtained by Arrhenius linear regression of  $\ln[k_{\text{NE} \rightarrow \text{E}}(T)]$

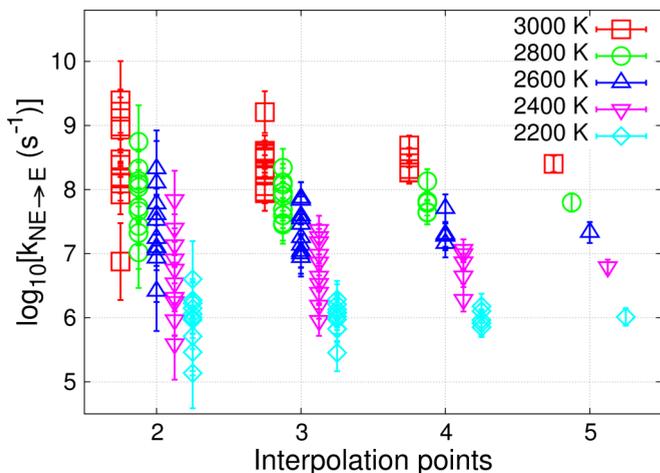


FIG. 6. Spread of extrapolated  $k_{NE \rightarrow E}$  values as a function of the number of interpolated  $k_{NE}$  data. The number of  $k_{NE \rightarrow E}$  results is equal to the binomial coefficient  $n_f! / [i_p! (n_f - i_p)!]$ , for which  $i_p$  is the number of interpolation points, while  $n_f$  is the total number, 5, of  $k_{NE}$  points available. For clarity, points corresponding to different temperatures are slightly laterally shifted.

versus  $1/T$  data yield  $E_a = (3.78 \pm 0.56)$  eV and  $A = 4.45(\times 13^{\pm 1}) \times 10^{14} \text{ s}^{-1}$  for gamma-distribution statistics and  $E_a = (3.77 \pm 0.64)$  eV and  $A = 4.65(\times 16^{\pm 1}) \times 10^{14} \text{ s}^{-1}$  for exponential-distribution statistics. Calculated  $E_a$  values are consistent with those  $(3.6 \pm 1.0 \text{ eV})$  experimentally determined for metal-atom migration across single-crystal  $B1$  TiN/ $B1$  NbN superlattice interfaces at temperatures ranging from 1100 to 1200 K (see Fig. 8 in Ref. [103]).

We suggest that the discrepancy between Ti vacancy migration energies extracted from linear regression of  $\ln[k_{NE \rightarrow E}(T)]$  versus  $1/T$  AIMD data ( $E_a = 3.78 \pm 0.56$  eV) and the one obtained from 0 K DFT+NEB calculations ( $E_a^{0K} = 4.26$  eV)

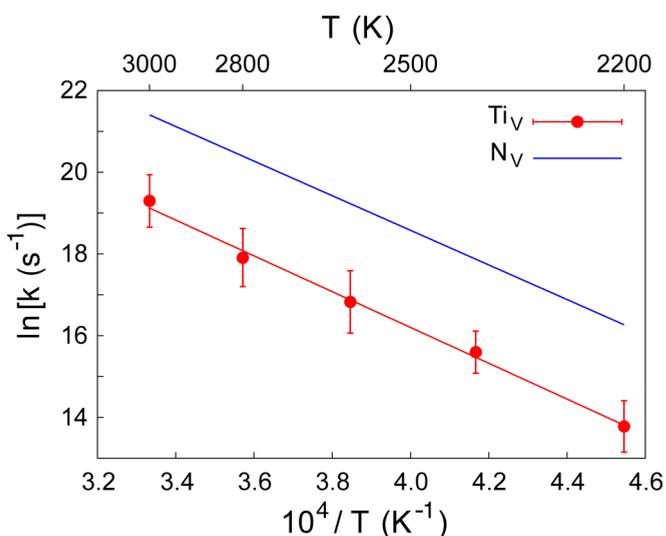


FIG. 7. Arrhenius plot of the migration of Ti (red, present work) and N (blue, from [54]) vacancies in TiN, i.e., the logarithm of the extrapolated jump rate as a function of the inverse of the temperature. Rate values are expressed as  $\ln[k] \pm 2 \cdot \ln[\sigma]$ , with  $\sigma$  equal to one standard deviation.

TABLE II. Comparison between Ti vacancy jump rates (obtained by gamma and exponential distribution statistics) and N vacancy jump rates from Ref. [54].

$T$ (K)	$k_{NE \rightarrow E}^{TiV} (\mu\text{s}^{-1})$		$k_{CMD}^{NV} (\mu\text{s}^{-1})$ [54]
	Gamma	Exponential	
2200	$1.03(\times 2.0^{\pm 1})$	$1.15(\times 2.0^{\pm 1})$	11.62
2400	$6.30(\times 1.7^{\pm 1})$	$6.12(\times 2.0^{\pm 1})$	57.79
2600	$15.4(\times 2.3^{\pm 1})$	$17.8(\times 2.0^{\pm 1})$	224.6
2800	$63.0(\times 2.0^{\pm 1})$	$64.5(\times 2.0^{\pm 1})$	719.1
3000	$253(\times 2.0^{\pm 1})$	$274(\times 2.0^{\pm 1})$	2370 ( $\sim 3000$ , AIMD)

is due to lattice vibrations. More specifically, the effect may be due to differences in vibrational frequencies (within the plane normal to the migration path) at the initial position and/or at the transition state of the diffusion reaction. This is indeed associated with temperature variations of the vibrational entropy in the plane orthogonal to the migration path. Lattice vibrations are known to affect the shape of the effective potential energy landscape, especially at temperatures close to melting [49,50,96].

A separate quantification of the effects of harmonic and anharmonic vibrations on temperature-induced variations in migration energies is a challenging research topic that deserves thoughtful investigation. Confining ourselves, for the moment, to estimating the contribution of thermal expansion, we carry out additional *static* DFT+NEB calculations of  $Ti_V$  migration energies, in which  $B1$  TiN is fixed at its equilibrium volumes at 2200 and 3000 K, i.e., lowest and highest AIMD simulation temperatures. The equilibrium lattice parameter at a temperature  $T$  is obtained as  $a_0(T) = a_0(0 \text{ K}) \times (\alpha_L \times T + 1)$ .

The NEB energy landscapes of *expanded* cells exhibit transition-states located midway between the initial and final positions of the diffusing Ti atom, as also seen for the energy-profile calculated for the 0-Kelvin equilibrium volume (see Fig. 4). The energy barriers  $E_a^{0K, a_0(2200 \text{ K})} = 3.67$  eV and  $E_a^{0K, a_0(3000 \text{ K})} = 3.50$  eV are both smaller than the finite-temperature  $E_a$  value, but yet well within the uncertainty range  $[E_a - 2 \cdot \sigma_{E_a}, E_a]$ . These results indicate that, in  $B1$  TiN, the modifications induced by lattice vibrations on the effective potential energy landscape stem in large part from thermal expansion effects, which reduce the migration energy. We cannot exclude, however, that part of the difference between finite-temperature versus 0-Kelvin results may arise from different k-space-sampling employed in AIMD versus DFT+NEB calculations (see Sec. II).

### C. Computational efficiency gain

The use of the CD algorithm, extended well beyond the linear-fitting rate vs. force regime [85], speeds-up AIMD times required to estimate vacancy migration rates by several orders of magnitude. Table III summarizes the gain in computational efficiency as a function of temperature. The gain factor  $t_E/t_{NE}$  is expressed as ratio between an estimate of the simulation

TABLE III. Comparison between computational times  $t_{\text{NE}}$  and  $t_{\text{E}}$  required for obtaining accurate estimates of Ti vacancy jump rates in TiN via non-equilibrium versus equilibrium AIMD. The CD method provides acceleration factors  $t_{\text{E}}/t_{\text{NE}}$ , which increase rapidly for decreasing simulation temperature  $T$ .

$T$ (K)	$t_{\text{NE}}$ (ns)	$t_{\text{E}}$ (ns)	Gain factor $t_{\text{E}}/t_{\text{NE}}$
3000	0.115	61.66	537
2800	0.174	257.1	1470
2600	0.246	1091	4430
2400	0.359	2984	8310
2200	0.587	19417	33100

time ( $t_{\text{E}}$ ) necessary to obtain *well-converged* equilibrium jump rates via *nonaccelerated* AIMD, and the total simulation time ( $t_{\text{NE}}$ ) employed in NE-AIMD at all force field intensities:  $t_{\text{NE}} = \sum_{l=1}^{n_f} \sum_{k=1}^{n(F_l)} t_k(F_l)$ .  $t_{\text{E}}(T)$  can be approximated as

$$t_{\text{E}}(T) = \bar{n} \cdot [k_{\text{NE} \rightarrow \text{E}}(T)]^{-1}, \quad (9)$$

where  $\bar{n}$  is the mean number of runs performed at each  $F$  and  $k_{\text{NE} \rightarrow \text{E}}(T)$  is the equilibrium jump rate extrapolated at temperature  $T$ . The use of a factor  $\bar{n}$  for the estimation of  $t_{\text{E}}$  in Eq. (7) is motivated by the fact that approximately  $\bar{n}$  vacancy migration events are to be recorded during nonaccelerated AIMD simulations to achieve average jump rates with confidence ranges  $\sigma_{\text{E}}$  comparable to  $\sigma_{\text{NE} \rightarrow \text{E}}$ .

As shown in Table III, NE-AIMD simulations are five hundreds times faster than equilibrium AIMD at 3000 K. The speed-up factor increases rapidly for decreasing temperatures, and reaches a value of  $\sim 3 \times 10^4$  at the lowest  $T$  (2200 K) considered in our investigations.

#### IV. CONCLUSIONS

Nonequilibrium *ab initio* molecular dynamics simulations based on the color-diffusion algorithm are used to calculate Ti monovacancy jump rates in *B1* TiN at temperatures ranging from 2200 to 3000 K. We propose a scheme based on gamma-distribution statistics which provides quantitatively reliable confidence-ranges on extrapolated vacancy migration rates. Within the investigated temperature range, migration frequencies are well described by an Arrhenius trend with activation energy  $E_a = (3.78 \pm 0.56)$  eV and attempt frequency  $A = 4.45(\times 13^{\pm 1}) \times 10^{14} \text{ s}^{-1}$ . We suggest that the difference between the finite-temperature  $E_a$  value and the DFT migration energy calculated at 0 K ( $E_a^{0\text{K}} = 4.26$  eV) is to be primarily attributed to thermal expansion effects. The results presented in this work demonstrate that the color-diffusion algorithm can be used to efficiently determine metal vacancy migration rates in binary systems characterized by inherently low mobilities of lattice defects. The gain in computational efficiency compared to equilibrium (non-accelerated) *ab initio* molecular dynamics increases for decreasing temperatures from a factor of five hundreds for  $T = 3000$  K up to a factor of 33 thousands for  $T = 2200$  K.

#### ACKNOWLEDGMENTS

This research was carried out using resources provided by the National Supercomputer Centre (NSC) in Linköping (Gamma supercomputer) and the Swedish National Infrastructure for Computing (SNIC): Triolith Cluster located at NSC and Beskow cluster located at the Center for High Performance Computing (PDC) in Stockholm, Sweden. Dr. Grisell Diaz Leines is gratefully acknowledged for useful discussions. W. Olovsson and P. Mürger at NSC and H. Leskelä and J. Vincent at PDC are acknowledged for assistance with technical aspects. We gratefully acknowledge financial support from the Swedish Foundation for Strategic Research (SSF) project SRL Grant No. 10-0026, the Swedish Research Council (VR) Grants No. 621-2011-4417 and No. 2015-04391 the Swedish Government Strategic Research Area Grant in Materials Science on Advanced Functional Materials (Grant No. MatLiU 2009-00971 through Sweden's innovation agency VINNOVA). I.A.A. is grateful for support from the Ministry of Education and Science of the Russian Federation (Grant No. 14.Y26.31.0005). Financial support by the Swedish Research Council (VR) through the International Career Grant No. 330-2014-6336, Marie Skłodowska Curie Actions, Cofund, Project INCA 600398, and the Swedish Foundation for Strategic Research through the Future Research Leaders 6 program is gratefully acknowledged by B.A. D.G.S. gratefully acknowledges financial support from the Stiftelsen Olle Engkvist Byggmästare.

#### APPENDIX A

The exponential PDF is commonly considered in the form

$$\text{PDF}_k(t) = k \cdot e^{-k \cdot t}, \quad (\text{A1})$$

where  $k$  is the rate of the process under consideration. The relation with the gamma distribution is easily obtained from Eq. (4), setting  $\lambda = 1$  and  $\theta = k^{-1}$ . This means that  $\theta$ , the sole adjustable parameter, determines the average occurrence time  $\langle t \rangle$ . For practical convenience, the expressions contained in this appendix employ a rate  $k$  in place of an occurrence-time  $\theta$ . The likelihood function [Eq. (5)] for an exponential distribution applied to a set  $i$  of uncensored data ( $m$  occurrence-times  $t_i$ ) and a set  $j$  of censored data ( $n-m$  occurrence-times  $t_j$ ) is

$$\begin{aligned} L(k) &= \prod_{i=1}^m [k \cdot \exp(-k \cdot t_i)] \\ &\cdot \prod_{j=m+1}^n \left[ 1 - \int_0^{t_j} k \cdot \exp(-k \cdot t) \cdot dt \right] \\ &= \prod_{i=1}^m [k \cdot \exp(-k \cdot t_i)] \cdot \prod_{j=m+1}^n [\exp(-k \cdot t_j)]. \quad (\text{A2}) \end{aligned}$$

Taking the logarithm of  $L(k)$  yields

$$\ln[L(k)] = \sum_{i=1}^m [\ln(k) - k \cdot t_i] - k \cdot \sum_{j=m+1}^n [t_j]. \quad (\text{A3})$$

$\ln[L(k)]$  is maximized for

$$\begin{aligned} \frac{\partial \ln[L(k)]}{\partial k} &= \frac{m}{k} - \sum_{i=1}^m [t_i] - \sum_{j=m+1}^n [t_j] = 0 \\ \Rightarrow \frac{m}{k} &= \sum_{i=1}^m [t_i] + \sum_{j=m+1}^n [t_j] \\ \Rightarrow k &= \frac{m}{\sum_{i=1}^m [t_i] + \sum_{j=m+1}^n [t_j]}, \quad (\text{A4}) \end{aligned}$$

where Eq. (A4) corresponds to the formula used in Ref. [85] to calculate  $k_{\text{NE}}$ . Thus, the procedure used in our previous work [85] corresponds to calculating average jump rates via exponential-distribution statistics, as also applied to analyze results of atomic diffusivities in Ref. [92].

### APPENDIX B

In this Appendix, we describe the performances of gamma versus exponential distribution statistics in estimating non-equilibrium rates (with corresponding error bars) as a function of the relative occurrence of unsuccessful runs (censored data). Especially important is to verify the fidelity of error bars for case studies in which accelerated rates can be obtained during feasible computational time for only two or three force field values close to the  $F_{\text{max}}$  limit (e.g., at low temperatures).

Our tests are carried out employing kinetic Monte Carlo (KMC) simulations. The accuracy of gamma- and exponential-distribution results is quantified by comparing the rates extrapolated to zero force  $F$  with equilibrium (nonaccelerated) rates.  $\text{Ti}_V$  jump activation energies and attempt frequencies obtained from Arrhenius linear fitting of molecular dynamics results (Sec. III B) are used as input for KMC employing a temperature-parameter of 3000 K. The dependence of activation energies and attempt frequencies on  $F$  is taken into account as described in the Supplemental Material of reference [85], imposing  $x_{\text{TS0}}(3000 \text{ K}) = 1.5 \text{ \AA}$ .

KMC simulations are used to calculate  $n = 20$  occurrence times  $t_i$  at each  $F$  applied in NE-AIMD simulations, as well as the equilibrium (nonaccelerated) rate. Average accelerated jump rates  $\langle t \rangle^{-1}$  are obtained as a function of  $F$  via both gamma and exponential distribution statistics by considering different numbers of censored data ( $n-m = 0, 4, 8, \text{ and } 12$ ). The 4, 8, and 12 longest occurrence times  $t_i$  are censored at each  $F$  according to the procedure described in Sec. II D.

KMC results demonstrate that, irrespective of the relative occurrence  $(n-m)/n$  of unsuccessful simulations (censored data), equilibrium rates extrapolated by both exponential and gamma distribution nonequilibrium rates calculated at five different  $F$  values are within, or close to, nonaccelerated rate uncertainty intervals (Fig. 8). As expected, the degree of agreement between extrapolated and nonaccelerated equilibrium jump rates decreases as the number of censored data increases. However, the uncertainty range on extrapolated values increases monotonically with  $(n-m)/n$ . Thus, even censoring 60% of the total number of runs, the error bars of extrapolated rates overlap with those of equilibrium values. Although  $(n-m)/n = 0.6$ , the margin in logarithm scale between extrapolated and non-accelerated rates is less than 3%. These results prove that the procedure used in this work (see,

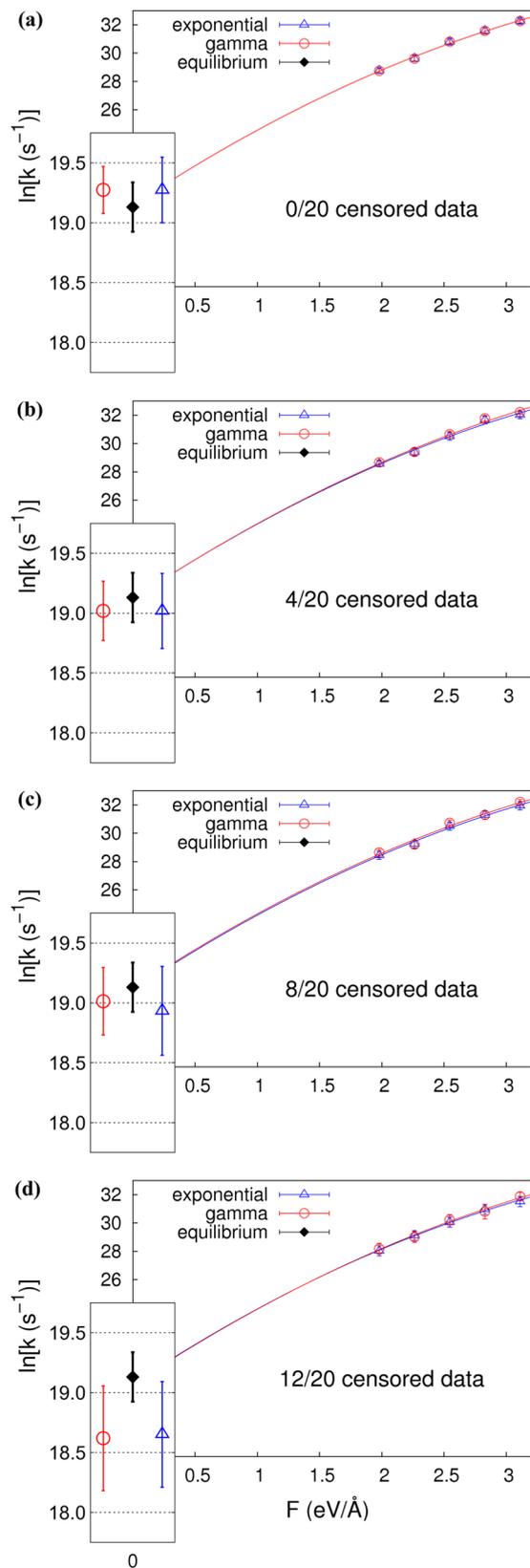


FIG. 8. Comparison between extrapolated and equilibrium jump rates for different numbers of censored data, as well as comparison between gamma and exponential fitting. Data here represented are produced with Kinetic Monte Carlo, taking into account the effect of the force field on activation energy and jump attempt frequency.

for example, requirement (8b)  $[(n-m)/n \leq 0.25]$  in Sec. IID) allows retrieving accurate equilibrium jump rate values.

The reliability of error bars for cases in which accelerated rates can be obtained only at force fields of high intensities is tested by fitting  $k_{\text{NE}}$  values corresponding to four different sets of forces:  $\{F_a^{(100)}, F_b^{(100)}, F_c^{(100)}\}$ ,  $\{F_a^{(100)}, F_b^{(100)}\}$ ,  $\{F_a^{(100)}, F_c^{(100)}\}$ , and  $\{F_b^{(100)}, F_c^{(100)}\}$ , where  $F_a^{(100)} = 1.8$ ,  $F_b^{(100)} = 2.0$ , and  $F_c^{(100)} = 2.2$  eV/Å. For each set of forces, we use  $k_{\text{NE}}$  values obtained by censoring 0, 4, 8, and 12 jump occurrence times. This yields a total of 16 combinations of data sets with corresponding extrapolated equilibrium rates. Fig. 9 shows that irrespective of the relative number of censored data and of the fact that only two or three nonequilibrium rate values are used in Eq. (3), the difference between extrapolated rates and equilibrium (nonaccelerated) rates is always smaller than one order of magnitude for both gamma and exponential statistics. In addition, equilibrium rates (horizontal solid black line in Fig. 9) are, in all cases, within or close to the uncertainty intervals of extrapolated rates.

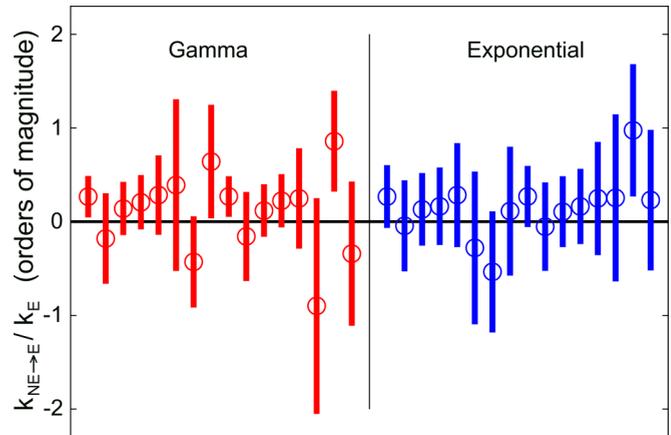


FIG. 9. Ratio between extrapolated and equilibrium rates obtained by fitting accelerated rates calculated for 2 or 3 different force field intensities close to the limit  $F_{\text{max}}$ . For convenience, the vertical axis is in logarithmic (base 10) scale.

- [1] P. H. Mayrhofer, C. Mitterer, L. Hultman, and H. Clemens, *Prog. Mater. Sci.* **51**, 1032 (2006).
- [2] G. Abadias, L. E. Koutsokeras, S. N. Dub, G. N. Tolmacheva, A. Debelle, T. Sauvage, and P. Villechaise, *J. Vac. Sci. Technol. A* **28**, 541 (2010).
- [3] S. Yu, Q. Zeng, A. R. Oganov, G. Frapper, and L. Zhang, *Phys. Chem. Chem. Phys.* **17**, 11763 (2015).
- [4] H. Kindlund *et al.*, *APL Mater.* **1**, 042104 (2013).
- [5] H. Kindlund, D. G. Sangiovanni, J. Lu, J. Jensen, V. Chirita, I. Petrov, J. E. Greene, and L. Hultman, *J. Vac. Sci. Technol. A: Vac. Surf. Films* **32** (2014).
- [6] D. G. Sangiovanni, L. Hultman, V. Chirita, I. Petrov, and J. E. Greene, *Acta Mater.* **103**, 823 (2016).
- [7] V. V. Uglov, V. M. Anishchik, S. V. Zlotski, G. Abadias, and S. N. Dub, *Surf. Coat. Technol.* **202**, 2394 (2008).
- [8] L. Hultman, *Vacuum* **57**, 1 (2000).
- [9] X. Z. Ding, A. L. K. Tan, X. T. Zeng, C. Wang, T. Yue, and C. Q. Sun, *Thin Solid Films* **516**, 5716 (2008).
- [10] Y. Kayali, *Protect. Metals Phys. Chem. Surf.* **50**, 412 (2014).
- [11] A. B. Mei, A. Rockett, L. Hultman, I. Petrov, and J. E. Greene, *J. Appl. Phys.* **114**, 193708 (2013).
- [12] L. Ning, S. C. Veldhuis, and K. Yamamoto, *Int. J. Mach. Tools Manufact.* **48**, 656 (2008).
- [13] J.-E. Ståhl, *Metal Cutting-Theories and Models* (SECO TOOLS AB, Elanders, Lund, Sweden, 2012).
- [14] J. L. Ruan, D. F. Lii, J. S. Chen, and J. L. Huang, *Ceramics Int.* **35**, 1999 (2009).
- [15] Y. L. Kuo, H. H. Lee, C. Lee, J. C. Lin, S. L. Shue, M. S. Liang, and B. J. Daniels, *Electrochem. Solid State Lett.* **7**, C35 (2004).
- [16] J.-E. Sundgren, B. O. Johansson, A. Rockett, S. A. Barnett, and J. E. Greene, *AIP Conf. Proc.* **149**, 95 (1986).
- [17] W. K. Schubert, R. N. Shelton, and E. L. Wolf, *Phys. Rev. B* **23**, 5097 (1981).
- [18] W. Lengauer, *J. Alloys Compd.* **186**, 293 (1992).
- [19] Z. Gu, C. Hu, H. Huang, S. Zhang, X. Fan, X. Wang, and W. Zheng, *Acta Mater.* **90**, 59 (2015).
- [20] H. S. Seo, T. Y. Lee, I. Petrov, J. E. Greene, and D. Gall, *J. Appl. Phys.* **97**, 083521 (2005).
- [21] C. S. Shin, D. Gall, Y. W. Kim, P. Desjardins, I. Petrov, J. E. Greene, M. Oden, and L. Hultman, *J. Appl. Phys.* **90**, 2879 (2001).
- [22] L. E. Koutsokeras, G. Abadias, C. E. Lekka, G. M. Matenoglou, D. F. Anagnostopoulos, G. A. Evangelakis, and P. Patsalas, *Appl. Phys. Lett.* **93**, 011904 (2008).
- [23] J. H. Kang and K. J. Kim, *J. Appl. Phys.* **86**, 346 (1999).
- [24] A. B. Mei, M. Tuteja, D. G. Sangiovanni, R. T. Haasch, A. Rockett, L. Hultman, I. Petrov, and J. E. Greene, *J. Mater. Chem. C* **4**, 7924 (2016).
- [25] T. Lee, K. Ohmori, C. S. Shin, D. G. Cahill, I. Petrov, and J. E. Greene, *Phys. Rev. B* **71**, 144106 (2005).
- [26] C. S. Shin, D. Gall, N. Hellgren, J. Patscheider, I. Petrov, and J. E. Greene, *J. Appl. Phys.* **93**, 6025 (2003).
- [27] H. Kindlund, D. G. Sangiovanni, J. Lu, J. Jensen, V. Chirita, J. Birch, I. Petrov, J. E. Greene, and L. Hultman, *Acta Mater.* **77**, 394 (2014).
- [28] C. Teichmann, W. Lengauer, P. Ettmayer, J. Bauer, and M. Bohn, *Metal. Mater. Trans. A: Phys. Metal. Mater. Sci.* **28**, 837 (1997).
- [29] F. Elstner, H. Kupfer, and F. Richter, *Physica Status Solidi A: Appl. Mater. Sci.* **147**, 373 (1995).
- [30] C. Hoglund, B. Alling, J. Birch, M. Beckers, P. O. A. Persson, C. Baecht, Z. Czigan, J. Jensen, and L. Hultman, *Phys. Rev. B* **81**, 224101 (2010).
- [31] P. H. Mayrhofer, A. Hörling, L. Karlsson, J. Sjölen, T. Larsson, C. Mitterer, and L. Hultman, *Appl. Phys. Lett.* **83**, 2049 (2003).
- [32] J. E. Sundgren, *Thin Solid Films* **128**, 21 (1985).
- [33] L. E. Toth, *Transition Metal Carbides and Nitrides* (Academic Press, New York, 1971).

- [34] G. Henkelman, B. P. Uberuaga, and H. Jonsson, *J. Chem. Phys.* **113**, 9901 (2000).
- [35] G. Henkelman and H. Jonsson, *J. Chem. Phys.* **113**, 9978 (2000).
- [36] E. Weinan, W. Ren, and E. Vanden-Eijnden, *Phys. Rev. B* **66**, 052301 (2002).
- [37] B. Peters, A. Heyden, A. T. Bell, and A. Chakraborty, *J. Chem. Phys.* **120**, 7877 (2004).
- [38] Y. Ren, X. J. Liu, X. Tan, and E. Westkamper, *Comput. Mater. Sci.* **77**, 102 (2013).
- [39] B. Alling, P. Steneteg, C. Tholander, F. Tasnadi, I. Petrov, J. E. Greene, and L. Hultman, *Phys. Rev. B* **85**, 245422 (2012).
- [40] L. Tsetseris, N. Kalfagiannis, S. Logothetidis, and S. T. Pantelides, *Phys. Rev. Lett.* **99**, 125503 (2007).
- [41] A. S. Bochkarev, M. N. Popov, V. I. Razumovskiy, J. Spitaler, and P. Puschnig, *Phys. Rev. B* **94**, 104303 (2016).
- [42] C. Tholander, B. Alling, F. Tasnádi, J. E. Greene, and L. Hultman, *Surf. Sci.* **630**, 28 (2014).
- [43] C. Mastail, M. David, F. Nita, A. Michel, and G. Abadias, *Appl. Surf. Sci.* **423**, 354 (2017).
- [44] G. H. Vineyard, *J. Phys. Chem. Solids* **3**, 121 (1957).
- [45] A. A. Maradudin, E. W. Montroll, and G. H. Weiss, *Theory of Lattice Dynamics in the Harmonic Approximation* (Academic Press, New York and London, 1963).
- [46] B. Fultz, *Prog. Mater. Sci.* **55**, 247 (2010).
- [47] E. I. Isaev, S. I. Simak, I. A. Abrikosov, R. Ahuja, Y. K. Vekilov, M. I. Katsnelson, A. I. Lichtenstein, and B. Johansson, *J. Appl. Phys.* **101**, 123519 (2007).
- [48] A. B. Mei, O. Hellman, N. Wireklint, C. M. Schlepütz, D. G. Sangiovanni, B. Alling, A. Rockett, L. Hultman, I. Petrov, and J. E. Greene, *Phys. Rev. B* **91**, 054101 (2015).
- [49] A. Glensk, B. Grabowski, T. Hickel, and J. Neugebauer, *Phys. Rev. X* **4**, 011018 (2014).
- [50] A. Glensk, B. Grabowski, T. Hickel, and J. Neugebauer, *Phys. Rev. Lett.* **114**, 195901 (2015).
- [51] H. M. Gilder and D. Lazarus, *Phys. Rev. B* **11**, 4916 (1975).
- [52] R. Car and M. Parrinello, *Phys. Rev. Lett.* **55**, 2471 (1985).
- [53] D. G. Sangiovanni, D. Edström, L. Hultman, I. Petrov, J. E. Greene, and V. Chirita, *Surf. Sci.* **624**, 25 (2014).
- [54] D. G. Sangiovanni, B. Alling, P. Steneteg, L. Hultman, and I. A. Abrikosov, *Phys. Rev. B* **91**, 054301 (2015).
- [55] D. G. Sangiovanni, A. B. Mei, L. Hultman, V. Chirita, I. Petrov, and J. E. Greene, *J. Phys. Chem. C* **120**, 12503 (2016).
- [56] S. Piscanec, L. C. Ciacchi, E. Vesselli, G. Comelli, O. Sbaizero, S. Meriani, and A. De Vita, *Acta Materialia* **52**, 1237 (2004).
- [57] D. Music and J. M. Schneider, *New J. Phys.* **15**, 073004 (2013).
- [58] D. G. Sangiovanni, D. Edström, L. Hultman, I. Petrov, J. E. Greene, and V. Chirita, *Surf. Sci.* **627**, 34 (2014).
- [59] C. Bennett, *Exact Defect Calculations in Model Substances; Diffusion in Solids: Recent Developments* (Academic Press, New York, San Francisco, London, 1975).
- [60] C. Dellago and P. G. Bolhuis, *Adv. Comput. Simul. Approaches Soft Matter Sci. III* **221**, 167 (2008).
- [61] E. Weinan, W. Q. Ren, and E. Vanden-Eijnden, *J. Phys. Chem. B* **109**, 6688 (2005).
- [62] E. Vanden-Eijnden and M. Venturoli, *J. Chem. Phys.* **130**, 194103 (2009).
- [63] A. Laio and M. Parrinello, *Proc. Natl. Acad. Sci. USA* **99**, 12562 (2002).
- [64] J. VandeVondele and U. Rothlisberger, *J. Phys. Chem. B* **106**, 203 (2002).
- [65] C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, *J. Chem. Phys.* **108**, 1964 (1998).
- [66] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, *Annu. Rev. Phys. Chem.* **53**, 291 (2002).
- [67] C. Dellago, P. G. Bolhuis, and P. L. Geissler, *Adv. Chem. Phys.* **123**, 1 (2002).
- [68] T. S. van Erp, D. Moroni, and P. G. Bolhuis, *J. Chem. Phys.* **118**, 7762 (2003).
- [69] T. S. van Erp and P. G. Bolhuis, *J. Comput. Phys.* **205**, 157 (2005).
- [70] S. E. Bouffelfel, A. R. Oganov, and S. Leoni, *Sci. Rep.* **2**, 471 (2012).
- [71] R. J. Allen, P. B. Warren, and P. R. ten Wolde, *Phys. Rev. Lett.* **94**, 018104 (2005).
- [72] R. J. Allen, D. Frenkel, and P. R. ten Wolde, *J. Chem. Phys.* **124**, 024102 (2006).
- [73] R. J. Allen, D. Frenkel, and P. R. ten Wolde, *J. Chem. Phys.* **124**, 194111 (2006).
- [74] D. Moroni, P. G. Bolhuis, and T. S. van Erp, *J. Chem. Phys.* **120**, 4055 (2004).
- [75] A. K. Faradjian and R. Elber, *J. Chem. Phys.* **120**, 10880 (2004).
- [76] A. M. A. West, R. Elber, and D. Shalloway, *J. Chem. Phys.* **126**, 145104 (2007).
- [77] K. M. Bal and E. C. Neyts, *J. Chem. Theory Comput.* **11**, 4545 (2015).
- [78] P. Tiwary and M. Parrinello, *Phys. Rev. Lett.* **111**, 230602 (2013).
- [79] A. F. Voter, *J. Chem. Phys.* **106**, 4665 (1997).
- [80] M. R. Sorensen and A. F. Voter, *J. Chem. Phys.* **112**, 9599 (2000).
- [81] P. C. Aeberhard, S. R. Williams, D. J. Evans, K. Refson, and W. I. F. David, *Phys. Rev. Lett.* **108**, 095901 (2012).
- [82] J. O. Nilsson, O. Y. Vekilova, O. Hellman, J. Klarbring, S. I. Simak, and N. V. Skorodumova, *Phys. Rev. B* **93**, 024102 (2016).
- [83] S. R. Williams and D. J. Evans, *Phys. Rev. Lett.* **96**, 015701 (2006).
- [84] A. F. Voter, *Phys. Rev. B* **57**, R13985 (1998).
- [85] D. G. Sangiovanni, O. Hellman, B. Alling, and I. A. Abrikosov, *Phys. Rev. B* **93**, 094305 (2016).
- [86] A. D. Mulliner, P. C. Aeberhard, P. D. Battle, W. I. F. David, and K. Refson, *Phys. Chem. Chem. Phys.* **17**, 21470 (2015).
- [87] P. E. Blöchl, *Phys. Rev. B* **50**, 17953 (1994).
- [88] R. Armiento and A. E. Mattsson, *Phys. Rev. B* **72**, 085108 (2005).
- [89] C. S. Shin, S. Rudenja, D. Gall, N. Hellgren, T. Y. Lee, I. Petrov, and J. E. Greene, *J. Appl. Phys.* **95**, 356 (2004).
- [90] H. J. Monkhorst and J. D. Pack, *Phys. Rev. B* **13**, 5188 (1976).
- [91] V. I. Razumovskiy, M. N. Popov, H. Ding, and J. Odqvist, *Comput. Mater. Sci.* **104**, 147 (2015).
- [92] J. O. Nilsson, M. Leetmaa, O. Y. Vekilova, S. I. Simak, and N. V. Skorodumova, *Phys. Rev. B* **94**, 085206 (2016).
- [93] N. Balakrishnan and A. C. Cohen, *Order Statistics and Inference: Estimation Methods* (Academic Press, San Diego, 1991).
- [94] W. Nelson, *Applied Life Data Analysis* (John Wiley & Sons, New York, 1982).

- [95] RStudio Team (2015). *RStudio: Integrated Development for R*. RStudio, Inc., Boston, MA. <http://www.rstudio.com/>
- [96] C. Freysoldt, B. Grabowski, T. Hickel, J. Neugebauer, G. Kresse, A. Janotti, and C. G. Van de Walle, *Rev. Mod. Phys.* **86**, 253 (2014).
- [97] L. Tsetseris, N. Kalfagiannis, S. Logothetidis, and S. T. Pantelides, *Phys. Rev. B* **76**, 224107 (2007).
- [98] A. Janotti and C. G. Van de Walle, *Phys. Rev. B* **76**, 165202 (2007).
- [99] N. I. Papanicolaou and H. Chamati, *Comput. Mater. Sci.* **44**, 1366 (2009).
- [100] M. W. D. Cooper, S. Kelly, and M. Bertolus, *Phys. Chem. Chem. Phys.* **18**, 16902 (2016).
- [101] J. Mulroue and D. M. Duffy, *Proc. Roy. Soc. A: Math. Phys. Eng. Sci.* **467**, 2054 (2011).
- [102] S. V. Divinski and C. Herzig, *Defects Diffus. Metals: Annu. Retrospect. Iv* **203-2**, 177 (2002).
- [103] L. Hultman, C. Engstrom, and M. Oden, *Surf. Coat. Technol.* **133**, 227 (2000).