Diffusion of two-dimensional epitaxial clusters on metal (100) surfaces: Facile versus nucleation-mediated behavior and their merging for larger sizes

King C. Lai,[1,2] Da-Jiang Liu,[2] and James W. Evans[1,2,3]
[1]*Department of Physics & Astronomy, Iowa State University, Ames, Iowa 50011, USA*
[2]*Ames Laboratory–USDOE, Iowa State University, Ames, Iowa 50011, USA*
[3]*Department of Mathematics, Iowa State University, Ames, Iowa 50011, USA*
(Received 19 June 2017; published 5 December 2017)

For diffusion of two-dimensional homoepitaxial clusters of N atoms on metal (100) surfaces mediated by edge atom hopping, macroscale continuum theory suggests that the diffusion coefficient scales like $D_N \sim N^{-\beta}$ with $\beta = 3/2$. However, we find quite different and diverse behavior in multiple size regimes. These include: (i) facile diffusion for small sizes $N < 9$; (ii) slow nucleation-mediated diffusion with small $\beta < 1$ for "perfect" sizes $N = N_p = L^2$ or $L(L+1)$, for $L = 3,4,\ldots$ having unique ground-state shapes, for moderate sizes $9 \leqslant N \leqslant \mathrm{O}(10^2)$; the same also applies for $N = N_p + 3$, $N_p + 4$, ... (iii) facile diffusion but with large $\beta > 2$ for $N = N_p + 1$ and $N_p + 2$ also for moderate sizes $9 \leqslant N \leqslant \mathrm{O}(10^2)$; (iv) merging of the above distinct branches and subsequent anomalous scaling with $1 \lesssim \beta < 3/2$, reflecting the quasifacetted structure of clusters, for larger $N = \mathrm{O}(10^2)$ to $N = \mathrm{O}(10^3)$; (v) classic scaling with $\beta = 3/2$ for very large $N = \mathrm{O}(10^3)$ and above. The specified size ranges apply for typical model parameters. We focus on the moderate size regime where we show that diffusivity cycles quasiperiodically from the slowest branch for $N_p + 3$ (not N_p) to the fastest branch for $N_p + 1$. Behavior is quantified by kinetic Monte Carlo simulation of an appropriate stochastic lattice-gas model. However, precise analysis must account for a strong enhancement of diffusivity for short time increments due to back correlation in the cluster motion. Further understanding of this enhancement, of anomalous size scaling behavior, and of the merging of various branches, is facilitated by combinatorial analysis of the number of the ground-state and low-lying excited state cluster configurations, and also of kink populations.

DOI: 10.1103/PhysRevB.96.235406

I. INTRODUCTION

Significant long-range diffusion of large two-dimensional (2D) homoepitaxial adatom clusters on single-crystal metal (100) surfaces with sizes on the order of hundreds or even thousands of atoms was studied by scanning tunneling microscopy (STM) as early as the mid-1990's [1,2] and also more recently [3]. It is generally accepted that cluster diffusion is mediated by periphery diffusion (PD), also described as edge diffusion, of adatoms along the steps at the periphery of the cluster. The STM studies prompted extensive atomistic lattice-gas modeling starting in the 1990's of epitaxial cluster diffusion [4–11] and of related reshaping phenomena [12–19]. This work supplemented limited earlier studies [20–22]. Mesoscale continuum Langevin theory for PD-mediated cluster diffusion has also been applied, and predicts that the diffusion coefficient for clusters of N atoms satisfies $D_N \sim \sigma_{\mathrm{PD}} N^{-\beta}$ with $\beta = 3/2$, where σ_{PD} denotes the mesoscale mobility for atoms at step edges [23,24]. Simple mean-field type atomistic-level theory for compact clusters also predicts the same size dependence as the continuum theory [25,26]. However, significantly, the experimentally observed size scaling exponent β for moderate cluster sizes, $N = \mathrm{O}(10^2)$ to $\mathrm{O}(10^3)$, is below the prediction of the continuum and mean-field theories [2,3].

Diffusion of smaller 2D clusters with less than ~10 atoms on metal (100) surfaces was also observed but instead by field ion microscopy [27–29], and has been interpreted with appropriate theoretical analyses [30–33]. However, diffusion of small sized clusters exhibits a distinctive irregular size dependence and Arrhenius energetics, which is readily understood, e.g., given the innate stability of 2×2 atom square clusters relative to two-atom dimers and three-atom trimers. We also mention that there have been multiple studies of 2D cluster diffusion for metal (111) and metal (110) homoepitaxial systems, and also for heteroepitaxial metal systems [34–37]. Theoretical studies, particularly for metal (111) systems, have explored concerted many-atom and off-lattice nonepitaxial mechanisms [38–41]. These latter systems are of less relevance for the current study, so we do not discuss them further.

For 2D cluster diffusion on metal (100) surfaces, there is naturally interest in the effective or overall activation barrier E_{eff} for the process where $D_N \sim \exp[-E_{\mathrm{eff}}/(k_B T)]$. Here, k_B denotes the Boltzmann constant, and T denotes the surface temperature. E_{eff} is related to the kinetic parameters in atomistic-level models including the barrier E_{e} to diffuse along close-packed ⟨110⟩ cluster step edges, and any additional barrier δ to round corners or kinks. E_{eff} also reflects thermodynamic parameters determined by adatom interactions, particularly the formation energy E_{form} to create a step edge atom from a kink atom. It was previously suggested that long-range cluster diffusion is limited by creation of edge atoms through their extraction from the core of the cluster or "core breakup" [1,20], so that $E_{\mathrm{eff}} = E_e + \delta + E_{\mathrm{form}}$ [19]. This perspective is consistent with the predictions of the mesoscale continuum Langevin theory where the activation energy for cluster diffusion corresponds to that for mobility of edge atoms E_{PD}, where $\sigma_{\mathrm{PD}} \sim \exp[-E_{\mathrm{PD}}/(k_B T)]$ with $E_{\mathrm{PD}} = E_e + \delta + E_{\mathrm{form}}$ [23,24]. The latter result for E_{PD} has been rigorously demonstrated in the absence of a corner or kink rounding barrier [42], but it is expected to apply more generally [43].

However, Mills *et al.* [10] noted that if cluster edges are effectively facetted, then cluster diffusion can be limited by nucleation of new edge layers on these facetted step edges.

This picture leads to higher values of $E_{\rm eff}$ than predicted above (see Sec. III), and also to a weaker dependence of D_N on N reminiscent of experimental observations. This facetted regime occurs for linear cluster sizes, $L \sim N^{1/2}$ (in units of surface lattice constant, $a = 1$), which are below the characteristic separation, $L_k \approx {}^1\!/_2 \exp[\varepsilon_k/(k_B T)]$, of kinks on close-packed $\langle 110 \rangle$ edges [44]. Here, ε_k denotes the kink creation energy. Another perspective on anomalous size scaling for diffusivity was provided by Pierre-Louis [45] who modified the continuum Langevin theory by introducing an additional diffusion field for edge atoms. This approach also recovered weaker size scaling.

Jensen *et al.* [15] adopted an analogous nucleation-mediated picture to describe the effective barrier and anomalous size scaling for the relaxation to equilibrium of convex nonequilibrium cluster shapes. Regarding the relationship between this shape relaxation process and the long-range diffusion of clusters, it should be noted that both require nucleation of new edge layers. Furthermore, a simple relationship was proposed between the size scaling exponents for cluster diffusion and relaxation of convex shapes [46]. It was later shown that further refinement to anomalous scaling could be induced in the presence of an additional kink or corner rounding barrier [14,17].

However, we show in this contribution that the above observations, while providing key insight into deviations from standard macroscale and mean-field theories, fall far short of providing a complete characterization of the full diversity of cluster diffusion behavior on the nanoscale. A comprehensive and precise characterization of the dependence of the cluster diffusion coefficient D_N on size N can be provided by analysis utilizing kinetic Monte Carlo (KMC) simulation of a stochastic atomistic-level lattice-gas model for cluster diffusion which incorporates an appropriate description of PD kinetics. Indeed, this approach is a key component of the current study, and reveals various size regimes with distinct behavior: (i) facile diffusion for small sizes $N < 9$; (ii) slow nucleation-mediated diffusion with weak size scaling $\beta < 1$ for "perfect" sizes $N = N_p = L^2$ or $L(L+1)$ with $L = 3,4,\ldots$ having unique square or near-square ground-state shapes, and also for sizes $N_p + 3$, $N_p + 4, \ldots$, versus facile diffusion with strong size scaling $\beta > 2$ for sizes $N_p + 1$ and $N_p + 2$ for moderate sizes $9 \leqslant N \leqslant {\rm O}(10^2)$; (iii) merging of these distinct branches and subsequent anomalous scaling with $1 \lesssim \beta < 3/2$, the latter reflecting the quasifacetted structure of clusters for larger $N = {\rm O}(10^2)$ to $N = {\rm O}(10^3)$; and (iv) classic scaling with $\beta = 3/2$ consistent with macroscopic or mean-field theories for very large $N = {\rm O}(10^3)$ and above. We mainly focus elucidation of behavior in regime (ii), and to some extent regime (iii). To this end, in addition to KMC analysis, we also develop and utilize results from combinatorial analysis of cluster configurations to provide deeper insight.

In Sec. II, we describe our stochastic lattice-gas model for PD-mediated cluster diffusion, and also various strategies for model analysis. In Sec. III, we discuss different possible types or branches of cluster diffusion, and Sec. IV present KMC results providing an overview of the variation of D_N versus N. A brief report of such behavior was recently provided for just one choice of adatom interactions and no kink rounding barrier, $\delta = 0$ [47]. Here, we consider different interactions, and finite $\delta > 0$ as well as $\delta = 0$. We also present a comprehensive analysis and interpretation of diverse aspects of this behavior, as detailed in the following sections. In Sec. V, we describe the variation of the effective diffusivity, $D_N(\delta t)$, for short time increments, δt, where $D_N = \lim_{\delta t \to \infty} D_N(\delta t)$. Characterization of the variation of $D_N(\delta t)$ with δt, which reflects a strong back correlation in cluster motion, is necessary for a reliable extraction of D_N. Additional elucidation of diverse size scaling and cyclic variation of diffusivity in regime (ii), and of intermingling and merging of diffusion branches by regime (iii), is provided in Sec. VI based on counting the number of ground- state and first-excited-state configurations of key classes of clusters. Conclusions are provided in Sec. VII.

II. ATOMISTIC MODEL FOR CLUSTER DIFFUSION

A. Tailored stochastic lattice-gas model

We adopt a tailored model for PD-mediated epitaxial cluster diffusion on metal (100) surfaces, which captures the key features of these systems [16]. In our stochastic lattice-gas model, clusters of adatoms reside on a square lattice of adsorption sites with lattice constant 'a' typically set to unity. Adatoms interact with just nearest-neighbor (NN) attractive lateral interactions of strength $\phi > 0$. They can hop to NN sites, and also to 2nd NN (2NN) sites, provided that hopping retains at least one NN adatom in the cluster. Thus this hopping dynamics preserves NN connectivity (and size) of the cluster. All hop rates have the Arrhenius form $h = \nu \exp[-E_{\rm act}/(k_B T)]$, where ν is a common attempt frequency for both NN and 2NN hops. Let $n_{\rm NN}$ denote the number of in-plane NN adatoms of the hopping adatom in its initial configuration. Then, the activation barrier $E_{\rm act}$, selected to be consistent with detailed-balance, satisfies

$$\begin{aligned} E_{\rm act} &= E_e + (n_{NN} - 1)\phi \text{ for NN hops and} \\ E_{\rm act} &= E_e + (n_{NN} - 1)\phi + \delta \text{ for 2NN hops.} \end{aligned} \tag{1}$$

In this model, the edge atom formation energy equals $E_{\rm form} = \phi$. It also follows that one has activation barriers of E_e for hopping of isolated adatoms along close-packed $\langle 110 \rangle$ edges via NN hops, $E_r = E_e + \delta$ for hopping around corners or kinks via 2NN hops, $E_k = E_e + \phi$ for kink escape via NN hops, and $E_c = E_e + \phi + \delta$ for "core breakup" via 2NN hops (cf. Sec. 1, see Fig. 1). The corresponding rates are denoted h_e,h_r,h_k, and h_c, respectively. The characteristic times associated with these various hop rates are denoted $\tau_e = 1/h_e$, $\tau_k = 1/h_k$, etc. An atom can also be extracted from a straight close-packed step edge with barrier

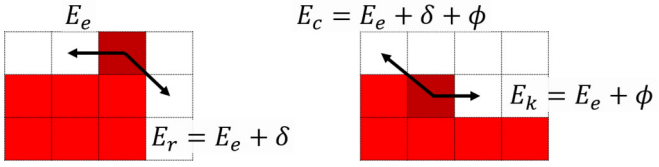

FIG. 1. Schematic of different hopping processes in our stochastic lattice-gas model. Atoms correspond to filled red squares and available adsorption sites to empty squares.

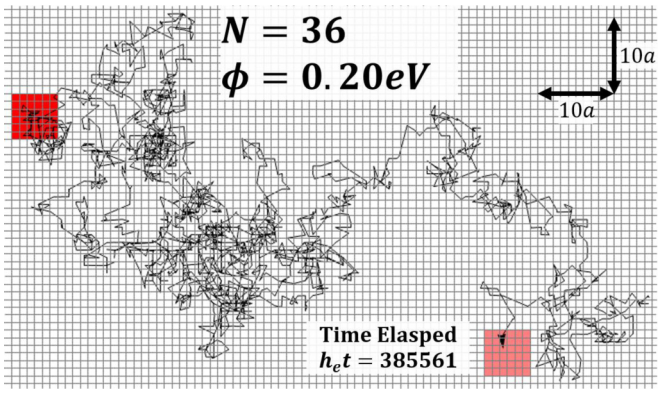

FIG. 2. Trajectory of CM of a diffusing cluster with $N = 36$ for $\phi = 0.20$ eV with $\delta = 0$ at 300 K. Start: red square. End: pink square.

$E_{\text{extract}} = E_e + 2\phi + \delta$, but this process is not prominent, and thus is not shown in Fig. 1.

B. Model analysis

Our focus is on analysis of the diffusion coefficient D_N for clusters of various sizes N (in atoms). To this end, it is appropriate to first define an effective time-dependent diffusion coefficient, $D_N(\delta t) = \langle[\delta r(\delta t)]^2\rangle/(4\delta t)$, where $\boldsymbol{\delta}\mathbf{r}(\delta t)$ is the displacement in the cluster center of mass (CM) in a time interval δt, and $\langle\,\rangle$ is an average of data over a long trajectory. Also we set $[\delta r]^2 = \boldsymbol{\delta}\mathbf{r}\cdot\boldsymbol{\delta}\mathbf{r}$. Comprehensive characterization of model behavior is naturally extracted from KMC simulation. (See Fig. 2 for a typical cluster CM trajectory extracted from such a simulation.) The algorithm used is a standard rejection-free Bortz type algorithm. Note that in contrast to a "pure" random walk, $D_N(\delta t)$ is not in general constant, but can vary for shorter δt due to correlations in the walk of the cluster CM [1,10,22,32]. However, $D_N(\delta t)$ plateaus for larger δt, and the conventional diffusion coefficient is obtained from $D_N = \lim_{\delta t\to\infty} D_N(\delta t) = D_N(\infty)$. Thus appropriate analysis of D_N must account for this transient behavior. For our model where $D_N(\delta t) \propto a^2 h_e$, one has that $D_N(\delta t)/D_N$ versus $h_e\,\delta t$, and $D_N/(a^2 h_e)$ are independent of our choice of E_e and ν, and thus h_e. For reference, choosing $E_e = 0.29$ eV and $\nu = 10^{12.5}\text{s}^{-1}$ mimicking Ag/Ag(100) yields $h_e = 10^{7.6}\text{s}^{-1}$ at 300 K.

We expect $D_N(\delta t)$ to have converged to its plateau value of D_N for $\delta t > \delta t_c$, where $\langle[\delta r(\delta t_c)]^2\rangle \sim a^2$, i.e., when the cluster of CM has moved about one lattice constant. To obtain precise D_N, we need the total length of the trajectory t_{tot} of at least $O(10^3\,\delta t_c)$. Then, $\langle[\delta r(\delta t_c)]^2\rangle$ can be estimated from $O(10^3)$ statistically independent values obtained from nonoverlapping time increments of length δt_c along the trajectory. Overlapping time increments can be used, although then the values of $[\delta r(\delta t_c)]^2$ are not completely independent. We choose $t_{tot} \sim 35000\,\delta t_c$.

It is appropriate to note that D_N can in principle be determined exactly for any cluster size N by analysis of the linear master equations for the stochastic lattice-gas model [30,32]. These master equations track the evolution of the probability of various cluster configurations for the infinite possible number of CM positions. Let Ω_N denote the total number of distinct configurations for a cluster of size N. Then, applying a discrete spatial Fourier transform to these master equations with respect to cluster position converts them into a finite-dimensional $\Omega_N \times \Omega_N$ matrix evolution equation in transform space. One then extracts D_N from analysis of the "acoustic" eigenmode of this evolution matrix, and specifically from its quadratic variation for small wavenumbers. It should also be noted that the transformed $\Omega_N \times \Omega_N$ matrix encodes connectivity between cluster configurations, i.e., indicating which configurations can be directly reached from other configurations by hopping of a single edge atom. Thus the behavior of D_N also reflects this connectivity, although in a nontrivial indirect way. Finally, we emphasize that an exact analysis utilizing this approach is only viable for relatively small clusters since Ω_N increases quickly with N. Nonetheless, it is useful to elucidate behavior in the small cluster size regime (i) (see Appendix A).

The relevance of the total number of cluster configurations Ω_N is already clear from the above discussion of exact analysis. However, one anticipates that not all configurations are equally relevant for the cluster diffusion processes, particularly at lower T. Thus, it is natural to separately analyze the number of ground-state configurations $\Omega_N(0)$, the number of first excited state configurations $\Omega_N(1)$, etc. This analysis involves nontrivial combinatorics exploiting results related to partitions of integers in number theory. Additional useful analysis will involve estimation of the number of kinks in ground state, etc., configurations. Details are provided in Appendices B–D. These results will be utilized to elucidate short-time transient behavior, anomalous scaling observed for moderate sizes, and intermingling and merging of different diffusion branches.

III. DISTINCT BRANCHES OF CLUSTER DIFFUSIVITY FOR MODERATE SIZES

First, we characterize of various branches or classes of cluster sizes for which distinct diffusion behavior is observed in regime (ii) of moderate clusters sizes $N = 9$ to $O(10^2)$. We close with comments on behavior for small clusters with $N < 9$.

A. Nucleation-mediated (NM) diffusion for "perfect" sizes

"Perfect" sizes $N = N_p = L^2$ or $L(L+1)$, with $L = 3,4,\ldots$, have unique nondegenerate ground-state shapes corresponding to perfect squares and near-square rectangles, respectively. This uniqueness does not apply for sizes $N = L(L+n)$ with $n \geqslant 2$ where the $L \times (L+n)$ rectangular configuration is either one of multiple ground states, or corresponds to an excited state. If $\phi/(k_B T)$ is not too small, clusters with $N = N_p$ primarily exist in their unique ground-state shapes, and are subject to "nucleation-mediated" diffusion. In this process, the first step is extraction of one of the four corner atoms onto a straight close-packed $\langle 110\rangle$ step edge, which raises the total energy by $\Delta E = +\phi$. However, typically, this atom will soon return to the more highly coordinated corner site. Thus, to initiate significant cluster restructuring leading to long-range diffusion, it is necessary that a second atom detaches from a corner and aggregates with the first atom before the first atom can return to the corner [9,14,16]. In this way, a step edge dimer is formed, thus potentially nucleating

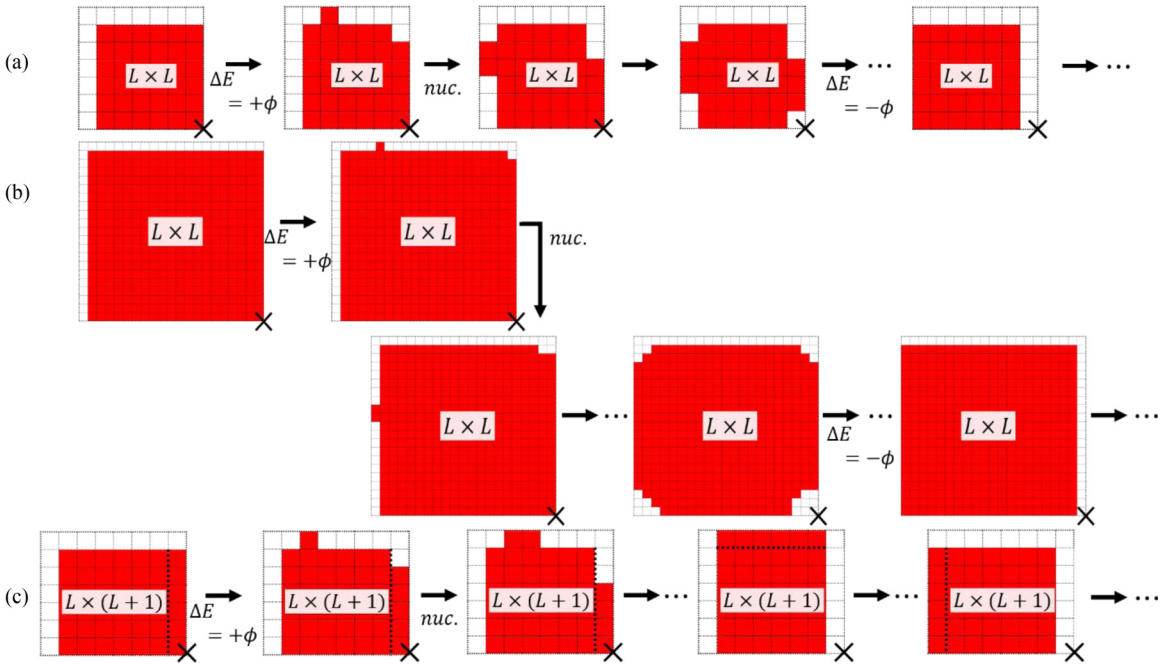

FIG. 3. Nucleation-mediated cluster diffusion for perfect sizes $N_p = L^2$: (a) direct and (b) indirect pathways. (c) Direct pathway for perfect sizes $N_p = L(L+1)$.

a new edge layer. Once this dimer is formed on one edge, subsequent atoms can migrate from kink or corner sites to complete that new edge layer.

The most direct pathway to facilitate translation of the unique ground state for $N_p = L^2$ to a different location, a key component of long-range diffusion, is shown in Fig. 3(a). In this case, two atoms are shifted from one side of the cluster to nucleate a dimer on the opposite side. Thereafter, atoms continue to be shifted from that same side to the opposite side. After each individual atom transfer is completed, the cluster is in a different first excited state configuration with energy $\Delta E = +\phi$ above the ground state. Only when the last atom is transferred does the energy decrease again by $\Delta E = -\phi$. However, we note that there are indirect pathways leading to long-range diffusion as shown in Fig. 3(b). Here, atoms shifted from multiple corners of the cluster whose configuration (after each atom transfer) wanders through a large number of first-excited state configurations. However, to achieve the translated ground state, multiple eroded corners must be largely reconstructed, so that, ultimately, atoms are only removed from a single side of the cluster. Significantly, we note that while long-range diffusion accesses many configurations isoenergetic with the first excited state, it requires repeatedly returning to the unique ground-state shape. Figure 3(c) shows the direct pathway for $N_p = L(L+1)$, which is analogous to that for $N_p = L^2$.

Finally, we comment on the effective barrier for nucleation-mediated diffusion of perfect clusters. An isolated edge atom extracted from the corner of a perfect core exists with low quasiequilibrium density, $n_{\text{eq}} = \exp[-\phi/(k_B T)]$. Mills *et al.* [10] argued that D_N should reflect the nucleation rate $k_{\text{nuc}} \sim n_{\text{eq}} h_c$ to create a dimer on an outer edge. k_{nuc} is the product of the density n_{eq} times the rate h_c of extracting a second atom at the core, as the extracted atom must meet the preexisting edge atom to nucleate a new step edge. Consequently, the effective barrier for cluster diffusion is given by $E_{\text{eff}} = E_e + 2\phi + \delta$ [10,15,17].

B. Facile (FA) cluster diffusion

For clusters of size $N = N_p + 1$ and $N = N_p + 2$, with either $N_p = L^2$ or $L(L+1)$, the edge dimer nucleation process described above for perfect clusters is not necessary for long-range cluster diffusion. For $N = N_p + 1$, we note the existence of a "special" ground-state configuration with an isolated adatom on the edge of a perfect square or rectangular core of N_p atoms. For these special configurations, the isolated edge adatom can readily diffuse around the entire cluster perimeter. For $N = N_p + 2$, "special" ground-state configurations involve an NN pair of edge atoms or edge dimer on a perfect core, where this edge dimer can dissociate and readily reform on another edge. Either process results in no net change of energy. After transferring the isolated edge atom or dimer to new edge of the core, atoms can be transferred from another edge to complete the new edge of the core. This again leaves an isolated atom or dimer on the edge of a perfect core with shifted location.

The above scenario for $N = N_p + 1$ with atoms transferred from a single edge corresponds to a direct pathway to facilitate translation of the special ground-state configuration to a different location. This direct pathway is shown in Fig. 4(a). However, there are indirect pathways leading to the same outcome. Analogous to the above case of perfect sizes, these indirect pathways involve shifting of atoms from multiple corners of the cluster as shown in Fig. 4(b) so the cluster wanders through a large number of ground-state configurations. However, to achieve the translated ground state, most of these eroded corners must be reconstructed so

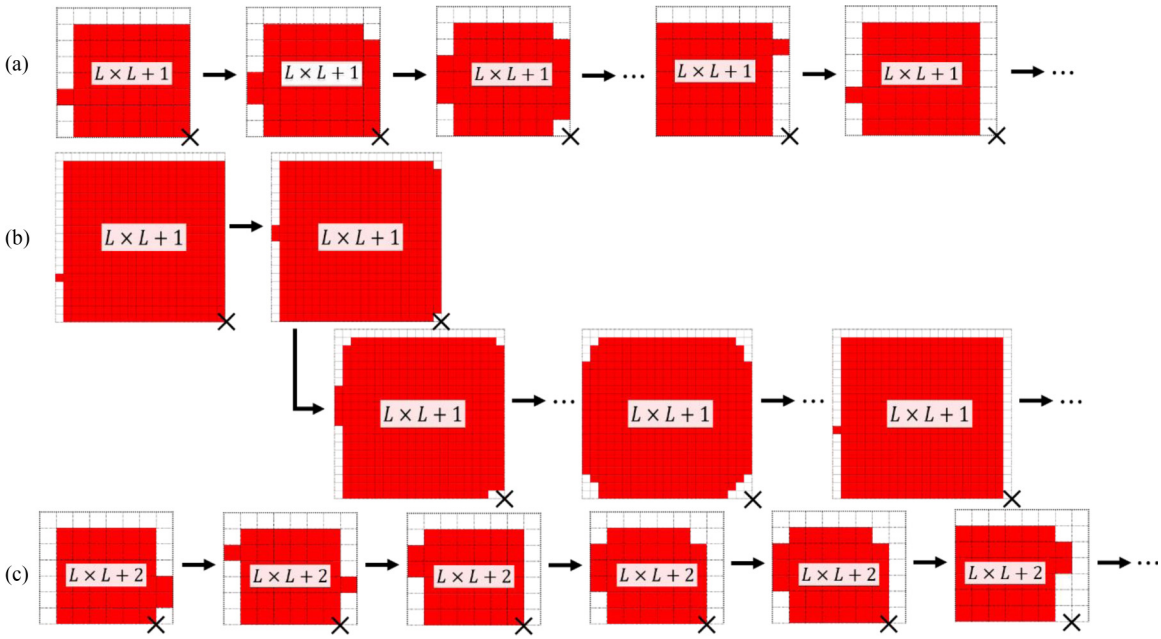

FIG. 4. Facile cluster diffusion for sizes $N = L^2 + 1$: (a) direct and (b) indirect pathways. (c) Direct pathway for sizes $N = L^2 + 2$.

that atoms are only shifted from a single side of the cluster. Shifting atoms from one kink to another does not change the energy after reattachment, so as a result for either direct or indirect pathways, after each atom transfer, the system evolves through a set of configurations isoenergetic with the special ground-state configurations. The direct pathway for $N = N_p + 2$ is shown in Fig. 4(c).

Finally, we emphasize that while the diffusing cluster can wander through many isoenergetic configurations, long-range diffusion (if restricted to these configurations) requires that the cluster repeatedly passes through a special configuration with an isolated atom or dimer at an edge of a perfect core. This is the only way to create a new complete edge on the original perfect core. Also, we note that since diffusion of facile clusters just involves breaking atoms out of kink sites and subsequent edge diffusion, the effective cluster diffusion barrier E_{eff} is simply given by $E_{\text{eff}} = E_e + \phi + \delta$.

C. Other cases of nucleation-mediated cluster diffusion

Clusters of size $N = N_p + n$ with $3 \leqslant n \leqslant L$, for either $N_p = L^2$ or $L(L+1)$, also exhibit nucleation-limited diffusion. The ground states for these sizes include the subclass of configurations with a linear triple or longer string of atoms at the edge of a perfect square or rectangular core. For these configurations, adatoms can readily transfer from the opposite complete edge to that on which the string of n adatoms reside (without raising the energy after transfer), thereby completing that edge. However, this leaves behind a triple or longer string of atoms which cannot readily be transferred to another edge. Certainly, the ground states are degenerate, as starting with the above subclass of configurations, atoms can be removed from multiple corners, and added to the above mentioned string with no net change in energy. However, in any case, nucleation of a dimer on a new outer edge (i.e., on an edge outside the rectangle inscribing the ground-state configurations) is always required to facilitate long-range diffusion of the cluster CM. The same argument as used for perfect clusters indicates that the effective barrier for cluster diffusion equals $E_{\text{eff}} = E_e + 2\phi + \delta$.

D. Facile behavior for small sizes $N < 9$

Diffusion for all small clusters with $N < 9$ is always facile (i.e., not nucleation-mediated). For $N = 2$ or 3, cluster diffusion does not even require breaking atoms out of kink sites, so the effective barrier is even lower than described above for facile diffusion of larger clusters. A dimer CM undergoes a pure random walk on a square grid rotated at 45° to the adsorption sites with lattice constant $a/\sqrt{2}$ hopping at rate h_r. Thus, one has $D_2 = D_2(\delta t) = {}^1\!/_2\, a^2 h_r$ and $E_{\text{eff}} = E_e + \delta$. For a trimer, $D_3(\delta t)$ generally decreases with increasing δt to its asymptotic value, and diffusion is controlled by corner rounding so that again $E_{\text{eff}} = E_e + \delta$ [32]. Cases $N = 5 = 2 \times 2 + 1$ and $N = 7 = 2 \times 3 + 1$ fit within the category $N_p + 1$. Cases $N = 4 = 2 \times 1 + 2$, $N = 6 = 2 \times 2 + 2$, and $N = 8 = 2 \times 3 + 2$ fit within the category $N_p + 2$. Thus all these cases with $4 \leqslant N \leqslant 8$ have $E_{\text{eff}} = E_e + \phi + \delta$, and they all exhibit nonconstant $D_N(\delta t)$. (See Appendix A for an exact master equation based analysis for some of these cases.)

IV. CLUSTER DIFFUSIVITY VERSUS SIZE: KMC RESULTS

A. Cluster diffusivity with no kink rounding barrier ($\delta = 0$)

We first present an overview of KMC results illustrating various size regimes and branches of D_N behavior focusing on the case $\phi = 0.20\,\text{eV}$ and $\delta = 0$ at 300 K. See Fig. 5. For small sizes $N = 4$ to 8, high facile values of D_N are evident. Even higher values for $N = 1$ to 3 are not shown. For moderate sizes, $N = 9$ to $O(10^2)$, we just show for clarity only four distinctive branches: facile $N_p + 1$, facile $N_p + 2$, perfect N_p, and slow $N_p + 3$. The following key features are present:

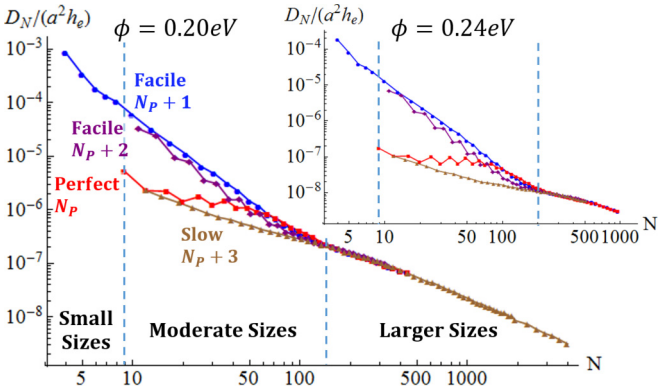

FIG. 5. KMC results for D_N vs N with $\delta = 0$ and $\phi = 0.20\,\text{eV}$ ($\phi = 0.24\,\text{eV}$ in the inset) at 300 K.

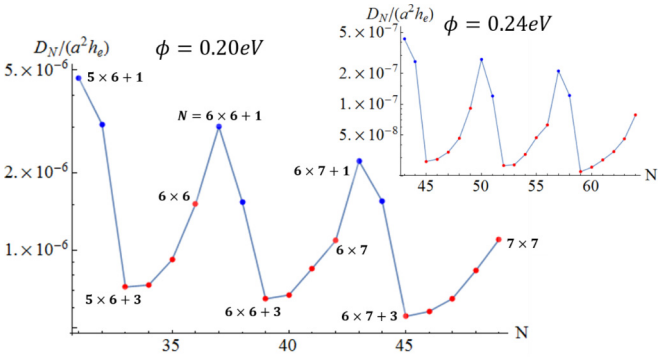

FIG. 7. Cyclical behavior of D_N vs N between minima ($N_P + 3$) and maxima ($N_P + 1$) for $\phi = 0.20$ eV and $\delta = 0$ at 300 K. Inset: $\phi = 0.24$ eV.

(a) initially high values and rapid decay of $D_N \sim N^{-\beta f}$ for facile $N_p + 1$ clusters up to $N \sim 82$ with large $\beta_f \approx 2.3$; similarly high D_N, but less regular decay for facile $N_p + 2$ clusters; (b) the lowest values and slow decay of $D_N \sim N^{-\beta s}$ for sizes $N_p + 3$ for $N \sim 39 - 103$ with small $\beta_s \approx 0.83$; (c) very weak size dependence of D_N for perfect N_p clusters up to $N \approx 81$; perfect N_p and slow $N_p + 3$ branches merge for small $N = 12$ (and $N = 9$); (d) intermingling of D_N for perfect N_p with facile branches for $N_{\text{mingle}} \approx 43$, and subsequent transition to a rapid decrease of D_N for perfect clusters; (e) near-merging of all branches for $N \approx N_{\text{merge}} \approx 150$. For larger sizes $N > N_{\text{merge}}$, if we write $D_N \sim N^{-\beta\text{eff}}$, the effective exponent varies slowly from $\beta_{\text{eff}} \approx 1.09$ for N just above N_{merge}, to $\beta_{\text{eff}} \approx 1.33$ for N from 500–1000, to $\beta = 1.50$ (the asymptotic value for compact clusters) for N from 2000–3600 (see Fig. 6). This latter result is consistent with a kink separation $L_k = {}^1\!/_2 \exp[{}^1\!/_2\, \phi/(k_B T)] \approx 24$ for $\phi = 0.20\,\text{eV}$, given that the asymptotic regime should apply for $N \gg (L_k)^2 \approx 570$.

It is instructive to contrast behavior for $\phi = 0.20\,\text{eV}$ with that for $\phi = 0.24\,\text{eV}$ retaining $\delta = 0$ at 300 K (see the insets for Figs. 5 and 6). All of the features described above are preserved qualitatively for $\phi = 0.24\,\text{eV}$. However, now the deviations between the different branches for moderate sizes are enhanced, which is a natural consequence of larger

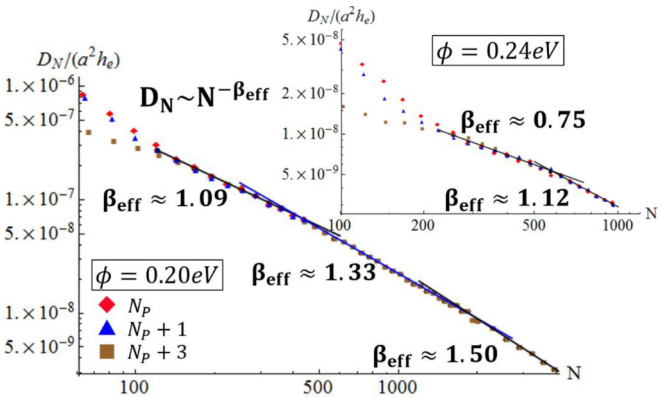

FIG. 6. Post-merging effective scaling behavior of D_N with N for $\phi = 0.20\,\text{eV}$ ($\phi = 0.24\,\text{eV}$ in the inset) and $\delta = 0$ at 300 K.

values of $\phi/(k_B T)$ producing a larger difference between E_{eff} for facile and nucleation-mediated branches. Also, the approach to asymptotic behavior is significantly delayed for larger $\phi/(k_B T)$, as expected given the larger values of L_k. Specifically, for $\phi = 0.24\,\text{eV}$, we find that $\beta_f \approx 2.6$ up to $N \sim 101$, $\beta_s \approx 0.53$ for $N \sim 67$–200, $N_{\text{mingle}} \approx 81$, and $N_{\text{merge}} \approx 200$–250. With regard to scaling for larger sizes, we find that $\beta_{\text{eff}} \approx 0.75$ just above N_{merge}, and $\beta_{\text{eff}} \approx 1.12$ for N from 500–1000. Now $L_k = 52$ for $\phi = 0.24\,\text{eV}$, so we do not access the asymptotic scaling for $N \gg (L_k)^2 \approx 2700$. Naturally, choosing $\phi < 0.20\,\text{eV}$ would minimize the difference between different branches for moderate sizes and accelerate the approach to asymptotic behavior. However, if $\phi/(k_B T)$ is too small, the cluster connectivity constraint becomes artificial. In the limit as $\phi/(k_B T) \to 0$, the clusters become "random animals" with perimeter length proportional to size. This also results in deviations from $\beta = 1.5$ [22].

Next, we consider in more detail diffusion behavior in the moderate size regime. Figure 7 reveals a quasiperiodic variation of D_N with $N = N_p + n$ within each cycle $n = 1$ to $n_{\max}$, where $n_{\max} = L$ for $N_p = L^2$ or $(L–1)L$. Specifically, D_N has a local maximum for $n = 1$, drops significantly for $n = 2$, and again for $n = 3$, where the latter corresponds to the lowest value within each cycle. D_N then increases within each cycle $N = N_p + n$ for increasing $n = 3,4,5,\ldots,n_{\max}$, where $N = N_p + n_{\max}$ recovers the next perfect size above N_p. For example, for N_p=30 (36), $n_{\max}$=6, and $N_p + n_{\max}$=36 (42). Note that the length of these cycles increases for larger N, and that $N = 15,24,35,\ldots$ is the smallest value of N for which one can realize $N_p + 3$, $N_p + 4$, $N_p + 5$, ...

Interestingly, D_N values for perfect sizes for $n = n_{\max}$ within each cycle can be comparable to those for facile clusters for $n = n_{\max} + 2$. On the other hand, they are often well above D_N for $n = 3$ (the slowest clusters). This contrasts a possible perception that perfect sizes should be the slowest. Thus one might question the assignment of nucleation-mediated diffusion for $n = n_{\max}$ versus facile diffusion for $n = n_{\max} + 1$. However, an Arrhenius plot for D_N versus $\phi/(k_B T)$ does show clearly the distinction between E_{act} for these classes. Typically, such Arrhenius plots plot $\ln[D_N]$ versus $1/(k_B T)$ for fixed ϕ, the slope corresponding to E_{eff}. Here, instead we plot $\ln[D_N/(a^2 h_e)]$ versus ϕ for fixed $T = 300\,\text{K}$ yielding a slope of $-n/(k_B T)$ with $n = 1$ ($n = 2$) for facile

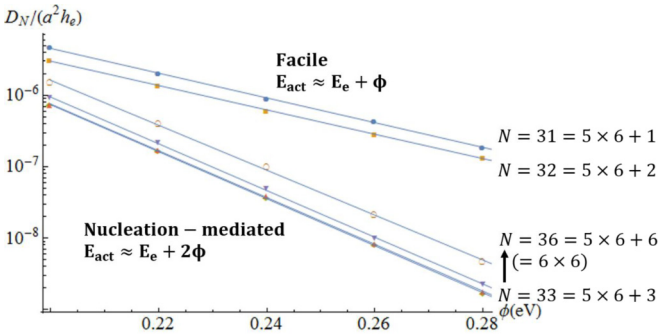

FIG. 8. Arrhenius analysis of D_N for facile ($N_p + 1$, $N_p + 2$) and nucleation-mediated ($N_p + n$ for $n = 3,4,\ldots,n_p$) sizes with $N_p = 30$ and $n_p = 6$. $T = 300$ K is fixed and ϕ is varied.

(nucleation-mediated) diffusion (see Fig. 8). This format is instructive for showing the extent of variation of D_N for the expected range of ϕ values for metal (100) homoepitaxial systems, and for a typical experimental temperature ($T = 300$ K).

B. Cluster diffusivity with a finite kink rounding barrier ($\delta = 0.1$)

The introduction of a significant kink rounding barrier, $\delta > 0$, reduces the magnitude of D_N as a result of the increased E_{eff} described in Sec. III. However, the qualitative features of the different diffusion branches for moderate sizes, and the variation of D_N versus N are the same as for $\delta = 0$. These features are shown in Fig. 9 for $\phi = 0.20$ eV and $\delta = 0.1$ eV at 300 K (and in the inset for $\phi = 0.24$ eV). A detailed characterization of the cyclical behavior of D_N versus N in the moderate size regime is shown in Fig. 10 where again the local maxima (minima) in D_N occur for $N = N_p + 1$ ($N = N_p + 3$). As for $\delta = 0$, D_N for $N = N_p + n$ for the case of perfect sizes with $n = n_{\max}$ is not so far below that for facile sizes with $n = n_{\max} + 2$, but well above that for $n = 3$. Again, we have performed an Arrhenius analysis to reveal that E_{eff} for $n = 3,4,\ldots$, and $n_{\max}$ (nucleation-mediated cases) are all similar, and are clearly above those for $n = n_{\max} + 1$ and $n = n_{\max} + 2$ (facile cases).

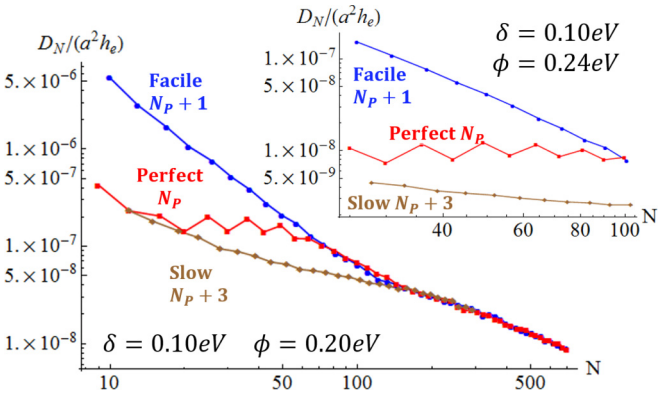

FIG. 9. KMC results for D_N vs N with $\delta = 0.10$ and $\phi = 0.20$ eV ($\phi = 0.24$ eV in the inset) at 300 K.

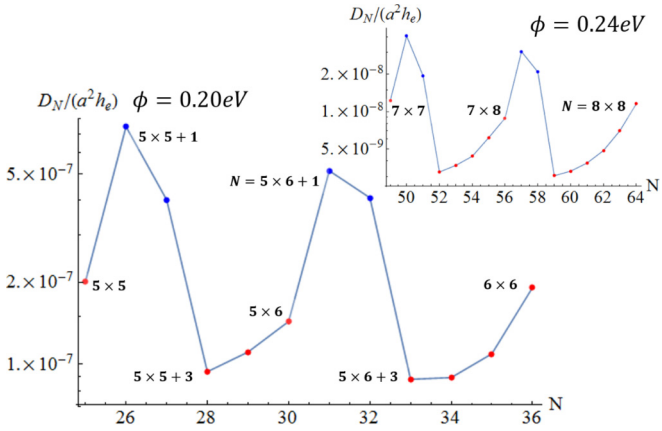

FIG. 10. Cyclical behavior of D_N vs N between maxima (N_P+1) and minima ($N_P + 3$) for $\phi = 0.20$ eV and $\delta = 0.1$ eV at 300 K. (Inset) $\phi = 0.24$ eV and $\delta = 0.1$ eV.

A previous study [17] indicated that introduction of a kink rounding barrier reduces the values of effective scaling exponents β_{eff}. Specifically, this should apply for regime (iii) where facile and nucleation-mediated branches have merged, but prior to the true asymptotic regime of large sizes. For $\phi = 0.20$ eV at 300 K, we find that just after merging, $\beta_{\text{eff}} \approx 0.86$ for $144 \leqslant N \leqslant 325$ when $\delta = 0.1$ eV (versus $\beta_{\text{eff}} \approx 1.09$ for $121 \leqslant N \leqslant 327$ when $\delta = 0$). We also find that $\beta_{\text{eff}} \approx 1.09$ for $361 \leqslant N \leqslant 677$ when $\delta = 0.1$ eV (versus $\beta_{\text{eff}} \approx 1.32$ for $364 \leqslant N \leqslant 2028$ when $\delta = 0$). For $\phi = 0.24$ eV, data are more limited for $\delta = 0.1$ eV as the simulation is more computationally demanding [48]. However, we estimate that just after merging, $\beta_{\text{eff}} \approx 0.71$ when $\delta = 0.1$ eV (versus $\beta_{\text{eff}} \approx 0.75$ when $\delta = 0$). These results confirm the proposal that increasing δ decreases β_{eff}.

V. TIME-DEPENDENT DIFFUSIVITY AND BACK-CORRELATION

The time-dependent diffusion coefficient, $D_N(\delta t) = \langle[\delta r(\delta t)]^2\rangle/4\delta t$, was introduced in Sec. IIB, where $\boldsymbol{\delta r}(\delta t)$ is the CM displacement in a time interval δt. The plateau value of $D_N(\delta t)$ corresponds to the conventional diffusion coefficient, $D_N = \lim_{\delta t \to \infty} D_N(\delta t) = D_N(\infty)$. Thus it is important to understand the transient behavior in order to reliably assess D_N. In fact, this was essential to obtain the smooth cyclical variation of D_N shown in Sec. IV. Here, we consider behavior only in the absence of a kink rounding barrier, $\delta = 0$, although the basic observations and strategies of analysis apply more generally. In Fig. 11, we show KMC simulation results for $\delta = 0$ for the behavior of $D_N(\delta t)/D_N(\infty)$ versus $h_e\delta t$ for sizes within a single cycle $N = N_p + 1$ to $N = N_p + n_{\max}$ (cf. Sec. IV). As noted in Sec. IIA, the form of these curves is independent of the choice of h_e. There is a strong decrease in $D_N(\delta t)$ to its plateau value $D_N = D_N(\infty)$. In Sec. VA, we estimate the short time-increment values, $D_N(\delta t \to 0)$, for special cases of perfect and facile sizes. Then, in Sec. VB, we provide further insight into the underlying back correlation in cluster motion.

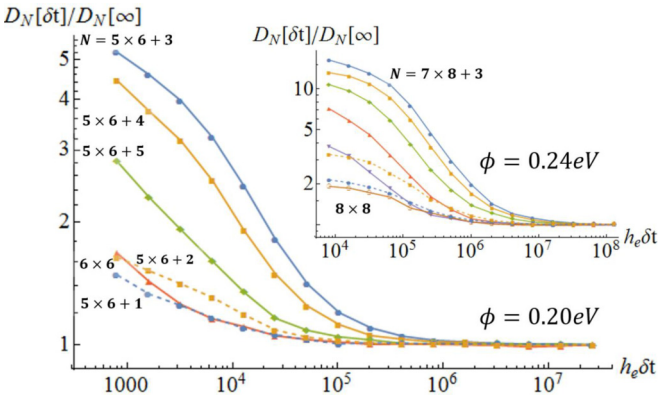

FIG. 11. Time-dependent diffusion coefficients reflecting backward correlation in the CM motion for various cluster sizes within a cycle (see text) with $\phi = 0.20$ eV (and $\phi = 0.24$ eV in the inset) for $\delta = 0$ at 300 K. Here $D_N(\infty) = \lim_{\delta\to\infty} D_N(\delta t) = D_N$.

A. Short-time behavior of $D_N(\delta t)$

Our estimate of the value of $D_N(\delta t \to 0)$ assumes independent contributions to the mean-square displacement of the cluster CM from the short-time motion of all isolated (singly coordinated) edge atoms and all doubly coordinated kink atoms. Thus, we sum over these contributions to obtain $D_N(\delta t \to 0)$. For short-time increments, δt, the mean-square displacement of isolated edge atoms (called "monomers" below) from their initial position satisfies $\langle \delta r_e(\delta t)^2\rangle \approx 2h_e\delta t$, $3h_e\delta t$, and $4h_e\delta t$ for atoms on straight close-packed steps that can make two NN hops, atoms at corners that can make one NN and one 2NN hop, and atoms that can make two 2NN hops, respectively. The latter case is rare for larger clusters, so effectively one has $2h_e\delta t \leqslant \langle \delta r_e(\delta t)^2\rangle \leqslant 3h_e\delta t$. The mean-squared displacement of kink atoms (just called "kinks" below) from their initial position satisfies $\langle \delta r_k(\delta t)^2\rangle \approx 3h_k\delta t$ for atoms that can make one NN and one 2NN hop, and $\langle \delta r_k(\delta t)^2\rangle \approx 4h_k\delta t$ for corner atoms that can make two 2NN hops. Thus one has that $3h_k\delta t \leqslant \langle \delta r_k(\delta t)^2\rangle \leqslant 4h_k\delta t$. To simplify the analysis below, we will not discriminate between the different categories of monomers and kink atoms, and will interpret $\langle \delta r_e(\delta t)^2\rangle$ and $\langle \delta r_k(\delta t)^2\rangle$ as suitable averages over all categories. Subsequently, we will just obtain upper and lower bounds for $D_N(\delta t \to 0)$ using the above upper and lower bounds on $\langle \delta r_{e,k}(\delta t)^2\rangle$.

Before presenting our approximation for $D_N(\delta t \to 0)$, we also note that when a periphery atom is shifted by one lattice constant in a certain direction, the CM of the cluster is shifted by $1/N$ in that direction. This will produce an additional factor of $1/N^2 = 1/L^4$ in our analysis of mean-squared cluster displacement. Thus our expression for $D_N(\delta t \to 0)$ becomes

$$D_N(\delta t \to 0) \approx \frac{1}{4N^2}\sum_i \left(n_{N,e}(i)\frac{\langle \delta r_e(\delta t)^2\rangle}{\delta t} + n_{N,k}(i)\frac{\langle \delta r_k(\delta t)^2\rangle}{\delta t}\right) \times \frac{\exp[-E_i/(k_BT)]}{Z}, \quad (2)$$

where $n_{N,e}(i)$ and $n_{N,k}(i)$ are the number of monomers and kinks in ith state with energy E_i, and $Z = \sum_i \exp[-E_i/(k_BT)]$ is the relevant partition function. We use this result to estimate $D_N(\delta t \to 0)$ focusing on two special cases. Further details are provided in Ref. [49].

Perfect sizes $N_p = L^2$. The ground state is unique, i.e., $\Omega_{L^2}(0) = 1$, and has a square shape with no monomers and four kinks. Thus the total contribution to $D_N(\delta t \to 0)$ from the ground state is of order h_k/N^2, denoted $O(h_k/N^2)$. There are $4\times(4L-2)$ first excited states where an atom is shifted from one of the four corners of the ground state and placed as a monomer on an edge, and $2\times4L$ first excited states with a monomer on an edge of a $(L-1)\times(L+1)$ rectangle. Thus the total number of first excited states with a monomer is $\Omega'_{L^2}(1) = (24L-8)$. The total contribution to $D_N(\delta t \to 0)$ from these states is dominated by monomer hopping and is $O(\Omega'_{L^2}(1)\, h_e \exp[-\phi/(k_BT)]/N^2) = O(Lh_k/N^2)$, which exceeds the contribution from the ground state.

The great majority of the $\Omega_{L^2}(1)$ first excited states have no monomers, but many kinks. If $n_{L^2,k}(1)$ denotes the number of kinks in such states, then one has that $3 \leqslant n_{L^2,k}(1) \leqslant 2(1+\sqrt{2L+1})$ (see Appendix B). Despite the penalty of a Boltzmann factor of $\exp[-\phi/(k_BT)]$, the total contribution of kinks in first excited states, $O(n_{L^2,k}(1)\Omega_{L^2}(1)h_k\exp[-\phi/(k_BT)]/N^2)$, becomes comparable to those above for moderate N due to the large number of first excited states $\Omega_{L^2}(1)$. Specifically, the contribution becomes comparable when $\Omega_{L^2}(1)\exp[-\phi/(k_BT)] \sim O(1)$, which occurs when $N \sim 49$ (81) for $\phi = 0.20$ eV (0.24 eV) (see Appendix C).

Finally, we find that it is also necessary to consider contributions from the subclass of second excited states, which include a monomer. We note that the number of such states, $\Omega'_{L^2}(2) \sim 4L\ \Omega_{L^2-1}(1)$ (see Appendix D for a more precise analysis) is somewhat larger than $\Omega_{L^2}(1)$ for $N \sim O(10^2)$. The total contribution of such states is of order $O(\Omega'_{L^2}(2)h_e\exp[-2\phi/(k_BT)]/N^2)$, which is of the same order as the above contributions for moderate cluster sizes if one accounts for this large $\Omega'_{L^2}(2)$ and for the high monomer hop rate h_e. Combining these four types of contributions (of which the last one dominates for moderate N) yields estimates for $D_N(\delta t \to 0)$ close to simulation values as shown in Fig. 12 for $h_e\delta t = 1$, $\phi = 0.20$ eV.

It is appropriate to note that the contributions explicitly included above correspond to exactly the configurations that arise in our picture of nucleation-mediated cluster diffusion for moderate sizes. The cluster primarily exists in the ground state, but must access first excited states in order to initiate motion. However, transitions between the numerous monomer-free first excited states involve second excited states with a monomer. We note that contributions from second excited states without a monomer and higher excited states are of lower order than those above since the number of relevant configurations is not substantially greater than $\Omega_{L^2}(1)$ or $\Omega'_{L^2}(2)$.

Facile clusters of sizes $N = L^2+1$. Here, we mimic the above analysis for perfect clusters. For $N = L^2+1$, there are $4L$ ground states with a monomer, i.e., $\Omega'_{L^2+1}(0) = 4L$, each of which provide a contribution $O(h_e\Omega'_{L^2+1}(0)) \sim O(h_e)$ dominated by monomer hopping. All ground states

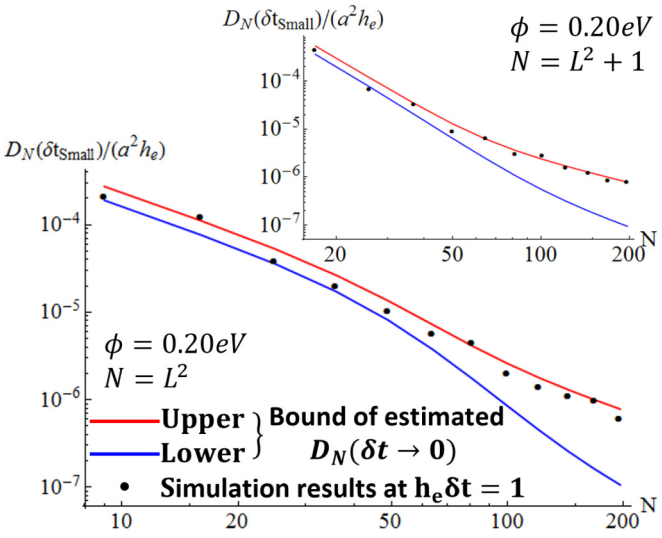

FIG. 12. Estimated upper and lower bounds of $D_N(\delta t \to 0)$ vs simulation results for $h_e \delta t = 1$ (black dots) for $N = L^2$ (inset $N = L^2 + 1$) for $\phi = 0.20$ eV and $\delta = 0$ at 300 K.

contribute by kink hopping with total contribution of order $O(n_{L^2+1,k}(0) h_k \Omega_{L^2+1}(0)) \sim O(h_e)$ for $N \geqslant 65(101)$ with $\phi = 0.20$ eV (0.24 eV), using the feature that $\Omega_{L^2+1}(0)$ grows far more quickly than $\Omega'_{L^2+1}(0)$. Note also that $n_{L^2+1,k}(0) \leqslant 2(1 + \sqrt{2L-1})$ (see Appendix B). The third contribution comes from the first excited states with a monomer, where the number of such states satisfies $\Omega'_{L^2+1}(1) \sim 4L\, \Omega_{L^2}(1)$ (see Appendix D for a more precise analysis). Thus the total contribution of first excited states is $O(h_e \Omega'_{L^2+1}(1) \exp[-\phi/(k_B T)]) \sim O(h_e)$, due to the large number of states considered, $O(\Omega'_{L^2+1}(1) \exp[-\phi/(k_B T)]) \sim O(1)$ for $N \geqslant 65(101)$ with $\phi = 0.20$ eV (0.24 eV). Combining these three contributions yields estimates for $D_N(\delta t \to 0)$ close to simulation values [see Fig. 12 (inset) for $h_e \delta t = 1$, $\phi = 0.20$ eV]. Note that the states explicitly included above are exactly those in our picture of facile diffusion for moderate sized clusters, and other states have a lower order contribution.

Other cases and further comparison. The above analysis readily extends to other cases. For the nucleation-mediated cases, $N = N_p + n$ with $n = 3,4, \ldots, n_{\max}$, we claim that $D_N(\delta t \to 0)$ will decrease from a local maximum for $N = N_p + 3$ to a local minimum for $N = N_p + n_p$ (corresponding to perfect clusters). Clusters within this class for $N = N_p + 3$ have the highest ground-state degeneracy and importantly also the highest number of kinks. Consequently, the contribution from the ground states $O(n_{L^2+3,k}(0) h_k \Omega_{L^2+3}(0))$ for $N = N_p + 3$ will exceed that for perfect clusters due to the substantial number of kink sites, $n_{L^2+3,k}(0) \leqslant 2(1 + \sqrt{2L-5})$. The larger factor $\Omega_{L^2+3}(0)$ versus $\Omega_{L^2+n_p}(0) = 1$ does not in itself boost $D_N(\delta t \to 0)$, as this factor also appears in the partition function denominator of (2). For $N = N_p + n$, as n increases from 3 towards n_p, the degeneracy of the ground-state and importantly the typical number of kinks decreases. Correspondingly, $D_N(\delta t \to 0)$ also decreases with increasing $n = 3,4, \ldots$. Finally, comparing the above analysis for perfect and facile clusters shows that $D_N(\delta t \to 0)$ for perfect clusters is smaller roughly by a Boltzmann factor of $\exp[-\phi/(k_B T)]$ than for facile clusters.

B. Further analysis of back correlation

The substantial characteristic time δt_c, associated with the transient short-time diffusion behavior of $D_N(\delta t)$, is evident from Fig. 11. These data suggest $h_e\, \delta t_c \sim 10^5$–$10^6$ (10^6–10^7) for $\phi = 0.20$ (0.24) eV at 300 K, at least for nucleation-mediated (NM) cluster diffusion, where the branch with $N = N_p + 3$ appears to have a larger δt_c than for $N = N_p + n$ with $n > 3$. This latter feature is confirmed by a suitably rescaled version of Fig. 11, which is shown in Ref. [49]. It is reasonable to expect that for NM diffusion, δt_c should reflect the characteristic time $\delta t_{\rm nuc} = 1/k_{\rm nuc}$ to nucleate a dimer on an outer edge. This implies that $h_e\, \delta t_c \sim h_e\, \delta t_{\rm nuc} \sim \exp[+2\phi/(k_B T)] \sim 10^{6.4}$ ($10^{8.0}$) for $\phi = 0.20$ eV ($\phi = 0.24$ eV) at 300 K with $\delta = 0$. These crude estimates at least roughly reflect those from Fig. 11, and also the feature that δt_c increases with ϕ. The larger δt_c for $N = N_p + 3$ plausibly reflects the larger degeneracy of the ground state and the larger typical number of kinks for that cluster size (see Sec. VII), which can inhibit nucleation of new outer edges.

For facile clusters with $N = N_p + 1$ or $N = N_p + 2$, Fig. 11 perhaps suggests a somewhat shorter δt_c although this is not evident in the further rescaled plots in Ref. [49]. One might expect a shorter δt_c based upon the feature that nucleation is not needed so correspondingly $E_{\rm eff}$ is lower, and the long-time diffusion coefficient is higher. However, other factors, such as the high degeneracy of the ground state (see Sec. VII), no doubt play a role in determining δt_c.

As noted previously, assessment of transient behavior in $D_N(\delta t)$ is essential for precise determination of D_N, where precise determination becomes more demanding for longer δt_c. Thus, accurate treatment of the case $N = N_p + 3$ is most demanding, failure to do so leads to a distorted representation of the cyclical behavior of D_N versus N (see Sec. VII). Practically, we initially estimate that the plateau in $D_N(\delta t)$ is achieved for $\delta t > \delta t^*$, where $\langle [\delta r(\delta t^*)]^2 \rangle$ is of the order of a^2 (where δt^* gives a measure of δt_c). The total length of the trajectories used to determine D_N is $t_{\max} \sim 35\,000\, \delta t^*$ where data is collected only for $\delta t \gg \delta t^*$. (For reference, choosing $E_e = 0.29$ eV and $\nu = 10^{12.5}$ s^{-1} for Ag/Ag(100) yields $h_e = 10^{7.6}$ s^{-1} at 300 K, and $t_{\max} \sim 70\,000$ s for $N = 59$.)

Finally, we elaborate on the interpretation of the decrease of $D_N(\delta t)$ to a plateau value as corresponding to a back correlation in the walk of the cluster. Consider the canonical model of a correlated walk with hops to NN sites on a lattice at total rate h. If $\mathbf{r_j}$ denotes the displacement of the jth hop, then the displacement of the jth hop is correlated to that of previous hops as quantified by $A(k) = \langle \mathbf{r_j} \cdot \mathbf{r_{j-k}} \rangle / \langle \mathbf{r_1} \cdot \mathbf{r_1} \rangle$, where $A(k) < 0$ for back correlation. Here $\langle \mathbf{r_j} \cdot \mathbf{r_j} \rangle = \langle \mathbf{r_1} \cdot \mathbf{r_1} \rangle$ for all j. Adapting results for the time-dependent diffusion coefficient $D(\delta t)$ for this system into a continuous-time framework for a large number of hops yields

$$D(\delta t)/D(\delta t \to 0) = 1 + 2 \int_{0 \leqslant u \leqslant h\delta t} du\, A(u), \quad \text{so that}$$

$$A(h\delta t) = \frac{1}{2} d/ds [D(s)/D(0)]|_{s = h\delta t}. \qquad (3)$$

TABLE I. Number of isoenergetic ground-state configurations $\Omega_N(0)$ and restricted isoenergetic configurations $\Omega_N^*(0)$ for $N=L^2+1$.

$N = L^2 + 1$	10	17	26	37	50	65	82	101
$\Omega_N(0)$	28	80	210	504	1148	2480	5160	10 360
$\Omega_N^*(0)$	28	80	202	464	988	1976	3748	6792

Note that the magnitude of cumulative (integrated) correlation is strictly bounded by $^1/_2$ in this formulation. Clearly, the decrease in $D_N(\delta t)$ with increasing δt shown in Fig. 11 corresponds to back-correlation $A(u) < 0$. One could extract an effective $A(u)$ from the form of $D_N(\delta t)$ after assigning an effective total hop rate.

VI. FURTHER ANALYSIS OF DIFFUSIVITY VIA CONFIGURATION COUNTING

Deeper insight into the diverse aspects of cluster diffusion behavior described in Sec. IV follows from exploiting results of a combinatorial analysis of cluster configurations corresponding to ground states and first excited states. This nontrivial analysis utilizes results related to (number theoretic) partitions of integers. Details are relegated to Appendix C.

A. Anomalous scaling for facile clusters

As noted in Sec. IV, for facile $N_p + 1$ clusters, one finds initially high values and rapid decay of $D_N \sim N^{-\beta f}$ with large $\beta_f \approx 2.3$ ($\beta_f \approx 2.6$) up to $N \sim 82$ (101) for $\phi = 0.20$ (0.24) eV at 300 K. These exponent values are far larger than any reported in previous studies. To elucidate this behavior, recall that long-range diffusion requires that the cluster repeatedly passes through a special configuration with one edge atom on a perfect core. We suggest that the behavior of D_N reflects the possibility to wander through a large number of isoenergetic ground-state configurations far removed from the special configuration, where the number $\Omega_N(0)$ of these states increases rapidly with increasing N. After the system leaves the special configuration, let $t_{\rm ret}$ denote the mean-time for the system to return, where one expects that $D_N \sim a^2/t_{\rm ret}$. A key result of Montroll and Weiss [50] for regular lattices is that this return time is directly proportional to the size of the system, independent of dimension. This in turn suggests that $D_N \sim a^2 h_c/\Omega_N(0)$. The results presented in Table I indicate that $\Omega_N(0) \sim N^\alpha$ with $\alpha \approx 2.6$ up to $N \sim 100$, reasonably consistent with the above large β_f values (see Appendix C).

For another perspective, note that all isoenergetic states have equal population. Thus the probability $P_{\rm ret}$ that the system is in a ground state, which can directly transition to (or "return to") the special configuration, scales like $P_{\rm ret} \sim 1/\Omega_N(0)$. Then, we claim that $D_N \sim a^2 h_c P_{\rm ret}$, which recovers the above result.

The exact behavior of D_N actually depends not just on the number of isoenergetic configurations, but on their connectivity to the special configuration [30,32]. Presumably, configurations more closely connected to the special configuration should play a more significant role. This motivates analysis of the number $\Omega_N{}^*(0)$ of restricted isoenergetic configurations where starting from the special configuration, additional atoms are shifted to the edge with the isolated atom from just the outermost layer of the other edges. Analysis of $\Omega_N *(0)$ data also in Table I produces a modified exponent of $\alpha \approx 2.4$, again reasonably consistent with the β_f values.

TABLE II. Values of $\Omega_{N_P}(1)$ for $N_p = L^2$.

$N = L^2$	25	36	49	64	81	100
$\Omega_{N_P}(1)$	496	1140	2472	5152	10352	20208

B. Intermingling of perfect and facile branches

While D_N for facile clusters decreases strongly with N for moderate sizes, the variation of D_N for perfect clusters is extremely weak. The latter behavior reflects the feature that diffusion of perfect clusters is largely controlled by the nucleation step, which depends weakly on N, and not so much on the subsequent transfer of atoms to complete the new edge. Thus the D_N in the facile branch, which are large for smaller sizes but rapidly decreasing naturally meet and "intermingle" with the D_N of the perfect branch, which are lower for small sizes but slowly decreasing. Since D_N for the $N_p + 3$ branch are even lower than for perfect clusters and decrease with increasing N, this branch remains separate from the facile and perfect clusters at the point of intermingling.

The distinction between perfect clusters and facile (or other) classes of clusters is predicated on the feature that the former primarily exist in their ground states. However, perfect N_P clusters would have a significant probability of being in the first excited state when $\Omega_{N_P}(1)/\Omega_{N_P}(0) \approx \Omega_{N_P}(1) \approx \exp[\phi/(k_B T)]$, where again $\Omega_N(n)$ gives the number of isoconfigurations for the nth excited state for a cluster of size N, and $\Omega_{N_P}(0) = 1$. Results for $\Omega_{N_P}(1)$ determined from combinatorial analysis in Appendix C are reported in Table II. For $\phi = 0.20$ eV (0.24 eV), the Boltzmann factor $\exp[\phi/(k_B T)] = 2290$ (10730), and thus intermingling perfect and facile branches should occur around $N = N_{\rm mingle} \sim 49$ (81). This prediction is consistent with the behavior shown in Fig. 13 where $N_{\rm mingle}$ is indicated by a dashed vertical line. Note that D_N for perfect (facile) clusters decreases more quickly (slowly) after intermingling.

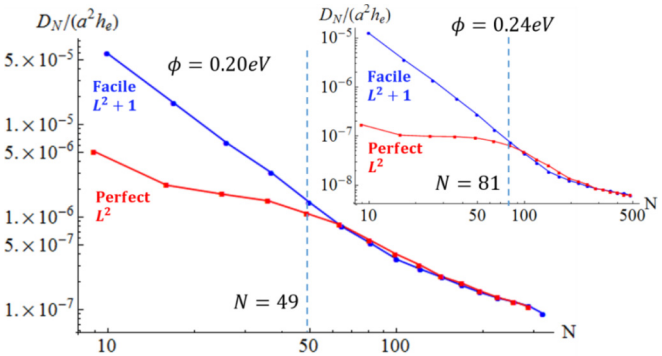

FIG. 13. The intermingling sizes of L^2 and $L^2 + 1$ branches predicted with thermodynamics for $\phi = 0.20$ eV (inset: $\phi = 0.24$ eV) with $\delta = 0$ at 300 K.

TABLE III. Values for $\Omega_{N_P+3}(0)$, $\Omega_{N_P+3}(1)$, and the ratio $\Omega_{N_P+3}(1)/\Omega_{N_P+3}(0)$ for $N = L^2 + 3$.

$N = L^2 + 3$	147	172	199	327	364	403
$\Omega_{N_P+3}(0)$	10360	20216	38416	407968	706034	1.20×10^6
$\Omega_{N_P+3}(1)$	1.53×10^7	3.95×10^7	9.86×10^7	2.86×10^9	6.25×10^9	1.34×10^{10}
$\frac{\Omega_{N_P+3}(1)}{\Omega_{N_P+3}(0)}$	1475	1955	2565	7002	8847	11116

C. Merging of all branches of cluster diffusivity

As noted above, the feature that D_N for the $N_p + 3$ branch are lower than those for perfect clusters and also that they decrease slowly with N delays merging with the perfect and facile branches. It is appropriate to note that while both $N_p + 1$ and $N_p + 3$ branches have a high ground-state degeneracy, this only produces strong size dependence of D_N for the former. Why? Long-range diffusion of clusters for sizes $N_p + 3$ does not require repeatedly passing through a single special configuration, unlike for $N_p + 1$. Thus the strong increase in the number of ground states with increasing N does not induce a strong reduction in D_N for $N = N_p + 3$.

Analogous to our assessment of intermingling and perfect branches, here we argue that the distinctive nature of $N_p + 3$ clusters (relative to $N_p + 1$) is lost when the ratio of the number of the first excited states $\Omega_{Np+3}(1)$ to the number of ground states $\Omega_{Np+3}(0)$ satisfies $\Omega_{Np+3}(1)/\Omega_{Np+3}(0) \approx \exp[\phi/(k_BT)]$. The method to count the number of isoenergetic states, $\Omega_{N_P+3}(1)$, $\Omega_{N_P+3}(0)$ is the same as that of counting $\Omega_{N_P}(1)$. Relevant results are presented in Table III (see Appendix C for details). The predicted sizes for merging, $N = N_{\text{merge}} \approx 199$ (403) for $\phi = 0.20$ eV (0.24 eV), are indicated by dashed vertical lines in Fig. 14.

D. Analysis of the cyclical variation of cluster diffusivity

It is clear from Fig. 5 that D_N actually increases with increasing size $N = N_p + n$, within each cycle $n = 3,4,5,\ldots,n_{\max}$, where $n_{\max} = L$ for $N_p = L^2$ or $(L–1)L$ recovers a perfect cluster. A local minimum (maximum) in D_N occurs for the $n = 3$ ($n = n_{\max} + 1$). We suggest that the key feature controlling this behavior is a strong decrease with increasing n in the degeneracy of the ground state from a maximum for $n = 3$ to a minimum for $n = n_{\max}$. The minimum is 1 for $N_p = L^2$, and 4 for $N_p = (L–1)L$. A larger number of degenerate ground states means a higher probability that the cluster is in a configuration with multiple atoms removed from the corners and thus many kink sites which can trap diffusing edge atoms. This makes nucleation of a new outer edge more difficult, as the lifetime of isolated atoms is reduced). Many kinks also inhibit transfer atoms to complete that new outer edge. Consequently, D_{Np+n} increases with increasing n. We remark that "oscillations" in D_N versus N were observed in previous simulation studies [7,9]. However, the analysis was limited [9], e.g., perhaps giving a misimpression that perfect clusters $N = N_p$ diffuse slowest, and not recognizing that $N = N_p + 2$ (as well as $N_p + 1$) are facile.

Finally, we emphasize the substantial computational challenge in obtaining precise values for D_N particularly for $N = N_p + 3$ or $N_p + 4$. This is evident from Fig. 11 where one must sample over substantially longer time intervals δt to obtain the correct asymptotic value of D_N. Lack of precision in analysis fails to produce the correct trend in D_N within each cycle. To illustrate this issue, in Fig. 15, we present results obtained for $D_N(\delta t)$ with a small $h_e\delta t = 811$ and with a large $h_e\delta t = 12970$ for $\phi = 0.20$ eV and $\delta = 0$ at 300 K (both well below $h_e\delta t_c = 10^5$–10^6). Even the latter is insufficiently large to recover the correct asymptotic behavior. Such analysis gives the misimpression that the slowest diffusion occurs not for $N = N_p + n$ with $n = 3$, but for somewhat larger n.

VII. CONCLUSIONS

Our precise KMC analysis of a tailored but effective model for cluster diffusion on metal (100) surfaces has revealed extraordinarily diverse behavior particularly for the regime of moderates sizes $9 \leqslant N \leqslant O(10^2)$. Perhaps unexpectedly, the slowest diffusion does not occur for perfect sizes

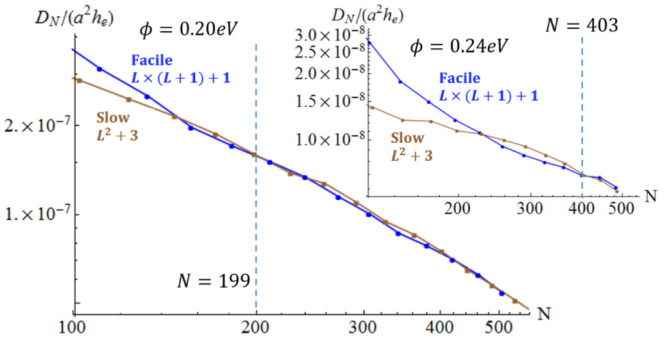

FIG. 14. The intermingling sizes of $L^2 + 3$ and $L \times (L + 1) + 1$ branches predicted with thermodynamics for $\phi = 0.20$ eV (inset $\phi = 0.24$ eV) with $\delta = 0$ at 300 K.

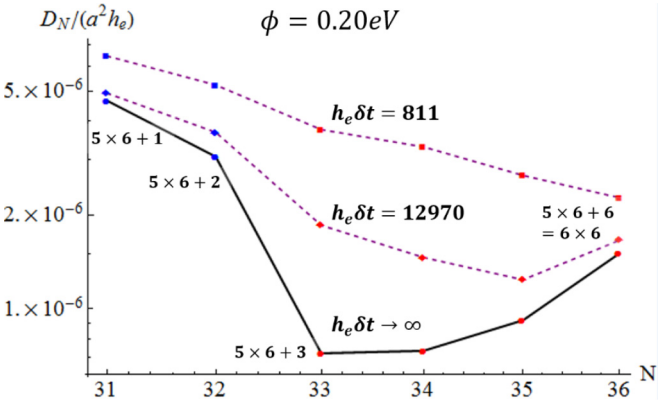

FIG. 15. Illustration of analysis with diffusion coefficients not converged for $\phi = 0.2$ eV with $\delta = 0$ at 300 K for $31 \leqslant N \leqslant 36$.

$N = N_p = L^2$ or $L(L+1)$ with unique square or near-square ground-state shapes, but rather for $N = N_p + 3$. However, the slowest short-time diffusivity does occur for perfect sizes. We are able to elucidate the distinct behavior of different branches (facile, perfect, and slow) in this regime, exploiting combinatorial analysis of the number of ground states, first excited states, etc.

Also of interest is the intermingling and merging of these branches for larger N. Combinatorial analysis was also utilized to provide insight into the intermingling and merging points essentially by determining at what cluster size thermal fluctuations or excitations smeared the distinction between various branches. As an aside, we note that another way to assess merging is based on the realization that the effective Arrhenius energy $E_{\rm eff}$ for cluster diffusion adopts a higher value, $E_{\rm eff} = E_e + 2\phi + \delta$, for nucleation-mediated diffusion for moderate sizes than in the asymptotic regime of large sizes where $E_{\rm eff} = E_e + \phi + \delta$. We have checked that for nucleation-mediated diffusion, the effective value of $E_{\rm eff}$ decreases with increasing N and is reduced to about $E_{\rm eff} = E_e + 1.5\phi + \delta$ at the point where merging occurs (see Ref. [49]).

We have not presented a comparison with experimental data. However, our results are particularly valuable in revealing the complexity of behavior for moderate sizes and the potential shortcomings in extracting size scaling exponents from data over a limited size range. We plan to apply our modeling to analyze the behavior for Ag clusters on Ag(100) where recent experimental analysis [3] has suggested somewhat lower exponent values from those determined previously [2] (but where in both cases the exponent is significantly below the classic value of $\beta = 3/2$). Also, with regard to experiment, we note that facile clusters of size $N = N_p + 1$ should be susceptible to dissociation of the isolated edge atom in the special ground-state configuration with this atom and a perfect core. However, this is only one of many isoenergetic ground states for larger N reducing this likelihood. For $N = N_p + 2$, there are no isolated edge atoms in the ground state, so this issue does not arise.

Finally, we note that basic features of results from our modeling should be more general than for cluster diffusion on metal (100) surfaces. Similar behavior is expected for metal (111) surfaces. The surprising feature that perfect clusters do not have the lowest diffusivity may even extend to supported 3D clusters. However, there are certainly other fundamental issues that remain to be addressed. For example, the degeneracy of the ground state is important in explaining various basic features of behavior. However, if one includes more lateral adatom interactions, degeneracies can be broken, so how does this change the behavior from that of our basic model?

ACKNOWLEDGMENTS

K.C.L. and J.W.E. were supported for this work by NSF grant CHE-1507223. They performed detailed KMC analysis and developed theoretical interpretations of the data. D.J.L. was supported by the USDOE, Office of Science, Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biological Sciences, and his work was performed at Ames Laboratory which is operated by Iowa State University for the USDOE under Contract No. DE-AC02-07CH11358. He performed the exploratory KMC simulations revealing diverse behavior together with preliminary analysis.

APPENDIX A: EXACT ANALYSIS FOR THE SMALL CLUSTER SIZE REGIME $N < 9$

Exploiting the exact master equation analysis discussed in the text, for dimers with two linear configurations (rotated by 90°), one finds that [32]

$$D_2 = D_2(\delta t) = (a^2/2)h_r \text{ so } E_{\rm eff} = E_e + \delta. \tag{A1}$$

For trimers with six distinct configurations (two linear and four bent), $D_3(\delta t)$ generally decreases with increasing δt to its asymptotic value [32]

$$D_3 = (a^2/3)h_r h_e/(h_r + h_e) \text{ so } 1/D_3 = 3a^{-2}(1/h_r + 1/h_e). \tag{A2}$$

The latter expression confirms the obvious feature that both edge diffusion and corner rounding are required for long-range diffusion. In this case, one does not in general have perfect Arrhenius behavior except for $\delta = 0$, where $E_{\rm eff} = E_e$. However, in practice, for typical nonzero δ, one has that $E_{\rm eff} = E_e + \delta$. For tetramers with 19 distinct configurations, $D_4(\delta t)$ generally decreases with increasing δt to its asymptotic value

$$D_4 = h_c h_r [6(h_e)^3 + 38(h_e)^2 h_r + 35 h_e (h_r)^2 + 6(h_r)^3]/ \{(18h_c + h_r)[(h_e)^3 + 10(h_e)^2 h_r + 24 h_e (h_r)^2 + 9(h_r)^3]\}. \tag{A3}$$

As expected, this result shows that core breakup is essential for long-range cluster diffusion. For typical values of parameters with nonzero δ, the effective barrier is given by $E_{\rm eff} = E_e + \phi + \delta$.

Previous analysis [32] also exploited the possibility of simplified (dimensionally reduced) analysis in the limit as $h_e \to \infty$ where various configurations convert infinitely quickly between each other and may be grouped into a smaller set of quasiconfigurations. For the trimer, there are two quasiconfigurations (two linear and a single quasibent configuration), and the above result reduces to $D_2 = (a^2/3)h_r$. For tetramers, there are five quasiconfigurations, and the above result reduces to

$$D_4 = 6a^2 h_c h_r/(18h_c + h_r), \quad \text{so that}$$
$$1/D_4 = (a^{-2}/6)(1/h_c + 18/h_r). \tag{A4}$$

Results are also available for pentamers.

APPENDIX B: ESTIMATING THE NUMBER OF KINKS n_k IN CLUSTER CONFIGURATIONS

Here, we obtain bounds on the number of kinks n_k for various cluster configurations. The lower bound can readily be determined for specific cases, and is O(1). Thus we focus on estimating the upper bound in this section. First, consider removing m_1 atoms from a single corner of an otherwise perfect rectangular cluster. The number of kinks n_k is maximized if the atoms are removed to create a vacancy region as close as possible to a triangle with a 45°

diagonal (corresponding to a perfect staircase of kinks each of height a). This can be achieved exactly if $m_1 = 1 + 2 + \cdots + (n_k - 1) = \frac{1}{2} n_k (n_k - 1)$, so that $n_k = (1 + \sqrt{1 + 8m_1})/2$.

Next, consider removing m_i atoms from the ith corner of a perfect rectangular cluster where $m_1 + m_2 + m_3 + m_4 = m$, and where m is less than either side length of the rectangle. Then, since the above expression for n_k with atoms removed from a single corner increases sublinearly with m_1, it follows that the total number of kinks can be maximized by removing roughly equal numbers of kinks from all corners, i.e., $m_1 \approx m_2 \approx m_3 \approx m_4 \approx m/4$. Consequently, for an upper bound on the total number of kinks n_k, we replace m_1 by $m/4$ in the above expression and multiply by 4 to obtain $n_k \leqslant 2(1 + \sqrt{1 + 2m})$. Considering the quantities relevant for the analysis of Sec. IV, we have that $m = L$ for $n_{L^2,k}(1)$, $m = L - 1$ for $n_{L^2+1,k}(0)$, and $m = L - 3$ for $n_{L^2+3,k}(0)$.

APPENDIX C: COUNTING OF ISOENERGETIC CLUSTER CONFIGURATIONS

In our representation of clusters as collections of atoms, themselves represented as contiguous red squares, the energy of the cluster corresponds to its perimeter length. Consider the cluster shapes that are obtained by starting with a fully populated rectangle and then removing atoms from each corner of the cluster to form a simple "staircase" (i.e., steps at each corner are of one sign, not both). Then, the energy of these configurations is determined exactly by the perimeter length of the smallest rectangle inscribing these clusters (which corresponds to the original rectangle from which atoms were removed). This follows since the perimeter length of the inscribing rectangle and the actual cluster are equal. These observations will be useful in the following analysis.

First, we consider ground-state configurations, which have the minimum perimeter length for the prescribed number, N, of atoms. For ground states, the inscribing rectangle is either a $L_i \times L_i$ square of occupied sites, or a near-square $L_i \times (L_i + 1)$ or $L_i \times (L_i + 2)$ rectangle. The unique ground state for $N = L^2$ is inscribed by a square with $L_i = L$. The ground states for $N = L^2 + m$ with $1 \leqslant m \leqslant L$ are inscribed by a $L_i \times (L_i + 1)$ rectangle with $L_i = L$. The ground states for $N = L(L + 1) + m$ with $1 \leqslant m \leqslant L$ are inscribed by $L_i \times L_i$ squares with $L_i = L + 1$ or by $L_i \times (L_i + 2)$ rectangles with $L_i = L$. Next, we consider nth excited state configurations where the perimeter length of the cluster is increased relative to the ground state by an amount $2n$ (in units of lattice constant $a = 1$). Thus the size of the inscribing rectangle must also be increased. Specifically, the side lengths are increased by amounts n_x and n_y, where $n_x + n_y = n$ to achieve the desired perimeter length.

Thus, to evaluate the number of convex isoenergetic nth excited state configurations of a size N cluster $\Omega_N(n)$, first, one determines the different possible inscribing rectangles for the ground states. Second, one expands the side lengths of these rectangles by amounts n_x and n_y, where $n_x + n_y = n$. Third, regarding all sites in this larger inscribing rectangle as initially populated, one considers all possible ways to remove the appropriate number of atoms from the four corners of the rectangle (making sure the cluster is touching all four edges of the rectangular frame), until the final number of atoms matches the cluster size N, which we are targeting. It is instructive to provide a few examples: (i) determination of $\Omega_{L^2+3}(0)$ requires counting different possible ways to remove L-3 atoms from an $L \times (L + 1)$ inscribing rectangle; (ii) determination of $\Omega_{L^2}(1)$ requires counting different possible ways to remove L atoms from an $L \times (L + 1)$ inscribing rectangle; and (iii) determination of $\Omega_{L^2+3}(1)$ requires counting different possible ways to remove L-3 atoms from $L \times (L + 2)$ and $(L + 1) \times (L + 1)$ inscribing rectangles.

Now, we describe in detail a systematic procedure to count the number of ways of removing the appropriate number of atoms from the inscribing rectangle. We start by considering removal of m_1 atoms from one fully populated corner. The number of possibilities is identical to the number of Young or Ferrers diagrams that represents integer partition of m_1. In number theory, this integer partition is traditionally denoted by $P(m_1)$ [51]. An example for $m_1 = 4$ where $P(m_1 = 4) = 5$ is shown in Fig. 16.

Next, we address the more complex challenge of counting the total number of configurations of the cluster, where one removes m_1, m_2, m_3, and m_4 atoms from each of the four corners of the inscribing rectangle, respectively, for a total of m atoms where $m = m_1 + m_2 + m_3 + m_4$. One constraint with this analysis is that removal of atoms from one corner does not interfere with removal from other corners, which requires that m is no larger than the side lengths of the inscribing rectangle. (We will comment further below on cases where this condition is not satisfied.) Subject to this constraint, the total number of configurations comes from considering the product of the corresponding integer partions, and then summing over all possible choices of m_i consistent with the constraint on the sum (and finally adjusting for any overcounting).

An example for $\Omega_{L^2}(1)$ is shown below where $m = L$ atoms are removed from an inscribing $L \times (L + 1)$ rectangle.

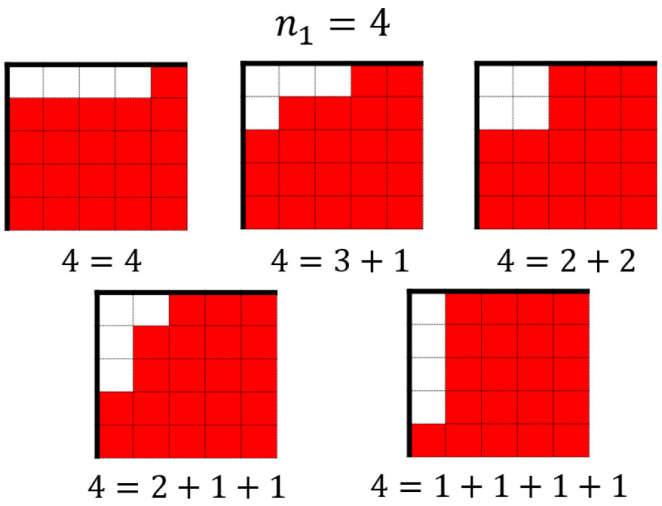

FIG. 16. Number $P(m_1 = 4) = 5$ of possible ways to remove $m_1 = 4$ atoms from a corner illustrated by Ferrers diagrams. Partitions of 4 into strings of integers indicate the number of atoms removed from each row starting with the top row.

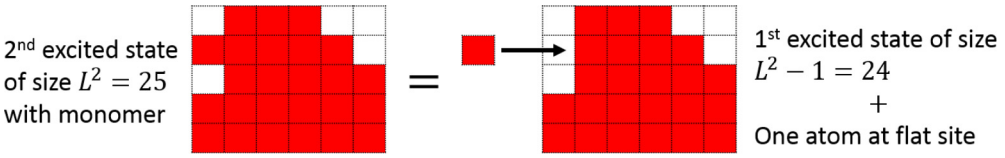

FIG. 17. Example of an excited state with one monomer.

Here, one has

$$\Omega_{L^2}(1) = 2 \times \sum_{m_1+m_2+m_3+m_4=L} P(m_1)P(m_2)P(m_3)P(m_4) - (\text{over counting})$$
$$= 2 \times \sum_{m_L=0}^{L} \left[\sum_{m_l=0}^{m_L} P(m_l)P(m_L - m_l) \sum_{m_r=0}^{L-m_L} P(m_r)P(L - m_L - m_r) \right] - (\text{over counting}). \quad \text{(C1)}$$

In the second sum, $m_L(m_R)$ gives the total number of atoms removed from the left (right) side on the inscribing rectangle, and $m_l(m_r)$ give the number of atoms removed from one corner on the left (right) side. The factor of 2 comes from a 90° rotation of the $L \times (L+1)$ rectangle, correponding to another set of discrete states. Note that "overcounting" in (8) includes the ground state being counted four times ($m_L = L$ and $m_l = 0$ or L) or ($m_L = 0$, $m_r = 0$ or L). If one wishes to consider just first excited states without any monomers, then it is also necessary to subtract $4 \times (4L-2)$ states where an atom is shifted from a corner of the $L \times L$ ground-state configuration and placed on a side. One must also subtract $4L$ configurations with a monomer on the edge of a completely populated $(L-1) \times (L+1)$ rectangle.

In addition, we have analyzed $\Omega_{L^2+1}(0)$ and $\Omega_{L^2+3}(0)$, where L-1 and L-3 atoms are removed from an $L \times (L+1)$ inscribing rectangle, respectively. In these cases, the procedure described above is directly applicable. Finally, we have also analyzed $\Omega_{L^2+3}(1)$, where $2L$-3 atoms are removed from $L \times (L+2)$ or $(L+1) \times (L+1)$ inscribing rectangles. In this case, since the number of removed atoms significantly exceeds side lengths of the inscribing rectangle, significant modification is required from the formulation (8) used to obtain $\Omega_{L^2}(1)$ and other quantities mentioned above.

Results reported in the text for $\Omega_{L^2}(1)$, $\Omega_{L^2+1}(0)$, and $\Omega_{L^2+3}(0)$ include all states, i.e., those with monomers and those without. (See Ref. [49] for corresponding results excluding states with monomers.)

APPENDIX D: COUNTING OF EXCITED STATE CONFIGURATIONS WITH ONE MONOMER

In Sec. V, we estimated number of configurations, $\Omega'_N(n)$, of clusters with N atoms corresponding to nth excited state, which include a single monomer. In some cases, this analysis was simple, e.g., $\Omega'_{L^2+1}(0) = 4L$. However, analysis of other cases including $\Omega'_{L^2}(2)$ and $\Omega'_{L^2+1}(1)$ is nontrivial, and is thus described in more detailed below.

To estimate $\Omega'_N(n)$, we first remove the monomer, and then count the number of states $\Omega_{N-1}(n-1 \text{ or } n)$, where the appropriate choice is discussed below. For the latter, we utilize the scheme introduced in Appendix C. Next, let n_f denote the number of empty edge sites n_f with only one neighbor, which could thus accommodate a monomer. Then, it follows that

$$\Omega'_N(n) = n_f \times \Omega_{N-1}(n-1 \text{ or } n). \quad \text{(D1)}$$

To determine n_f, we note that each kink roughly contributes two units of perimeter; it follows that the total perimeter length for clusters of size $N-1$ in the $(n-1)$th excited state is given by the sum $n_f + 2n_{N-1,k}(n-1)$, where $n_{N-1,k}(n-1)$ denotes the number of kinks in these clusters (see Appendix B).

To determine $\Omega'_{L^2}(2)$, we note that first excited states for clusters of size $N = L^2$ have configurations within a $L \times (L+1)$ inscribing rectangle. For second excited states with a single monomer, this monomer is located at the perimeter of a cluster of size L^2–1 with no monomers, but still with an $L \times (L+1)$ inscribing rectangle and which thus corresponds to a first excited state (see Fig. 17 for an example). Thus, one has that

$$\Omega'_{(L^2)}(2) \approx \Omega_{(L^2-1)}(1)[4L + 2 - 2n_{L^2-1,k}(1)]. \quad \text{(D2)}$$

To determine $\Omega'_{L^2+1}(1)$, we note that ground states for clusters with size $N = L^2 + 1$ have configurations within a $L \times (L+1)$ inscribing rectangle. For first excited states with a single monomer, this monomer is located at the perimeter of a cluster of size L^2 with no monomers, but still with an $L \times (L+1)$ inscribing rectangle. The latter thus also corresponds to a first excited state. In conclusion, one has that

$$\Omega'_{(L^2+1)}(1) \approx \Omega_{(L^2)}(1)[4L + 2 - 2n_{L^2,k}(1)]. \quad \text{(D3)}$$

[1] J.-M. Wen, S.-L. Chang, J. W. Burnett, J. W. Evans, and P. A. Thiel, Phys. Rev. Lett. **73**, 2591 (1994).
[2] W. W. Pai, A. K. Swan, Z. Zhang, and J. F. Wendelken, Phys. Rev. Lett. **79**, 3210 (1997).
[3] X. Ge and K. Morgenstern, Phys. Rev. B **85**, 045417 (2012).
[4] D. S. Sholl and R. T. Skodje, Phys. Rev. Lett. **75**, 3158 (1995).
[5] C. DeW. Van Siclen, Phys. Rev. Lett. **75**, 1574 (1995).
[6] J. M. Soler, Phys. Rev. B **53**, R10540(R) (1996).

[7] K. A. Fichthorn and S. Pal, Bull. Am. Phys. Soc. **43**, 850 (1998).
[8] A. Bogicevic, S. Liu, J. Jacobsen, B. Lundqvist, and H. Metiu, Phys. Rev. B **57**, R9459 (1998).
[9] J. Heinonen, I. Koponen, J. Merikoski, and T. Ala-Nissila, Phys. Rev. Lett. **82**, 2733 (1999).
[10] G. Mills, T. R. Mattsson, L. Mollnitz, and H. Metiu, J. Chem. Phys. **111**, 8639 (1999).
[11] A. Lo and R. T. Skodje, J. Chem. Phys. **111**, 2726 (1999).
[12] H. Shao, S. Liu, and H. Metiu, Phys. Rev. B **51**, 7827 (1995).
[13] S. Liu and H. Metiu, Surface Sci. **405**, L497 (1998).
[14] C. R. Stoldt, A. M. Cadilhe, C. J. Jenks, J.-M. Wen, J. W. Evans, and P. A. Thiel, Phys. Rev. Lett. **81**, 2950 (1998).
[15] P. Jensen, N. Combe, H. Larralde, J. L. Barrat, C. Misbah, and A. Pimpinelli, Eur. Phys. J. B **11**, 497 (1999).
[16] N. Combe and H. Larralde, Phys. Rev. B **62**, 16074 (2000).
[17] D. J. Liu and J. W. Evans, Phys. Rev. B **66**, 165407 (2002).
[18] J. L. Iguain and L. J. Lewis, Phys. Rev. B **68**, 195407 (2003).
[19] Y. Han, C. R. Stoldt, P. A. Thiel, and J. W. Evans, J. Phys. Chem. C **120**, 21617 (2016).
[20] A. F. Voter, Phys. Rev. B **34**, 6819 (1986).
[21] A. F. Voter, Proc. SPIE **821**, 214 (1987).
[22] H. C. Kang, P. A. Thiel and J. W. Evans, J. Chem. Phys. **93**, 9018 (1990).
[23] S. V. Khare, N. C. Bartelt, and T. L. Einstein, Phys. Rev. Lett. **75**, 2148 (1995).
[24] S. V. Khare and T. L. Einstein, Phys. Rev. B **57**, 4782 (1998).
[25] K. Binder and M. Kalos, J. Stat. Phys. **22**, 363 (1980).
[26] M. Rao, M. H. Kalos, J. L. Lebowitz, and J. Marro, Phys. Rev. B **13**, 4328 (1976).
[27] G. L. Kellogg, Phys. Rev. Lett. **73**, 1833 (1994).
[28] G. L. Kellogg, Surf. Sci. Rep. **21**, 1 (1994).
[29] G. Antczak and G. Ehrlich, *Surface Diffusion: Metals, Metal Atoms, and Clusters* (Cambridge University Press, Cambridge, 2010).
[30] U. M. Titulaer and J. M. Deutsch, J. Chem. Phys. **77**, 472 (1982).
[31] Z.-P. Shi, Z. Zhang, A. K. Swan, and J. F. Wendelken, Phys. Rev. Lett. **76**, 4927 (1996).
[32] J. R. Sanchez and J. W. Evans, Phys. Rev. B **59**, 3224 (1999).
[33] O. S. Trushin, P. Salo, M. Alatalo, and T. Ala-Nissila, Surf. Sci. **482–485**, 365 (2001).
[34] K. Morgenstern, G. Rosenfeld, B. Poelsema, and G. Comsa, Phys. Rev. Lett. **74**, 2058 (1995).
[35] D. C. Schlößer, K. Morgenstern, L. K. Verheij, G. Rosenfeld, F. Besenbacher, and G. Comsa, Surf. Sci. **465**, 19 (2000).
[36] K. Morgenstern, E. Lægsgaard, and F. Besenbacher, Phys. Rev. Lett. **86**, 5739 (2001).
[37] C. Zaum, M. Rieger, K. Reuter, and K. Morgenstern, Phys. Rev. Lett. **107**, 046101 (2011).
[38] J. C. Hamilton, Phys. Rev. Lett. **77**, 885 (1996).
[39] O. U. Uche, D. Perez, A. F. Voter, and J. C. Hamilton, Phys. Rev. Lett. **103**, 046101 (2009).
[40] A. Karim, A. Kara, O. Trushin, and T. S. Rahman, J. Phys.: Condens. Matter **23**, 462201 (2011).
[41] S. R. Acharya, S. I. Shah, and T. S. Rahman, Surf. Sci. **662**, 42 (2017).
[42] J. Krug, H. T. Dobbs, and S. Majaniemi, Z. Phys. B: Condens. Matter **97**, 281 (1995).
[43] D.-J. Liu, C. R. Stoldt, P. A. Thiel, and J. W. Evans, Mater. Res. Soc. Symp. Proc. **749**, W2.8.1 (2003).
[44] H.-C. Jeong and E. D. Williams, Surf. Sci. Rep. **34**, 171 (1999).
[45] O. Pierre-Louis, Phys. Rev. Lett. **87**, 106104 (2001).
[46] M. Eßer, K. Morgenstern, G. Rosenfeld, and G. Comsa, Surf. Sci. **402–404**, 341 (1998).
[47] K. C. Lai, J. W. Evans, and D.-J. Liu, J. Chem. Phys. **147**, 201101 (2017).
[48] For $f = 0.24$ eV at 300 K, we also find that $\beta_s = 0.37$ for $45 \leqslant N \leqslant 103$ when $\delta = 0.1$ (versus $\beta_s = 0.55$ for $67 \leqslant N \leqslant 200$ for $\delta = 0$).
[49] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.96.235406 for further analysis of time-dependent diffusivity, of Arrhenius behavior of D_N, and for tabulated D_N values.
[50] E. W. Montroll and G. H. Weiss, J. Math. Phys. **6**, 167 (1965).
[51] G. E. Andrews and K. Eriksson, *Integer Partitions* (Cambridge University Press, Cambridge, 2004).