Capture zone area distributions for nucleation and growth of islands during submonolayer deposition

Yong Han
Iowa State University, y27h@iastate.edu

Maozhi Li
Renmin University of China

James W. Evans
Iowa State University, evans@ameslab.gov

Follow this and additional works at: http://lib.dr.iastate.edu/physastro_pubs

Part of the Biological and Chemical Physics Commons

The complete bibliographic information for this item can be found at http://lib.dr.iastate.edu/physastro_pubs/453. For information on how to cite this item, please visit http://lib.dr.iastate.edu/howtocite.html.

This Article is brought to you for free and open access by the Physics and Astronomy at Iowa State University Digital Repository. It has been accepted for inclusion in Physics and Astronomy Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Capture zone area distributions for nucleation and growth of islands during submonolayer deposition

Abstract
A fundamental evolution equation is developed to describe the distribution of areas of capture zones (CZs) associated with islands formed by homogeneous nucleation and growth during submonolayer deposition on perfect flat surfaces. This equation involves various quantities which characterize subtle spatial aspects of the nucleation process. These quantities in turn depend on the complex stochastic geometry of the CZ tessellation of the surface, and their detailed form determines the CZ area distribution (CZD) including its asymptotic features. For small CZ areas, behavior of the CZD reflects the critical island size, i. For large CZ areas, it may reflect the probability for nucleation near such large CZs. Predictions are compared with kinetic Monte Carlo simulation data for models with two-dimensional compact islands with $i = 1$ (irreversible island formation by diffusing adatom pairs) and $i = 0$ (adatoms spontaneously convert to stable nuclei, e.g., by exchange with the substrate).

Keywords
Magnetic islands, Data analysis, Stochastic processes, Aggregation, Homogeneous nucleation

Disciplines
Biological and Chemical Physics | Physics

Comments
This article is published as Han, Yong, Maozhi Li, and James W. Evans. "Capture zone area distributions for nucleation and growth of islands during submonolayer deposition." The Journal of Chemical Physics 145, no. 21 (2016): 211911, doi:10.1063/1.4961264. Posted with permission.

This article is available at Iowa State University Digital Repository: http://lib.dr.iastate.edu/physastro_pubs/453
Capture zone area distributions for nucleation and growth of islands during submonolayer deposition

Yong Han,1 Maozhi Li,2 and James W. Evans3
1Department of Physics and Astronomy, Iowa State University, Ames, Iowa 50011, USA
2Department of Physics, Renmin University of China, Beijing 100872, People’s Republic of China
3Department of Physics and Astronomy, Department of Mathematics, and Ames Laboratory–USDOE, Iowa State University, Ames, Iowa 50011, USA

(Received 25 May 2016; accepted 6 August 2016; published online 6 September 2016)

A fundamental evolution equation is developed to describe the distribution of areas of capture zones (CZs) associated with islands formed by homogeneous nucleation and growth during submonolayer deposition on perfect flat surfaces. This equation involves various quantities which characterize subtle spatial aspects of the nucleation process. These quantities in turn depend on the complex stochastic geometry of the CZ tessellation of the surface, and their detailed form determines the CZ area distribution (CZD) including its asymptotic features. For small CZ areas, behavior of the CZD reflects the critical island size, i. For large CZ areas, it may reflect the probability for nucleation near such large CZs. Predictions are compared with kinetic Monte Carlo simulation data for models with two-dimensional compact islands with i = 1 (irreversible island formation by diffusing adatom pairs) and i = 0 (adatoms spontaneously convert to stable nuclei, e.g., by exchange with the substrate). Published by AIP Publishing. [http://dx.doi.org/10.1063/1.4961264]

I. INTRODUCTION

Homogeneous nucleation and growth of islands during submonolayer deposition on perfect low-index extended surfaces have been explored for decades as a route for self-assembly of arrays of supported nanoclusters or “islands.”1–3 This process involves a competition between simultaneous nucleation and growth of stable islands mediated by terrace diffusion of deposited atoms. Nucleation is often characterized by specifying a critical size i (measured in atoms) above which islands are stable and effectively immobile on the timescale of island formation. Thus, i = 1 corresponds to irreversible island formation where dimers and larger clusters are stable. For i > 1, island formation is reversible as sub-stable clusters of i or less atoms can dissociate. For i = 0, one regards isolated diffusing adatoms as being able to spontaneously convert to stable nuclei (e.g., by exchange with the substrate) about which stable islands then grow.4

Theoretical and experimental analyses have focused on behavior of the mean island density, N_is, and of the island size distribution (ISD).1–3 However, there has also been a long-standing appreciation of the significance of capture zones (CZs) associated with individual islands.5,6 The key idea underlying the CZ concept is that the surface can be tessellated into a distribution of CZs, one per island, where the area of the CZ for each island determines its growth rate.3,5–11 Recently, attention has turned to analysis of the CZ area distribution (CZD).12–21 A heuristic formulation by Pimpinelli and Einstein suggested that the CZD has a Gaussian tail for large areas, and that the small area behavior is related to the critical size, i, characterizing nucleation.14

Our strategy for analysis of these surface deposition systems is to develop exact beyond-mean-field evolution equations for the basic quantities of interest: the ISD, the CZD, or related quantities such as the joint probability distribution (JPD) for island sizes and CZ areas.9,12,13,16 While exact analysis is not possible, these equations generally provide a solid foundation for reasonable approximate treatments. Also from this approach, one immediately recognizes the importance of characterizing spatial aspects of the nucleation process and of the resulting island distribution. This includes description of correlations between island size and island separation (or, equivalently, the dependence of CZ area on island size). Critically, here for analysis of the CZD, it is essential to provide a detailed description of various subtle spatial aspects of nucleation. Indeed, the central importance of the latter has been noted previously both in analytic theories based on the JPD,12,13 and in the development of an efficient “geometry based simulation” (GBS) procedure to describe island formation.22–24 Of particular significance is the feature that nucleation is enhanced by a higher adatom density. It therefore occurs preferentially far away from stable islands, the edges of which are sinks for diffusing adatoms, and thus nearby CZ boundaries.3,22–24

In Sec. II, we describe our model for island formation, develop the fundamental equation for evolution of the CZD, and indicate expected scaling behavior of the solution. In Sec. III, we describe in detail key quantities characterizing the spatial aspects of nucleation. The scaling form of the fundamental evolution equation for CZD is described in Sec. IV, as well as the anticipated asymptotic form of its solutions for small and large CZ area. Scenarios for creating CZs with small areas are described in detail in Sec. V. Simulation results for i = 1 and i = 0 are presented and analyzed in Sec. VI. Further discussion and conclusions are presented in Sec. VII. A brief sketch of the basic formulation
presented here and limited results were provided in a previous Comment.16

II. EVOLUTION OF CZ AREA DISTRIBUTION (CZD)

A. Model prescription and basic behavior

Our atomistic model for island formation during submonolayer deposition includes the following basic ingredients: (i) Random deposition at rate \( F \) per site on a square lattice of adsorption sites. (ii) Hopping of isolated adatoms to empty nearest-neighbor (NN) sites with rate \( h \) per direction. (iii) Aggregation of adatoms into two-dimensional (2D) islands, which are stable above a critical size, \( i \) (measured in atoms). Nucleation requires aggregation of \( i + 1 \) atoms for \( i \geq 1 \). Nucleation is mediated by exchange of single adatoms with the substrate for \( i = 0 \). Sub-stable islands can be regarded as in quasi-equilibrium with isolated adatoms. (iv) Periphery diffusion is facile so that islands relax to achieve compact square shapes during growth. (v) Atoms deposited on top of islands are instantaneously transported to the island edge and hop down. See Appendix A for further details.

We will sometimes comment on behavior for analogous “point-island” models where all islands occupy just a single lattice site but carry a label to indicate their size.23 These models avoid some complications seen in the above more realistic models for which the “footprint” of the 2D compact islands expands during growth. Basic features of the ISD, CZD, and JPD are similar for point and compact islands at lower coverage, although we find that some quantities characterizing the details of nucleation differ significantly.

We define the coverage, \( \theta \), from the deposition time, \( t \), via \( \theta = Ft \) in monolayers (ML). Then, the theory below is applicable to the regime of lower coverage, \( \theta \leq 0.1 \) ML, where the typical linear dimension of islands is well below the mean island separation. Let \( N_i \) denote the density of islands of \( s \) adatoms, where densities are measured per adsorption site. Thus, \( N_1, N_i, \) and \( N_{isl} = \sum_{s > i} N_s \) denote the densities of adatoms, critical clusters, and stable islands with size \( s > i \), respectively. The Walton relation for quasi-equilibrated critical clusters implies that \( N_i \sim (N_1)^{-3/2,26} \). For \( i = 0 \), one naturally sets \( N_0 \approx 1 \). For \( i \geq 1 \), in terms of the nucleation rate, \( K_{agg} \propto hN_1N_0 \), and the rate of aggregation, \( K_{agg} \propto hN_1N_{isl} \), of adatoms with stable islands, one has that1–3

\[
\frac{dN_i}{dt} = F - K_{agg} - (i + 1)K_{nuc} \quad \text{and} \quad \frac{dN_{isl}}{dt} = K_{nuc}.
\]  

(1)

Integration of (1) for \( i \geq 1 \) reveals a “transient regime” of rapidly growing \( N_i \) for \( \theta \) below a crossover value \( \theta^* \sim (h/F)^{-2/(i+3)} \) and a subsequent steady-state regime where adatom gain due to deposition roughly balances the loss due to aggregation so that \( F \approx K_{agg} \). The island density scales like \( N_{isl}^* \sim (h/F)^{-x} \) with \( x = (i + 1)/(i + 3) \) at crossover and like \( N_{isl} \sim (h/F)^{-x} \) with \( x = 1/(i + 2) \) in the steady-state regime.1–3

Since \( x \sim 1 \leq x = 2(i + 2)^{-1} (i + 3)^{-1} < 0 \), nucleation in the steady-state regime dominates that in the transient regime for large enough \( h/F \) when \( i \geq 1 \). Indeed, this dominance underlies selection of scaling forms for the ISD and CZD.

For \( i = 0 \), one must also specify an exchange probability, \( \sigma \), per hop, so that the nucleation rate becomes \( K_{nuc} \propto \sigma hN_1 \).4

In this case, there are also transient and steady-state regimes, although the characterization of behavior differs.27 Our version of this model is tailored to represent a “pure” \( i = 0 \) mechanism with no nucleation by aggregation of multiple diffusing adatoms. See Appendix A.

B. CZ area distribution (CZD)

Let \( N_A \) denote the density of stable islands with CZ area \( A \), where areas will be measured in units of the surface unit cell area (and distances will be measured in units of surface lattice constant). Also let \( A_{av} \) denote the mean CZ area. Then, one has that

\[
\sum_{A > 1} N_A = N_{isl} = \sum_A AN_A = 1 \quad \text{and} \quad A_{av} = \sum_A ANA/\sum_A N_A = 1/N_{isl}.
\]  

(2)

The mean island separation can be defined as \( r_{isl} = (A_{av})^{1/2} = (N_{isl})^{-1/2} \). To assess the evolution of \( N_A \), we introduce the quantity \( P_A \) denoting the probability that the CZ of a just-nucleated island overlaps that of a pre-existing island with CZ area \( A \) (and thus destroys a pre-existing CZ of area \( A \)). Similarly, we introduce \( P^*_A \) denoting the probability that formation of the CZ of a just-nucleated island reduces to \( A \) the area of the CZ of a pre-existing island which was larger than \( A \) (and thus creates a new CZ of area \( A \)). We also let \( P_A^* \) denote the probability that the CZ of a just-nucleated island has area \( A \). Then, one obtains16 the exact fundamental evolution equation

\[
dN_A/dt = (-P_A + P^*_A + P^*_A) dN_{isl}/dt.
\]  

(3)

Fig. 1 illustrates schematically the terms in this equation. We shall see in Sec. III A that \( P_A \) and \( P_A^* \) are related, and one has that \( \sum_A P_A = \sum_A P^*_A = M \), where \( M \) denotes the mean number Czs of pre-existing islands overlapped by the CZ of a just-nucleated island. This equality, together with the normalization condition, \( \sum_A P^*_A = 1 \), on \( P_A \), ensures consistency of (3) after applying \( \sum_A \) to both sides, and noting the identity \( N_{isl} = \sum_A N_A \). We remark that there are some subtleties in the definition of \( P_A \), etc., for smaller \( A_{av} \) where \( A \) are strongly discrete, and thus there is a chance that the CZ of a just-nucleated island can overlap two or more existing Czs of the same size. These issues are briefly discussed in Appendix B.

Note that the mean number of Czs overlapped by that of a just-nucleated island is \( M \approx 5.5 \) for point-island models with \( i = 1 \),13 and we will find a similar value for compact islands. We emphasize that this result demonstrates the failure

\[
\frac{d}{dt} N_A = -\left[ (P_A \propto N_AQ_A) + (P^*_A \propto N_{isl}Q_A) \right] + (P^*_A) dN_{isl}/dt.
\]

FIG. 1. Schematic of gain and loss terms in the fundamental evolution equation for \( N_A \). Nucleation locations are denoted by \( \times \).
of a simple fragmentation picture where each nucleation event is regarded as simply fragmenting an existing CZ into two parts (corresponding to \( M = 1 \)).\textsuperscript{28} The reason for the high \( M \)-values is that, as noted above, nucleation tends to occur near existing CZ boundaries where overlap with multiple CZs is expected.\textsuperscript{22–24}

Our focus will be on the asymptotic scaling regime of large \( A_{\text{sub}} \) or small \( N_{\text{isl}} \) (corresponding to \( h/F \gg 1 \) for \( i \geq 1 \)) where \( N_{\Lambda} \) is proposed to have the form\textsuperscript{12–21}

\[
N_{\Lambda} \approx (N_{\text{isl}}/A_{\text{sub}})g(A/A_{\text{sub}}) \text{ with } \int d\alpha g(\alpha) = \int d\alpha a\gamma(\alpha) = 1.
\]

(4)

Unless stated otherwise, \( \int d\alpha \) ranges over \( 0 \leq \alpha \leq \infty \). Further comment on two aspects of this scaling are appropriate. First, we should note that deviations from the limiting scaling form can be significant for \( h/F \sim 10^0-10^7 \) considered here.\textsuperscript{3,29,30} The true limiting or asymptotic shape, \( g \), of the distribution may only be achieved for much higher \( h/F \). Second, the formulation (4) leaves implicit any dependence of \( g \) on coverage, \( \theta \), given that in this study we just focus on a single \( \theta = 0.1 \text{ ML} \). However, there certainly is a dependence on \( \theta \), although we do not expect this dependence to be strong for models incorporating compact islands in the regime for small \( \theta \) (specifically in the regime \( \theta \leq 0.1 \) which we consider). See Ref. \textsuperscript{30} for a detailed discussion and analysis of \( \theta \)-dependence in the context of the ISD. Regarding the CZD, a weak \( \theta \)-dependence of \( g \) will require that key quantities characterizing nucleation also have weak dependence on \( \theta \). In Secs. III and V, we will offer some comments related to the origin of the \( \theta \)-dependence for compact islands. From these observations, it is possible that basic behavior of the CZD for small CZ areas could depend on \( \theta \), but this is not expected of behavior for large areas. As an aside, we remark that \( \theta \)-dependence should strictly only be absent in asymptotic scaling functions for point-island models.\textsuperscript{13,25,29,30}

Our main goal is to provide insight into the form of the scaling function \( g(\alpha) \) based on simulation studies, and also on analysis of (3). The complexity of (3) undoubtedly precludes simple forms for this solution, providing a significant caveat for most existing analyses of simulation data for \( g(\alpha) \). However, just as for the ISD, there is a natural appeal in developing simple fitting-forms which can be conveniently used for experimental data analysis. In this context, generalizing a strategy used in the field of stochastic geometry for analysis of tessellations,\textsuperscript{31–33} it is natural to consider fits to the generalized gamma (GG) distribution of the form\textsuperscript{16}

\[
g_{\text{GG}}(\alpha) \approx \exp(\beta \alpha^\gamma),
\]

(5a)

where

\[
b = \left( \Gamma\left(\frac{\beta}{2} + 1\right) / \Gamma\left(\beta + 1\right) \right)^\gamma
\]

and

\[
a = \frac{\exp(b(1+\gamma)/\gamma)}{\Gamma\left(\beta + 1\right)/\gamma}.
\]

(5b)

The value of “b” is determined by the unit mean constraint, and that of “a” from normalization. The special case where \( \gamma = 2 \) corresponds to the so-called Generalized Wigner Surmise (GWS) choice in Ref. 14. Another significant component of Ref. 14 was to suggest that \( \beta \) is determined by the critical size \( i \) for nucleation, and specifically that \( \beta_{\text{GWS}} = i + 1 \)\textsuperscript{14} (later changed to \( \beta_{\text{GWS}} = i + 2 \)) for island formation in a 2D surface deposition system. It should also be noted that traditional applications to analysis of tessellations in stochastic geometry usually set \( \gamma = 1.31–33 \).

III. DETAILED CHARACTERIZATION OF THE NUCLEATION PROCESS

A. Characterization of key quantities describing nucleation

Application of (3) to determine the CZ area distribution, \( N_{\Lambda} \), or the associated scaling function, \( g(\alpha) \), requires characterization of the quantities \( P^*_A \), \( P_A \), and \( P_{\Lambda} \) describing nucleation. For analysis of \( P^*_A \) and \( P_A \) and the relationship between them, we introduce a more detailed quantity, \( P_A(A_{\text{sub}}) \), giving the probability that the CZ of a just-nucleated island overlaps that of a pre-existing island with CZ area \( A \) by an amount \( A_{\text{sub}} \), where \( A_{\text{sub}} < A \).\textsuperscript{34} See Fig. 1. Then, it follows that

\[
P_A = \sum_{A_{\text{sub}}} P_A(A_{\text{sub}}) \text{ and } P_{\Lambda} = \sum_{A_{\text{sub}}} P_A(A_{\text{sub}}). \tag{6}
\]

Summing both quantities over \( A \) immediately yields the identity \( \sum_A P^*_A = \sum_A P_A (= M) \) mentioned in Sec. II B. See Appendix B.

Next, we note that since \( P_A \) should be proportional to \( N_{\Lambda} \), it is natural to write

\[
P_A = M(N_{\Lambda}/N_{\text{isl}})Q_A, \text{ where } \sum_A(N_{\Lambda}/N_{\text{isl}})Q_A = 1.
\]

It is also convenient to write \( P_A(A_{\text{sub}}) \) in the factorized form

\[
P_A(A_{\text{sub}}) = P_A E_A(A_{\text{sub}}) \quad \text{with } \sum_A E_A(A_{\text{sub}}) = 1. \tag{7}
\]

We regard \( Q_A \) as an “intrinsic” overlap probability, and \( E_A(A_{\text{sub}}) \) as characterizing the distribution of overlap areas for CZs of area \( A \). In addition, it will be appropriate to introduce the quantity \( A_{\text{sub} \text{nuc}}(A) \) giving the mean area of the portion of the CZ of a just-nucleated island which overlapped that of a pre-existing island with CZ area \( A \) (or equivalently, the mean area of the portion of the existing CZ of area \( A \) which is overlapped by the new CZ). It then follows that

\[
A_{\text{sub} \text{nuc}}(A) = \sum_{A_{\text{sub}}} A_{\text{sub}} E_A(A_{\text{sub}}). \tag{8}
\]

Averaging \( A_{\text{sub} \text{nuc}}(A) \) over all CZ areas, \( A \), yields the quantity

\[
A_{\text{sub} \text{av}} = \sum_A A_{\text{sub} \text{nuc}}(A) N_{\Lambda}/N_{\text{isl}} \equiv \mu_{\text{av}} A_{\text{sub}}. \tag{9}
\]

defining \( \mu_{\text{av}} \) as the average fractional overlap with each existing CZ. In fact, previous analyses for point islands with \( i = 1 \) found that \( \mu_{\text{av}} \approx 0.17 \) and also indicated that \( A_{\text{sub} \text{nuc}}(A) \approx \mu_{\text{av}} A \) for all \( A \).\textsuperscript{12,13} We will find a significantly lower value of \( \mu_{\text{av}} \) for square islands at \( \approx 0.1 \text{ ML} \).

Finally, it is also appropriate to introduce a quantity \( A_{\text{av} \text{sub}} = \sum_A P_{\Lambda} A \) corresponding to the average area for the CZs of just-nucleated islands. Note that \( A_{\text{av} \text{sub}} = A_{\text{av} \text{nc}}/A_{\text{av}} \approx 0.97 \) for point islands with \( i = 1,13 \) but we will find a significantly smaller value for compact islands at \( \approx 0.1 \text{ ML} \). It is tempting to equate \( A_{\text{av} \text{nc}} \) with the product \( M A_{\text{av} \text{sub}} = M \mu_{\text{av}} A_{\text{av}} \),\textsuperscript{12,13} i.e., to identify \( A_{\text{av} \text{nc}} \) with \( M \mu_{\text{av}} \).
\[ P^* \approx (A_{av})^{-1} p(A/A_{av}), \quad \text{where } \int d\alpha p^*(\alpha) = 1, \quad (10) \]

\[ P_A \approx M(A_{av})^{-1} p(A/A_{av}), \quad \text{where } \int d\alpha p(\alpha) = 1, \quad (11) \]

\[ Q_a \approx q(A/A_{av}), \quad \text{so } p(\alpha) = g(\alpha)q(\alpha) \]

In Appendix B, we confirm that these expressions preserve the equality \( \sum A \approx \sum P_A \). Since \( A_{av} \) is large, one can replace the sum \( \sum A_{sn} \) in (15) with an integral \( \int dA_{sn} \). After making the change of variable to the fractional overlap \( \mu_{sn} = A_{sn}/(A + A_{sn}) \), so that \( (A + A_{sn})/A \) \( d\mu_{sn} = dA_{sn}/(A + A_{sn}) \), one obtains

\[ P^*_A \approx M(A_{av})^{-1} \sum A_{sn} (A + A_{sn})^{-1} g((A + A_{sn})/A_{sn})q((A + A_{sn})/A_{sn}) \]

\[ = (A_{av})^{-1} \{ (A + A_{sn})/A_{sn} \} g((A + A_{sn})/A_{sn}). \quad (16) \]

where \( \langle \ldots \rangle_{sn} = \int \mu_{sn} e(\mu_{sn}) \ldots \) denotes a weighted average over fractional overlaps integrating over \( 0 \leq \mu_{sn} \leq 1 \).

Given the somewhat obscure nature of the average in (16), it is instructive to consider a simple, although very crude approximation which greatly simplifies its form. If one assumes that the overlap area \( A_{sn} \) with a pre-existing CZ of area \( A \) is always close to \( A_{submuc}(A) \approx \mu_{av}A \), then the distribution \( e(\mu) \) is sharply peaked around \( \mu = \mu_{av} \). In the simplest approximation, one could set \( e(\mu) \approx \delta(\mu - \mu_{av}) \), where \( \delta \) is the Dirac delta function. Then, (16) reduces to

\[ P^*_A \approx M(A_{av})^{-1} \{ 1 - \mu_{av} \}^{-1} g(1 - \mu_{av})^{-1} A/A_{av} \]

\[ \times q(1 - \mu_{av})^{-1} A/A_{av}, \quad (17) \]

which imparts a simpler but non-local nature to the fundamental evolution equation analogous to that seen in evolution equations for fragmentation and coalescence processes. It is also readily confirmed that (17) preserves the equality \( \sum A \approx \sum P_A \). However, again we emphasize the crude nature of the approximation yielding (17).

IV. SCALING FORM OF THE EVOLUTION EQUATION FOR CZ AREAS

A. Scaled evolution equation

Given the forms in Sec. III of various quantities describing nucleation, we recast (3) as an equation for the scaling function, \( g(\alpha) \), of the CZ area distribution. From (4) and using that \( N_{av} = 1/A_{av} \), one obtains

\[ dN_A/dt \approx (A_{av})^{-1} (dN_{av}/dt)(2g + \alpha dg/d\alpha). \quad (18) \]

Then substituting (18) and the scaling relations from Sec. III B for \( P_A \), \( P^*_A \), and \( P^* A \) into (3), we obtain the scaled evolution equation

\[ 2g(\alpha) + adg(\alpha)/d\alpha = -M g(\alpha)q(\alpha) + M\alpha^{-1}(\alpha + \alpha_{sn}) \]

\[ \times g(\alpha + \alpha_{sn})q(\alpha + \alpha_{sn}) + p'(\alpha). \quad (19) \]

Again \( \langle \ldots \rangle_{sn} = \int \mu_{sn} e(\mu_{sn}) \ldots \) denotes a weighted average over fractional overlaps \( \mu_{sn} = \alpha_{sn}/(\alpha + \alpha_{sn}) \) which can be simplified using (17), but this crude approximation is not pursued here. As a consistency check, integration \( \int d\alpha \) of the left-hand side (LHS) of (19) using integration by parts yields \( \int d\alpha g(\alpha) = 1 \). See Appendix B. Integrating the right-hand side (RHS), the first and second terms cancel exactly, and the last term yields \( \int d\alpha p'(\alpha) = 1 \) matching the LHS.

Of course, the solution \( g(\alpha) \) to (19) depends on the detailed form of the scaling functions for the intrinsic overlap probability, \( q(\alpha) \), the CZ areas of just-nucleated islands, \( p'(\alpha) \),
and the distribution of fractional overlaps, \(e(\nu)\). The complex form of (19) precludes a simple form such as GWS or GG for \(g(\alpha)\). This is analogous to the situation for the scaling function for the ISD where the complex form of its exact evolution equation also strictly precludes commonly adopted simple forms.\(^3\)

B. Asymptotic forms for solution

Despite the difficulty of exact analysis for (19), it is instructive to explore some approximations which might provide insight into asymptotic behavior of the solution \(g(\alpha)\). First, for large \(\alpha\), one might anticipate that the major impact on large CZs is reduction in their size due to overlap by CZs of just-nucleated islands. (Such large CZs are perhaps created early at the end of the transient regime, and surviving into the steady-rate regime.) In this case, it is plausible that the first loss term on the RHS of (19) dominates the gain terms on the RHS, so that

\[
\{2 + M q(\alpha)\} g(\alpha) + adg(\alpha)/d\alpha \approx 0, \text{ for large } \alpha, \text{ so (20)}
\]

However, our picture for the formation of small CZs is quite different from Ref. 14 (see Sec. V), and furthermore is actually more complicated than suggested above. New CZs with small area can also be created “indirectly” by the CZs of just-nucleated islands mostly overlapping those of existing islands with typical areas. This corresponds to the second term on the RHS of (18) with small \(\alpha\) and \(\alpha_m = O(1)\), and thus does not involve \(p^*(\alpha)\) for small \(\alpha\). More generally, suppose that the overall rate of creation of new CZs with small area \(\alpha\) combining contributions from the second and third terms on the RHS of (19), has the scaling form \(p_{\text{new}}(\alpha) \sim \alpha^\beta\). Then, it still follows from (19)) that \(g(\alpha) \sim \alpha^\beta\), as \(\alpha \to 0\).

V. SCENARIOS FOR CREATING SMALL CZS

The scenario considered by Pimpinelli and Einstein for creating small CZs involves nucleation of new islands at a location on the surface surrounded by multiple pre-existing CZs having small size.\(^{14,17}\) However, such regions are extremely rare, their population scaling like a higher power of the population of small CZs, a feature not analyzed in these previous studies.\(^{36}\) Instead, we claim the dominant pathway for creating small CZs does not involve pre-existing small CZs, i.e., that new small CZs are primarily created in the absence of pre-existing small CZs.\(^{16}\) Below, we present and analyze various key scenarios, although a comprehensive analysis of all possibilities is difficult.

A. Direct creation of new CZs with small area

One scenario is where a small CZ is created by nucleation of an island in between the sides of two \((m = 2)\) closely spaced near-square islands. See Fig. 2(a). We note that the likelihood of this scenario is dependent on the structure of the islands. Here, the finite extent of islands is key in facilitating creation of a small CZ between a nearby existing pair. Also higher coverage, \(\theta\), will facilitate this scenario. Another scenario is nucleation of an island in the middle of a triangle of three \((m = 3)\) closely spaced islands. See Fig. 2(b). In both these cases, the small CZ is associated with
the just-nucleated island, and thus with $P_A^∗$ or $p^∗(α)$ described in Sec. III. Of course, one could also consider nucleation in the middle of larger collections of nearby islands ($m > 3$). However, analysis below indicates that the likelihood of such events is negligible. A semi-quantitative analysis of the $p^∗(α)$ associated with these scenarios has two components. First, we must determine the small population, $P_m$, of the above “rare” CZ configurations. Second, we must determine the probability, $P_{muc}$, that nucleation occurs in the appropriate restricted sub-region as needed to create a desired small CZ.

For the first part of the analysis, consider the density, $N_{isl}(r)$, of islands with their centers separated by a distance $r$ from the center of a specified island by a vector separation $r$. One has that $N_{isl}(r) \approx C_{isl}(r)N_{isl}$, where the island-island correlation function satisfies $C_{isl}(r) \sim [N_i(r)/N_i]^{−1}$, and $N_i(r)$ represents the adatom density at position $r$ from the center of an island. $N_i(r)$ decreases linearly to zero at the island edge, so we expect that $N_i(r)/N_i \sim r/r_{isl}$ for closely spaced islands, where $r = |r|$ and again $r_{isl}$ is the mean island separation. We conclude that the relative probability to find a pair ($m = 2$) or trio ($m = 3$) of closely spaced islands with separations $r \ll r_{isl}$ should scale like

$$P_m \sim (r/r_{isl})(C_{isl}^{m−1} \sim (r/r_{isl})^{(m−1)(i+1)+1}).$$  \hspace{1cm} (24)

The factor of $(C_{isl})^{m−1}$ gives the probability for a specific orientation of the m-tuple of islands, and the additional factor of $(r/r_{isl})$ accounts for the volume of phase-space for arbitrary rotations.

The second factor is the relative probability for nucleation in an appropriate restricted subregion, as indicated above, separated by island edges by a distance of order $r$. First, we note that the total rate of nucleation, $K_{CZ}$, for a typical CZ scales like

$$K_{CZ} \sim h(N_i)^{i+1}A_{av} \sim h(F/h)^{i+1}(r_{isl})^{2i+4}$$

using

$$N_i \sim (F/h)(r_{isl})^{-1} = (F/h)r_{isl}^2,$$

which follows from the steady-state results from Sec. II B.

This type of scaling of nucleation rates with linear size was a key component of previous geometry-based simulations of island nucleation and growth.\(^{22-24}\) Next, we consider the total nucleation rate for the above-mentioned restricted sub-region of area $\sim r^2$ denoted by $K_{SUB}$. Let $N_i(loc)$ denote the typical value of the adatom density in this subregion. Then, it follows that $K_{SUB}$ scales like

$$K_{SUB} \sim h[N_i(loc)]^{i+1}r^2 \sim h(F/h)^{i+1}r^{2i+4}$$

using

$$N_i(loc) \sim (F/h)r^2.$$  \hspace{1cm} (26)

Thus, one has that

$$P_{muc} = K_{SUB}/K_{CZ} \sim (r/r_{isl})^{2i+4}. \hspace{1cm} (27)$$

Combining the results (24) and (27), and using that $α \sim (r/r_{isl})^2 ≪ 1$, the likelihood to create a new small CZ by the above mechanisms scales like

$$p^∗(α) \sim P_m P_{muc} \sim α^6,$$

where

$$β = β_m(i) \approx (m + 1)(i + 1)/2 + 3/2.$$  \hspace{1cm} (28)

Thus, one has $β_2 = 3i/2 + 3$ and $β_3 = 2i + 7/2$. For compact islands with finite extent, one expects that contributions from $m = 2$ (i.e., the smaller $β_m$) will dominate.

B. Major reduction in size of existing CZs

Another scenario can generate island and CZ configurations similar to those described in Sec. V A, but with a different order of island nucleation. This scenario will be most viable for $m = 3$, i.e., a triangular triple of nearby islands surrounding a fourth, rather than for $m = 2$. The idea for $m = 3$ is that islands could first nucleate creating two corners of the triangle, then a third island nucleates which will end up in the middle, and finally a fourth nucleates constituting third corner. See Fig. 3. In this case, the CZ of the last just-nucleated island is not generally small. However, the CZ of the third nucleated island, which was generally large upon formation, is greatly reduced (i.e., mostly covered) upon nucleation of the fourth island. In this case, the formation of a small CZ is associated with the term $P_A^∗$ in $dN_A/dt$. Specifically, it reflects the small but non-zero weight in the upper end of the fractional overlap distribution, $e(η)$, with $μ$ close to unity.

For a semi-quantitative analysis of the scenario in Fig. 3, the probability to nucleate the first three islands in the above configurations should scale as $P_{3}^∗ \sim (r/r_{isl})^{2i+3}$ just as does $P_3$. However, the last “outer” island will nucleate in a region where $N_i(r)/N_i \sim r/r_{isl}$ noting the linear increase of adatom density moving away from the first three islands. This suggests a relative nucleation probability of $P_{muc}^{3}$.
\( \sim (r/r_{\text{th}})^2 \left[ N_1(r)/N_1 \right]^{\alpha+1} \sim (r/r_{\text{th}})^{\alpha+3} \). Using that \( \alpha \sim (r/r_{\text{th}})^2 \), the overall rate for formation of the small CZ scaling like

\[ p_{\text{new}}(\alpha) \sim P_3^{3\text{r}}\text{mac} \sim \alpha^0 \]  

(versus \( \beta_3(i) = 2i + 7/2 \) in Sec. V A). Since \( \beta_3 \) has the same value as \( \beta_2 \) given in Sec. V B, this process should also make a significant contribution to the measured \( \beta \). In fact for point islands, where the \( m = 2 \) pathway of Sec. V A is not effective, one expects this alternative pathway for \( m = 3 \) to dominate.

VI. KINETIC MONTE CARLO (KMC) SIMULATION RESULTS FOR SQUARE ISLANDS

Efficient KMC simulation for large values of \( h/F \) utilizes a standard Bortz-algorithm to enable collection of extensive statistics on island configurations. To assess quantities such as \( g(\alpha) \), CZ tessellations are constructed which are taken as edge cells, so that CZ boundaries are equidistant from island edges.\(^{10-12,24} \) However, to collect adequate statistics at a selected coverage for quantities associated with new islands such as \( p(n) \), the following refined approach is implemented:\(^{9,12,13} \) (i) we perform a standard simulation to generate an island distribution at the desired coverage; (ii) we continue simulation until one new island is nucleated and measure desired quantities associated with this new island; (iii) we remove the new island and continue simulation until another new island is nucleated and perform desired measurements; (iv) we repeat this process for \( \sim 100 \) nucleations per configuration generated in (i). This approach is far more efficient than collecting information on just one new island from each simulated configuration.

A. Behavior for \( i = 1 \)

We focus below on analysis of data from simulations with \( i = 1 \) and \( h/F = 10^6 \) at \( \theta = 0.1 \) ML. The total number of islands and thus CZs used to assess \( g(\alpha) \) is \( \sim 268 \) 500, with average area \( A_{\text{av}} = 372.4 \). With regard to detailed analysis of quantities such as \( p(\alpha) \), \( p(n) \), \( e(\eta) \), etc., we consider a total of 50 000 new CZs which have an average area of \( A_{\text{synce}} = 201.6 \). Key parameters extracted from this analysis include \( a_{\text{sync}} = A_{\text{sync}}/A_{\text{av}} = 0.54 \), \( \mu_{\text{av}} = 0.099 \), and \( M = 4.65 \). Note that the \( a_{\text{sync}} \) and \( \mu_{\text{av}} \) values are significantly lower than for point islands. This feature is a natural consequence of the finite extent of the square islands at 0.1 ML reducing the relevant full or partial CZ areas. Also, for point islands, \( a_{\text{sync}} \) is somewhat above \( M \mu_{\text{av}} = 0.46 \).

We start by performing least-square fits of GG forms to \( g(\alpha) \).\(^{37} \) See Fig. 4. First, this is done by selecting the reasonable choices of \( n \)-value and fitting \( \beta \). For \( h/F = 10^6 \), this yields \( \beta = 3.05 \) for \( n = 2 \), \( \beta = 3.48 \) for \( n = 1.75 \), and \( \beta = 4.07 \) for \( n = 1.5 \). One can also simultaneously adjust both \( \beta \) and \( n \). For \( h/F = 10^5 \), this yields \( \beta = 3.84 \) and \( n = 1.59 \). As an aside, from more limited simulation data, we obtain \( \beta = 3.55 \) and \( n = 1.65 \) for \( h/F = 10^7 \), and \( \beta = 3.36 \) and \( n = 1.70 \) for \( h/F = 10^8 \).\(^{38} \)

The advantage of above fitting approach, which assumes that \( g(\alpha) \) is described by a GG form, is that one utilizes simulation data for all \( \alpha \). These data are extensive for CZ areas around the average value (\( \alpha \approx 1 \)), thereby compensating for limited data in the upper and lower tails of the distribution. However, there are also significant disadvantages. First, the fit is expected to be sensitive to deviations for finite \( h/F \) from asymptotic behavior as \( h/F \to \infty \). Variation with \( h/F \) in the peak if the CZD has been assessed previously and significantly impacts estimates of \( \beta \).\(^{15} \) Second, a more fundamental concern is that \( g(\alpha) \) will not be exactly described by a GG form (as indicated above). This feature can introduce some bias and correlation in the resulting values of \( \beta \) and \( n \) (e.g., higher \( n \) produces smaller \( \beta \)).

The above observations motivate direct unbiased fitting of the lower tail of \( g(\alpha) \) to assess \( \beta \), and thus the proposed relationship to \( i \). Here, one must choose the range of \( \alpha \) over which to fit. Lower (and narrower) ranges better describe asymptotic behavior as \( \alpha \to 0 \), but these suffer from limited data. Thus, we give \( \beta \)-values from fitting over different ranges with results shown in Table I. The decrease in estimated \( \beta \) with broader range of \( \alpha \) just reflects non-asymptotic behavior as \( \log g(\alpha) \) versus \( \log \alpha \) is convex. Thus, extrapolating values

![Fig. 4.](image-url)

**FIG. 4.** \( g(\alpha) \) for \( i = 1 \) at 0.1 ML. (a) Linear plot for \( h/F = 10^6 \) (inset: behavior for \( h/F = 10^7 \)). Fitted curve is for \( n = 1.75, \beta = 3.5 \). (b) Log-log plot for \( h/F = 10^6 \). GG curves are shown for \( n = 2 \) with \( \beta = 2.5, 3, 3.5 \), and for \( n = 1.5 \) with \( \beta = 3.5, 4, 4.5 \).

<table>
<thead>
<tr>
<th>( \alpha ) (h/F = 10^6)</th>
<th>0.04-0.3</th>
<th>0.04-0.4</th>
<th>0.04-0.5</th>
<th>0.04-0.6</th>
<th>0.07-0.3</th>
<th>0.07-0.4</th>
<th>0.07-0.5</th>
<th>0.07-0.6</th>
<th>0.1-0.3</th>
<th>0.1-0.4</th>
<th>0.1-0.5</th>
<th>0.1-0.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta ) (h/F = 10^6)</td>
<td>3.37</td>
<td>3.34</td>
<td>3.28</td>
<td>3.20</td>
<td>3.30</td>
<td>3.28</td>
<td>3.19</td>
<td>3.09</td>
<td>3.49</td>
<td>3.29</td>
<td>3.14</td>
<td>2.98</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \alpha ) (h/F = 10^5)</th>
<th>0.07-0.3</th>
<th>0.07-0.4</th>
<th>0.07-0.5</th>
<th>0.07-0.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta ) (h/F = 10^5)</td>
<td>3.30</td>
<td>3.28</td>
<td>3.19</td>
<td>3.09</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \alpha ) (h/F = 10^8)</th>
<th>0.1-0.3</th>
<th>0.1-0.4</th>
<th>0.1-0.5</th>
<th>0.1-0.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta ) (h/F = 10^8)</td>
<td>3.49</td>
<td>3.29</td>
<td>3.14</td>
<td>2.98</td>
</tr>
</tbody>
</table>
for broader ranges to narrower ranges should give the most reliable estimates. Results from analysis of extensive data for \( h/F = 10^6 \), and more limited data for higher \( h/F \), consistently suggest that \( \beta \) is around 3.5.

As indicated in the theoretical development of Secs. II–V, \( g(\alpha) \) behavior is controlled by subtle spatial aspects of the nucleation process. Thus, it is appropriate to also examine simulation results for the key scaling functions \( p^*(\alpha) \), \( p(\alpha) \), \( q(\alpha) \), and \( e(\mu) \). See Fig. 5. Consistent with the (over)simplified analysis of (22) and (23), the distribution, \( p^*(\alpha) \), of areas of new CZs appears to scale in the same way as \( g(\alpha) \) for small \( \alpha \). Also, \( g(\alpha) \) is naturally smaller than \( p^*(\alpha) \) in the small-\( \alpha \) range. However, direct creation of new small CZs associated with \( p^*(\alpha) \), as described in Sec. VA and as assumed in (23), is not necessarily the dominant pathway. Rather, examination of simulation movies reveals a significant fraction of cases

\[ \text{FIG. 5. (a) } p^*(\alpha) \text{ compared with } g(\alpha); \text{ (b) } e(\mu); \text{ (c) } p(\alpha); \text{ and (d) } q(\alpha) = p(\alpha)/g(\alpha). \]

\[ \text{FIG. 6. CZ distribution for } i = 1, \ h/F = 10^6 \text{ and } \theta = 0.1 \text{ ML (500} \times 500 \text{ system). Also indicated for small CZs is the mechanism of formation: yellow)} \text{ red circle = direct formation; light blue/blue circle = indirect formation.} \]
where such small CZs are formed indirectly from a substantial reduction in area of existing CZs, as described in Sec. V B. See Fig. 6. While the associated values of the fractional overlap scaling function, $e(\mu)$, for larger $\mu$ are small, so are values of $p'(\alpha)$ for correspondingly small $\alpha \sim 1 - \mu$. (We should also note that $e(\mu)$ is far from the delta function form corresponding to a narrow distribution of fractional overlaps where typically $\mu \approx \mu_{av}$. Instead, small fractional overlaps are common.)

With regard to behavior of $g(\alpha)$ for large $\alpha$, as anticipated, the intrinsic overlap probability, $q(\alpha)$, increases non-linearly with $\alpha$. The precise form is not clear from the limited data, but a simple fit to $q(\alpha) \sim \alpha^n$ yields $n \approx 1.5$ with considerable uncertainty.

**B. Behavior for $i = 0$**

For the $i = 0$ model, the island density scales like $N_{isl} \sim \sigma^{1/2}$ independent of $h/F$. Most studies considered the regime of “rapid conversion” with $\sigma \gg (h/F)^{-1/2}$ yielding persistent nucleation in the steady state and a monotonically decreasing ISD. A crossover occurs reducing $\sigma$ to $\sigma = (h/F)^{-1/2}$ and below to growth behavior avoiding persistent nucleation. Thus, in this work, we select $\sigma = (h/F)^{-1/2}$ which leads to relatively low $N_{isl} \sim (h/F)^{-1/4}$. Our construction and analysis of CZs includes single converted adatoms as well as larger clusters, but not isolated diffusing adatoms. Even at a low $0 \sim 0.1$ ML, there is significantly more island

<table>
<thead>
<tr>
<th>$\alpha$ (0 = 0.1 ML)</th>
<th>0.057-0.3</th>
<th>0.057-0.4</th>
<th>0.057-0.5</th>
<th>0.057-0.6</th>
<th>0.056-0.3</th>
<th>0.056-0.4</th>
<th>0.056-0.5</th>
<th>0.056-0.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$ (0 = 0.1 ML)</td>
<td>2.74</td>
<td>2.64</td>
<td>2.53</td>
<td>2.42</td>
<td>2.75</td>
<td>2.76</td>
<td>2.71</td>
<td>2.64</td>
</tr>
</tbody>
</table>

**TABLE II.** Estimates of $\beta$ from direct fitting to $g(\alpha)$ data for $i = 0$ and $h/F = 10^6$ for ranges of small $\alpha$ (shown) for both $\theta = 0.1$ ML and $\theta = 0.04$ ML.
coalescence for \( i = 0 \) than for \( i = 1 \) consistent with the larger island-island correlation function, \( C_{\text{ii}}(r) \sim (r/r_0)^{i+1} \), for small \( r \). Since our theory is geared to the pre-coalescence regime, we also consider behavior for lower \( \theta = 0.04 \) ML to limit coalescence.

We present a more limited analysis of behavior for \( i = 0 \) than for \( i = 1 \) focusing on the case \( h/F = 10^6 \) and \( \sigma = 10^{-3} \) for \( \theta = 0.1 \) ML. The total number of islands and thus CZs used to assess \( g(\alpha) \) is \( \approx 346 \) 500 with an average CZ area \( A_{\text{av}} = 192.71 \). Least-squares fitting of a GG form adjusting both \( \beta \) and \( n \) yields \( \beta = 2.80 \) and \( n = 1.42 \) for \( \theta = 0.1 \) ML.

One obtains \( \beta = 3.46 \) and \( n = 1.42 \) for \( \theta = 0.04 \) ML. Given the limitations or bias in least square fitting to GG forms as discussed for \( i = 1 \), we do not give a more extensive analysis here. The reader is referred to Fig. 7 where GG forms are compared with simulation data for 0.1 ML for various choices of \( \beta \) and \( n \). Choosing \( n = 2 \) requires \( \beta \) values above 2.5 for a reasonable fit. It appears that the data are better fit by lower \( n \) values and higher \( \beta \) closer to 3 which is consistent with \( \beta \approx \beta(0) = 3 \) for \( i = 0 \) in (27). As for \( i = 1 \), direct fitting of the lower tail of \( g(\alpha) \) to assess \( \beta \) avoids bias of an assumed GG form, and is thus most reliable provided there is sufficient simulation data. Results are shown in Table II which suggest again that \( \beta \) is just below 3.0.

Finally, we remark that analysis of simulation movies suggests that as for \( i = 1 \), a significant fraction of small CZs is created indirectly rather than directly. See Fig. 8.

VII. DISCUSSION AND CONCLUSIONS

Recent extensive interest in the capture zone area distribution (CZD) has been prompted by the proposal of Pimpinelli and Einstein (PE) that the CZD has a GWS form, and furthermore that the exponent, \( \beta \), describing behavior for small areas is directly related to the critical size.\textsuperscript{14,17,21} Considering \( 2D \) systems for \( i \geq 1 \), Shi, Shim, and Amar (SSA) utilized point-island models to critique the original proposal of PE that \( \beta_{\text{WS}} = i + 1 \) for \( i = 1 \).\textsuperscript{13} Their simulation data for typical \( h/F = 10^6 \) fitted to a GWS form instead suggested that \( \beta \approx 3 \). Furthermore, SSA demonstrated that key features of the scaled CZD, particularly the peak height, had a significant dependence on \( h/F \). Also, slightly different versions of the point-island model had significantly different peak heights for smaller \( h/F \). However, extrapolation to \( h/F \to \infty \) removed discrepancies between different versions of the model, as might be expected. This analysis again indicated that \( \beta \approx 3 \) assuming that the CZD, and particularly its peak, is described by a GWS form. This analysis appears consistent with the modified proposal of PE that \( \beta_{\text{WS}} = i + 2 \) (= 3 for \( i = 1 \)).\textsuperscript{17} However, as discussed above, there are limitations to an analysis based on fitting to a GWS (or GG) form which does not correspond to the exact form of the CZD.

Additional studies by Oliveira and Aarão Reis\textsuperscript{18–20} considered the CZD for point, square, and fractal island models utilizing a modified scaling formulation to reduce dependence of the scaled CZD upon \( h/F \). However, this modified formulation was not designed to highlight or analyze behavior for small CZ areas. Point-island behavior for \( i = 1 \) is fit reasonably but not perfectly by \( \beta \approx 3 \) and \( n = 2 \), or by \( \beta \approx 4 \) and \( n = 1.5 \). Square island behavior is also fit reasonably by \( \beta \approx 3 \) and \( n = 2 \), but with significant discrepancies in the tails of the CZD. A key focus of these studies was on the form of the right tail of the CZD for large areas. For \( i \geq 1 \), simulation data for point islands and also for square islands at pre-coalescence coverages appear consistent with a Gaussian tail, i.e., with \( n = 2 \), or at least with \( n \) close to but probably not above 2. For square islands at higher coverages where coalescence is significant, there was a clear deviation from Gaussian towards exponential tails for large CZ areas.

Our own analysis for \( 2D \) systems with \( i \geq 1 \) is motivated by our development of an exact evolution equations for the CZD. This equation suggests (but does not rigorously prove) the possibility of a non-Gaussian tail for large areas. It also provides and incorporates a different picture from PE of behavior for small areas. Specifically for the latter, appropriately identifying and analyzing the dominant pathways for creation of small CZ areas (which in our picture do not require pre-existing small CZs), we predict larger \( \beta \) values than the refined PE prediction of \( \beta_{\text{WS}} = i + 2 \). Furthermore, we argue that the most reliable assessment of \( \beta \) comes from direct fitting of CZD data for small \( \alpha \). This contrasts the strategy of previous studies of fitting to a GWS or GG form (which can bias the outcome since the CZD is not exactly described by these forms). Such direct fitting does indicate somewhat larger values of \( \beta \) than \( \beta_{\text{WS}} = i + 2 \) for \( i = 1 \), although the values are below those for the pathways which we identified for creation of new small CZs. We note one limitation of our analysis is that data are obtained for relatively small \( h/F \) (as in analyses of PE). Data for higher \( h/F \) are preferred to avoid possible discrete lattice effects for small \( \alpha \), but its acquisition is computationally more demanding. With regard to behavior for large areas, our simulation data are not sufficiently extensive to make a definitive assessment. Behavior is consistent with Gaussian tails (\( n = 2 \)) but could also be described well by somewhat smaller \( n \) for \( i = 1 \). Also taking into consideration the analysis of Refs. 18 and 19, one concludes that \( n \) is close to 2 (and that it does not exceed 2).

Considering \( 2D \) systems for \( i = 0 \), our analysis indicates \( \beta \) values significantly above the refined PE prediction of \( \beta_{\text{WS}} = i + 2 \) for \( i = 0 \). It also seems likely that the large area tail of the CZD is described by \( n < 2 \) rather than by a Gaussian tail. Previous studies for \( i = 0 \) with larger exchange rates, \( \sigma >> (h/F)^{-1/2} \), and thus persistent nucleation, indicated exponential tails for large areas.\textsuperscript{15} This is certainly reasonable since persistent random nucleation produces quasi-random island distribution for which one expects a CZD closer to that of a Poisson point process. For the latter, a Gamma function form with exponential tail is expected.\textsuperscript{29–31} However, for our choice of lower exchange rate \( \alpha = (h/F)^{-1/2} \) which eliminates such persistent nucleation,\textsuperscript{3} the form of the large area tail is less clear a priori.

In summary, our exact evolution equation for the CZD provides a valuable additional formulation and perspective beyond a heuristic treatment for analysis and understanding of this distribution. In particular, our formulation highlights the importance of subtle spatial details of the nucleation process, and specifically of related quantities such as the intrinsic overlap probabilities and fractional overlap distribution.
These quantities have received limited attention and are not considered in heuristic treatments, but their detailed form impacts that of the CZD. Given the complexity of the process of homogeneous island nucleation and growth of islands during deposition, there does not yet exist a definitive characterization and understanding of either the ISD or the CZD. However, this goal is facilitated with the aid of detailed analytic theory such as that developed here, together with more extensive simulation data are most readily obtained for smaller h/F = 10^6-10^7, but one is primarily interested in the scaling limit h/F → ∞, so reliable extrapolation is also required.29

ACKNOWLEDGMENTS

We thank Tiago Oliveira for valuable comments on the manuscript. Y.H. and J.W.E. were supported for this work by NSF Grant Nos. CHE-1111500 and CHE-1507223. Computations utilized NSF-supported XSEDE resources. M.L. was supported by NSF of China under Grant No. 51271197. The work was performed at Ames Laboratory which is operated for the USDOE by Iowa State University under Contract No. DE-AC02-07CH11358.

APPENDIX A: ATOMISTIC MODELS

FOR ISLAND FORMATION

Our tailored models for island formation with i = 1 involve: (i) Random deposition at rate F on a square lattice of adsorption sites with periodic boundary conditions. (ii) Isolated adatoms hop to a NN site with hop rate h per direction. (iii) Typically, singly coordinated “edge atoms” can hop at rate h to either NN or next-NN edge sites provided they are still coordinated (i.e., they cannot detach from islands). (iv) Multiply coordinated adatoms are immobile. (v) Atoms deposited on top of islands immediately undergo a random walk until they reach the edge of the island and hop down to the substrate. Also, atoms landing on isolated adatoms immediately hop down.

Incorporation of facile diffusion of atoms along the island periphery both along straight edges and around corners or kinks suffices to achieve compact islands.3 Implementation of rapid transport of atoms landing on top of islands to the first layer avoids second layer population.

One extra constraint not described above is that singly coordinated adatoms in clusters of size s = 2 or 3 are not allowed to undergo hops to next-NN edge sites. This constraint excludes long-range diffusion of these small clusters, which if present would have the effect of changing the scaling behavior of N_{isl}3. For clusters of 4 or more atoms, restructuring will produce a rectangular core of doubly coordinated adatoms which is permanently frozen, thus excluding long-range diffusion.

Our version of a “pure” i = 0 model excludes nucleation by aggregation of multiple diffusing adatoms. This is achieved by preventing isolated diffusing adatoms from reaching NN sites, and also exclude deposition on top of or next to isolated adatoms. Island nucleation occurs exclusively by isolated adatoms converting to permanently immobile adatoms at rate oh. Diffusing adatoms aggregate irreversibly with converted immobile adatoms and with clusters of adatoms which form around these, i.e., detachment from such islands is prohibited. Edge diffusion is included as for i = 1 ensuring the formation of compact islands. Atoms deposited on top of islands are also treated as for i = 1. In our construction and analysis of CZs for i = 0, we will count as islands single converted adatoms in addition to larger clusters, but not isolated diffusing adatoms.

We will not present results for i > 1, however we mention that the model for i = 1 can be naturally extended to i = 2 or 3. For example, for i = 2, singly coordinated edge atoms can only detach at rate h if their unique neighbor is also singly coordinated (i.e., if they are part of an ad-dimer). Otherwise detachment is excluded. Doubly coordinated adatoms are again frozen. In these models with i ≥ 1, edge diffusion, as described above, ensures the formation of compact islands. For KMC images of compact islands in such models with i = 1-3, see Refs. 23 and 24.

APPENDIX B: OVERLAP PROBABILITIES

AND CONSISTENCY RELATIONS

First, we comment on the regime of smaller Aav so that CZ areas A are discrete. Then, there is a significant probability that the CZ of a just-nucleated island can overlap k = 2 (or more) pre-existing CZs of the same size, A. Then, in determining PA, one considers a large number, j = 1 – N, of nucleation events and counts the number k of overlaps in each case, setting PA ≈ j/k/∑j = 1 j = 1/k/N. A similar treatment is implemented for related quantities P^+ A and P A(A_{isl}).

Second, we show that ∑A P^+ A = ∑A P A exploiting a change of variable B = A + A_{isl},

∑A P^+ A = ∑A P A = ∑A ∑A_{isl} P A(A_{isl}) = ∑B ∑A_{isl} P B(A_{isl}) = ∑B P B. (B1)

The same change of variable B = A + A_{isl} applies to simplify

\[ \sum_A P^+ A \approx M(A_{av})^{-1} \sum_A \sum_{A_{isl}} (A + A_{isl})^{-1} g({A + A_{isl}}/{A_{av}}) q({A + A_{isl}}/{A_{av}}) e(A_{isl}/{A + A_{isl}}). \]

\[ \approx M(A_{av})^{-1} \sum_B g(B/{A_{av}}) q(B/{A_{av}}) \sum_{A_{isl}} (B)^{-1} e(A_{isl}/B) \]

\[ \approx M(A_{av})^{-1} \sum_B g(B/{A_{av}}) q(B/{A_{av}}) = \sum_B P_B. \] (B2)

Third, we comment on the consistency of the evolution Equation (19) for the scaling function which follows from integrating over a. After integration, the second term on the RHS of (19) yields
which cancels the first term given the normalization of $e(\mu_{sn})$. Finally, we remark that for the most simple $\delta$-function choice of $e(\mu_{sn}) \approx \delta(\mu_{sn} - \mu_{av})$ (which again has limitations), (19) reduces to the non-local equation

$$
M \int d\alpha \int d\mu_{sn} e(\mu_{sn}) \alpha^{-1}(\alpha + \alpha_{sn}) g(\alpha + \alpha_{sn}) q(\alpha + \alpha_{sn}) = M \int d\mu_{sn} e(\mu_{sn}) \int d\alpha (1 - \mu_{sn})^{-1} g(1 - \mu_{sn})^{-1} (\alpha + \alpha_{sn}) q(1 - \mu_{sn})^{-1} \\
= M \int d\mu_{sn} e(\mu_{sn}) \int d\alpha' g(\alpha') q(\alpha'). \quad (B3)
$$

8P. A. Mulheran and J. A. Blackman, Phys. Rev. B 53, 10261 (1996). This study and also Ref. 7 inappropriately identify the ISD with the CZD (at least for $i > 1$).
29Deviations from the limiting scaling form clearly occur when characteristic lengths (e.g., the mean island separation) become comparable to the lattice constant. For another perspective ISD studies suggest that these deviations decrease with increasing ratio of the number of islands nucleated in the steady state relative to the transient state. See Y. Han, E. Gaudry, T. J. Oliveira, and J. W. Evans, J. Chem. Phys. 145, 211904 (2016).
34$A_{\text{ad}}$ corresponds to $A_{\text{ad}}$ in previous papers.
35If $\frac{d}{dt} A = F(A) + \eta(t)$ with $F(A) \sim -kA$ for large $A$ and noise $\eta$ satisfying $\langle \eta(t) \rangle = 0$ and $\langle \eta(t) \eta(t') \rangle = \Lambda \delta(t-t')$, then one has that $P_A \sim \exp[\langle A \rangle \sim -kA^2/2\Lambda]$.
36Ref. 14 used the local nucleation rate $k_{\text{loc}}(\Lambda^{-1}) \sim A^{-1}$ to estimate the rate of formation of CZs with small area $A$ in regions of the surface with other CZs of that size. Ref. 17 instead used the total nucleation rate, $k(\Lambda^{-1}) \sim A^{1/2}$ integrated over an area $A$. The latter can be obtained by integrating the local adatom density $\int_0^1 \text{dr} N(r)\Lambda^{1/2}$, but such analysis is not necessary as $N(r)$ adopts similar values over most of the area. These analyses did not consider the probability of finding the configurations of existing CZs in which they regarded formation of new small CZs as occurring.
37Fitting of a GG form to $p^*(\alpha)$ yields $\beta = 3.52$ selecting $n = 2$ and $\beta = 4.01$ selecting $n = 1.75$, and $\beta = 4.69$ selecting $n = 1.5$. Adjusting both $\beta$ and $n$ to fit $p^*(\alpha)$ yields $\beta = 3.70$ and $n = 1.90$.
38For $h/F = 10^7$ we also obtain $\beta = 2.94$ for $n = 2$, and $\beta = 3.92$ for $n = 1.5$. For $h/F = 10^8$, we obtain $\beta = 2.87$ for $n = 2$ and $\beta = 3.83$ for $n = 1.5$. 

$$
M \int d\alpha \int d\mu_{sn} e(\mu_{sn}) \alpha^{-1}(\alpha + \alpha_{sn}) g(\alpha + \alpha_{sn}) q(\alpha + \alpha_{sn}) = M \int d\mu_{sn} e(\mu_{sn}) \int d\alpha (1 - \mu_{sn})^{-1} g(1 - \mu_{sn})^{-1} (\alpha + \alpha_{sn}) q(1 - \mu_{sn})^{-1} \\
= M \int d\mu_{sn} e(\mu_{sn}) \int d\alpha' g(\alpha') q(\alpha'). \quad (B3)
$$