LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

UCRL-JRNL-221029

Equilibrium Distributions and the Nanostructure Diagram for Epitaxial Quantum Dots

R. E. Rudd, G. A. D. Briggs, A. P. Sutton, G. Medeiros-Ribeiro, R. S. Williams

May 2, 2006

Journal of Computational and Theoretical Nanoscience

Equilibrium Distributions and the Nanostructure Diagram for Epitaxial Quantum Dots

Robert E. Rudd[a*], G. A. D. Briggs[b], A. P. Sutton[c],

G. Medeiros-Ribeiro[d] and R. Stanley Williams[e]

[a] Lawrence Livermore National Lab., L-045, Livermore CA 94551 USA

[b] Dept. of Materials, University of Oxford, Parks Road, Oxford OX1 3PH UK

[c] Dept. of Physics, Imperial College, Exhibition Road, London SW7 2AZ UK

[d] Laboratório Nacional de Luz Síncrotron, P.O. Box 6192, Campinas, SP 13084-971 Brazil

[e] Hewlett-Packard Laboratories, 1501 Page Mill Road, Palo Alto, CA 94304 USA

July 2, 2006

Abstract

We present in detail a thermodynamic equilibrium model for the growth of nanostructures on semiconductor substrates in heteroepitaxy and its application to ger-

*robert.rudd@llnl.gov

manium deposition on silicon. Some results of this model have been published previously, but the details of the formulation of the model are given here for the first time. The model allows the computation of the shape and size distributions of the surface nanostructures, as well as other properties of the system. We discuss the results of the model, and their incorporation into a nanostructure diagram that summarizes the relative stability of domes and pyramids in the bimodal size distributions.

Keywords: Ge/Si Islands, Stranski-Krastanov, Quantum Dot, Nanostructure Diagram

1 Introduction

Surface nanostructures that form through self-assembly during semiconductor heteroepitaxy hold the promise of a variety of important applications.[1] Their nanoscale size leads to quantum confinement and hence they function as quantum dots (QDs) with size-dependent energy levels. They provide preferential recombination sites for electrons and holes enhancing optical properties. Furthermore, because they self-assemble on semiconductor substrates, there is a natural match with the existing technology of semiconductor fabrication facilities. As a result, epitaxial quantum dots have been proposed for applications ranging from solid-state lasers to integrated circuit transistors to quantum computation q-bits.

Despite the intense effort to study the self-assembly of quantum dots, many of the factors controlling the formation of these islands in semiconductor heteroepitaxial growth

remain poorly understood. These issues are at the heart of the nascent field of nanoprocessing of materials, and they must be resolved before this important field can advance. How, for example, can the distribution of sizes and shapes of the nanostructures be controlled? Uniformity in size and shape is crucial to the majority of the envisaged QD applications, and yet there remains no complete, predictive model to guide control of their growth, let alone facilitate and optimize the design of nanostructure systems. One step toward predictive nanostructure design was the introduction of the nanostructure diagram,[2] a map of QD growth properties as a function of the amount of material deposited and the growth temperature, which is similar in spirit to a phase diagram. The goal of this article is to elaborate its theoretical and computational underpinnings.

The literature on island size distributions is extensive.[1] During the deposition of germanium on silicon the surface is observed to undergo a roughening transition.[3] A series of surface reconstructions is observed, and then after a wetting layer has formed, a variety of different nanostructures are observed to form: pyramids, huts, domes, and superdomes.[4] The bimodal size distribution of coherently strained islands comprised of smaller islands (pyramids) and larger islands (domes) (see Fig. 1) have received considerable attention. Bimodal distributions are seen in nanostructure systems other than Ge/Si. InAs/GaAs(001) exhibits well-defined distributions;[5–7] InP/GaInP shows characteristic bimodal size distributions, with at least two types of islands,[8] as does GaN/AlN.[9]

The basic picture of the self-assembly of Ge islands on a Si(001) surface is clear and uncontroversial: it is a Stranski-Krastanov process.[4,10] During deposition the Ge atoms initially form a flat wetting layer, but the wetting layer is strained due to the 4% mismatch

of the lattice constants and is thus in a state of plane stress with substantial elastic energy. Above a critical thickness of about 3 monolayers the strain is relieved by island formation. The increased total surface energy of the system with islands is more than compensated by the reduction in bulk elastic energy as the three-dimensional islands are unconstrained laterally and able to relax.

The elastic energies of heteroepitaxial islands have been calculated using analytic and numerical models based on a variety of assumptions. Shchukin et al.[11] using analytic elastic Green functions derived an expression for the energy of a single island and a pair of islands in the small slope regime. Atomistic calculations of the island energies provide further insight into the physics.[12] Recently, the small slope assumption has been relaxed in the analytic calculations of Gill and Cocks.[13] Daruka and Barabási[14] used the Shchukin expression to determine a map of the surface growth mode as a function of lattice mismatch and the amount of material deposited at zero temperature, and we note that the results of several experimental studies of the surface growth mode have been summarized in maps as a function of temperature and the amount of material deposited.[15,16] The analysis of the island energetics was extended to a thermodynamic equilibrium model by Medeiros-Ribeiro et al.[17]. In their model the system is essentially at equilibrium during the growth, and the distribution in sizes results from thermal fluctuations about a minimum in the island elastic energy per atom as a function of size. There is experimental evidence that supports the idea of an equilibrium state, or at least a metastable state. After long anneals of Ge/Si(001) at 550 °C the island distribution reached a stationary state, and islands of significantly disparate sizes were observed next to each other.[18]

Nevertheless, island growth often displays kinetic effects, and there has been controversy over the relative importance of thermodynamic and kinetic constraints.[19] A different model of island growth due to Ross et al.[20–22] is that the system is far from equilibrium, even in a meta-stable sense, and it undergoes Ostwald ripening in which the large islands grow at the expense of the small islands. In this picture, the distribution of island sizes is an artifact of the nucleation process: it is not predicted within the model. The formation of distinct population of islands (pyramids, domes, etc) is attributed to differing internal chemical potentials. Other kinetic studies have taken these ideas further.[23] One investigation recently has studied metastability due to the kinetics governed by the chemical potential as the system evolves toward a state determined by the energy per atom in the islands.[24]

We have developed the equilibrium model of island growth to include a more robust description of the statistical mechanics of the system, including the dependence on the chemical potential.[2] While it is clear that the island systems never strictly attain equilibrium, it was shown that not only does the distribution of island sizes agree with the Gibbs distribution (confirming similar findings by the Williams group and others), but that the distributions directly determine the diagram of phases of surface populations. This added further evidence that, at least for Ge islands on Si over a range of material deposition and a narrow range of temperatures, the system is to a large extent in equilibrium. In the process of analyzing this model we also begin to address deviations from the model that are due to kinetic effects. The model was solved to determine the distributions of domes and pyramids in the Ge/Si system. The results were compiled in the nanostructure diagram.

[2] The model has also been applied to GaN/AlN growth, where bimodal distributions are also observed.[9] The results were encouraging for this nitride system, too.

In this article we present the statistical mechanics model in much greater detail. We show explicitly how the model was constructed, parameterized and solved. The previous letter only presented the model in the broadest terms. Much of the mathematics and numerics had to be omitted due to space restrictions. Here we present those details. We also describe the results of the model in greater detail. For example, we present a formula for the chemical potential from which the island distributions and related properties of the system can be derived.

2 Construction of the thermodynamic model

We begin with the construction of the model describing the thermodynamic equilibrium of a system of islands governed by elastic energies. The system will be taken to be at a given temperature and equivalent coverage. Here and throughout this article we use the term coverage to denote the amount of material deposited, as measured in equivalent monolayers (MLs). A sketch of the formulation of the model was presented previously;[2] the first detailed account is given here. The elastic internal energy of the system is dominated by contributions from the elastic strain of each individual island a and that of pairs of islands, a and b,

$$E_{\text{elastic}} \approx \varepsilon_0 + \sum_a \varepsilon_a^{(1)} + \sum_{a<b} \varepsilon_{a,b}^{(2)} \tag{1}$$

where these are the first terms in a cluster expansion of the energy valid for sufficiently well separated islands. Many-island interactions would enter only at very high coverages. The term ε_0 is the energy of the reference state: the substrate without islands. The single island energy $\varepsilon_a^{(1)}$ represents the reduction in the strain energy of the compressed wetting layer as atoms are rearranged from the two dimensional surface into three dimensional islands. The absence of surrounding material allows the islands to relax laterally, lowering the strain energy. The pair interaction, $\varepsilon_{a,b}$, results from the strain field emanating from each island.[25]

Shchukin and coworkers[11] have used fairly general arguments in continuum mechanics to show that the internal energy of an individual island, $\varepsilon^{(1)}$, takes on a form with bulk, surface and edge terms,

$$\varepsilon^{(1)} = A_X\nu + B_X\nu^{2/3} + C'_X\nu^{1/3}\log\nu + D'_X\nu^{1/3} \quad (2)$$

$$\approx A_X\nu + B_X\nu^{2/3} + C_X\nu^{1/3} + D_X \quad (3)$$

where X labels the type of the island, ν is the number of atoms in the island and A_X, B_X, C'_X and D'_X are coefficients that determine the bulk, surface, surface stress and edge energies, respectively, relative to the system with no island present. Following Williams et al.,[26] in going from Eq. (2) to Eq. (3) we have suppressed the weak logarithmic dependence of the $\nu^{1/3}\log\nu$ term expressing the elastic relaxation energy due to the lines of force at the edges. More specifically the logarithm is eliminated by expanding it in a Taylor series about a value ν_{X0} near the peak of the corresponding island type. The original coefficients are re-expressed in terms of A_X, B_X, C_X and D_X, where $C_X = D'_X + C'_X\,(3 + \log\nu_{X0})$ and

$D_X = -3C'_X \nu_{X0}^{1/3}$.

The pair interaction energy is more complicated, and we need to make an approximation in order to develop a workable model. It is clearly a function of several properties of the two interacting islands, $\varepsilon^{(2)}_{a,b}(r_{ab}, \nu_a, \nu_b, \ldots)$. At the very least it depends on the number of atoms in each island and the separation between them, r_{ab}, as well as materials properties such as the elastic constants and the intrinsic mismatch strain. To a lesser extent it depends on the types of islands, their orientation and any internal degrees of freedom such as the degree of Ge-Si alloying,[27,28] the length of a hut or the relaxation of a superdome, if present.

The island-island interactions will be treated in mean field, reducing the two-body term to a coverage-dependent single-body potential energy. The coverage dependence will be treated self-consistently. Specifically, we define

$$\bar{\varepsilon}^{(2)}_a = \frac{1}{2}\sum_{b\neq a} \varepsilon^{(2)}_{a,b} \tag{4}$$

$$\approx \lambda\, \epsilon_0^2\, \nu_a^\beta\, \bar{\nu}^\gamma \theta^\delta \tag{5}$$

where the second line is the mean-field approximation: λ is a combination of elastic constants and geometrical factors, ϵ_0 is the intrinsic mismatch strain, $\bar{\nu}$ is the mean island size and θ is the coverage in equivalent monolayers. The parameters β, γ and δ are exponents.

We have assumed that the elementary two-body potential is isotropic and blind to the types of surrounding islands, which is justified at the low-to-moderate coverages we will consider where the ordering does not change with θ. We will take $\beta = \gamma = \frac{2}{3}$ following Shchukin et al.[11] who argued that the interaction should scale as the footprint of the island, which grows with $\nu^{2/3}$. Further, Shchukin et al.[11] have argued that the leading term in

the elastic interaction between islands is of a dipole-dipole form, and that the coverage dependence of the energy should go like $\theta^{3/2}$, where θ is the coverage. Medeiros-Ribeiro et al.[17] have found to the contrary that the energy is linear in the coverage ($E \sim E_0 + E_1\theta$). The coefficient λ in principle depends on the elastic constants, the radial distribution function of the islands, and if the system is ordered, the angular distribution function. It could be computed self-consistently based on a statistical model, but for the purposes of this article it is a fit parameter. In practical terms, it is useful to rewrite the mean-field island-island interactions in terms of a single fit parameter, α:

$$\bar{\varepsilon}_i^{(2)} \approx \alpha\, \nu_i^{2/3}\, \theta \tag{6}$$

where $\alpha = \lambda\, \epsilon_0^2\, \bar{\nu}^{\gamma}$, and the total elastic energy for an island is given by

$$E^I(n;\theta) = A_I \nu + (B_I + \alpha_I\, \theta)\nu^{2/3} + C_I \nu^{1/3} + D_I. \tag{7}$$

Thus in the mean-field approximation the energy of an island is determined by its type I, the number of atoms in the island, ν, and the coverage, θ, and the island system reduces to a single particle model. At fixed coverage the islands obey the usual statistical mechanics of a system of molecules that can exchange atoms; for instance, hydrocarbon molecules exchanging dimers. Of course, the specific island energetics are quite different, but the thermodynamic considerations are similar.

Now consider a system of islands in equilibrium on a surface. The system consists of N atoms distributed among M islands containing $\nu_1, \cdots, \nu_M$ atoms such that there are n_ν islands of size ν. Our goal is to compute the number of islands n_ν of a given size ν. The

configurational partition function for the system is given by

$$Z(n_1, n_2, \cdots) = \int d^{2N}x\, \Delta(\vec{x}_a; \nu)\, S_N\, e^{-\beta E} \tag{8}$$

$$= C \prod_{\nu} \left(\frac{A^{n_\nu}}{n_\nu!} \right) e^{-\beta E^{(1)}(\nu)} \tag{9}$$

where E is the total internal energy of system and $\vec{x}_a$, $a = 1, \ldots, M$ are the locations of the islands. As usual $\beta = 1/(kT)$ and A is the area of the system. The function $\Delta(\vec{x}_a; \nu)$ is a product of delta functions specifying that the atoms are in specific islands and S_N is a symmetry factor to eliminate double-counting of identical particles. The second line is obtained from the first by integrating over the spatial coordinates. Since the mean-field energy is independent of the positions of the islands, the result is one factor of area for each island, divided by $n_\nu!$ to account for the symmetry of identical islands. The prefactor C is a function of temperature that contains factors that are independent of the island distribution, and hence irrelevant to the analysis that follows. By focusing on the configurational partition function we have suppressed the kinetic part of the partition function, which gives a negligible contribution to the free energy. The energy $E^{(1)}(\nu)$ is the mean-field energy for a single island of size ν given in Eq. (7). We have written the partition function for a single type of island for notational simplicity. This expression is generalized to the case of multiple island types below.

The configurational partition function given above (8) assumes that the sizes of the islands are known. In reality during growth the island sizes fluctuate, as indeed does the total number of atoms in the islands, N, as atoms move to and from the wetting layer. A better thermodynamic variable is the chemical potential. We introduce it as a conjugate

variable to $N = \sum_\nu \nu n_\nu$ in the grand partition function,

$$
\begin{aligned}
Z &= \sum_{n_1,n_2,\cdots} Z(n_1,n_2,\cdots) e^{\beta\mu N} && (10)\\
&= C \sum_{n_\nu} \prod_\nu \left(\frac{A^{n_\nu}}{n_\nu!} \right) e^{-\beta \sum_\nu \left[n_\nu E^{(1)}(\nu) - \mu\nu n_\nu \right]} && (11)\\
&= C \exp\left\{ A\, e^{-\beta \sum_\nu \left[E^{(1)}(\nu) - \mu\nu \right]} \right\} && (12)
\end{aligned}
$$

The chemical potential is chosen to give the appropriate amount of material in the islands, i.e. the correct equivalent coverage. The coverage is related to the expectation value of the island size distribution by

$$
\begin{aligned}
\theta &= \langle N/A \rangle && (13)\\
&= \sum_\nu \nu \langle n_\nu \rangle / A. && (14)
\end{aligned}
$$

We now derive an expression for the expectation value of the island distribution. It is given by the weighted sum

$$
\begin{aligned}
\langle n_\nu \rangle &= Z^{-1} \sum_{n_1,n_2,\cdots} n_\nu\, Z(n_1,n_2,\cdots) e^{\beta\mu N} && (15)\\
&= -\beta^{-1} \partial_{E^{(1)}(\nu)} \log Z && (16)\\
&= A\, e^{-\beta \left[E^{(1)}(\nu) - \mu\nu \right]} && (17)
\end{aligned}
$$

This expression is a very important part of the model. We have used a mathematical trick to derive it. The formal derivative with respect to $E^{(1)}(\nu)$ brings down a factor of n_ν from the argument of the exponential, which gives exactly the weighted average we need for the coverage. Since the partition function Z is an exponential (12) , we can evaluate its logarithm and we arrive at an exact expression for the island distribution at a given

chemical potential μ. The generalization of Eq. (17) to the case of more than one type of islands is

$$\langle n_\nu^I \rangle = A\, e^{-\beta\left[E_I^{(1)}(\nu)-\mu\nu\right]} \tag{18}$$

where I labels the type of island.

The island distribution gives us the expression we need to compute the coverage from Eq. (14),

$$\theta = \sum_{\nu,I} \nu\, e^{-\beta\left[E_I^{(1)}(\nu;\theta)-\mu\nu\right]} \tag{19}$$

where again we have restored the dependence on island type, summing over the contribution of the different types to the coverage. Since the mean-field island energy $E_I^{(1)}(\nu;\theta)$ depends on the coverage, this equation must be solved self-consistently in order to determine the chemical potential. The equation is non-linear, so in practice it must be solved numerically through iteration until a converged result is obtained.

It is also possible to derive expressions for fluctuations in the island distributions. For example, the variance of the distribution is given by

$$\begin{aligned}
\langle \delta(n_\nu^I)^2 \rangle &= \langle (n_\nu^I)^2 \rangle - \langle n_\nu^I \rangle^2 && (20)\\
&= \beta^{-2} \partial^2_{E_I^{(1)}(\nu)} \log Z && (21)\\
&= A\, e^{-\beta\left[E_I^{(1)}(\nu)-\mu\nu\right]} && (22)\\
&= \langle n_\nu^I \rangle && (23)
\end{aligned}$$

This expression relates the variance at fixed chemical potential to the island distribution itself.

We end this section calculating the free energy of the ν-atom islands, F_ν:

$$\begin{aligned} F_\nu &= n_\nu E^{(1)}(\nu) - T\, S_\nu && (24)\\ &= n_\nu E^{(1)}(\nu) + kT\, \log n_\nu! && (25)\\ &\approx n_\nu \left[E^{(1)}(\nu) + kT\, \log n_\nu\right] && (26)\\ &= n_\nu\, \nu\, \mu && (27) \end{aligned}$$

where S_ν is the contribution to the entropy coming from the ν-atom islands (and we have suppressed a term in the entropy of the form $\log \sum_\nu n_\nu!$ that is very weakly dependent on n_ν). We have used the leading term in Stirling's approximation to derive the third line, and inserted the island distribution function (18) to obtain the fourth. The free energy per atom of the ν-atom islands is then

$$F_\nu/(\nu n_\nu) = \mu. \quad (28)$$

Thus the entropic contribution has balanced the inherent differences in the energy per atom of the islands so that they are at the same chemical potential in equilibrium.

3 Solving the model analytically

The statistical mechanics framework for the model was detailed in the previous section, but there are still a few steps to reach the solution. The parameters must be determined from experiment, and the chemical potential needs to be computed from equation (19). Ideally, it is desirable to have the parameters of the model determined from first principles, or at least from experiments that are quite different in nature from those to be analyzed. At this

point we do not have a means to derive the parameters from independent experiments and calculations, so we must resort to fitting the model parameters to a subset of the island distributions measured by AFM. We discuss the procedure in section 4. First, however, we need to be able to compare the model with experiment, and for that we need an expression for the chemical potential. It turns out that the chemical potential (and hence the solution of the model) can be expressed analytically to a great extent.

The chemical potential is determined by the condition for self-consistency (19)

$$\tilde{\theta} = \langle \sum_{\nu,I} \nu n_\nu^I \rangle = \sum_{\nu,I} \nu e^{-\beta\,[E^I(\nu;\theta)-\mu\nu]} \tag{29}$$

where we have introduced the notation $\tilde{\theta}$ in recognition of the fact that these terms do not sum to θ unless the correct value for μ is chosen. The correct μ leads to self-consistency: $\tilde{\theta} = \theta$. This sum could be computed numerically–indeed, we have written a computer code to do just that–but it is possible to produce high quality analytic approximations to the sum as well. The analytic results have the advantage that they are much faster to compute. In principle, the dependence on parameters is more evident in the analytic expression, although the robust expressions are sufficiently complicated to offer little advantage in this regard.

The basic approach is to observe that the sum in the self-consistency equation (19) may be approximated by an integral, and the integral may be computed in the saddle point approximation. In the first step the sums for N and M^I are approximated by the integrals

$$N \approx \sum_I \int_0^\infty d\nu\, \nu\, e^{-\beta[E^I(\nu)-\mu\,\nu]} \tag{30}$$

$$M^I \approx \int_0^\infty d\nu\, e^{-\beta[E^I(\nu)-\mu\,\nu]} \tag{31}$$

In practice, the island number distribution is more peaked than the island volume distribution, so equation (31) is better behaved than equation (30).

Accurate analytic formulas can be derived using saddle point approximations, as is shown in detail in the next section in the discussion leading to Eq. (73). The results are to leading order (typically 0.1 - 1% accuracy):

$$\tilde{\theta}(\theta,\mu,T) \approx \sqrt{\frac{\pi}{B'}}\, e^{-D''} \left(\frac{15}{4}\frac{x_*}{B'^2} + 5\frac{x_*^3}{B'} + x_*^5 \right) \tag{32}$$

$$M(\theta,\mu,T) \approx \sqrt{\frac{\pi}{B'}}\, e^{-D''} \left(\frac{1}{2\,B'} + x_*^2 \right) \tag{33}$$

$$\partial_\mu\tilde{\theta}(\theta,\mu,T) \approx \sqrt{\frac{\pi}{B'}}\, e^{-D''} \left(\frac{105}{16\,B'^4} + \frac{105\,x_*^2}{2\,B'^3} + \frac{105\,x_*^4}{2\,B'^2} + 14\frac{x_*^6}{B'} + x_*^8 \right) \tag{34}$$

where B', D'' and x_* are defined in Eqs. (52), (59) and (57), respectively in the following section. Their values differ in the 3 equations. We include the expression for $\partial_\mu\tilde{\theta}$ for reference, but we do not actually make use of it. At high temperatures the peaks broaden and the accuracy of the analytic expressions degrades, and we must resort to numerical integration. The chemical potential is determined by solving the non-linear equation $\tilde{\theta}(\theta,\mu,T) = \theta$ using Eq. (32) for $\tilde{\theta}$. In practice the Brent method[29] was used to find the root on a regular grid of points covering the region of interest in θ-T space, and starting the iteration with the closest previous solution in all but the first case.

3.1 Derivation of the analytic sums

In this section we present the derivation of the analytic expressions approximating the sums in the formulas for the coverage and the number of islands. Readers should feel free to skip to the next section if they are not interested in the mathematical derivation. The

strategy is to use a saddle-point approximation to integrate the peaked distributions. The difficulty is that the peaks are rather broad and asymmetric, so a conventional saddle-point approximation is not very accurate. A careful series of approximations is needed to produce an accurate formula. In particular, we need to evaluate integrals of the form:

$$I(A,B,C,D;p) = \int_0^\infty dx\, x^p\, e^{-(Ax^3+Bx^2+Cx+D)} \tag{35}$$

where $A > 0$ and $f(x) = Ax^3+Bx^2+Cx+D$ has two real local extrema, so that the island size distribution will have an equilibrium peak. This kind of integral appears in Eqs. (30), (31) and (34) with $x = \nu^{1/3}$, where $p = 5, 2$ and 8, respectively. The latter condition is equivalent to positivity of the discriminant of $f'(x)$: $(2B)^2 - 4(3C)A > 0$; i.e. $B^2 > 3AC$.

We use a saddle point approximation for Eq. (35) as follows:

$$I(A,B,C,D;p) = \int_0^\infty dx\, e^{-(Ax^3+Bx^2+Cx+D)+p\,\log x} \tag{36}$$

so the saddle point is given by

$$f'(x) = \left.\frac{p}{x}\right|_{x=x_{sp}} \tag{37}$$

which comes from the derivative of the argument of the exponential. Upon expansion, we have

$$3Ax_{sp}^3 + 2Bx_{sp}^2 + Cx_{sp} - p = 0 \tag{38}$$

which is a cubic equation and therefore has a solution in closed form for its three roots.

The most positive, real root is given by

$$x_{sp} = \frac{-2\,B}{9\,A} + \frac{\left(1 - i\,\sqrt{3}\right)\,\left(-4\,B^2 + 9\,A\,C\right)}{9\,2^{\frac{2}{3}}\,A\left(-16\,B^3 + 54\,A\,B\,C + 243\,A^2\,p + \chi\right)^{\frac{1}{3}}} - \tag{39}$$

$$\frac{\left(1 + i\,\sqrt{3}\right)\,\left(-16\,B^3 + 54\,A\,B\,C + 243\,A^2\,p + \chi\right)^{\frac{1}{3}}}{18\,2^{\frac{1}{3}}\,A} \tag{40}$$

$$\chi = \sqrt{4\left(-4\,B^2 + 9\,A\,C\right)^3 + \left(-16\,B^3 + 54\,A\,B\,C + 243\,A^2\,p\right)^2} \tag{41}$$

which may be written more compactly and avoiding complex numbers in the form:

$$x_{sp} = \frac{-2\,B}{9\,A} + \frac{2\,\sqrt{\Delta}}{9\,A}\cos\left[\tfrac{1}{3}\cos^{-1}\left(1 - \xi\right) + \phi\right] \tag{42}$$

$$\Delta = 4\,B^2 - 9\,A\,C \tag{43}$$

$$\xi = 1 - \left(\tfrac{1}{2}\left|-16\,B^3 + 54\,A\,B\,C + 243\,A^2\,p\right|/\Delta^{3/2}\right) \tag{44}$$

$$\phi = 0 \tag{45}$$

where $\phi = \pm 2\pi/3$ gives the other two roots, but $\phi = 0$ is the most positive root, and hence the saddle point of interest. Because A is small, ξ is small, and the following expansion converges very rapidly:

$$\cos\left[\tfrac{1}{3}\cos^{-1}\left(1 - \xi\right)\right] = 1 - \frac{1}{3^2}\xi - \frac{4}{3^5}\xi^2 - \frac{28}{3^8}\xi^3 - \frac{80}{3^{10}}\xi^4 - \cdots . \tag{46}$$

The converged result is found even more quickly through iteration of the following formula:

$$y_{i+1} = y_i + \frac{\varepsilon_i}{3 - 12y_i^2} \tag{47}$$

$$\varepsilon_i = \xi - 1 - 3y_i + 4y_i^2 \tag{48}$$

where ε_i is the error: $\varepsilon_i = \cos\left[\tfrac{1}{3}\cos^{-1}\left(1 - \xi\right)\right] - y_i$. Use of this formula avoids the numerically phlegmatic trig functions, achieving an increase in speed by a factor of 10

million. The formula

$$\sqrt{\Delta} = 2|B| \left[1 - \frac{1}{2} \frac{9\,A\,C}{4\,B^2} - \frac{1}{8} \left(\frac{9\,A\,C}{4\,B^2} \right)^2 - \frac{1}{16} \left(\frac{9\,A\,C}{4\,B^2} \right)^3 - \frac{5}{128} \left(\frac{9\,A\,C}{4\,B^2} \right)^4 + \cdots \right] \quad (49)$$

offers similar advantages.

We now expand $f(x)$ about x_{sp}:

$$\begin{aligned} f(x) &= A[(x - x_{sp}) + x_{sp}]^3 + B[(x - x_{sp}) + x_{sp}]^2 + C[(x - x_{sp}) + x_{sp}] + D \quad (50) \\ &= A(x - x_{sp})^3 + B'\,(x - x_{sp})^2 + C'\,(x - x_{sp}) + D' \quad (51) \end{aligned}$$

where the coefficients are given by

$$\begin{aligned} B' &= 3Ax_{sp} + B \quad (52) \\ C' &= 3Ax_{sp}^2 + 2Bx_{sp} + C \quad (53) \\ D' &= Ax_{sp}^3 + Bx_{sp}^2 + Cx_{sp} + D. \quad (54) \end{aligned}$$

We will treat the term proportional to $(x - x_{sp})^3$ as small.

We take as the first approximation

$$f(x) \approx \tilde{f}(x) \equiv B'\,(x - x_{sp})^2 + C'\,(x - x_{sp}) + D' \quad (55)$$

and rewrite $\tilde{f}(x)$ in the form

$$\tilde{f}(x) = B'(x - x_*)^2 + D'' \quad (56)$$

where x_* and D'' are given by

$$x_* = x_{sp} + \delta \tag{57}$$

$$D'' = \tilde{f}(x_*) = B'\,\delta^2 + C'\,\delta + D' \tag{58}$$

$$= -\frac{C'^2}{4B'} + D' \tag{59}$$

$$\delta = x_* - x_{sp} = -\frac{C'}{2B'}. \tag{60}$$

So we have the saddle point approximation

$$I(A,B,C,D;p) \approx \int_0^\infty dx\, x^p\, e^{-(B'(x-x_*)^2+D'')} \tag{61}$$

$$\approx \int_{-\infty}^\infty dx\, x^p\, e^{-(B'(x-x_*)^2+D'')} \tag{62}$$

$$= \int_{-\infty}^\infty dy\, (y+x_*)^p\, e^{-(B'y^2+D'')} \tag{63}$$

where we have made the substitution $y = x - x_*$.

The final expression is a Gaussian integral which may be evaluated immediately using

$$\int_{-\infty}^\infty dy\, y^m\, e^{-\alpha y^2} = \alpha^{-(m+1)/2}\Gamma\left(\frac{m+1}{2}\right) = (2\alpha)^{-(m+1)/2}\,\sqrt{2\pi}\,(m-1)!! \tag{64}$$

so that

$$I(A,B,C,D;p) \approx 2\sqrt{\pi}\,e^{-D''} \sum_{\mathrm{even}m=0}^{p} \frac{p!}{(m/2)!\,(p-m)!}\,(4B')^{-(m+1)/2}\,x_*^{p-m} \tag{65}$$

$$= \sqrt{\frac{\pi}{B'}}\,e^{-D''}\;\partial_k^p \exp\left(\frac{k^2}{4B'} + k\,x_*\right)\Big|_{k=0} \tag{66}$$

where we have used $m!/(m-1)!! = 2^{m/2}(m/2)!$. This is the leading order in an asymptotic expansion of $I(A,B,C,D;p)$. It is typically accurate to a few percent or better.

The polynomials that appear are the elementary Schur polynomials, related to certain homogeneous polynomials that arise in the theory of symmetric functions (where for example they express homogeneous product sums in terms of power sums)[30] defined according to

$$\exp(\sum_{i=1}^{\infty} x_i\, z^i) = \sum_{n=0}^{\infty} P_n(x_1, \cdots, x_n) z^n \tag{67}$$

where P_n is the n$^{\text{th}}$ polynomial. The first few elementary Schur polynomials are:

$$P_0 = 1 \tag{68}$$

$$P_1 = x_1 \tag{69}$$

$$P_2 = x_2 + \frac{1}{2}x_1^2 \tag{70}$$

$$P_3 = x_3 + x_1 x_2 + \frac{1}{6}x_1^3 \tag{71}$$

$$P_4 = x_4 + x_1 x_3 + \frac{1}{2}x_2^2 + \frac{1}{2}x_1^2 x_2 + \frac{1}{24}x_1^4 \tag{72}$$

In terms of the Schur polynomials we have

$$I(A, B, C, D; p) \approx \sqrt{\frac{\pi}{B'}}\, e^{-D''}\, p!\, P_p(x_*, [4B']^{-1}, 0, \ldots, 0) \tag{73}$$

This is the leading saddle point approximation. The explicit formulas are given above in Eqs. (32) – (34).

There are additional refinements of this approximation that can be made. In practice, the first order approximation works well for most of the temperature/coverage range of interest. The errors are greatest at high temperatures beyond the yellow curve in Fig. 2 where the local minimum in the island distribution function is not sufficiently close to zero. It these cases we must impose that the integrals are cut off at the island size corresponding

to the local minimum, and then we compute those integrals numerically. Since this is only a small fraction of the total evaluations, the analytic formulas still lead to a large improvement in speed.

4 Parameterization

Next we obtain parameters for the model and compare the predictions of the model to experimental data. The test is the island distributions themselves. We do not yet have a theoretical calculation of the Shchukin energy parameters, so we determine them through a best fit of the model to the island distributions.

The basic experimental data are in the form of the number of islands observed on a fixed region of the substrate surface that have sizes falling within a certain bin; i.e. there may be 200 islands on a 9 μm^3 region of the surface with sizes between 7500 and 8500 nm^3. For the fit we need to know both the data values and their uncertainty. There are two sources of error in these data. First, the island size may not be measured accurately, due both to noise in the measurements resulting from the resolution of the AFM, and to systematic errors such as the effect of the finite radius of the AFM tip. These errors tend to broaden (through Gaussian convolution) and shift the distribution, respectively, at the level of a few percent. The resolution of the AFM is to a large extent determined by the size of the pixel: 5 nm $\times$ 5 nm. The vertical accuracy is better than 1% and may be neglected apart from discretization error. Thus, a typical pyramid with a 25 nm $\times$ 25 nm base may have a 40% error in its measured volume; a typical dome with a 60 nm $\times$ 60 nm base may

have a 15% error. Second, there is statistical noise due to the finite number of counts per bin. This noise results in a binomial distribution of the data, which is characterized by the variance:

$$\sigma^2_{\text{Binomial}} = \bar{m}_i \left(1 - \frac{\bar{m}_i}{N_{\text{count}}}\right), \tag{74}$$

where $\bar{m}_i$ is the expected number of counts in bin i, and N_{count} is the total number of counts. When $\bar{m}_i \ll N_{\text{count}}$, the distribution becomes a Poisson distribution.

The goal is to fit the data with our model distribution, and to do this in a way that will give us a measure of the reliability of the model. Since we have a model of the expected error–basically characterized by the variance (74)–this can be done as follows.

The figure of merit is given by the weighted square error:

$$\chi^2 = \sum_i \frac{\left(f(n_i) - m_i\right)^2}{\sigma_i^2} \tag{75}$$

where $f(n)$ is the model island distribution, n_i is the size at the center of bin i, and m_i is the measured number of islands in bin i. This function is minimized with respect to the parameters of the model in order to produce what is known as the best chi-squared fit. The effect of the weighting is to place more emphasis on the peaks and less on the tails.

The model distribution is given by (18)

$$\langle n^I_\nu / A \rangle \quad = \quad e^{-\left(a\nu + b\nu^{2/3} + c\nu^{1/3} + d\right)} \tag{76}$$

where the parameters are related to the Shchukin energy parameters expressed in units of kT: $a = A_I - \mu$, $b = B_I + \epsilon_I\,\theta$, $c = C_I$ and $d = D_I$.

As written, the chi-squared fit implied by Eqs. (75) and (76) is a non-linear minimization problem. We have used a linearized least square technique[32] to obtain the best fit, and

Table 1: Parameters for the single-island (3) and mean-field, island-pair (6) contributions to the energy (1) as a function of island type and size. For comparison, the elastic strain energy in unrelaxed epitaxial Ge/Si would be 0.038 eV/atom. The value of the parameters A may be changed by an arbitrary constant provided the chemical potential is changed by the same constant. Regardless of the number of digits given none of the parameters is reliable to more than two significant figures.

Parameter	Pyramids	Domes	Units
A	-9.60×10^{-6}	-5.137×10^{-6}	eV/atom
B	9.0286×10^{-4}	1.0834×10^{-3}	eV/atom$^{2/3}$
C	-0.04517	-0.1396	eV/atom$^{1/3}$
D	0.1422	4.1625	eV
α	2.305×10^{-5}	1.974×10^{-5}	eV/atom$^{2/3}$/ML

then checked it in a few cases with a full non-linear chi-squared minimization starting from the linearized fit values. The parameters were deduced from χ^2 fitting to data from tens of thousands of islands grown at 600 °C with coverages $\theta = 2.2$, 4.8 and 9.1 ML. Although we have up to four parameters for each type of island and one more for the mean-field interaction, the parameters do provide a good fit to this much larger data set, consisting of over 100 data points and many thousand islands. The results are shown in Table 1.

The linearized chi-squared fitting is very fast, and it has been our experience in the island systems that the solution gives a value of χ^2 that is reasonable; i.e. the fit is good and the Binomial distribution of errors is accurate. To further quantify the accuracy of the

fit, we calculate the probability Q that this or a worse value of χ^2 should occur by chance

$$Q = \Gamma([N_{\rm bin} - N_{\rm params}]/2, \chi^2/2)/\Gamma([N_{\rm bin} - N_{\rm params}]/2] \tag{77}$$

where $\Gamma(n, x)$ is the incomplete Gamma function. We have typically found Q ranges from 0.1 to 0.8 in the linearized fitting procedure. These are good fits. The high coverage distributions typically were not fit as well as the low coverage.

4.1 The physical basis for the parameter values

We now consider whether the fit values of the parameters are physically sensible. It is the energy with respect to a surface without islands that enters the formula for the island distributions. Thus, for example, it is not the bulk elastic energy, which would be positive, but the amount the bulk energy is relaxed by the formation of the islands that enters the formula. To the best of our knowledge, there has been no independent measurement of the individual terms that enter the energy, so we can only use general considerations to guide the decision about whether the size of a particular parameter is physically reasonable or not.

Consider first the size of the surface energy parameters. The B parameters are relative surface energies, and for pyramids are dominated by the difference between the (001) and (105) surface energies. We are not aware of a calculation or measurement of the Ge(105) surface energy, but some information is available. Ab initio calculations estimate that the RS reconstruction of the Ge(105) surface is lower in energy than the PD-reconstructed surface by 4.3 meV/Å^2 [33] or 6.2 meV/Å^2,[34] and empirical potential calculations estimate

that the formation energy for an island is about 0.01 eV/atom. The energy difference between the (111) and (001) reconstructed Ge surfaces is calculated to be 0.01 J/m^2.[35] The quantity relevant to our model, the B parameter, is the difference between the (001) and (105) surface energies suitably weighted by geometrical factors. In particular, the formula for a pyramid is

$$B_p = \left(\frac{\gamma_{facet}}{\hat{n}_{facet} \cdot \hat{n}_{wetting}} - \gamma_{wetting} \right) v_0^{2/3} \frac{18^{1/3}}{\left| \frac{\hat{n}_{facet}}{\hat{n}_{facet} \cdot \hat{n}_{wetting}} - \hat{n}_{wetting} \right|^{2/3}} \tag{78}$$

$$= 3.83 \left(1.0198 \gamma_{(105)} - \gamma_{(001)} \right) \quad \mathrm{eV/atom}^{2/3} \tag{79}$$

where γ_{facet} and $\gamma_{wetting}$ are the surface energies of the facet and wetting layer, respectively, $\hat{n}_{facet}$ and $\hat{n}_{wetting}$ are the corresponding unit normal vectors and v_0 is the atomic volume. On the second line (79) we have shown the values of the parameters for a (105)-faceted Ge pyramid, where the surface energies $\gamma_{(105)}$ and $\gamma_{(001)}$ are expressed in J/m^2. In the case of domes, the relevant facet energies are $\gamma_{(113)}$ and $\gamma_{(15\,3\,23)}$, and the coefficient multiplying the facet energy is larger, roughly 1.2 instead of 1.0198. The surface energies should be appropriately renormalized for strain[36] and alloying.[37] Neglecting these effects for the moment, Stekolnikov et al.[35] give the Ge $\gamma_{(001)}$ as 1.00 J/m^2. Our model value of $B_p = 9.03 \times 10^{-4}$ eV/atom$^{2/3}$ then corresponds to a $\gamma_{(105)}$ of 0.98 J/m^2. We have not found a value for $\gamma_{(105)}$ in the literature, even for pure, unstrained Ge, but to get an estimate of the range of reasonable values, its difference from $\gamma_{(001)}$ is expected to be significantly less than the energy difference between the (110) and (001) surfaces (cf. for silicon Eaglesham et al.[38] and for germanium Stekolnikov et al.[35]), i.e. about 0.17 J/m^2. The magnitude of the difference in our case is 0.02 J/m^2 which appears to be reasonable.

There are two points about B_p that deserve more comment, however. First, the fact that $\gamma_{(105)}$ must be less than $\gamma_{(001)}$ is somewhat surprising. For pure Ge, it would be unexpected, but taking the strain renormalization of the surface energy[36] and the alloying [37] into account, it it possible. For example, an alloy concentration gradient at the surface of an island contributes to the the B parameter.[37] Second, the magnitude of the splitting of the $\gamma_{(105)}$ and $\gamma_{(001)}$ must be very precise. For example, taking $\gamma_{(105)} = \gamma_{(001)}$ in Eq. (79) gives a value of B_p 84 times larger than our value. Such fine tuning of a parameter is not common, and it suggests that the alloying may be evolving in a way to promote island formation. The ability of alloying to destabilize island growth has been emphasized previously.[39] A careful calculation of the energetics of alloying is needed, but it is beyond the scope of this article. As an estimate of whether it can be important, we can consider the free energy of mixing

$$\Delta\gamma = -n_{layer}\, kT\, \left[c_{Ge} \log c_{Ge} + (1 - c_{Ge}) \log(1 - c_{Ge})\right] / a_0 \tag{80}$$

where c_{Ge} is the number fraction of Ge in the alloy.[40] The factor n_{layer} is the thickness in atomic monolayers of the alloyed region at the surface and a_0 is the atomic area. This function is maximized at $c_{Ge} = 1/2$, although we note that a more precise calculation including the Ge-Si interaction energies would have the maximum shifted from this value. Using $c_{Ge} = 1/2$, $T = 600$ °C, $n_{layer} = 5$ and $a_0 = 0.16$ nm^2, we find $\Delta\gamma = -0.26$ J/m^2. This rough estimate indicates that the free energy of mixing can affect the B parameter, and its effect is to lower the facet energies (as needed to account for the value of B). The sensitivity of the shape and size of strained Ge islands on Si(001) to the surface energy

is seen experimentally when P is co-deposited with Ge to form a surfactant layer.[41–43] There are still two different metastable island shapes, but they are qualitatively different from pyramids and domes (in that they display completely different facets) and they are significantly smaller in size as well.

There is substantially more information available about the bulk term in the island energy (the term proportional to ν). It corresponds to the drop in the bulk elastic energy of the island compared to that of the wetting layer. The strain energy per atom in the wetting layer is given by $U_0 = [(C_{11} + C_{12}) - 2(C_{12}^2/C_{11})]\ \epsilon_0^2\, v_0 = 0.038$ eV/atom where v_0 is the atomic volume and C_{ij} are the cubic elastic constants. The values of the A_I represent this relaxation energy with respect to an arbitrary zero of the chemical potential. The values of the A_I are much smaller than 38 meV/atom, but because of the arbitrary zero, it is only their difference that matters. The difference of roughly 1/8000 of the wetting layer energy is less than we expected, but again alloying of Si in the Ge islands may account for the small difference. There are some relevant data from synchrotron experiments that measure differences in the strain,[44,45] and they imply a larger difference in the bulk strain energies.

Next consider now the term that scales like $\nu^{1/3}$. It contains a contribution that scales with the square of the surface stress.[11] There are also several other contributions. The one that has been discussed most is the elastic relaxation edges that scales like $\nu^{1/3}\log(\nu)$. [11] The elastic relaxation lowers the energy. Another effect is the formation of a trench or moat seen around islands.[4,45]. This effect has not been modeled, but if it is energetically driven, then it too would give a contribution that scales like $\nu^{1/3}$ and reduces the energy.

The rounding of edges on the pyramid also gives a contribution that scales like $\nu^{1/3}$ and reduces the energy. The size of these effects has not been calculated in the literature, to the best of our knowledge, and it is beyond the scope of this article to do so.

The Ge/Si system is sufficiently complex that the parameters are not tightly constrained. In the fitting process, it was possible to obtain a comparably good fit with a set of parameters in which the signs of the parameters A, B and C were reversed. As our ability to model the complex growth processes advances, we should be able to pin down the values better.

5 The nanostructure diagram

The principal result of the model is the calculation of the relative populations of domes and pyramids, their sizes and the widths of their distributions. It is convenient to summarize those results in a map in temperature-coverage space that we have called the nanostructure diagram. It is shown in Fig. 2.

The nanostructure diagram contains a wealth of information. The coloring indicates the number fraction of nanostructures that are pyramids. The distribution could equally well be shown as the volume fraction. At relatively low temperatures and coverages pyramids dominate, as is well known from experiment; at high temperatures and coverages the domes are dominant. During growth at constant temperature there is a transition from pyramids to domes. The contour lines indicate the widths of the distributions relative to the mean value. Narrow widths are preferable for applications, and the dome distributions

are narrower than the pyramid distributions throughout the diagram. The narrow peak is a characteristic of the larger islands in bimodal distributions. There are few domes that are as small as pyramids, so their size distribution is well separated from zero leading to a narrower peak.

For sufficiently large islands, the term proportional to the island volume dominates and causes an indefinite growth in the island size. The temperatures and coverages above which this leads to a breakdown of the model are indicated by the yellow curve. In our model the competition between the different contributions to the energy leads to a local minimum in the energy separated by a local maximum from the monotonically decreasing range at large ν. The magnitude of the dome distribution is not zero at the local maximum, which corresponds to a critical size for unstable ripening. In classical nucleation theory, [46] the rate of nucleation depends on the population of embryos, clusters of barely sub-critical size, together with the rate at which an embryo fluctuates to attain the critical nucleus size. Assuming that the fluctuation rate is not a strong function of temperature and coverage, within our model the population of embryos can be calculated as an estimate of the importance of kinetic effects. We note that a moment analysis may also be used for this purpose.[18] We have plotted a yellow contour line where the ratio of the dome distribution at the critical size relative to the peak of the distribution is equal to 10^{-5} to indicate that the description of the system in the high temperature region is incomplete and should be supplemented with a robust kinetic model.

6 Chemical potential

The vast majority of the effort in solving the model involves the computation of the chemical potential $\mu(T,\theta)$ for the range of temperature and coverage of interest. Once a table of values for $\mu(T,\theta)$ is constructed, it is straightforward to compute all of the quantities comprising the nanostructure diagram as well as even more detailed information such as the size distributions. For example, the equation for the full island distribution (18) involves only elementary functions once the chemical potential has been calculated:

$$\langle n_\nu^I/A\rangle \;=\; e^{-\beta\left[E_I^{(1)}(\nu)-\mu(T,\theta)\nu\right]}. \tag{81}$$

Thus, the full physical content of the model is encoded in a single function $\mu(T,\theta)$. The values of the chemical potential over a range of temperatures and coverages are plotted in Fig. 3. The chemical potential varies to a greater extent over the range of coverages than the range of temperatures. In general, the chemical potential of an island is also a function of its composition (e.g. the mole fraction of Ge in the island), since some alloying takes place even at low growth temperatures. This composition dependence has been suppressed in Fig. 3, which represents the chemical potential of islands that follow a particular kinetic trajectory of island composition determined by the growth conditions, such as Ge arrival flux, etc.

We have found an approximate function for the chemical potential over the range

450 °C $< T <$ 950 °C and $\theta <$ 14 ML. The approximate expression is

$$\begin{aligned}\mu(T,\theta) &= 0.587 - 6.15\,e^{-T/100} - 1.36\times10^{-4}\,T - 0.0305\,\theta \\ &\quad + \left(6.07\,e^{-T/75} + 5.87\times10^{-6}\,T\right)\,\theta + e^{-\theta/2.9}\left(0.108 + 739\,e^{-T/75}\right)\end{aligned} \tag{82}$$

The formula is given in units A_p, and T is the absolute temperature in Kelvin and θ is the equivalent coverage in monolayers. The complicated functional dependence on temperature was introduced in order to capture the behavior of the chemical potential in the non-linear regime at higher temperatures (up to 950 °C) than those shown in Fig. 3. This expression is accurate to better than 5% throughout the range 450 °C $\leq T \leq$ 950 °C and $2 \leq \theta \leq 16$ ML, and is typically accurate to better than 1%. We emphasize, however, that that the form of this approximation is ad hoc, not based on physics. The factors in the exponentials have not been optimized. The parameters in this approximation would change if the energy parameters were changed; this approximation to the resulting $\mu(T,\theta)$ does not introduce any additional parameters. It just summarizes the result.

7 Island size distributions

Once the chemical potential is determined, the island size distributions immediately follow from Eq. (18). The results are shown in Fig. 4. The figure shows the island distribution functions for the pyramids, the domes and both combined. The experimental data are plotted, with error bars, and it may be observed that the agreement is quite good. A vertical dashed line indicates where the pyramid distribution achieved its local minimum. We have truncated the distribution at that point, as described above.

With the approximate expression for the chemical potential, it is straightforward to calculate the island distributions and related properties such as the peak values and the widths. For example, according to the numerical solution of the model, $\mu(T = 600\ ^\circ\mathrm{C}, \theta = 9.0\ \mathrm{ML}) = -2.34 \times 10^{-6}$ eV/atom; using the approximate formula we find a value that is 0.6% less. The formula for the distribution of domes at this coverage and temperature is then

$$n_\nu^{\mathrm{dome}}(T = 600\ ^\circ\mathrm{C}, \theta = 9.0\ \mathrm{ML}) = \exp\left(-55.3 + 6.56\,\nu^{1/3} - 0.209\,\nu^{2/3} + 0.00164\,\nu\right) \quad (83)$$

with ν in nm^3. The positive coefficient of ν in the argument of the exponential causes the distribution to diverge for large ν; we truncate the distribution at the local minimum (here at $\nu = 616000$ nm^3) based on kinetic constraints as explained above. The pyramid distributions use different values for the energy parameters, but the same chemical potential.

In principle, there are a large number of different island shapes that could occur for the Ge on Si(001) and other systems (and others do occur when surfactants are present). Why then does it seem that researchers observe two (and only two) metastable island shapes for these semiconductor on semiconductor systems? The answer may come from the lever rule commonly used to determine the two stable phases in a binary solid state system. By drawing a line tangent to the lowest available free energy curves (or shapes) available in a phase (shape) diagram, one sees that a mixture of the two phases (or shapes) that share the tangent will always have a lower free energy than any combination of other possibilities. A system purely determined by kinetics would demonstrate a wide range of shapes with

increasing size as the chemical potential of one shape intercepted another.

8 Beyond Thermodynamic Equilibrium

While we have presented evidence that the Ge/Si island system attains a quasi-equilibrium distribution and we have constructed a model to describe this equilibrium, it may be advantageous in practice that kinetic effects are apparent in these systems under many growth conditions.[20] Materials processing in general thrives on the interplay of equilibrium and non-equilibrium processes.

Also it may be helpful to keep in mind that the island systems are very far from the true equilibrium state of the system, i.e. the state in which the deposited Ge is homogeneously alloyed with the Si substrate[18] and the surface is flat. The equilibrium description of island formation is only justified to the extent that the relaxation time for alloying is much longer than the deposition time. This is the case for chemical vapor deposition (CVD) of Ge at 600°C, but by 700°C the alloying time is comparable to the deposition time. The other kinetic constraint comes from the mobility of the Ge on the surface, which must be large enough that the island formation rate is fast compared to the deposition rate. Again this is true for CVD of Ge at 600°C, but by 500°C the rates are comparable.

9 Conclusions

We have developed a thermodynamic equilibrium model of the self-assembly of epitaxial nanostructures. In the model the surface islands form in order to reduce their bulk elastic strain energy at the cost of additional surface energy. This much is standard, but then the model departs from earlier work in two important ways. First, the edge energies act to create a metastable equilibrium at the observed nanostructure sizes, and secondly the chemical potential has been introduced to allow for fluctuations in the number of atoms in the islands as they diffuse back and forth between the islands and the wetting layer. In this article we have described in detail how the model has been constructed and solved. The solution for the Ge/Si system was summarized in the nanostructure diagram.[2] Each point of the nanostructure diagram corresponds to a full solution of the nano-thermo-mechanical model. Here we have gone further and presented an approximate analytic formula for the chemical potential throughout the range of interest in temperature and coverage. Once the chemical potential is known, the formula for the island distributions involves only elementary functions, and hence all of the predictions of the model are easily accessed.

The model has been applied with encouraging results to a nitride (GaN/AlN) nanostructure system.[9] Certainly one direction in which this work could be extended is the application to other systems exhibiting Stranski-Krastanov growth. The principal challenge is the collection of AFM data for the island distribution functions with sufficient statistics to construct and validate the model. Another extension would be a more precise test of the temperature dependence of the island distributions. Again, obtaining the ex-

perimental data is a challenge because very precise temperature control during growth is needed. The range of temperatures over which experiments may be conducted is limited. At 550 °C the approach to equilibrium is quite slow,[47] and at 650 °C alloying takes place rapidly. Temperature control would need to be maintained finely enough to distinguish systems within this narrow temperature range. Finally, the most interesting theoretical development would be an extension of the model that transitions smoothly into a kinetic model at high temperature and coverage where ripening and alloying become very important. The recent work of Vine et al.[24] is a step in this direction, but clearly there is much to be done.

Acknowledgments

R.E.R.'s work was performed in part under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory, under Contract No. W-7405-Eng-48, funded by Laboratory Directed Research and Development (04-ERD-043). G.A.D.B. thanks EPSRC for a Professorial Research Fellowship (GR/S15808/01).

References

[1] J. Stangl, V. Holý, and G. Bauer, *Rev. Mod. Phys.* 76, 725 **(2004)**.

[2] R. E. Rudd, G. A. D. Briggs, A. P. Sutton, G. Medeiros-Ribeiro, and R. S. Williams, *Phys. Rev. Lett.* 90, 146101 **(2003)**.

[3] Y.-W. Mo, D. E. Savage, B. S. Swartzentruber, and M. G. Lagally, *Phys. Rev. Lett.* 65, 1020 **(1990)**.

[4] I. Goldfarb, P. T. Hayden, J. H. G. Owen, and G. A. D. Briggs, *Phys. Rev. Lett.* 78, 3959 **(1997)**.

[5] R. Leon, C. Lobo, J. Zou, T. Romeo, and D. J. H. Cockayne, *Phys. Rev. Lett.* 81, 2486 **(1998)**.

[6] B. A. Joyce, D. D. Vvedensky, G. R. Bell, J. G. Belk, M. Itoh, and T. S. Jones, *Mat. Sci. Eng.* B67, 7 **(1999)**.

[7] U. W. Pohl, K. Potschke, A. Schliwa, F. Guffarth, D. Bimberg, N. D. Zakharov, P. Werner, M. B. Lifshits, V. A. Shchukin, and D. E. Jesson, *Phys. Rev. B* 72, 245332 **(2005)**.

[8] C. M. Reaves, V. Bressler-Hill, S. Varma, W. H. Weinberg, and S. P. DenBaars, *Surf. Sci.* 326, 209 **(1995)**; *Surf. Sci.* 341, 29 **(1995)**.

[9] C. Adelmann, B. Daudin, R. A. Oliver, G. A. D. Briggs, and R. E. Rudd, *Phys. Rev. B* 70, 125427 **(2004)**.

[10] R. S. Williams, G. Medeiros-Ribeiro, T. I. Kamins, and D. A. A. Ohlberg, *Ann. Rev. Phys. Chem.* 51, 527 **(2000)**.

[11] V. A. Shchukin, N. N. Ledentsov, P. S. Kop'ev, and D. Bimberg, *Phys. Rev. Lett.* 75, 2968 **(1995)**; V. A. Shchukin and D. Bimberg, *Rev. Mod. Phys.* 71, 1125 **(1995)**.

[12] W. Yu and A. Madhukar, *Phys. Rev. Lett.* 79, 905 **(1997)**.

[13] S. P. A. Gill and A. C. F. Cocks, "An analytical model for the thermodynamics of quantum dots: beyond the small slope assumption," submitted to *Proc. Roy. Soc. A* **(2006)**.

[14] I. Daruka and A.-L. Barabasi, *Appl. Phys. Lett.* 72, 2102 **(1998)**.

[15] M. Tomitori, K. Watanabe, M. Kobayashi, and O. Nishikawa, *Appl. Surf. Sci.* 76, 322 **(1994)**.

[16] I. Goldfarb, J. H. G. Owen, P. T. Hayden, D. R. Bowler, K. Miki, and G. A. D. Briggs, *Surf. Sci.* 394, 105 **(1997)**.

[17] G. Medeiros-Ribeiro, A. M. Bratkovski, T. I. Kamins, D. A. A. Ohlberg, and R. S. Williams, *Science* 279, 353 **(1998)**.

[18] T. I. Kamins, G. Medeiros-Ribeiro, D. A. A. Ohlberg, and R. S. Williams, *J. Appl. Phys.* 85, 1159 **(1999)**.

[19] M. Meixner, E. Schöll, V. A. Shchukin, and D. Bimberg, *Phys. Rev. Lett.* 87, 236101 **(2001)**.

[20] F. M. Ross, J. Tersoff, and R. M. Tromp, *Phys. Rev. Lett.* 80, 984 **(1998)**.

[21] F. M. Ross, R. M. Tromp, and M. C. Reuter, *Science* 286, 1931 **(1999)**.

[22] R. M. Tromp, F. M. Ross, and M. C. Reuter, *Phys. Rev. Lett.* 84, 4641 **(2000)**.

[23] A. Rastelli, M. Stoffel, J. Tersoff, G. S. Kar, and O. G. Schmidt, *Phys. Rev. Lett.* 95, 026103 **(2005)**.

[24] D. J. Vine, D. E. Jesson, M. J. Morgan V. A. Shchukin, and D. Bimberg, *Phys. Rev. B* 72, 241304(R) **(2005)**.

[25] J. A. Floro, G. A. Lucadamo, E. Chason, L. B. Freund, M. Sinclair, R. D. Twesten, and R. Q. Hwang, *Phys. Rev. Lett.* 80, 4717 **(1998)**.

[26] R. S. Williams, G. Medeiros-Ribeiro, T. I. Kamins, and D. A. A. Ohlberg, *Acc. Chem. Res.* 32, 425 **(1999)**.

[27] X. Z. Liao, J. Zou, D. J. H. Cockayne, J. Qin, Z. M. Jiang, X. Wang, and R. Leon, *Phys. Rev. B* 60, 15605 **(1999)**.

[28] X. Z. Liao, J. Zou, D. J. H. Cockayne, Z. M. Jiang, X. Wang, and R. Leon, *Appl. Phys. Lett.* 77, 1304 **(2000)**.

[29] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes*, Cambridge Univ. Press, Cambridge UK **(1992)**, Section 9.3.

[30] I. G. MacDonald, *Symmetric Functions and Hall Polynomials* Clarendon Press, Oxford, **(1979)**, Ch. 1.

[31] P. Bouwknegt, J. McCarthy, and K. Pilch, *Comm. Math. Phys.* 145, 541 **(1992)** Appendix B.

[32] P. R. Bevington, *Data Reduction and Error Analysis for the Physical Sciences* McGraw-Hill, New York, **(1969)**, Ch. 11.

[33] P. Raiteri, D. B. Migas, L. Miglio, A. Rastelli, and H. von Känel *Phys. Rev. Lett.* 88, 256103 **(2002)**.

[34] Y. Fujikawa, K. Akiyama, T. Nagao, T. Sakurai, M. G. Lagally, T. Hashimoto, Y. Morikawa, and K. Terakura, *Phys. Rev. Lett.* 88, 176101 **(2002)**.

[35] A. A. Stekolnikov, J. Furthmüller, and F. Bechstedt, *Phys. Rev. B* 65, 115318 **(2002)**

[36] T. P. Munt, D. E. Jesson, V. A. Shchukin, and D. Bimberg, *Appl. Phys. Lett.* 85, 1784 **(2004)**.

[37] G. Medeiros-Ribeiro, A. Malachias, S. Kycia, R. Magalhaes-Paniago, T. I. Kamins, and R. S. Williams, *Appl. Phys. A* 80, 1211 **(2005)**.

[38] D. Eaglesham, A. E. White, L. C. Feldman, N. Moriya, and D. C. Jacobson, *Phys. Rev. Lett.* 70, 1643 **(1993)**.

[39] R. S. Williams and G. Medeiros-Ribeiro, in *Nanostructures: Synthesis, Functional Properties and Applications*, ed. by T. Tsakalakos, I. A. Ovid'ko, and A. K. Vasudevan, Kluwer Academic Publishers, Dordrecht **(2003)** pp. 81-94.

[40] P. Haasen, *Physical Metallurgy*, 3rd ed., Cambridge University Press, Cambridge UK **(1996)**, Ch. 5.

[41] T. I. Kamins, D. A. A. Ohlberg, and R. S. Williams, *Appl. Phys. Lett.* 78, 2220 **(2001)**.

[42] T. I. Kamins, G. Medeiros-Ribeiro, D. A. A. Ohlberg, and R. S. Williams, *J. Appl. Phys.* 94, 4215 **(2003)**.

[43] T. I. Kamins, G. Medeiros-Ribeiro, D. A. A. Ohlberg, and R. S. Williams, *J. Appl. Phys.* 95, 1562 **(2004)**.

[44] A. Malachias, S. Kycia, G. Medeiros-Ribeiro, R. Magalhaes-Paniago, T. I. Kamins, and R. S. Williams, *Phys. Rev. Lett.* 91, 176101, **(2003)**.

[45] M. S. Leite, J. L. Gray, R. Hull, J. A. Floro, R. Magalhaes-Paniago, and G. Medeiros-Ribeiro, *Phys. Rev. B* 73, 121308 **(2006)**.

[46] R. Becker and W. Döring, *Ann. Phys.* 24, 719 **(1935)**.

[47] G. Medeiros-Ribeiro, T. I. Kamins, D. A. A. Ohlberg, and R. S. Williams, *Phys. Rev. B* 58, 3533 **(1998)**.

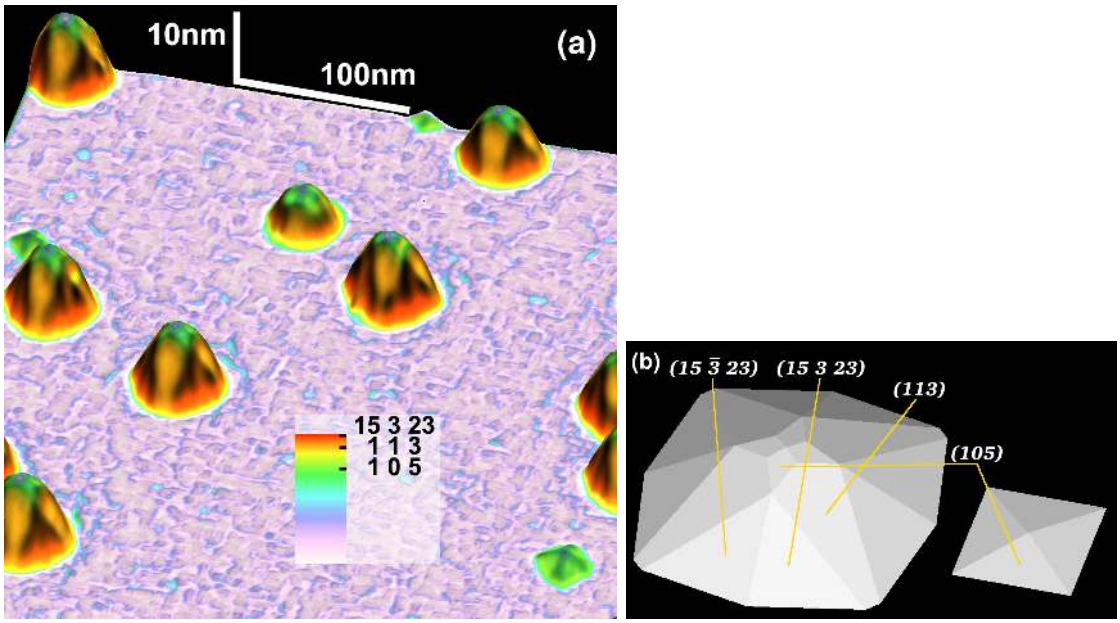

Figure 1: (a) The scanning tunneling micrograph shows Ge/Si pyramids and domes of different sizes coexisting in equilibrium next to each other. The nanostructures self-assembled during Ge molecular beam epitaxy on Si(001) at 3 ML/min at 600 °C to give an equivalent coverage $\theta = 8$ ML. The coloring provides an indication of the surface orientation, differentiating the (105) facets of the pyramids from the higher index facets of the domes (see inset for facet color coding). (b) An idealized representation of the dome and pyramid at their true aspect ratio with representative facets labeled.

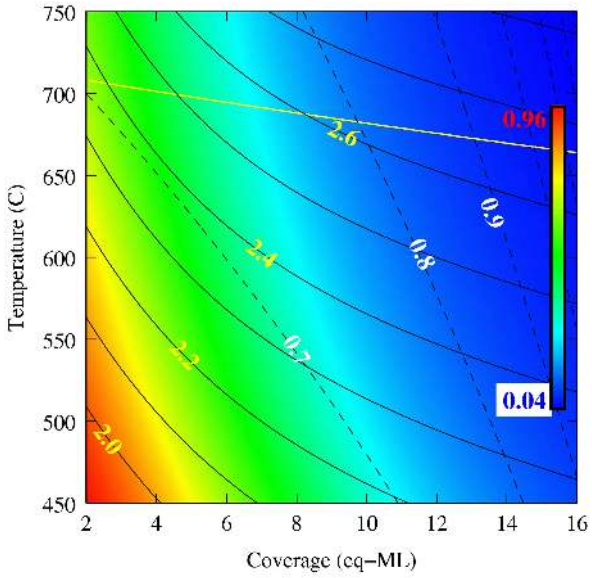

Figure 2: The nanostructure diagram for Ge/Si(001), after Rudd et al.[2] The color scale indicates the fraction of nanostructures that are pyramids (4%–96%). The solid yellow curve gives an indication of the onset of unstable ripening. It is a contour of the ratio of the dome population to the total island population restricted to islands whose size corresponds to the local minimum of the dome distribution function above the peak (see text). The other contours indicate peak widths (relative to the mean island size), with the solid black curves indicating the peak widths for pyramids and the dashed curves, those for domes.

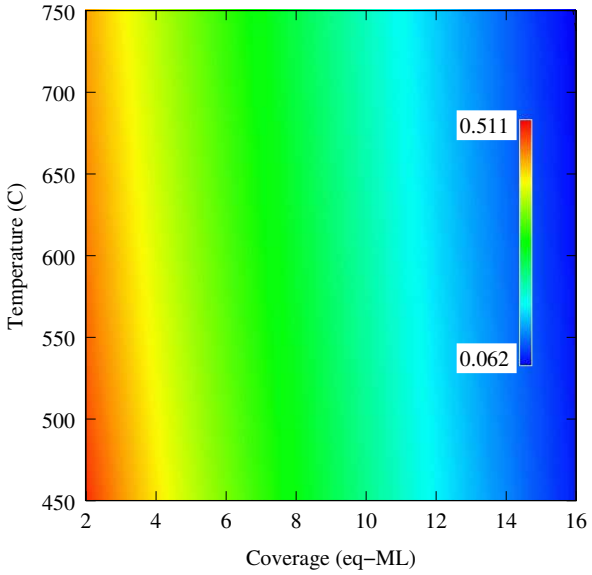

Figure 3: Chemical potential as a function of coverage and temperature, as calculated in our equilibrium statistical mechanics model. The values of the color map are given by the legend, ranging from $0.062\,A_p$ (red) to $0.511\,A_p$ (blue).

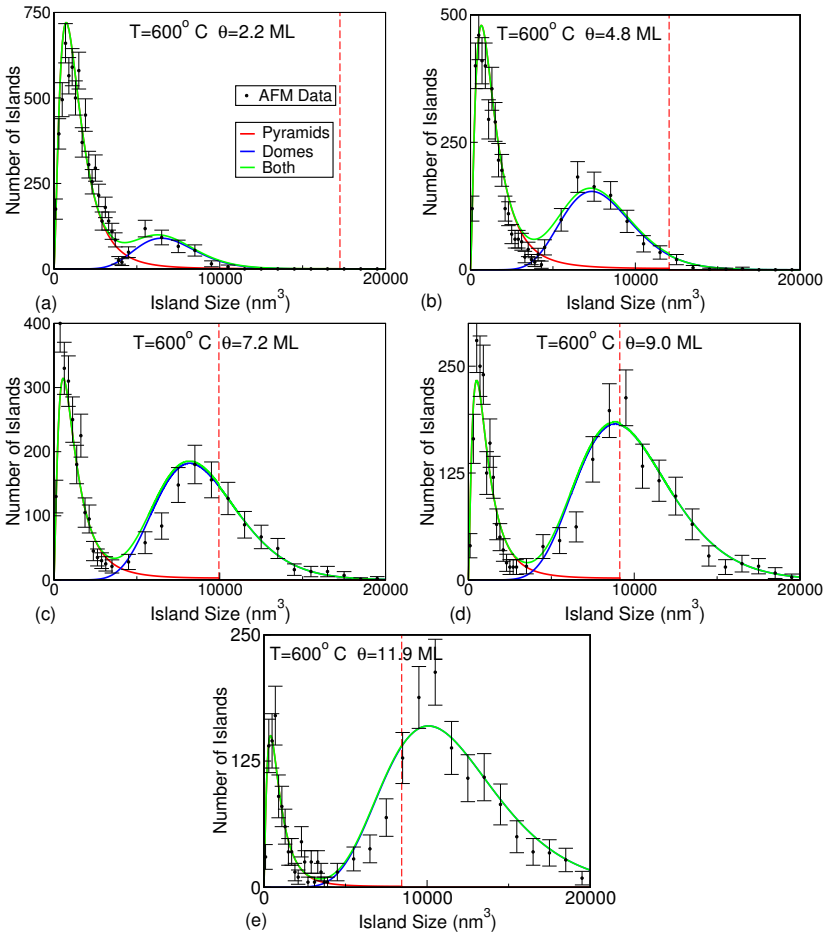

Figure 4: Island size distributions showing results from both the model (solid curves) and the experiments (dots with error bars). The curves give the numbers of islands of each type from Eq. (18), and the total number at $T = 600\,^{\circ}\mathrm{C}$ and (a) $\theta = 2.2$ ML Ge/Si(001); (b) $\theta = 4.7$ ML Ge/Si(001); (c) $\theta = 7.2$ ML Ge/Si(001); (d) $\theta = 9.0$ ML Ge/Si(001); (e) $\theta = 11.8$ ML Ge/Si(001). The vertical scales are different. The plots for $\theta = 4.8$ ML Ge/Si(001) and $\theta = 11.9$ ML were given in Rudd et al.[2] The bimodal size distribution is characteristic of two island types. The broken vertical line indicates the turning point in the pyramid energy at which the distribution was truncated; in both cases the equivalent point for domes is beyond 100,000 nm^3.