Ion-induced nanopatterning of silicon: Toward a predictive model ⓕ

View Online   Export Citation   CrossMark

Scott A. Norris[1] ⬤ and Michael J. Aziz[2] ⬤

AFFILIATIONS

[1]Department of Mathematics, Southern Methodist University, Dallas, Texas 75275, USA

[2]School of Engineering and Applied Science, Harvard University, Boston, Massachusetts 02138, USA

ABSTRACT

We review recent progress toward the development of *predictive* models of ion-induced pattern formation on room-temperature silicon, with a particular emphasis on efforts to eliminate fit parameters in the linear regime by means of experimental measurements or atomistic simulations. Analytical approaches considered include "mechanistic" models of the impact-induced collision cascade, the Crater Function Framework, and continuum treatments of ion-induced stress and viscous flow. Parameter evaluation methods include molecular dynamics and binary collision approximation simulations, as well as wafer curvature measurements and grazing incidence small-angle x-ray scattering. Mathematical detail is provided in the context of key results from pattern formation theory, which are also briefly summarized.

Published under license by AIP Publishing. https://doi.org/10.1063/1.5043438

TABLE OF CONTENTS

I. INTRODUCTION

The fabrication and tuning of materials at the nanometer lengthscale is an emerging foundational topic in materials science research. A growing number of experimental results demonstrate that nanoscale or nanostructured materials can exhibit very different properties than their bulk counterparts, suggesting promising applications in many areas including electronics, photonics, and energy generation, storage, and conversion. However, to reliably commercialize these technologies, it is necessary to have a sound theoretical understanding of the fundamental processes at work. A potentially powerful agent for nano-scale material control, which illustrates both the promise and the need for basic understanding, is the process of ion-beam irradiation. Attractive because of its existing widespread use in industry for doping, smoothing, and surface hardening, this technology has been shown capable of various feats at the nanoscale (see Fig. 1).

One regime of interest (and our exclusive focus here) occurs when nearly flat solids are irradiated by broad ion beams. Here, irradiation can lead either to surface ultra-smoothing[1,7] or to the spontaneous, self-organized formation of regular patterned structures.[8] The desirability of such patterning depends on the context. On the one hand, observations of periodic patterns including ripples and high-aspect ratio quantum dots,[3] with occasional long-range order[2,9,10] and characteristic spacing as small as 7 nm,[11] have stimulated interest in self-organized pattern formation as a means of sub-lithographic nanofabrication,[12] especially given the existing ubiquity of ion beams within the semiconductor industry. On the other hand, extended exposure to energetic particle irradiation can lead to the structural degradation of fission and fusion reactor components.[4,13] Hence, it is important to be able to predict–in advance–the expected behavior within a given environment, and such predictions are the ultimate goal of a complete theory.

Unfortunately, the physics of ion bombardment is complex, with several physical mechanisms being active simultaneously. These include (a) the sputtering of some surface atoms,[14–16] (b) the relocation of many others,[7,17,18] (c) surface diffusion,[19,20] (d) the creation of stress within the film,[21–25] and (e) the relaxation of strain and surface energies via viscous flow.[26–28] Moreover, the list of tunable environmental parameters is large, comprising the ion species, ion energy, ion flux, ion incidence angle, target species, and target temperature. A complete theory should ideally be able to predict the dependence of the strength of each physical mechanism in the first list on each of the environmental parameters in the second. Although progress is continual, understanding at this level remains elusive.

For this reason, much of the literature in the field has focused on trying to understand a subset of the mechanisms listed above, with respect to a subset of the accompanying environmental parameters. In particular, for many ion/target combinations, experiments have been conducted at ion energies in the range of 200 eV–2 keV, which span the set of incidence angles from 0° (normal incidence) to 90° (grazing incidence). The results of a representative study are shown in Fig. 2, which
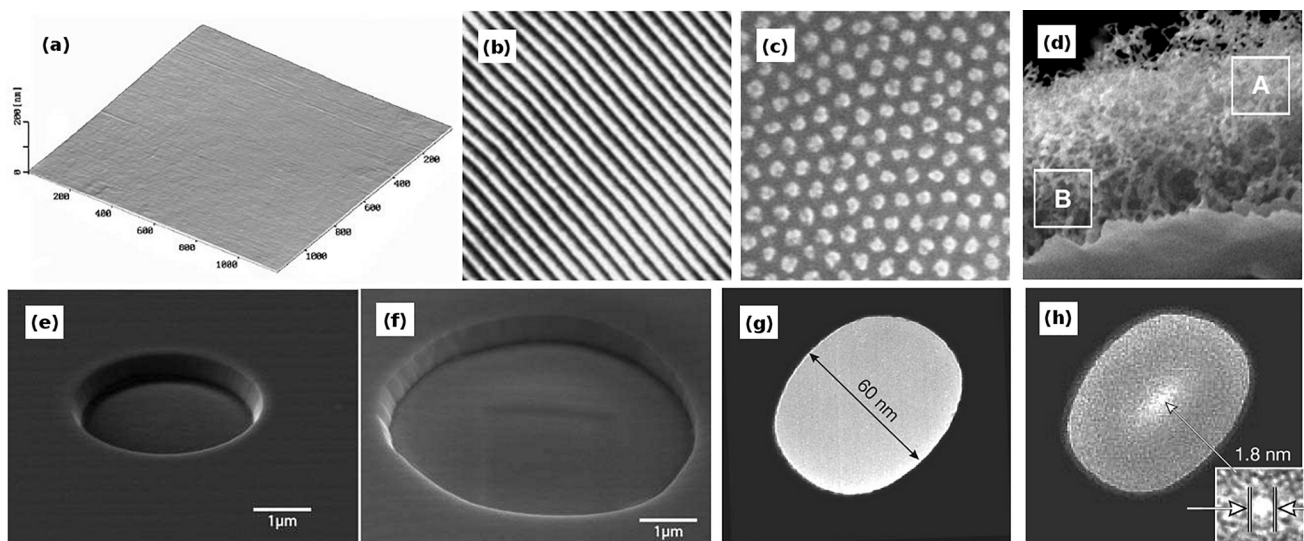
FIG. 1. A variety of uses of ion beam irradiation. (a) A polycrystalline diamond surface smoothed to a roughness less than 1 nm.[1] (b) Spontaneously formed, highly ordered ripples.[2] (c) A self-organized ordered hexagonal array of quantum dots.[3] (d) Degraded surface of a tungsten fusion-reactor wall.[4] (e) and (f) Shock-like pit walls open steadily under uniform irradiation.[5] (g) and (h) The controlled closing of a nanopore produces an opening small enough for use as a translocation DNA sequencer.[6]
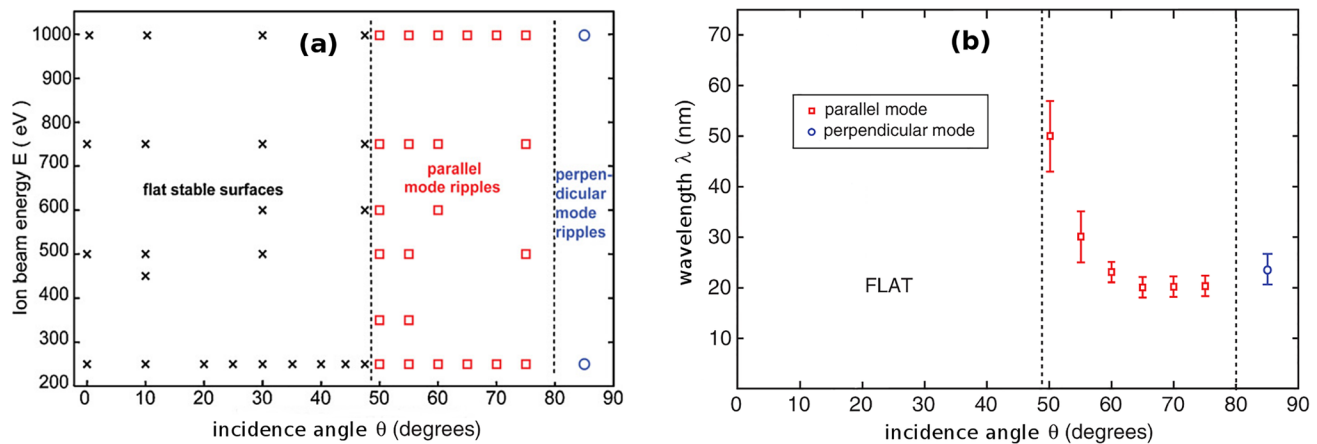
FIG. 2. (a) Observed pattern types as a function of incidence angle and ion energy. (b) Pattern wavelengths as a function of incidence angle for an ion energy of 250 eV [i.e., more detail on the bottom row of (a)]. Adapted from Ref. 29 to reflect results in Refs. 30 and 31.

illustrates the typical behavior of monatomic targets irradiated at such energies. At small angles of incidence, ion bombardment can have a smoothing effect, resulting in surfaces with a RMS roughness less than 1nm.[1,7] For intermediate angles of incidence, ripples will form, with the wavevector parallel to the ion beam−i.e., oriented as water waves are when driven by a strong wind. Ripples in this configuration are known as "parallel-mode" ripples. Finally, as the angle of incidence increases toward 90°, the ripples are often observed to change their orientation so that their wavevector is perpendicular to the ion beam−i.e., oriented as a corrugated roof is relative to falling rain. Ripples in this configuration are known as "perpendicular-mode" ripples. The presence and type of ripple can be recorded over a range of incidence angles and ion energies, leading to a "phase diagram" of outcomes shown in Fig. 2(a). In addition, the wavelength of each pattern can be measured, providing additional data of the type in Fig. 2(b). Much of this paper reviews efforts to provide a quantitatively accurate explanation for these behaviors with as few free parameters as possible−i.e., a model that is "as simple as possible, but not simpler."

This work covers the same general set of topics as the recent review by Munoz-Garcia *et al.*,[32] and our goal has been to complement that review with deeper coverage of a narrower set of topics, inevitably influenced by questions of interest in our own recent work. We concentrate on pattern formation on monatomic semiconductors such as silicon which develop a thin amorphous film under room temperature ion bombardment,[33] focus mostly on linear models valid at early times, and restrict our attention to models associated with identifiable physical mechanisms. In addition, this work emphasizes the following three themes.

1. *Mathematical details.* We are writing for a target audience of non-mathematicians but will include significant mathematical details. Our aim is to present the mathematics plainly and clearly enough to make readers with an experimental background comfortable with the basic results of stability and pattern formation theory as they pertain to ion-beam nanopattern formation.

2. *Parameter evaluation.* Ion-induced pattern formation involves many competing or cooperating mechanisms operating simultaneously. In such an environment, the use of fitting parameters can easily give a false sense of confidence in the explanatory power of a model that is actually too simple. Because a significant focus of our own work over the past decade has been the evaluation of such parameters, we will include in many sections a discussion of methods used to this end.

3. *Critical assessment.* Although much progress has been made in understanding ion-induced pattern formation, a complete predictive understanding remains to be achieved. Therefore, we will conclude each section with a discussion of the limitations of existing approaches and our thoughts on important questions relevant to each line of inquiry.

This review will proceed as follows: In Sec. II, we will give a brief overview of key concepts and results from pattern formation theory, which will then be used throughout the rest of this manuscript. In Secs. III–V, we will review three different approaches to understanding pattern formation: "mechanistic" linear models of the collision cascade, the Crater Function Framework, and continuum models of ion-induced viscous flow. Each of these sections will contain subsections associated with the three themes just described. In Sec. VI, these themes are then integrated in a brief review of recent work that directly compares scattering spectra measured by Grazing Incidence Small-Angle X-ray Scattering (GISAXS) with the predictions of various linear theories, to assess the relative magnitude of the mechanisms from Secs. III–V over a range of angles, and suggests that none of them can be neglected if a predictive understanding of ion-induced nanopattern formation is sought. Finally, in Sec. VII, we will provide concluding thoughts on the overall state of the field and suggest important outstanding questions for future study.

II. REVIEW OF GENERIC PATTERN FORMATION THEORY

In this section, we review general results on pattern formation and the tools used in its study from the seminal review by

Cross and Hohenberg and the recent text by Cross and Greenside.[34,35] The study of pattern formation can be usefully divided according to the stage of the pattern formation process itself. This includes (1) precisely characterizing the *steady state* that existed prior to the patterned state, (2) identifying how noise-induced perturbations to that state grow or decay at early times during the *linear regime*, and (3) analyzing the continued evolution of those perturbations as they grow larger within the *weakly nonlinear* regime. We will briefly review the approaches used at each stage.

A. The steady state

The study of pattern formation begins with the identification of a simple, uniform, and unpatterned configuration, in which the system could in principle exist forever. Examples across materials and scales could include (a) a nanometer-scale solid film under uniform compressive stress, (b) a centimeter-scale fluid film heated from below with a uniform temperature gradient, (c) a kilometer-scale cloud cover under uniform shear stress, or (d) a light-year scale hydrogen cloud with uniform density. In ion-induced pattern formation, the steady state is typically a flat surface receding under erosion at uniform velocity, with potentially nontrivial (but unchanging) stresses, velocities, or concentrations existing in an amorphous film persisting some depth below the surface.

Regardless of the system, the steady state is characterized by its *uniformity* in space and time, which immediately suggests symmetry arguments employed in its analytical identification as a solution to the relevant governing equations. Most obviously, because the steady state is unchanging in time, time derivatives can be set to zero (time invariance): $\frac{\partial}{\partial t} \to 0$. Next, because it is assumed to be uniform in space, lateral spatial derivatives can also be set to zero (translation invariance): $\frac{\partial}{\partial x} \to 0$, $\frac{\partial}{\partial y} \to 0$. In the simplest case, when a single equation governs the height $z = h(x, y, t)$ of a material interface, the resulting steady state is typically just a flat surface, possibly advancing or receding in the normal direction at a constant speed. In more complex cases, when multiple equations govern additional quantities such as pressure and velocity below the material interface, one typically must solve a system of boundary value problems in the vertical variable z, describing vertical profiles in these various fields.

B. The linear regime

Having identified a steady state, one next seeks to determine whether that state is stable to small perturbations, as will be present in any noisy system. We therefore seek to understand the behavior of the system—in particular the material interface—if placed in a configuration very near the steady state. For narrative simplicity, we first focus on the simplest case when a single equation governs the evolution of the surface profile $h(x, y, t)$. To explore the behavior of perturbations in this case, we simply write

$$h(x, y, t) = h_0(x, y, t) + \varepsilon h_1(x, y, t), \qquad (1)$$

where h_0 is the (possibly translating) flat steady surface, h_1 is a perturbation applied to that surface, and ε is a small

dimensionless parameter. This expansion is inserted into the governing equation, and the resulting expressions are then expanded, using Taylor series, in powers of ε. In the limit $\varepsilon \to 0$ of small perturbations, terms at higher than first order in ε are discarded, which leaves a linear equation for the behavior of the perturbation described within h_1.

The resulting linearized equation is then explored by probing the behavior of a single Fourier mode in the x–y plane. Mathematically, we look for a solution of the form

$$h_1(x, y, t) = h_{11} \exp\left[\sigma t + i(q_x x + q_y y)\right]. \qquad (2)$$

Here, h_{11} is simply a constant, $\mathbf{q} = \langle q_x, q_y \rangle$ is a Fourier wavevector, and σ is its growth rate. When this ansatz is inserted into the linearized equation, the solution of the result requires a specific value of σ for each $\mathbf{q}$; the resulting functional relationship $\sigma(\mathbf{q})$ is called the *dispersion relation*. [We note that although the two steps in the procedure just described are technically distinct, in practice, they are often performed simultaneously by combining Eqs. (1) and (2).] In systems where the governing equation is a simple partial differential equation, one may obtain the relationship $\sigma(\mathbf{q})$ quite directly: for instance, inserting Eq. (2) into the 1D diffusion equation $h_t = \alpha \frac{\partial^2 h}{\partial x^2} - \beta \frac{\partial^4 h}{\partial x^4}$ immediately yields the dispersion relation $\sigma(q) = -\alpha(q_x)^2 - \beta(q_x)^4$.

In general, the dispersion relation $\sigma(\mathbf{q})$ has both real and imaginary parts: $\sigma(\mathbf{q}) = R(\mathbf{q}) + i\psi(\mathbf{q})$. This allows the linear stability ansatz to be re-arranged into the form

$$h_1 = h_{11} \exp\left[R(\mathbf{q})t + i\mathbf{q} \cdot \left(\mathbf{x} + \frac{\psi(\mathbf{q})\mathbf{q}}{|\mathbf{q}|^2} t\right)\right]. \qquad (3)$$

This describes a sinusoidal waveform, with wavevector $\mathbf{q}$, growing at a rate $R(\mathbf{q})$ called the *amplification rate* and moving in direction $\frac{-\mathbf{q}}{|\mathbf{q}|}$ with speed $\frac{\psi(\mathbf{q})}{|\mathbf{q}|}$, called the *phase velocity*. One is often interested in one quantity or the other. For example, the study of wave propagation is often concerned with the phase velocity $\frac{\psi(\mathbf{q})}{|\mathbf{q}|}$. In contrast, the study of pattern formation is more often concerned with the amplification rate $R(\mathbf{q})$. If $R(\mathbf{q}) < 0$ for all the values of $\mathbf{q}$, then all Fourier modes decay exponentially over time, and the associated steady state is deemed *stable*. Conversely, if $R(\mathbf{q}) > 0$ for any values of $\mathbf{q}$, then some Fourier modes grow in time, rendering the steady state *unstable*.

Finally, we note that in general, if the system contains additional equations governing z-dependent quantities below the free surface, then the above procedure must also account for perturbations in these additional fields. For concreteness, suppose our problem includes, in addition to a surface height $h(x, y, t)$, a pressure field $p(x, y, z, t)$ and a velocity field $\vec{\mathbf{v}}(x, y, z, t)$ which extend below the surface. Then, the stability is investigated by the expressions

$$h(x, y, t) = h_0 + \varepsilon h_{11} \exp\left[\sigma t + i(q_x x + q_y y)\right],$$
$$p(x, y, z, t) = p_0(z) + \varepsilon p_{11}(z) \exp\left[\sigma t + i(q_x x + q_y y)\right],$$
$$\mathbf{v}(x, y, z, t) = \mathbf{v}_0(z) + \varepsilon \mathbf{v}_{11}(z) \exp\left[\sigma t + i(q_x x + q_y y)\right],$$

where h_0, $p_0(z)$, and $\vec{v}_0(z)$ describe the steady state, and the second terms are small perturbations which all have the form described above. These expressions are inserted into the

governing equations, the entire system is linearized (including boundary conditions on p and $\vec{v}$, which require special care), and h_{11}, $p_{11}(z)$, and $\vec{v}_{11}(z)$ must be obtained simultaneously by solving a system of boundary-value problems in z.

1. Classification

Unstable systems are further characterized according to which wavenumbers are unstable and which are stable. Any well-posed continuum model should be stable to sufficiently high-wavenumber perturbations: $\lim_{|\mathbf{q}|\to\infty} R(\mathbf{q}) < 0$. (Otherwise, the system would be expected to exhibit a structure on arbitrarily small scales, violating the continuum hypothesis.) Furthermore, any *unstable* model must transition, as $|\mathbf{q}|$ decreases, from stability to instability. The remaining question, then, is whether this instability persists to arbitrarily small wavenumbers or whether small wavenumbers once again become stable.

In the so-called "Type I" or "finite-wavelength" instability, only a limited range of wavenumbers is unstable, bounded on both sides by stable wavenumbers. In such an instability, patterns are dominated by the most unstable wavenumber and can exhibit a significant amount of order. In contrast, the so-called "Type II" or "longwave" instability exhibits the growth of structures at all wavenumbers below a critical value. Hence, although a most-unstable wavenumber exists, the broad range of unstable wavenumbers allows considerably more variety in structure lengthscale, typically leading to patterns with much less order.

We note that these classifications are related to the concept of a *bifurcation* or transition from stability to instability as a system parameter γ crosses some critical value γ_c. From this perspective, a Type I system is one in which the amplification rate $R(\mathbf{q})$ first becomes positive at a critical wavenumber $\mathbf{q}^\star$ that is nonzero: $|\mathbf{q}^\star| > 0$. In contrast, a Type II system is one in which the amplification rate first becomes positive at $\mathbf{q}^\star = 0$. These distinctions are illustrated in Fig. 3.

C. Nonlinear regime

In the linear regime, unstable perturbations exhibit exponential growth according to Eqs. (2) and (3). Obviously, such growth cannot continue forever, and after some time, the system reaches a state where the fundamental assumption of the linear regime—i.e., the perturbations to the steady state are small—is no longer valid. Consequently, nonlinear terms, which were previously discarded, become just as important as linear terms in determining the overall surface evolution. Numerical simulation is of course an option, but if further analytical progress is sought, practitioners introduce some alternative simplifying assumptions. In general, one focuses on the system behavior near a bifurcation, where amplification rates are still small, and only a limited range of wavenumbers is unstable (see Fig. 3). This works differently depending on the nature of the linear instability identified above, but the general goal is to gain validity at larger amplitudes in exchange for reduced validity outside the chosen wavenumber range.

1. Type I case

As discussed above, in a Type I bifurcation, when some model parameter γ crosses a critical value γ_c, an instability occurs with finite wavenumber $\mathbf{q} = \mathbf{q}^\star$. For nearby parameter values $\gamma = \gamma_c + \varepsilon$, a small range of wavenumbers near $\mathbf{q} = \mathbf{q}^\star$ grow until amplitudes are no longer "small," and nonlinear terms become important. A typical approximation in this case is to write $h(\mathbf{x}, t) = A(\mathbf{x}, t)e^{i\mathbf{q}^\star \cdot \mathbf{x}}$, where A is an amplitude of the perturbation (also called an "envelope"), and the spatial dependence, in principle, allows the formation and interaction of defects. When this approximation is inserted back into the governing equations under appropriate multiple-scale assumptions (for instance, A should vary slowly in $\mathbf{x}$ compared to the wavelength of the perturbation), then regardless of the specific nonlinearities in the equation, one typically obtains a slow evolution of the amplitude A called the Ginzburg-Landau equation

$$\frac{dA}{dt} = \left(D_1 \frac{\partial^2 A}{\partial x^2} + D_2 \frac{\partial^2 A}{\partial y^2} \right) + \sigma(\mathbf{q}^*)A - A|A|^2. \quad (4)$$

By working to eliminate rapid changes in amplitude via the diffusive terms, this equation tends to push the system toward fairly uniform patterns. Moreover, because the amplification rate $\sigma(\mathbf{q}^*)$ in this regime is expected to be proportional to ε [i.e., proportional to the amount by which γ exceeds its critical value, as suggested by Fig. 3(a)], it can be seen that at long times, the

(a)

R(q) ... γ increasing ... q ... (longwaves stable)

(b)

R(q) ... γ increasing ... q ... (longwaves unstable)

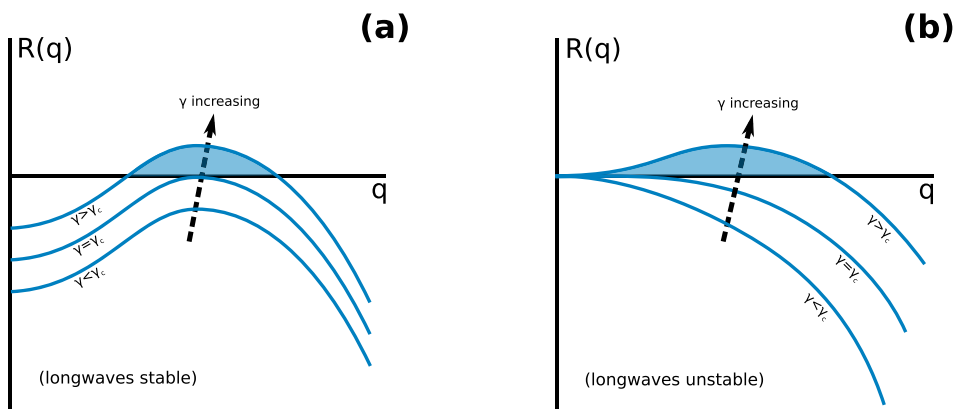**FIG. 3.** A schematic comparison of transitions from stability to the two most common instability types. (a) In "Type I" instabilities, when a model parameter crosses a critical threshold, a finite band of wavenumbers becomes unstable, but longwave modes remain stable. (b) In "Type II" instabilities, when the parameter crosses its critical value, all wavenumbers less than a maximum value become unstable, including longwave modes.

saturated amplitude $\lim_{t\to\infty} A$ of such patterns should be proportional to $\sqrt{\varepsilon}$. If the system is anisotropic, with a single oriented fastest-growing mode $\mathbf{q}^*$, one expects to see fairly regular ripple arrays, whereas if the system is isotropic with only the magnitude $|\mathbf{q}^*|$ of the fastest growing mode specified, it is possible for combinations of ripples aligned in certain compatible orientations to reinforce each other's growth, leading to square or hexagonal arrays of dots or pits (see Refs. 34 and 35)

2. Type II case

In contrast, for a Type II bifurcation, when some model parameter crosses its critical value in the form $\gamma = \gamma_c + \varepsilon$, one often observes a maximum amplification rate at a wavenumber $\mathbf{q}^* \sim \sqrt{\varepsilon}$; moreover, wavenumbers with smaller magnitude–i.e., perturbations with longer wavelengths–are *also all unstable*. Therefore, although perturbations at the critical wavelength grow the fastest, perturbations at longer wavelengths also grow, leading to increasing disorder with time. Indeed, depending on the nature of the nonlinear terms, the longer, slower-growing wavelengths can have a larger saturation amplitude, leading to coarsening or roughening over time. Hence, in contrast to the Type I case, the behavior at long times can depend sensitively on the specific form of the nonlinear terms. Therefore, in this case, a typical analytical strategy is to return to the full equation and identify the most important nonlinear terms in the limit where the wavenumber is small–the so-called "longwave approximation" (lubrication theory is a canonical example of this strategy[36]). This approach leads, in general, to one of a family of model equations, such as the Edwards-Wilkinson,[37] Kardar-Parisi-Zhang,[38] or Kuramoto-Sivashinsky[39,40] equations; for an isotropic system, the superset of terms in these equations is

$$h_t = -\alpha\nabla^2 h - \beta\nabla^4 h - \frac{\gamma}{2}|\nabla h|^2 + \eta(\mathbf{x}, t), \qquad (5)$$

where η is a noise term. Because of the longwave approximation, higher-order derivatives appear at increasingly high powers of the small parameter ε, and linear terms above fourth order are often discarded. Similarly, although different systems may contain a wide variety of nonlinearities, only the lowest-order in ε appears here.

D. Discussion

1. Instability types on monatomic targets

A great deal of interest in the field has surrounded the formation of well-ordered 2D hexagonal arrays of hills or "dots" under normal-incidence irradiation [see Fig. 1(c)]. Following Facsko's initial observation of these structures on irradiated GaSb[3] and the observation of similar structures on a variety of other III-V semiconductors,[41] they were also observed on irradiated monatomic Si.[42–44] This was perplexing because theoretical models of irradiated silicon have consistently yielded only Type II instabilities, whereas the formation of ordered structures is a hallmark of Type I systems. This discrepancy motivated a significant amount of theoretical exploration in a search for

mechanisms that could possibly generate an instability of the required type.[18,29,45,46]

However, it was gradually realized that the dot arrays only appeared on Si surfaces when significant amounts of metal were also present,[47–50] leading to the conclusion that the latter structures were due to metallic impurities.[51] It is now generally assumed that ordered dot arrays appear spontaneously only on targets containing two different atomic species. In such systems, additional physical mechanisms unique to multi-component systems arise (i.e., preferential atomic displacements and ion beam mixing), and much theoretical work has explored these systems in search of Type I bifurcations.[52–60] The most recent work suggests that the chemical interaction between the species plays an important role in the formation of these ordered structures,[58–60] which are therefore expected to be precluded in spontaneous patterns on monatomic targets. [We note several creative approaches for inducing ordered patterns on monatomic targets by other means, including the use of templates,[12] ultra-thin metal films which enable Si-metal interactions for only a limited time,[61,62] and regular oscillation of the ion beam incidence angle.[63]]

2. "Linear" vs. "longwave"

Because pattern formation on monatomic amorphous targets seems generically to exhibit Type II behavior, there is a long history of longwave models in the literature. In particular, the linearization of Eq. (5) produces an amplification rate given by a quartic polynomial

$$R(q) = Sq^2 - Bq^4, \qquad (6)$$

where the term Sq^2 describes a roughening mechanism associated with ion bombardment, and the term $-Bq^4$ describes a competing smoothing mechanism of surface energy relaxation. Given the ubiquity of this form in the literature, it is worth emphasizing that it represents an assumption of both small amplitudes (linear approximation) and small wavenumbers (longwave approximation). It can be viewed either as the longwave approximation of an amplification rate $R(q)$ with the arbitrary function of q–we will use the phrase "*full-spectrum*" to denote this general case–or the linearization of a longwave weakly nonlinear model.

The discussion above has highlighted three related, but distinct, mathematical frameworks for studying pattern formation, which we illustrate schematically in Fig. 4. Each quadrant of the diagram corresponds to a different combination of the presence or absence of the linear (small amplitude) and longwave (small wavenumber) approximations, and each offers a different combination of advantages and disadvantages. A polynomial amplification rate (south quadrant) most easily admits the development of simple, intuitive understanding of pattern formation and wavelength selection, given the relative ease of manipulating polynomials. If more detailed understanding is sought, longwave nonlinear terms (east quadrant) provide information over a greater range of amplitudes (i.e., more knowledge over time), whereas a full-spectrum linear theory (west quadrant) provides information over a greater range of wavenumbers
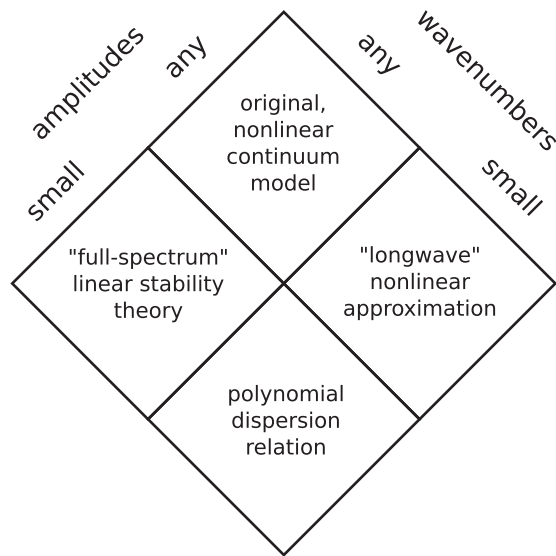
FIG. 4. Schematic illustration of analytical strategies encountered in the study of pattern formation. The two simplifications most commonly used are an assumption of either small amplitudes (a *linear* model) or small wavenumbers (a *longwave* model). Often, both quantities are assumed small, leading to an especially simple result.

(i.e., more knowledge over space). It is important to keep in mind the benefits and limitations of each quadrant, and we will refer frequently to this diagram in the remainder of this review.

3. "Top-down" vs. "bottom-up"

There are, broadly speaking, two complementary approaches to understanding pattern formation in a given system. Because all pattern-formation systems can be analyzed using the common framework just summarized, and the range of outcomes therein is finite, pattern-forming systems tend to exhibit similar physical characteristics, just on different length-scales. (Hence, the nanoscale ripples observed during ion bombardment look quite similar to ripples seen on wind- or ocean-driven sand.) One reasonable approach is therefore to begin with a knowledge of the generic governing equations and seek to apply or fit them to the behavior observed in a new system. The advantage is that one can exploit general knowledge to quickly build a nonlinear model that reasonably describes a physical experiment, often over several stages of evolution. As an example of this approach, a series of works based on the so-called "two-field" models of ion-induced pattern formation[64–67] has found success fitting an extended version of the longwave Eq. (5) to experimental data, with very impressive results.

The disadvantage of top-down models is that it is difficult to ascribe meaningful physical interpretations to the fitted model and its coefficients, making it difficult or impossible to predict parameter values from first principles. Hence, the complementary approach is to start from basic physical principles, identifying physical mechanisms, translating those into mathematical models, and then progressing through analyses of the

steady state, linear stability, and weakly nonlinear regimes. Critically, at each stage, the results of theory should be confirmed by experiment before proceeding to the next stage. Although this approach takes much longer and the benefits of universality are lost, one should obtain in the end a model whose parameter values are pre-determined in terms of the physical constants present in the system. Such a model, in principle, can then be queried to understand the behavior of new parameter regimes. A truly *predictive* model is one that can correctly identify the behavior of a system at a given combination of parameter values without the adjustment of free parameters.

4. Implications for tests of theory and parameter evaluation

In this context, and especially in light of the large number of relevant physical mechanisms described above (and in Secs. III–V below), a particular focus of our work over the past decade has been identifying the relative strengths of competing mechanisms, especially when those mechanisms have the same general mathematical form. For this specific goal, we have found that

1. the bottom-up approach is essential,
2. full-spectrum models are often helpful, and
3. the linear regime is often sufficient.

Notably, despite years of steady progress, new and surprising results on the relative importance of different mechanisms continue to emerge, which means that even the linear regime is still not fully understood from a truly predictive standpoint. Therefore, in what follows, we will spend most of our attention focusing on the linear regime, with particular attention given to methods of and progress toward parameter evaluation and the relative strength of various mechanisms.

III. MECHANISTIC MODELS OF THE COLLISION CASCADE

When an ion arrives at the surface of a solid target with an energy between 10^2 eV and 10^4 eV, it typically penetrates some distance into the solid before passing close enough to a target atom to initiate a nuclear collision cascade.[14,15] Displaced atoms that reach the surface with enough kinetic energy to leave are permanently sputtered away; all other displaced atoms come to rest within the solid or on the surface. By "mechanistic models," we mean analytical approaches built directly on an understanding of this physical mechanism.

A. Theory

We begin with a brief illustration, in Fig. 5, of the geometric conventions used in studies of ion bombardment (these apply to each of Secs. III–V). When a target is irradiated at other than normal incidence, the projected direction of the ion beam is usually taken to lie along the x-axis, with the direction of the ions corresponding to the positive x direction in the majority of published works. The incidence angle between the ion beam and the vertical axis is denoted θ. However, if the surface $z = h(x, y, t)$ is not flat, then the surface normal points in a direction other
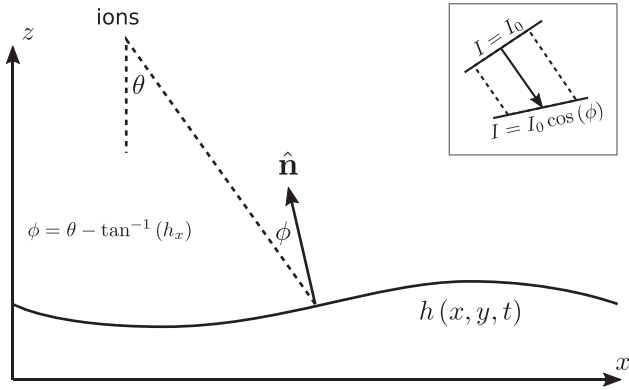
FIG. 5. Main figure: an illustration of the geometric conventions used in most studies of ion-induced pattern formation, including the direction of irradiation, the incidence angle in the lab frame θ, and the local incidence angle at the surface ϕ (note that for small slopes, $\phi - \theta \approx -\frac{\partial h}{\partial x}$). Inset: an illustration of flux dilution. If the flux through a plane normal to the ion beam is I_0, then the flux through a plane normal to the surface is $I_0 \cos(\phi)$.

than vertical. In that case, the local angle of incidence ϕ describes the angle between the ion beam and the surface normal. This angle is needed to describe the phenomenon of *flux dilution*, in which a flux of strength I_0 through a plane normal to the ion beam produces a flux through the plane normal to the surface of only $I_0 \cos(\phi)$.

1. Erosion/sputtering

The most obvious consequence of ion beam irradiation is its mathematical steady state, which consists of the gradual recession of the surface as incoming ions erode or "sputter" surface atoms. Consequently, when the spontaneous formation of ripples upon irradiated surfaces was first observed,[68] it was naturally assumed that some feature of the sputtering process might explain the ripple formation. Attempts to mathematically describe

the effect of sputter erosion on the nanopattern-formation process date to the work of Sigmund, who performed a detailed study of the collision cascade process resulting from an energetic ion impact.[14,15] Sigmund reasoned that the sputter rate at any point on the surface would be proportional to the power density at that point due to energy released by nearby ion impacts. This leads to an integral formulation for the surface normal velocity of the form

$$v_{\hat{\mathbf{n}}}(\mathbf{x}) = \int I(\mathbf{x}')E(\mathbf{x};\mathbf{x}')d\mathbf{x}', \qquad (7)$$

where $v_{\hat{\mathbf{n}}}(\vec{\mathbf{x}})$ is the normal velocity at a point $\vec{\mathbf{x}}$ on the surface, $I(\vec{\mathbf{x}}')$ is the local flux at points near $\vec{\mathbf{x}}$, and $E(\vec{\mathbf{x}};\vec{\mathbf{x}}')$ is a kernel describing the energy reaching a point $\vec{\mathbf{x}}$ due to an impact at $\vec{\mathbf{x}}'$. In addition, Sigmund introduced a simple model for the kernel $E(\vec{\mathbf{x}};\vec{\mathbf{x}}')$, proposing that the kinetic energy of each incoming ion is distributed into the target according to a Gaussian ellipsoid centered at the mean ion stopping point.[15] Under this assumption, it was shown qualitatively that indeed, when ions fall uniformly on a contoured surface, concave regions ("valleys") should receive more total energy than convex regions ("hills"), leading to a faster sputter erosion in the former and thereby to an instability [see Fig. 6(a)].

The first true linear stability analysis of Sigmund's model was performed by Bradley and Harper,[16] who exploited the experimental observation that the wavelength of ripples is often much larger than the size of an individual collision cascade. This allowed them to perform a longwave linear expansion of Eq. (7), resulting in a partial differential equation of the form

$$\frac{\partial h}{\partial t} = -v_0(\theta) + v_0'(\theta)\frac{\partial h}{\partial x} + \Omega I_0 \left[\tilde{\Gamma}_1(\theta)\frac{\partial^2 h}{\partial x^2} + \tilde{\Gamma}_2(\theta)\frac{\partial^2 h}{\partial y^2} \right] - B\nabla^4 h. \qquad (8)$$

Here, $v_0(\theta)$ is the base erosion rate of a flat surface under irradiation at an incidence angle of θ, Ω is the atomic volume of the target species, I_0 is the base flux (i.e., through a plane normal to the
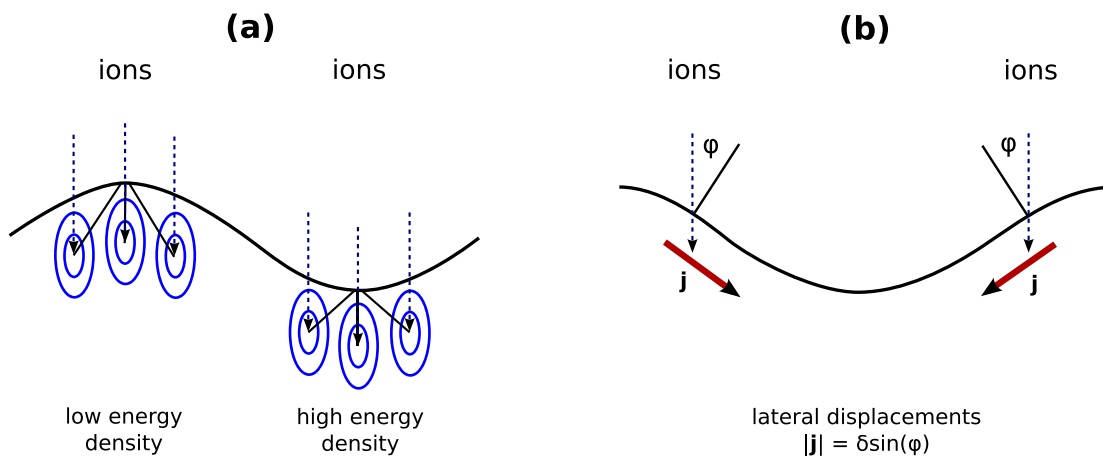
FIG. 6. Illustration of erosive and redistributive mechanisms at normal incidence. (a) Under the Sigmund model of erosion, impacts near valleys deposit more total energy in the valley than impacts near hilltops deposit on the hilltop. This leads to a faster sputtering rate in the valley, which is destabilizing. (b) Under the Carter-Vishnyakov model of mass redistribution, mass is driven downbeam from the impact. For normal incidence, this drives mass from hilltops into the valleys, which is stabilizing.

ion beam as illustrated in Fig. 5), and $\tilde{\Gamma}_1(\theta)$ and $\tilde{\Gamma}_2(\theta)$ are angle-dependent coefficients arising in the leading-order linearization of the sputtering process for a curved [to simplify discussion of competing physical mechanisms, we have used a slightly different notation than Ref. 16−our $\tilde{\Gamma}_i(\theta)$ is equal to their $\frac{a}{n\Omega}Y_0(\theta)\tilde{\Gamma}_i(\theta)$]. Because these coefficients can be negative, an additional, higher-order term is required to regularize the equation (i.e., to ensure that R < 0 in the small wavelength-limit $|q| \to \infty$). For this purpose, the authors chose to add the additional term $-B\nabla^4 h$, with angle-independent coefficient B > 0, which is the linearization of isotropic Mullins surface diffusion.[20,69]

We note that Eq. (8) represents both linear and longwave approximations (south quadrant of Fig. 4), and consequently, a stability analysis of Eq. (8) about the flat steady state produces a linear amplification rate with quartic polynomial form

$$R(q_x, q_y) = -\Omega I_0 \left[\tilde{\Gamma}_1(\theta)q_x^2 + \tilde{\Gamma}_2(\theta)q_y^2 \right] - B|\mathbf{q}|^4, \tag{9}$$

where again $\vec{q} = \langle q_x, q_y \rangle$ is the wavevector of a perturbation. We see that if either $\tilde{\Gamma}_1(\theta)$ or $\tilde{\Gamma}_2(\theta)$ is negative, then the amplification rate R(**q**) is positive for some wavevectors. The surface diffusion term, which is always stabilizing, selects one of these wavevectors as the most unstable. If $\tilde{\Gamma}_1(\theta) < 0$, the system is unstable to ripples in the x-direction via a Type-II instability with the most unstable wavelength of $q_x^* = \sqrt{\frac{-\Omega I_0 \tilde{\Gamma}_1(\theta)}{2B}}$; similarly, if $\tilde{\Gamma}_2(\theta) < 0$, the system is unstable to ripples in the y-direction via a Type-II instability with the most unstable wavelength of $q_y^* = \sqrt{\frac{-\Omega I_0 \tilde{\Gamma}_2(\theta)}{2B}}$. If both $\tilde{\Gamma}_1(\theta)$ and $\tilde{\Gamma}_2(\theta)$ are negative, it is presumed that one will observe ripples oriented in whichever direction exhibits a more negative value of $\tilde{\Gamma}$. Consistent with experiment, the model predicted parallel-mode ripples at low incidence angles, followed by a transition to perpendicular-mode ripples at incidence angles near 90°.

a. Nonlinear and higher-order terms. Given the apparent success of the combination of erosion and surface diffusion in the Bradley-Harper model,[16] many subsequent studies focused on the nonlinear extension of the model as described in Sec. II, producing weakly nonlinear models occupying the east quadrant of Fig. 4. For instance, the erosion rate $v_0(\theta)$ can be expanded as a Taylor series to higher powers of the surface slope.[70] In one dimension, expanding to second order yields a nonlinearity of the form $-\frac{1}{2}v_0''(\theta)\left(\frac{\partial h}{\partial x}\right)^2$ to Eq. (8), and as shown by Cuerno and Barabasi,[71] the equivalent process in two dimensions leads to the anisotropic Kuramoto-Sivashinski equation

$$\frac{\partial h}{\partial t} = -v_0(\theta) + v_0'(\theta)\frac{\partial h}{\partial x} + \Omega I_0 \left[\tilde{\Gamma}_1(\theta)\frac{\partial^2 h}{\partial x^2} + \tilde{\Gamma}_2(\theta)\frac{\partial^2 h}{\partial y^2} \right] - \frac{\gamma_1}{2}\left(\frac{\partial h}{\partial x}\right)^2$$

$$- \frac{\gamma_2}{2}\left(\frac{\partial h}{\partial y}\right)^2 - B\nabla^4 h.$$

Later works extended the expansion of Eq. (7) to even higher order. As the focus here is on linear theories, we refer readers

to two other reviews that cover this specific line of inquiry,[32,72] as well as more recent work specifically exploring the behavior of nearly vertical, shock-like features on otherwise flat surfaces,[5,73–75] and sawtooth-like structures featuring alternating facets of different slopes.[76–79]

Of particular interest in our context, however, we note that this approach also yielded additional *linear* terms at third and fourth orders. As noted in the original work of Bradley and Harper,[16] thermally activated Mullins surface diffusion is unlikely to be very strong at room temperature, and an alternate regularization mechanism should be identified for low temperature irradiation. The fourth-derivative terms obtained in the higher-order expansion of Eq. (7) had a negative sign, which suggested that even when sputtering produced a positive amplification rate at small wavenumbers, it could exhibit a negative amplification rate at higher wavenumbers. Therefore, it was proposed that these terms provided the needed low-temperature substitute for thermally activated surface diffusion, under the name "ion-induced effective surface diffusion."[80]

As noted above, the need for regularization arises because any model consistent with the continuum hypothesis must suppress structure formation as $|\mathbf{q}| \to \infty$. However, longwave expansions are technically only valid in the opposite limit, $|\mathbf{q}| \to 0$. As with any truncated Taylor series, these expansions can be highly inaccurate when the wavenumber is no longer small, so the use of a term obtained in the latter limit to regularize a model in the former limit is risky. [It is important to note that the term $-B|\mathbf{q}|^4$ in Eq. (9) is **not** a longwave truncation of some more-accurate full-spectrum amplification rate. Instead, the Mullins mechanism−whose nonlinear form is $v_{\hat{\mathbf{n}}} \propto \nabla_S^2 \mathcal{K}$, where $\mathcal{K}$ is the curvature and ∇_S is the surface Laplacian−*simply happens* to have a quartic amplification rate in the full-spectrum linear regime.] Indeed, in this case, it was later shown by Bradley[81] that for angles of incidence $\theta < 45°$, the *full-spectrum* amplification rate associated with the Sigmund kernel is *positive for all sufficiently large* $|\mathbf{q}|$, approaching a constant value of R as $|\mathbf{q}| \to \infty$ and implying that low temperature regularization could not come from the Sigmund kernel itself (a similar result was also obtained slightly later by More and Kree[82]). A physically consistent mechanism providing the needed regularization was not identified until sometime later (see Sec. V).

2. Displacements/redistribution

Despite the success of the Bradley-Harper model, and the subsequent effort spent extending it to higher order, a fundamental inconsistency remained between theory and experiment. Namely, the Sigmund mechanism should induce ripples at any angle of incidence, when in fact it has long been known that monatomic amorphous targets can be free of ripples if irradiated at near-normal angles of incidence.[17] Clearly, either the Sigmund model was fundamentally wrong in its description of erosion or some other mechanism was suppressing the Sigmund instability at these angles. Indeed, such a mechanism is readily apparent in simulations of ion impacts−although many atoms are displaced in the collision cascade, only a relatively small number of these are actually sputtered away from the

surface. The others remain in the target at new locations, and the effect of this *redistribution* of non-sputtered target atoms is ignored by the Sigmund model of sputtering.

The foundational work investigating the effect of redistributed atoms was performed by Carter and Vishnyakov,[17] who reasoned that when irradiated from off-normal incidence, the net displacement of redistributed atoms would occur in the direction of the ion beam as projected onto the surface. In aggregate, under many impacts, the effect could therefore be modeled as an atomic flux in the projected ion beam direction. In particular, under normal incidence bombardment, atoms on hilltops would tend to be driven downhill into valleys, stabilizing the surface [see Fig. 6(b)]. More generally, for nearly flat surfaces (i.e., in the linear regime), Carter and Vishnyakov proposed the presence of a lateral flux of target atoms proportional both to the local projected ion flux $I_0 \cos(\phi)$ and to the sine of the local incidence angle ϕ relative to the surface normal. In 1D, this gives

$$j(\theta) = \delta I_0 \cos(\phi)\sin(\phi), \tag{10}$$

where δ is a proportionality constant determined empirically. Then, applying conservation of mass, one has in the linear regime

$$\frac{\partial h}{\partial t} = -\Omega \frac{\partial j}{\partial x} \approx \Omega \delta I_0 \cos(2\theta)\frac{\partial^2 h}{\partial x^2}, \tag{11}$$

where we have employed the relation $\phi \approx \theta - \frac{\partial h}{\partial x}$ described in the caption of Fig. 5. The original work of Carter and Vishnyakov considered a one-dimensional surface (i.e., parallel-mode perturbations), and the effect of such an atomic flux on perturbations aligned in the perpendicular-mode direction was worked out later by Davidovitch,[18] who showed that a second term of the form $\Omega \delta I_0 \cos^2(\theta)\frac{\partial^2 h}{\partial y^2}$ appeared in Eq. (11).

Combining the results of Carter and Vishnyakov and Davidovitch with those of Sigmund and Bradley and Harper, one obtains a composite evolution equation of the form

$$\frac{\partial h}{\partial t} = \Omega I_0 \left\{ \left[\tilde{\Gamma}_1(\theta) + \delta \cos(2\theta)\right]\frac{\partial^2 h}{\partial x^2} \right.$$
$$\left. + \left[\tilde{\Gamma}_2(\theta) + \delta \cos^2(\theta)\right]\frac{\partial^2 h}{\partial y^2} \right\} - B\nabla^4 h, \tag{12}$$

with the corresponding linear dispersion relation

$$R(q) = -\Omega I_0 \left[(\tilde{\Gamma}_1(\theta) + \delta \cos(2\theta))q_x^2 \right.$$
$$\left. + (\tilde{\Gamma}_2(\theta) + \delta \cos^2(\theta))q_y^2 \right] - B|\mathbf{q}|^4. \tag{13}$$

Carter and Vishnyakov reasoned that, because both $\tilde{\Gamma}_1(\theta) < 0$ and $\tilde{\Gamma}_2(\theta) < 0$ for small angles of incidence, redistributive effects must be responsible for the observed stability near normal incidence (although they did not have the form of the redistributive coefficient for the y-direction, rotational symmetry at normal incidence dictates that its value be equal to that of the coefficient for the x-direction when $\theta = 0$). Similarly, Davidovitch reasoned that, because $\cos^2(\theta) > 0$ for all angles of incidence, erosive effects must dominate near grazing incidence where perpendicular-mode ripples are observed. Hence, the composite model (12) seemed potentially consistent with observations on the angle-dependence.

B. Parameter evaluation

Because irradiation by ion beams has long been used in industry as a means of implanting material surface layers with desired concentrations of dopants, the behavior of high-velocity ions interacting with solids has a long history of numerical simulation. In addition to being an early application of molecular dynamics (MD),[83–86] the nature of the collision cascade inspired the development of the simpler Binary Collision Approximation (BCA)[87,88] specifically to study ion-solid interaction. In particular, packages such as SRIM ("Stopping and Range of Ions in Matter")[89,90] and many subsequent derivatives (such as SDTRIM.SP[91] and TRI3DST[92]) have become widely used and include within their default output quantities such as the average ion penetration depth and the lateral and longitudinal straggles of these ions. These quantities are often taken to equal the parameters present in the Sigmund ellipsoidal model of energy release, and consequently, BCA simulations have long been used to estimate parameters in the Bradley-Harper model of sputtering and its extensions.

After the introduction of the study of displaced/redistributed atoms by Carter and Vishnyakov,[17] it was observed that the magnitude of the surface flux presumed by that model could also be obtained from MD or BCA simulations. This was first done by Moseler *et al.*, who studied the ultra-smoothness of carbon coatings deposited at low energies ($\sim$30 eV).[7] Independent of the work described above, these authors obtained Carter and Vishnyakov's 1D results and determined that the coefficient δ should equal the magnitude of the vector sum of all atomic displacements resulting from the collision cascade. They then used molecular dynamics simulations to obtain δ and compared the resulting models with experimental smoothing rates.

As it became clear that both erosion and redistribution contribute to the final patterning behavior and because the former is destabilizing while the latter is stabilizing near normal incidence, it was natural to begin probing the relative magnitudes of the two effects. The first attempts to measure both sets of parameters from a single simulation were performed by Zhou *et al.*,[93] who found that Eq. (12) with the simulation-informed values of both Sigmund's ellipsoidal parameters and Carter and Vishnyakov's δ, provided good agreement with experimental results on the irradiation of Al_2O_3 by 600 eV Ar^+ ions.

C. Discussion and open questions

The approaches described in this section have a number of attractive properties. They are based on simple and accessible models of the collision cascade process which can be communicated visually as in Fig. 6, and they exhibit reasonable agreement with experiment over a range of irradiation incidence angles. Nevertheless, there are several concerns and open issues pertaining to their use.

1. For the Sigmund model of erosion

The fundamental assumptions of the Sigmund model[14,15] in its simplest and most common form are that (a) on average, energy is deposited below the surface in a distribution with the Gaussian ellipsoidal shape, centered at the mean ion penetration

depth; (b) this distribution rotates about the impact point with the incidence angle of the incoming ions; and (c) the erosion rate is proportional to the energy deposited at the surface. Although plausible, these assumptions have undergone only limited verification, with mixed results.

The possibility that the energy distribution could be poorly fit by a Gaussian ellipsoid has been anticipated since the earliest studies,[14,94] and advances in computing resources now readily enable visualizations of cases where ellipses fit the observed energy distribution quite well[95] or rather poorly.[96] Such studies suggest that the fidelity of this approximation decreases with decreasing ion mass and/or increasing ion energy, in which case significant skewness appears in the distribution. Results on the relationship between the surface energy density and the erosion rate are very limited, but in molecular dynamics simulations of a variety of common experimental conditions (several noble gas ions, on both Si and Ge surfaces, at both 250 and 1500 eV), Hossain *et al.* observed that the relationship between the deposited energy and the sputter rate was in fact nonlinear and also depended significantly on ion-induced variations in the atomic number density at the surface.[95]

A particular source of concern is the issue of the angle-dependence of the energy distribution. In Sigmund's original model, the distribution was assumed to rotate upward toward the surface with increasing angle of incidence of the ion beam. As the angle θ approaches grazing incidence, a significant portion of the distribution exists *above* the surface and is ignored. Although non-physical, this correlates with the observation that sputtered atoms and backscattered ions carry away an increasing fraction of the ion energy in this limit. Nevertheless, at grazing incidence, the highest energy density occurs right at the surface, leading to maximum predicted sputter yield at grazing incidence, when in fact the sputter yield is observed to reach a peak at a smaller angle (often between 70° and 80°), before rapidly decreasing to nearly zero as ion backscattering becomes the dominant behavior.

This behavioral discrepancy has been addressed in a number of ways. Sigmund explained the downturn as the result of the incomplete development of the entire collision cascade due to the missing scattering from target atoms that would have been above the surface, had the solid extended in all directions.[97] In contrast, studying the dependence upon the incidence angle of the scattering events occurring near the surface−not the full spatially resolved distribution of released energy−Yamamura *et al.* proposed the following empirical expression based on an increased reflection probability from beneath the surface at off normal incidence and, near grazing incidence, surface channeling, which is a reduced probability of penetrating the first layer[98]

$$Y(\theta) = Y(0)\left(\frac{1}{\cos(\theta)}\right)^{f}\exp\left(-\Sigma\left(\frac{1}{\cos(\theta)}-1\right)\right),$$

where Y (0) is the sputter yield at normal incidence, and f and Σ are empirical parameters. While not based on an underlying theory of energy release, this form resembles the prediction of the Sigmund model for low angles but peaks at an intermediate angle before decaying rapidly to zero at high angles as the

$\exp\left(-\Sigma \sec(\theta)\right)$ term dominates. The resulting shape enables very good empirical agreement with experimental observations.[5,98,99] More recently, Bradley and Hofsäss proposed a modification to Sigmund's spatially resolved model, in which the ellipses rotate faster than the angle of incidence.[100] For any incidence angle, this places even more of the distribution above the surface than in the Sigmund model. Notably, the distribution's center rises above the surface at a critical angle $\theta_c < 90°$, producing a peak sputter yield at that angle, followed by a decrease at even higher angles. This approach was reported to produce somewhat better agreement with simulated sputter yields over a wider range of irradiation angles;[101] however, the observed strong trend toward zero sputter yield as the irradiation angle reaches grazing incidence remains difficult to reproduce with a spatially resolved model.

2. For the Carter-Vishnyakov model of redistribution

The issue of the angle-dependence that plagues the Sigmund model of erosion also affects the Carter-Vishnyakov model for redistribution. In Eq. (10), the assumption is that the magnitude of the net displacement vector for a single ion impact is a constant δ for any angle of incidence−the factor of $\sin(\theta)$ is simply the trigonometric projection of this vector in the lateral direction. However, if ion backscattering becomes dominant at grazing incidence, then in fact, δ depends on the incidence angle θ, with $\delta \rightarrow 0$ as $\theta \rightarrow 90°$. Thus, in this limit, the CV model, like the Sigmund model, exhibits significant quantitative error.

Of even greater concern, however, is that Carter and Vishnyakov's model of redistribution is fundamentally less comprehensive than the Sigmund model of erosion. Whatever its limitations, Sigmund's model is fully "spatially resolved"−erosion is driven by an energy distribution that is fully specified in three-dimensional space. Although early modeling focused on the limit of surface modulation with small amplitudes and wavenumbers,[16] the same underlying mechanism was later used to obtain weakly nonlinear longwave equations,[71,72,80] and full-spectrum linear amplification rates.[81] (Indeed, without any changes, the Sigmund model can also be applied in the fully nonlinear case of large amplitudes and slopes, leading to work on the behavior of steep-walled shock fronts[5,73–75] and on the behavior of alternating sawtooth-like structures.[76–79])

In contrast, the CV model does not give a spatially resolved picture of atomic redistribution in 3D. Instead, it represents a hypothesized behavior of an *average* quantity−the net atomic displacement vector. The trigonometric projection of this vector in the lateral direction [via the $\sin(\theta)$ term] is valid only when the amplitude is small (i.e., the linear regime). Moreover, its projection of the ion beam flux onto the local surface normal [via the $\cos(\theta)$ term] implicitly assumes that the surface normal is nearly constant at any given point on the surface, at least over the scale of the collision cascade. It is therefore fundamentally (even if implicitly) also a longwave model that only applies when the scale of the surface features is much larger than the scale of a single ion impact. Because of these limitations, it cannot be studied in more detail to uncover nonlinear terms in the

longwave limit nor the behavior of high-wavenumber perturbations in the linear regime.

IV. THE CRATER FUNCTION FRAMEWORK

By the mid-2000s, with the work described above becoming widely known, and increasing interest in simulation-based parameter evaluations, there were also being performed careful observations of experimental results over a range of irradiation parameters, and comparisons to the predictions of theory over the same ranges (see Ref. 8 for a review). In particular, the experiments of Madi et al.[29,30] on Ar$^+$ → Si − some results from which are reproduced in Fig. 2 above − seemed an ideal testbed for careful testing of the angle- and energy-dependence of theoretical predictions because of their use of a monatomic silicon target in a carefully shielded chamber designed to eliminate the possibility of metallic impurities influencing the observed patterns. As initially reported, these experiments were mostly consistent with the results expected from the composite model (12), with one notable exception–the presence of "holes" and perpendicular-mode ripples at small angles ($\theta < 20$) and energies ($E < 600$ eV), when Eq. (12) was expected to predict flat surfaces. Although they were eventually shown to be an artifact,[31] the search for additional physical mechanisms to explain these structures motivated significant advances in our knowledge in subsequent years.

One proposed candidate for the "missing" mechanism was the formation of small rims around the craters arising from single ion impacts. In sufficiently controlled conditions, these rimmed craters could be observed experimentally [Fig. 7(a)];[1,102] they were also readily observed during molecular dynamics simulations of impacts at sufficiently high energies [Fig. 7(b)].[86,103,105] Later, it was established by MD simulation that even at much lower energies, where neither experiment nor simulation exhibits obvious single-impact damage signatures, a clear rimmed crater shape is nevertheless observable when the surface change is averaged over many impacts [Fig. 7(c)].[104] A crater rim is clearly a redistributive effect, but it is clearly not captured by Carter and Vishnyakov's surface flux. This naturally raised questions on the effect of these shapes on surface dynamics and whether this phenomenon could be causing the unexplained structures at a low angle and energy. The study of these questions led to an approach known as the "Crater Function Framework" that allows a more direct connection between simulation and theory than was possible before, and additionally offers a potential remedy for some of the concerns with the mechanistic models described above in Sec. III C.

A. Theory

1. Integro-differential equation

If the dominant effects of the impact-induced collision cascade can be assumed to take place near the surface of the evolving film, then the normal surface velocity of an ion-bombarded surface can be represented by an integro-differential equation of the form[18,106,107]

$$\frac{\partial h}{\partial t}(\mathbf{x}) = \int I(\mathbf{x}')\Delta h(\mathbf{x} - \mathbf{x}'; \mathcal{S}(\mathbf{x}'))\,\mathrm{d}\mathbf{x}', \qquad (14)$$

where $I(\mathbf{x}') = I_0 \cos(\phi(\mathbf{x}'))$ is the projected ion flux depending on the local angle of incidence $\phi(\mathbf{x}')$, Δh is the "crater function" describing the average surface response at a point $\mathbf{x}$ due to single ion impacts at a point $\mathbf{x}'$, and $\mathcal{S}$ describes an arbitrary parametric dependence of the crater function on the surface shape at the point of impact. This is essentially just Sigmund's equation (7), with the kernel based on an ellipsoidal energy release replaced by the actual change in the surface height. Besides admitting rimmed craters, this form potentially has more fundamental advantages over traditional treatments of irradiation-induced morphology evolution. Instead of separate, simplified models of the processes of sputter erosion[15,16] and mass redistribution[17,108]–both of which break down as the angle of incidence approaches grazing–the crater function Δh naturally includes components due to both sputtered atoms and redistributed atoms (thus unifying the two approaches) and can in principle be obtained empirically (thus avoiding model inaccuracy at high angles of incidence).

Two initial attempts to study Eq. (14) should be noted. First, Davidovitch explored the question by imagining that the kernel Δh was formed by the difference of two Gaussian ellipsoids.[18] While not based on a physical mechanism, this work showed that in principle, craters with rims could produce a wider variety of surface dynamics than previously observed. This approach had the advantage of being compatible with analytical coarse-

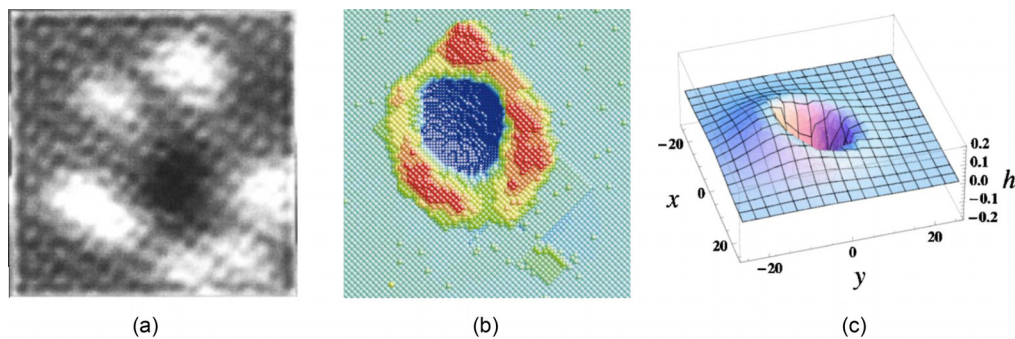**FIG. 7.** Examples of single impact craters. (a) An experimental image of a single impact of 1 keV Ne$^+$ → Ag, from Ref. 102. (b) A molecular dynamics simulation of a single impact of 100 keV Xe$^+$ → Au, from Ref. 103. (c) The average of many separate impacts of 500 eV Ar$^+$ → Si, from Ref. 104.

graining methods to produce coefficients of a partial differential equation, but was not ultimately obtained from a physical or numerical model of ion impacts. Second, the effect of rims was studied at a more direct level by Kalyanasundaram,[109] who designed a numerical algorithm in which at random intervals, individual crater shapes discovered in Ref. 104 were "stamped" on a surface, which was allowed to relax by simulated surface diffusion in between impacts. This approach had the advantage of directly incorporating MD data, but as a strictly numerical exploration, it did not reveal the conceptual connection between the crater shape and surface dynamics.

2. A generic framework

Additional progress in understanding Eq. (14) was made using the same approach as Bradley and Harper,[16] by exploiting typical experimental observations of a separation of spatial scales between the size of the impact (direct spatial dependence of Δh) and the much larger typical size of emergent structures (spatial dependence of ϕ and $\mathcal{S}$). Put differently, the local surface shape $\mathcal{S}$ and the local angle of incidence ϕ vary slowly in space compared to individual craters Δh. In general, then, Eq. (14) can be re-written as

$$\frac{\partial h}{\partial t}(\mathbf{x}) = \int I(\varepsilon\mathbf{u})\Delta h(-\mathbf{u}, \mathcal{S}(\varepsilon\mathbf{u}))\, d\mathbf{u}, \qquad (15)$$

where ε is a small parameter describing the ratio of crater to pattern lengthscales, and $\mathbf{u} = \mathbf{x}' - \mathbf{x}$ is the distance between the point $\mathbf{x}$ and the impact point $\mathbf{x}'$ in units associated with the scale of an individual impact. In this form, Δh is now seen to be a multiple-scale function with both a "fast" and a "slow" dependence on $\mathbf{u}$ (the projected ion flux varies only slowly on the local angle of incidence). Moreover, it suggests a multiple-scale analysis in which the integral in Eq. (15) is expanded in powers of ε into an infinite series of terms involving the moments of Δh.[107] In the linear regime, this process yields (see Appendix A)

$$\frac{\partial h}{\partial t} = \left[IM^{(0)}\right] - \nabla \cdot \left[IM^{(1)}\right] + \frac{1}{2}\nabla \cdot \nabla \cdot \left[IM^{(2)}\right] - \cdots, \qquad (16)$$

where the $M^{(i)}$ are the moments of the crater function Δh in increasing tensor order

$$M^{(i)} = \int \mathbf{u}^{(i)}\Delta h(\mathbf{u}; \mathcal{S}(0))\, d\mathbf{u}. \qquad (17)$$

We note that $\mathbf{u}^{(i)}$ denotes an i-th order tensor: $\mathbf{u}^{(1)} = \mathbf{u}$, $\mathbf{u}^{(2)} = \mathbf{u} \otimes \mathbf{u}$, $\mathbf{u}^{(3)} = \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{u}$, ... where $\otimes$ is the outer product, and ∇ are divergences which reduce these tensor moments back to scalars before summation.

Equation (16) represents a generic multiple-scale expansion of Eq. (14) in the sense that it should apply for any parametric dependencies of the crater function on the surface shape $\mathcal{S}$. To apply this result to a specific system, then, one must (a) define the crater function Δh and (b) insert it into Eqs. (16)–(17). In general, to leading order, one obtains the expression

$$\frac{\partial h}{\partial t} = c_0(\theta) + c_1(\theta)\frac{\partial h}{\partial x} + c_{11}(\theta)\frac{\partial^2 h}{\partial x^2} + c_{22}(\theta)\frac{\partial^2 h}{\partial y^2}, \qquad (18)$$

where we have adopted the convenient notation introduced in Ref. 110. Note that this form is the same as that in Eq. (8).

3. From moments to coefficients

The functional form of the coefficients $c_{...}(\theta)$ depends on the specific form of the crater function. For instance, in Ref. 107, it was shown that for the crater function associated with Sigmund's model of erosion, Eq. (16) recovers the results of Bradley and Harper for the coefficients $c_{11}(\theta)$ and $c_{22}(\theta)$. However, the original goal of this approach was to accommodate more general (and potentially empirical) crater shapes, such as those featuring rims. Therefore, instead of worrying about the specific form of Δh, one can specify only its parametric dependence and then linearize the resulting specific instance of Eq. (16) about a flat surface. This produces formulae for $c_{...}(\theta)$ with the same parametric dependence, which can both inform and exploit simulation results (as discussed in more detail below).

Because the first two terms

$$c_0(\theta) = v_0(\theta) = -\Omega I(\theta)Y(\theta),$$
$$c_1(\theta) = -v_0'(\theta) = \frac{\partial}{\partial\theta}[\Omega I(\theta)Y(\theta)], \qquad (19)$$

where Ω is the atomic volume, $I(\theta) = I_0 \cos(\theta)$ is the flux through a plane normal to the surface, and $Y(\theta)$ is the atomic sputter yield, are widely known and do not affect the surface stability (indeed, they can be eliminated by a the adoption of a suitable vertically and laterally advecting frame of reference), most attention has focused on obtaining the angle-dependent coefficients $c_{11}(\theta)$ and $c_{22}(\theta)$. This process was first demonstrated in Ref. 111, where for simplicity and consistency with available simulation data, the crater function

$$\Delta h = g(\mathbf{x} - \mathbf{x}'; \phi)$$

was chosen, which depends parametrically only on the local angle of incidence ϕ. Inserting this expression into (16), one finds that the angle-dependent coefficients $c_{11}(\theta)$ and $c_{22}(\theta)$ are related to the crater function via the expressions

$$c_{11}(\theta) = \frac{\partial}{\partial\theta}\left[I_0 \cos(\theta)M_x^{(1)}(\theta)\right],$$
$$c_{22}(\theta) = \left[I_0 \cos(\theta)\cot(\theta)M_x^{(1)}(\theta)\right], \qquad (20)$$

where $M_x^{(1)}$ is the component of the (vector) first moment in the projected direction of the ion beam.

More generally, as noted in Refs. 18, 107, and 111, the crater function in principle also depends on higher-order properties of the surface such as its curvature, via an expression such as[18]

$$\Delta h = g\left(\mathbf{x} - \mathbf{x}'; \phi, \frac{\partial^2 h}{\partial x^2}, \frac{\partial^2 h}{\partial x \partial y}, \frac{\partial^2 h}{\partial y^2}\right)$$

depending, in addition to the angle of incidence ϕ, on the second derivatives of the surface shape at the point of impact. It was anticipated that the inclusion of such a dependence would provide more accurate values of $c_{11}(\theta)$ and $c_{22}(\theta)$, enabling better agreement with

experimental tests of the crater function framework.[112] Recently, the theoretical implications of this dependence have been worked out—Harrison and Bradley found that including this dependency within the crater function reveals additional terms in the coefficient values, which take the revised form[110]

$$c_{11}(\theta) = \frac{\partial}{\partial \theta}\left[I_0 \cos(\theta) M_x^{(1)}\right] + \frac{\partial}{\partial K_{11}}\left[I_0 \cos(\theta) M^{(0)}\right],$$
$$c_{22}(\theta) = \left[I_0 \cos(\theta) \cot(\theta) M_x^{(1)}\right] + \frac{\partial}{\partial K_{22}}\left[I_0 \cos(\theta) M^{(0)}\right], \quad (21)$$

where $K_{11} = \frac{\partial^2 h}{\partial x^2}$ and $K_{22} = \frac{\partial^2 h}{\partial y^2}$.

4. Higher-order terms

The progression implied above can be carried further, yielding higher-order linear coefficients in Eq. (18). For instance, let us consider a 1D crater function of the form

$$\Delta h = g\left(\mathbf{x} - \mathbf{x}'; \phi, \frac{\partial^2 h}{\partial x^2}, \frac{\partial^3 h}{\partial x^3}, \frac{\partial^4 h}{\partial x^4}, \ldots\right),$$

i.e., in which the dependence on the surface shape now extends to an arbitrary number of derivatives of $h(x)$. Because the moments $M^{(i)}$ inherit the crater function's parametric dependence on the surface shape through Eq. (17) and, in the linear regime, all of the derivatives of h can be considered very small, each moment can be expanded to leading order in h via

$$M^{(i)}\left(\phi, \frac{\partial^2 h}{\partial x^2}, \frac{\partial^3 h}{\partial x^3}, \ldots\right) \approx M^{(i)} + \frac{\partial}{\partial \phi} M^{(i)}|_{(\text{flat})} \cdot (\phi - \theta)$$
$$+ \frac{\partial}{\partial\left(\frac{\partial^2 h}{\partial x^2}\right)} M^{(i)}|_{(\text{flat})} \cdot \frac{\partial^2 h}{\partial x^2} \ldots$$

Inserting this expression into the generic result (16) produces a linearized governing equation of the form

$$\frac{\partial h}{\partial t} = c_0(\theta) + c_1(\theta)\frac{\partial h}{\partial x} + c_{11}(\theta)\frac{\partial^2 h}{\partial x^2} + c_{111}(\theta)\frac{\partial^3 h}{\partial x^3} + c_{1111}(\theta)\frac{\partial^4 h}{\partial x^4} + \cdots,$$

where

$$c_0 = \left[I M^{(0)}\right],$$
$$c_1 = -\frac{\partial}{\partial \phi}\left[I M^{(0)}\right],$$
$$c_{11} = \frac{\partial}{\partial \phi}\left[I M^{(1)}\right] + \frac{\partial}{\partial K_{11}}\left[I M^{(0)}\right],$$
$$c_{111} = -\frac{\partial}{\partial \phi}\left[I M^{(2)}\right] - \frac{\partial}{\partial K_{11}}\left[I M^{(1)}\right] + \frac{\partial}{\partial K_{111}}\left[I M^{(0)}\right],$$
$$c_{1111} = \frac{\partial}{\partial \phi}\left[I M^{(3)}\right] + \frac{\partial}{\partial K_{11}}\left[I M^{(2)}\right] - \frac{\partial}{\partial K_{111}}\left[I M^{(1)}\right] + \frac{\partial}{\partial K_{1111}}\left[I M^{(0)}\right].$$
$$\ldots \quad (22)$$

Here, $I = I_0 \cos(\theta)$, $K_{1(n)} = \frac{\partial^{(n)} h}{\partial x^{(n)}}$, and all derivatives of bracketed terms are evaluated for a flat surface. (We note that the coefficients involving the y direction for a 2D expansion would require more care to obtain due to the common convention of creating

local co-ordinates aligned with the projected ion beam direction at the point of impact).

We caution, however, that the inclusion of such higher-order terms may be of limited value. First, higher-order moments are more sensitive to any inaccuracy in the simulated impact. Moreover, as described above, much effort was expended to obtain higher-order expansions of the Sigmund model of erosion, and for a time it was assumed that while erosion could be destabilizing at small wavenumbers through the second-order terms, it could also be stabilizing at higher wavenumbers through the fourth-order terms. However, the later study of the exact dispersion relation for the Sigmund model showed that this was incorrect.[81] Because the crater function framework is simply a generalization of the results in Sec. III, the same concerns apply here.

5. Implications

Conceptually, the crater function framework enables a concise explanation of precisely how an arbitrary collision cascade affects the surface stability—through the first few moments of the associated crater function. For instance, the zeroth moment $M^{(0)}(\theta) = -\Omega Y(\theta)$ is simply the negative product of the atomic volume Ω and the sputter yield Y. This means that the familiar expressions (19) for c_0 and c_1 have long expressed the erosion and lateral advection rates generically in terms of the zeroth moment, and Eqs. (20)–(22) can be seen as simply extending that genericism to the higher-order coefficients. Moreover, expression in terms of moments greatly simplifies an eventual quest to obtain an empirically informed model—it implies that statistically reliable data are not needed on the entire crater function shape (which would require many samples) but rather only for its first few moments (which requires far fewer samples).

This approach also directly addresses two of the main concerns with mechanistic methods described above in Sec. III. First, if the empirical, statistically converged values of the first few moments can be obtained over the whole range of incidence angles $\theta \in [0°, 90°]$, the resulting model should be much more accurate at large incidence angles than a model based on simple geometric arguments (i.e., the rotation of Sigmund's ellipses and the trigonometric projection of Carter and Vishnyakov's surface flux). Second, it places erosion and redistribution on completely equal theoretical footing. Assuming that the contributions of erosion and redistribution to the crater function Δh can be distinguished, i.e.,

$$\Delta h = \Delta h_{\text{eros.}} + \Delta h_{\text{redist.}}, \quad (23)$$

(this is fairly trivial in an atomistic simulation), then each of the moments in Eq. (16)—and therefore each of the coefficients in Eqs. (20)–(22)—can be similarly split into erosive and redistributive parts (with the exception of $M^{(0)}$, which is purely erosive). Comparing Eqs. (10) and (11) to Eqs. (20) and (21) and noting that $M^{(0)}$ is purely erosive, we see that Carter and Vishnyakov's surface flux model can be efficiently re-cast as an assumption that the redistributive contribution to $M^{(1)}$ is $M_{\text{redist.}}^{(1)} = \delta \sin(\theta)$. Moreover, this result immediately answers one of the questions that motivated its development—isotropic crater rims would first appear in the second moment $M_{\text{redist.}}^{(2)}$ and therefore have no leading-order effect on surface stability through $c_{11}(\theta)$ or $c_{22}(\theta)$.
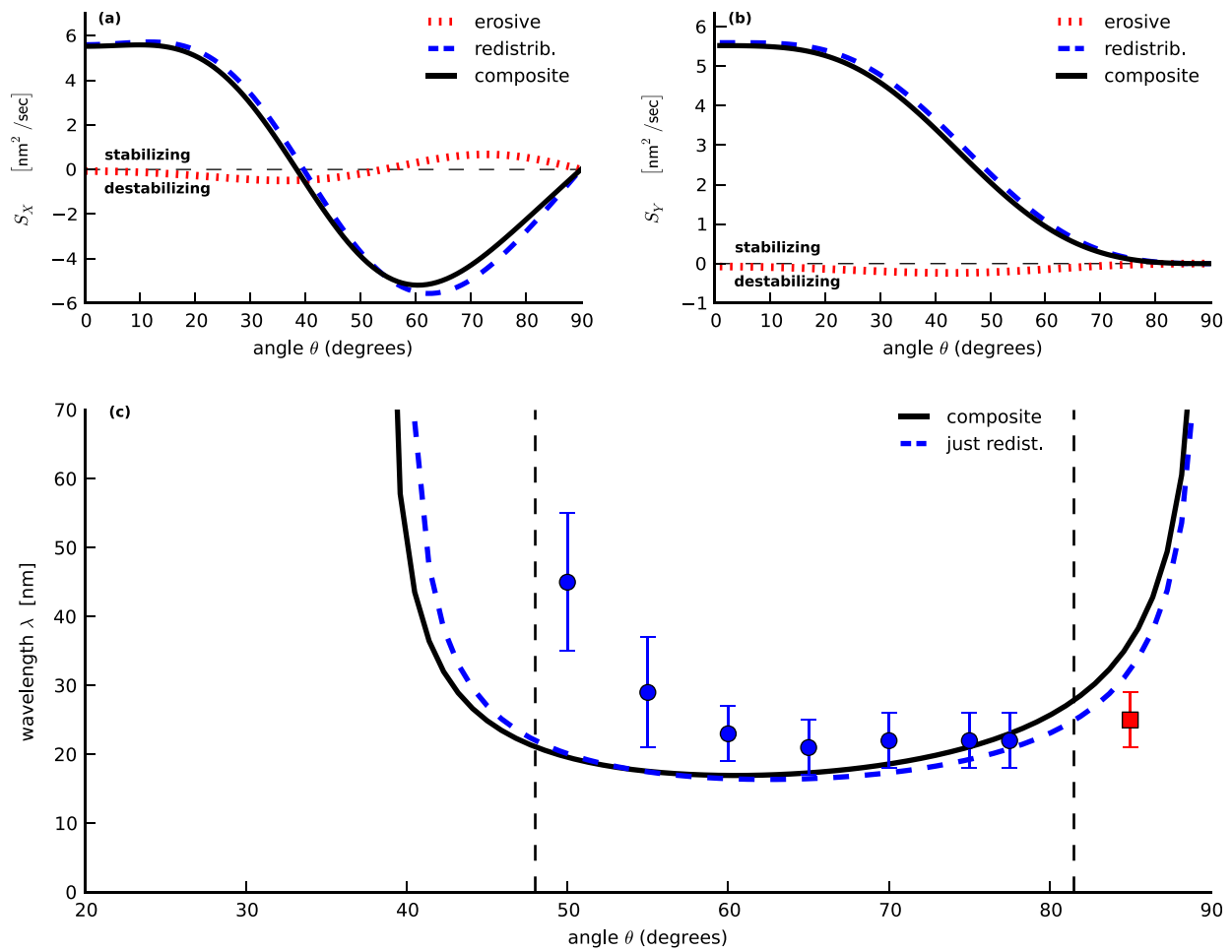
FIG. 8. First use of the Crater Function Framework, applied to the system 250 eV Ar$^+$ → Si, from Ref. 111. (a) The coefficient $S_X(\theta) = c_{11}(\theta)$ governing the leading order behavior of parallel-mode perturbations. (b) The coefficient $S_Y(\theta) = c_{22}(\theta)$ governing the leading order behavior of perpendicular-mode perturbations. (c) A prediction of the pattern wavelengths based on an additional estimate of the surface-relaxation coefficient B. [Note: although at the time of this work, the additional terms $\frac{\partial M^{(0)}}{\partial K_{ii}}$ from Eq. (21) were unknown, subsequent re-analysis has confirmed that, because $M^{(0)}$ is purely erosive and erosion is weak at 250 eV, these terms are essentially negligible.].

B. Parameter evaluation

The results described above immediately suggest a strategy for building a governing equation having the form of Eq. (18) directly from atomistic simulations, by means of the moments appearing in Eq. (16) and specifically their derivatives which appear in Eqs. (20)–(22). In general, we must

(1) create or select a simulation tool to perform many single-impact simulations,
(2) obtain statistically converged moments at various required surface geometries,
(3) estimate derivative values using data from adjacent parameter combinations,
(4) smooth the data to prevent uncertainties at step (2) from being amplified in step (3).

An approach incorporating such steps was first demonstrated in Ref. 111, where molecular dynamics simulations using the PARCAS code[113–115] were performed for irradiation of Si by 250 eV Ar$^+$ at 5-degree increments between 0° and 90°. Smoothing was accomplished by a weighted fitting of the simulation results to a truncated Fourier series, and the fitted values of $M_x^{(1)}(\theta)$ were then inserted into Eq. (20). Analyzing the resulting linear PDE with the form of Eq. (18) (with additional terms describing ion-enhanced thin film viscous flow, described below in Sec. V), the most-unstable wavelengths at each angle were compared to the wavelengths of experimentally observed structures, with reasonable agreement (see Fig. 8). In the process, the relative sizes of the effects of erosion and redistribution were directly obtained and compared, and the effects of redistribution were unexpectedly found to be dominant for the chosen system, not only causing the regime of smooth surfaces below 40° but also driving the instability to parallel-mode surfaces at higher angles (only the perpendicular-mode instability at 85° was not accounted for).

Since its introduction, the Crater Function Framework has been extended in a number of ways. As described above, Harrison and Bradley extended Eq. (20) to include the effect of curved surfaces through Eq. (21). Because the additional terms describe the effect of surface curvature on the purely-erosive zeroth moment $M^{(0)}$, adoption has tended to depend on the energy studied and simulation method used. Studies at higher energies using the BCA now regularly include the terms,[92,116–119] because they seem to contribute meaningfully to the coefficients $c_{11}(\theta)$ and $c_{22}(\theta)$ at those energies, and some BCA codes readily allow the simulation of ion impacts on curved targets. On the other hand, studies at lower energies using MD may still omit the terms,[111,115,116] because of the complexity of creating relaxed amorphous targets with varying curvatures, and because erosive effects are not particularly strong at such energies.

Another notable addition to the framework is the inclusion of the effect of implanted ions. Although implanted noble gas ions have often been ignored, Bradley and Hofsäss[122] and Hofsäss et al.[123] reasoned that when a target is irradiated with atoms of its own species, then implanted atoms would also contribute meaningfully to the crater function. Accordingly, they added a term $\Delta h_{\text{implant.}}$ to Eq. (23) which, when added to the integrand in Eq. (14), yields contributions due to implantation to the moments in Eq. (16) and hence to the coefficients $c_{11}(\theta)$ and $c_{22}(\theta)$. More recently, Hobler et al. discovered that even noble gas ions make such a contribution.[118] Although, in the steady state, these tend to sputter at approximately the same rate they are implanted, and therefore do not contribute to $M^{(0)}$, these authors found that the re-sputtered ions tend to originate farther "downstream" than the final resting place of the implanting ions that effectively replace them. This yields a meaningful contribution $M^{(1)}(\theta)$ even though the noble gas ions tend to remain embedded in the solid surface only a short while before they are eventually re-sputtered.

C. Discussion and open questions

As hinted above, the Crater Function Framework can be seen as a generalization of the mechanistic approaches described in Sec. III—it can reproduce the seminal results associated with those models, while providing insights into exactly how the moments of the crater function Δh affect morphology evolution. However, it can also be paired with empirical moments obtained from atomistic simulations, sidestepping potential model errors and thereby providing more accurate predictions. This also provides unprecedented granularity of understanding, as the crater function, and hence, the moments, can be broken down into contributions from sputtered atoms, redistributed atoms, and implanted ions, each of which can be further divided by atomic species.[118,124,125] Nevertheless, despite these advantages, the crater function framework has important limitations of its own.

1. Defining shape

One of the most basic needs of the framework is to convert changes in the positions of discrete atoms into the change in a surface height $\Delta h(x, y)$, so that the moments of that change can

be computed. At least three different methods have been used to do this in the literature.

1. The first attempt to analyze single-impact crater functions defined the surface $h(x, y)$ as the height for which a "virtual silicon probe atom" experienced zero force.[104] This approach is probably most consistent with how a "surface" is physically defined; however, it experiences the concavity bias associated with AFM tips and is more generally over-sensitive to local dynamics. For instance, sub-surface displacements may not register accurately, and therefore, there is no guarantee that the zeroth moment equals the sputter yield times the atomic volume.

2. The presentation of the generic crater function framework[107] proposed defining Δh as a sum of delta functions with the magnitude equal to the atomic volume, located at the pre- and post-impact atomic positions. Exploiting the integral properties of the delta function, this allows one to directly obtain the moments from the atomic positions, without first calculating the surface height. Moreover, it records deep displacements as happily as shallow ones and guarantees agreement between sputter yield and the zeroth moment.

3. The convenience provided by method #2 requires assuming that vacancies and interstitials immediately migrate vertically to the surface, which is unlikely to be true, especially at higher energies. In Ref. 126, the authors propose a method to avoid this assumption. After each simulated impact, they place a vacancy at the pre-impact point of all displaced atoms and an interstitial at the post-impact point of redistributed atoms. Then, these defects are allowed to diffuse under a KMC simulation until annihilating or reaching the surface, at which point the final crater shape is recorded. This approach seems likely to be more realistic, but also considerably more time consuming by requiring the extra KMC step.

2. Inherently longwave

A second important limitation of the crater function framework is that it is an inherently long-wavelength approximation—any expansion of Eq. (14) into a form like Eq. (16) is founded on the assumption that the crater is small compared to the pattern wavelength. As the wavelength shrinks to the scale of the individual craters, this assumption becomes invalid, and so the crater function framework can only be considered reliable in the "longwave" limit described above. As we will see below, this limits the crater function framework's utility in detailed comparisons between competing physical mechanisms. This should be contrasted with, for example, the full-spectrum linear amplification rate for the Sigmund model of sputtering,[81] which is valid for all wavenumbers but inherits any fundamental inaccuracy of the Sigmund model, in particular its breakdown at large angles of incidence. Because even higher-order expansions of the framework are cumbersome and do not mitigate its failure at high wavenumbers, it appears that quantitative accuracy at both large angles and large wavenumbers remains elusive.

3. MD vs. BCA

An important question for the use of the framework concerns the use of Molecular Dynamics vs the Binary Collision Approximation for the collection of crater function moments. On the one hand, the primary advantage of the BCA is obviously its speed and ready availability to researchers without access to extensive computing power. By design, however, BCA codes are incapable of observing the many-body interactions that dominate at the end of the collision cascade, which MD simulations reveal to be non-trivial.[127] On the other hand, most implementations in MD are performed in a rather small periodic computational cell to limit complexity; if the cell is too small, such simulations are subject to global shears, which must be addressed in post-processing[111] (fixed boundaries are vulnerable to different errors in a too-small cell). The presence of "small displacement" error in the BCA and "small cell" error in MD means that both approaches are likely to remain important checks on each other—most likely the BCA will become the standard for quick estimates and broad parameter sweeps, with MD used to answer more detailed questions or inform BCA parameter choices.

4. Comments on a proposed dependence on film thickness

We note that Refs. 120 and 128 reported the derivation of an additional effect within the Crater Function Framework due to variations in the thickness of the amorphous film, having the form $h_t \propto D_{I1}(\theta) \frac{\partial^2 h}{\partial x^2}$, where $D_{I1}(\theta)$ is positive for all θ. However, in our opinion, these derivations contain modeling errors which, if corrected, cause the postulated mechanism to vanish. In Ref. 120, Eq. (17) describing a "modified" kinematic condition that depends on film thickness contradicts the inherently local nature of the kinematic condition, which is simply a statement of conservation of mass applied at the film-vacuum interface. If the kinematic condition is restored to its proper form, the term D_{I1} obtained in Ref. 120 does not appear. In Ref. 128, a different derivation of the term D_{I1} is presented. There, the correct kinematic condition is used in Eq. (3), and the direct integration of the continuity equation leads correctly to Eq. (12). However, an arbitrary dependence on the film thickness is again added in Eq. (13), which is inconsistent with the boundary conditions used to obtain Eq. (12). Again, if this contribution is removed, the term D_{I1} does not appear. Hence, although an unconditionally positive coefficient resembling $D_{I1}(\theta)$ would seem to significantly improve agreement with experiment as shown in Refs. 121 and 129, at present, there does not seem to be physical justification for such a term. [We would also argue that the redistributive component of morphology evolution is unlikely to have a depth dependence for intuitive reasons. Both the crater function moments and the film depth are simply statistical averages associated with the collision cascade, and so, a direct dependence of the former upon the latter seems unlikely.]

D. The PyCraters library

This section has outlined the general features of the Crater Function approach, including the potential promise for the general problem of coefficient evaluation, and also technical hurdles associated with its use. However, while the process of simulation, statistical analysis, fitting, and differentiation is time-consuming, it is also in principle mechanical, suggesting the utility of an open-source library to centralize the best practices and avoid repeated re-implementations. We have recently developed such a library, which aims to be accessible to new users, compatible with various simulation codes and extensible by advanced users with needs not covered by the basic functionality. This suggests implementation as a layered framework, in which

- different simulation codes are "wrapped" by a common Python interface;
- statistical information is converted to a standard, internal format;
- smoothing and fitting are performed automatically using external libraries;
- commonly sought quantities, such as $c_{...}(\theta)$, are automatically calculated.

This library is freely available on the project-hosting site GitHub.[130]

V. STRESS AND VISCOUS RELAXATION

In addition to causing sputter erosion and atomic redistribution through the collision cascade, ion bombardment has been observed to induce significant amounts of stress in the amorphized layer of monatomic targets and has long been theorized to enhance the fluidity of that layer, allowing it to flow like a highly viscous fluid. Here, we review recent models focused on these effects and their interaction.

A. Theory

1. Ion-enhanced viscous flow

The original Bradley-Harper model assumed that the relaxation of surface energy occurred by surface self-diffusion as described by Mullins.[20,69] However, within the experimental community, an alternative mechanism of "ion-enhanced viscous flow" (IVF) has long been discussed. The essential idea for the cause of IVF is that the collision cascade creates[21] a supersaturation of vacancy-like[131,132] and interstitial-like[133] point defects which, after re-thermalization of the atomic velocities, migrate in directions so as to relieve the stress.[134] The ion-enhanced fluidity provides an additional mechanism of relaxing the excess surface free energy associated with non-planar surface geometries. Determining the precise contribution to the dispersion relation requires solving the Stokes equations (the limit of the Navier-Stokes equations in which the Reynolds number $Re \ll 1$). This was done by Mullins[135] for the case of a semi-infinite viscous fluid, and the resulting contribution to the linear amplification rate $R \propto -F_b q$ was first added to models of ion-induced nanopatterning by Mayer, Chason, and Howard.[136]

However, the semi-infinite viscous fluid studied by Mullins represents the limit $Q = h_0 q \to \infty$, i.e., Mullins's result applies only when the "fluid" layer thickness is much larger than the wavelength $\lambda = \frac{2\pi}{q}$ of the perturbation under consideration. In many pattern-forming systems, however, this is not true. To determine the effect of viscous flow in a film of *finite* thickness

requires solving the Stokes equations in a domain bounded both above and below. This was first accomplished by Orchard,[26] who found that

$$R_{\text{Orchard}}(Q) = -\frac{\gamma}{2\eta h_0} \frac{Q[\sinh(2Q) - 2Q]}{[1 + 2Q^2 + \cosh(2Q)]},$$ (24)

where again $Q = h_0 q$ is the viscous film thickness times the wavenumber, γ is the surface free energy, and η^{-1} is the ion-enhanced fluidity. In the limit $q \to \infty$, one recovers Mullins's result $R \approx -\frac{\gamma}{2\eta} q$. However, during low-energy ion bombardment, the opposite limit $q \to 0$ is much more relevant, and in that limit, one obtains a very different approximation of the form $R \approx \frac{\gamma h_0^3}{3\eta} q^4$ (it is worth noting briefly that these two approximations should never appear together in the same model). Notably, the dispersion relation in this limit of "surface-confined" viscous flow exhibits the same power dependency as that of surface diffusion for the small values of q, but in contrast to the mechanism of surface diffusion, it is only weakly dependent on temperature. These observations were first made by Umbach,[27] who found that replacing surface diffusion with surface-confined IVF enabled much better agreement with experimental data at low temperatures. These findings have been confirmed by several subsequent studies (see, e.g., Refs. 28 and 137)

2. Response to stress

It has long been known that in addition to enhancing the fluidity of the target surface layer, ion irradiation also creates significant stress within that layer. This phenomenon has been observed for both high-energy[21–24,138] and low-energy[25,49,139–141] irradiation. If IVF is an effective or even dominant means of relaxing surface free energy, how might it interact with the presence of these significant stresses? Two early studies of this type deserve mention. First, it was proposed by Alkemade that the different *local* ion incidence angles on the front and back of a ripple would lead to differential heating and hence a slope-dependent thermal expansion,[142] influencing the propagation direction of surface ripples formed through other processes. At about the same time, Davidovitch *et al.* proposed that stress could drive an instability as in the Asaro-Tiller-Grinfeld-Srolovitz (ATGS) mechanism,[143–145] in an attempt to explain anomalous pattern formation near normal incidence.[18] Although neither of these mechanisms was supported by later studies (and the anomalous patterns were found to be an artifact[31]), they pioneered the idea that nanopatterns on ion-irradiated targets could potentially be driven by processes beyond the collision cascade.

Since these early attempts, several works have appeared treating the amorphous film as a very viscous fluid, on which the ion beam acts through mechanical means. For instance, Castro and co-workers proposed a mechanism in which the ion beam was proposed to act as an "effective body force," exerting a gravity-like force in the ion beam direction throughout the amorphous film.[146,147] With appropriate selection of the irradiation angle dependence of the body force coefficient, ripples were predicted to appear at angles above 45°, and with

appropriate choice of force coefficient, good agreement with experimentally observed wavelengths was reported. More recently, a model more closely tied to the underlying physical processes was proposed by Norris.[46,148] Drawing on earlier studies on keV-range ion bombardment[138,139,149–151] and applications of those ideas to lower energies,[152,153] he proposed that, rather than acting as a body force on the film, the atomic re-arrangements of the collision cascade induce a "stress-free strain," adding not a force term but rather an additional contribution to the stress-strain relationship, describing so-called "anisotropic plastic flow" (APF). As described in Ref. 148, this model has a dispersion relation of the form

$$R_{\text{APF}}(Q) = -6fA \frac{\cos(2\theta)(h_0 q_x)^2 + \cos^2(\theta)(h_0 q_y)^2}{[1 + 2Q^2 + \cosh(2Q)]},$$ (25)

where f is the flux (through a plane normal to the ion beam) and A is a coefficient describing the induced plastic deformation per unit of flux. This result exhibits similar agreement with experiments using a more traditional continuum modeling approach, more direct coupling to the underlying physical mechanisms, and fewer assumptions. Indeed, all parameters needed for a wavelength prediction were measured experimentally in the same chamber, resulting in a fully parameter-free prediction of the pattern wavelength that also agreed very well with experimental results (see Fig. 9).

Subsequent work in this direction has shown great promise in predicting observed behaviors over a range of experimental parameters. For example, it was found that the body force model, when coupled with a longstanding model of the dependence of the in-film stress on the ion energy,[154] predicted remarkably well the scaling of the ripple wavelength as a function of ion energy observed in the linear regime.[155] More recently, several of the same authors showed that, when coupled to estimations of the stress tensor for different materials
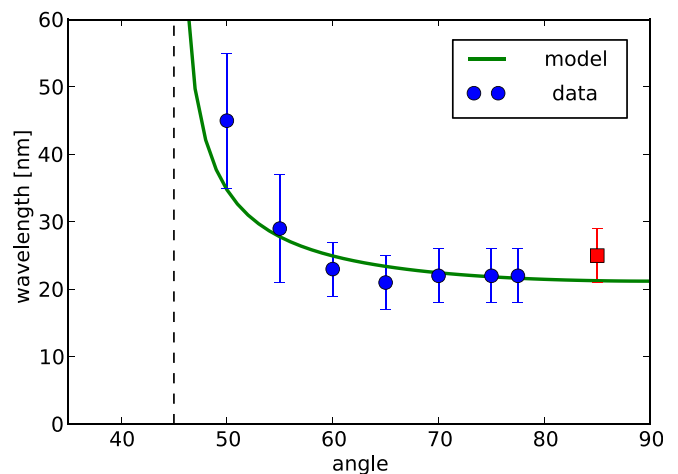

FIG. 9. Comparison of data and the predictions of the anisotropic plastic flow model of Ref. 148 for the system 250 eV Ar$^+$ → Si, with all required model parameter groups determined experimentally. Agreement with parallel-mode ripple wavelengths is excellent, but as in Fig. 8, perpendicular-mode ripples are not predicted.

obtained by molecular dynamics, the model was also consistent with the material dependence of the angular response.[156]

B. Parameter evaluation

The two parameters most relevant to the mechanisms described here are the ion-enhanced fluidity η^{-1} and the coefficient of strain per unit of flux A. In Ref. 148, direct measurements of these parameters were avoided because predicted pattern wavelengths could be expressed solely in terms of the magnitude of the steady stress $|\mathbf{T}_0| = 6\eta fA$. Measurements of $|\mathbf{T}_0|$ significantly pre-date the continuum models described above, and this quantity can be obtained either experimentally by wafer curvature measurements[21,23,134,140,141,157–159] or numerically using molecular dynamics simulations.[160–163] The former approach was used in Ref. 148 to determine all parameters needed for wavelength predictions to several significant digits. Consequently, the agreement between the model and data in Fig. 9 is obtained without *any* parametric adjustments.

Regardless of these considerations, however, direct values of η^{-1} and A are obviously highly desirable. Unfortunately, measurements or estimates of these parameters are non-trivial. Otani *et al.* obtained a value of A by comparing Scanning Electron Microscopy (SEM) images of irradiated trench structures to a simulation of the same using a continuum viscoelastic model. Because they were able to achieve very good agreement with experimental results by tuning the value of A, they presented this method as a means of evaluating A.[138] Mayr *et al.* obtained a value of η^{-1} by fitting a model of target evolution to the behavior of molecular dynamics simulations.[28,164] Neither of these methods have seen widespread adoption. In particular, most later works requiring an estimate of η^{-1} have simply re-scaled the results in Ref. 28 to different ion energies.[111,120,165]

In most of the work described above, it has been customary to assume that the mechanisms driving the ion-induced pattern formation process depend on the ion flux only linearly within rate constants–in other words, increasing the flux speeds up the pattern formation process but does not otherwise change the properties of the irradiated system. In the context of continuum models described in this section, this view leads to the common assumption that the fluidity η^{-1} (effectively a rate constant in viscous flow) is proportional to the flux,[21,164] which in turn implies that the steady stress $|\mathbf{T}_0| = 6\eta fA$ would be independent of flux. Recently, however, Ishii *et al.* performed a detailed study on the flux-dependence of the stress using wafer curvature measurements and found that these assumptions are not valid.[141] Instead, they found good agreement between their measurements and a model in which the fluidity is proportional to a concentration of "flow defects," which itself evolves according to a differential equation containing a bi-molecular recombination mechanism. For a given ion energy, a single fitted version of these equations was able to explain observed stresses over a wide variety of fluxes. The fit parameters included quantities sufficient to calculate both η^{-1} and A, making this a promising method for the direct experimental evaluation of these parameters.

C. Discussion and open questions

1. Predicted transition angle

The agreement between predicted and observed patterns in Fig. 9 obscures an important problem–although the dispersion relation (25) predicts a transition from flat surfaces to parallel-mode ripples at 45°, silicon transitions near this angle only by chance, and in general, the transition angle varies by ion/target combination. Therefore, some mechanism enabling different transition angles for different materials is required. One possibility is described in a recent study by Moreno-Barrado *et al.*,[156] who introduced a depth-dependence to the magnitude of the stress field and found that this dependence could modify the transition angle; with simulation data for impacts on both Si and Ge, these authors report good agreement with observations. A second possibility is that the assumptions of a rotating, divergence-free stress tensor, although confirmed in limited testing in Ref. 156, could nevertheless be violated–more studies to confirm this finding would be a welcome addition to the literature. Finally, although the works described here did not need to include the effects of erosion and redistribution to obtain good agreement with experimental results, such effects are undoubtedly still present and may shift the transition angle from the value predicted by stress effects alone.

2. Relationship to mass redistribution

The stress-driven viscous flow mechanism exhibits several similarities with the mechanism of mass redistribution described above in Secs. III–IV. Both mechanisms cause a net flux of mass in the "downbeam" direction. In both cases, the strength of that flux depends on the incidence angle, and indeed, the angle dependence of the dispersion relation (25) for the stress mechanism in both x- and y-directions is identical to that seen for the redistributive terms in the dispersion relation (13) combining the results of Bradley and Harper with those of Carter and Vishnyakov and Davidovitch. It is tempting, therefore, to conclude that perhaps the two approaches are simply different ways to describe the same underlying phenomenon. However, such a conclusion does not seem justified. First, the striking similarity follows from the assumptions of each approach–both suggest a mechanism with strength proportional to the incoming ion flux [which varies as $\cos(\theta)$] and also to the magnitude of the projection of the ion beam direction onto the surface [which varies as $\sin(\theta)$]. Second, the two mechanisms act on very different time scales. Mass redistribution occurs in the "prompt" regime of impact lasting only picoseconds, during which the atom velocity distribution has not fully relaxed to the Maxwell distribution. This impact leaves behind significant stresses, which then relax over the "gradual" regime lasting seconds.

3. Ripple rotation at grazing incidence

One of the fundamental successes of the original Bradley-Harper model was its ability to predict the oft-observed transition from "parallel mode" to "perpendicular mode" ripples as $\theta \to 90°$.[16] None of the alternative mechanisms described above,

whether mass redistribution or stress-driven viscous flow, has been able to explain this transition. Again, it is very likely that erosive effects, although neglected in this section, are nevertheless still critical at grazing angles and therefore must be included to explain the full range of angle-dependent behavior. This will be explored in more detail below.

4. Ripple disappearance at higher energies

As suggested by Fig. 2, pattern formation at ion energies below 1 keV seems to depend very little on the irradiation energy E.[31] However, experiments suggest that a significant energy dependence emerges in the tens of keV range and that at some energies, all incidence angles are stable.[121] This is somewhat puzzling, as all the physical mechanisms described above are unstable for at least some values of θ. This suggests the possible existence of some unconditionally stable mechanism that becomes important as energy increases. The term $D_{11}(\theta)\frac{\partial^2 h}{\partial x^2}$ proposed in Refs. 120 and 128 would seem to provide the needed stability,[121,129] but given the concerns about its derivation discussed above and the successes of stress-based models described here, it is natural to inquire whether some mechanism associated with stress could also supply such a term. Recently, inspired by images of expanding bubbles in cross-sectional TEM images such as those found in the work of Chini et al.,[33] Swenson and Norris studied the effect of ion induced *swelling*, an effect omitted from prior treatments of (purely shear) stress.[166] Under appropriate simplifying assumptions, these authors found that swelling could indeed be stabilizing at all angles for wavenumbers typically observed in experiments. This mechanism may therefore be useful in reconciling experiments with theory at ion energies above 1 keV.

VI. COMPARING SCATTERING SPECTRA TO LINEAR DISPERSION RELATIONS

In Secs. III–V, we have discussed several distinct approaches to understanding ion-induced pattern formation. Historically, there has been a tendency–whether in the high-order expansions of the Sigmund model of sputter erosion,[72,80] the early enthusiasm surrounding the discovery that mass redistribution could be the dominant contributor to the crater function,[111,165] or increasingly sophisticated models of stress-driven viscous flow[155,156]–to seek a complete explanation of observed pattern formation phenomena from a single physical mechanism. Such explanations are often sought after some notable early success of a new approach, such as the Bradley-Harper theory's ability to account for wavelength selection and ripple rotation,[16] the Crater Function Framework's revelation that mass redistribution could be a dominant mechanism at low energies,[111] or the realization of quantitatively accurate, parameter-free wavelength predictions through models of ion-induced stress.[147,148]

Although simple explanations for observed phenomena are highly desirable, unfortunately, no single approach has succeeded in explaining the richness of observed behaviors in ion-induced nanopatterning systems–even the relatively simple variety of behaviors observed in Fig. 2. The unconditional instability suggested by the Sigmund model is contradicted by the presence of flat surfaces near normal incidence.[17] When the surfaces are unstable as predicted, the observed growth rate can significantly exceed theoretical predictions.[167] The temperature-dependence of a wavelength selected by surface diffusion is inconsistent with observations at low temperatures.[8] The dominance of redistribution at low energies does not continue inevitably as the ion energy increases.[120] Models of ion-induced stress have been unable to accommodate ripple rotation near grazing incidence.[148] Except in limiting regimes (high temperature, low energy, normal incidence, grazing incidence, etc.), no single mechanism seems to dominate the pattern formation process.

A full accounting of observed behavior, then, must require the presence of multiple competing (or cooperating) mechanisms, and a central question is then how to distinguish and measure the relative contributions of each. A potentially powerful approach to distinguishing between different physical mechanisms is to compare the amplification rate $R(\mathbf{q})$ predicted by a linear theory to measurements of surface scattering obtained via Grazing Incidence Small-Angle X-ray Scattering (GISAXS), which correlate with the surface structure factor $S(\mathbf{q}, t)$ when the surface structure amplitudes are small.[165,168] More specifically, given an evolving height profile $h(x, y, t)$, the linearization of the governing equations would predict that the Fourier transform $\tilde{h}(q_x, q_y, t)$ of the surface height profile satisfies the differential equation

$$\frac{\partial \tilde{h}(\mathbf{q}, t)}{\partial t} = R(\mathbf{q})\tilde{h}(\mathbf{q}, t) + \beta(\mathbf{q}, t), \qquad (26)$$

where $\mathbf{q} = \langle q_x, q_y \rangle$ is the wavevector; the first term on the right describes the exponential growth associated with the linear amplification rate $R(\mathbf{q})$ and $\beta(\mathbf{q}, t)$ is the Fourier transform of an added stochastic noise term. As described in Refs. 165 and 168, this assumption leads to a structure factor $\langle S(\mathbf{q}, t) \rangle$ with the form

$$\langle S(\mathbf{q}, t) \rangle = \left[S(\mathbf{q}, 0) + \frac{\alpha(\mathbf{q})}{2R(\mathbf{q})} \right] \exp\left[2R(\mathbf{q})t \right] - \frac{\alpha(\mathbf{q})}{2R(\mathbf{q})}, \qquad (27)$$

where $\alpha(\mathbf{q})$ is the magnitude of the noise. Therefore, given GISAXS measurements of the structure factor $\langle S(\mathbf{q}, t) \rangle$ at early times, one can fit Eq. (27) to the data to reconstruct the linear amplification rate $R(\mathbf{q})$ and then compare this to linear models of the pattern-formation process. This process provides a rigorous means of model testing and validation, and if the model is faithful enough to the actual physical processes, it allows *experimental extraction* of model parameters.

In practice, however, this approach must overcome an important obstacle. GISAXS spectra exhibit a meaningful variation in time–necessary for a good fit of Eq. (27) to the measured behavior at a single wavenumber–for only a limited range of relatively small wavenumbers. At first, this seems fortuitous, as in this wavenumber range, linear amplification rates are reasonably well-approximated by convenient quartic polynomials with the form of Eq. (6). Unfortunately, however, in this range, both collisional models and stress-based models first produce terms at order q^2, whereas both surface diffusion and surface-confined

viscous flow appear at order q^4. These similarities mean that GISAXS spectra in this wavenumber range cannot distinguish between these sets of mechanisms, which frustrates efforts to compare their relative strengths.

A solution, introduced in Ref. 137 and made available to the community through the open-source library PyGLIDRE ("A Python library for GISAXS LInear Dispersion Relation Extraction"),[169] is to fit Eq. (27) to all wavenumber series at once, with $\alpha(\mathbf{q})$ assumed to have a single constant value across all q (i.e., the experimental noise is assumed to be white noise). With careful attention to noise reduction in both the experimental and data analysis stages, it was demonstrated that meaningful measurements of the dispersion relation could be extracted which extended to high values of q and clearly revealed that the decay rate of high-wavenumber modes grew at most linearly in q, invalidating the use of quartic polynomials of the form (6) as a tool for comparison with the data. Instead, the data require "full-spectrum" models retaining the full wavenumber dependence for comparison.

Fortunately, although they have received less attention than their longwave approximations, full-spectrum models exist for many of the physical mechanisms discussed above. As noted briefly, Bradley recently obtained the full-spectrum dispersion relation associated with Sigmund's model of sputtering.[81] In addition, continuum results on stress-driven viscous flow[147,148] and Orchard's result for viscous relaxation of surface energy[26] also retain the full wavenumber dependence. Therefore, the authors of Ref. 137 compared measured GISAXS spectra to the composite linear dispersion relation

$$R(q) = \left[R_{\text{displ.}}(q) + R_{\text{stress}}(q)\right] + R_{\text{Orch.}}(q), \qquad (28)$$

where

$$R_{\text{displ.}}(q) = -C_{\text{X,Y}}^{\text{displ.}}(\theta)\left(1 - \exp\left(-\frac{1}{2}(Dq)^2\right)\right),$$

$$R_{\text{stress}}(q) = -6fAC_{\text{X,Y}}^{\text{stress}}(\theta)\frac{Q^2}{[1 + 2Q^2 + \cosh(2Q)]}, \qquad (29)$$

$$R_{\text{Orch.}}(q) = -\frac{\gamma\eta^{-1}(\theta)}{2h_0}\left(\frac{Q[\sinh(2Q) - 2Q]}{1 + 2Q^2 + \cosh(2Q)}\right).$$

Here, $R_{\text{displ.}}(q)$ is a slight simplification of the dispersion relation associated with the Sigmund model obtained in Ref. 81 [it was hypothesized that although erosion and redistribution exhibit a very different dependence on the angle of incidence θ, they might exhibit the same dependence on the wavenumber q, allowing both effects to be included in this term], $R_{\text{stress}}(q)$ is the dispersion relation associated with anisotropic flow, obtained in Ref. 148, and $R_{\text{Orch.}}(q)$ is again Orchard's seminal result for the dispersion relation associated with surface leveling of a viscous film due to surface tension.[26]

The results of this approach were remarkable. Fitting the composite model (28) and (29), via the relation (27), to the expanded GISAXS data produced very good agreement with the original GISAXS spectra over the entire range of wavenumbers and irradiation angles studied (see Fig. 10). Moreover, for each of the model parameters in Eq. (29), this procedure produced estimates that were in reasonable to excellent agreement with available external simulations, experiments, or theoretical predictions of the model parameters. For instance, the fitted value of the film thickness $h_0(\theta)$ was identical to predictions based on BCA simulations, the fitted value of the stress coefficients $C_{\text{X,Y}}^{\text{stress}}(\theta)$ closely tracked the $\{\cos(2\theta),\ \cos^2(\theta)\}$ form predicted by Eq. (25), the inferred steady stress exactly matched the value measured experimentally by wafer curvature studies, and the fitted coefficients $C_{\text{X,Y}}^{\text{displ.}}(\theta)$ were a very reasonable match to the values obtained through the crater function framework.

The agreement in so many cases between the fitted and measured values of the model parameters suggests that Eq. (28) is a reasonably complete description of the dynamics of the irradiated surface and that the process of fitting Eq. (28) to GISAXS spectra may be seen as a means of *experimentally extracting* model parameters. This is especially important in the case of parameters for which external estimates are unavailable. In particular, Ref. 137 produced the first direct comparison of the relative magnitudes of stress and displacement effects (now distinguishable at higher values of q). Indeed, it was found that the magnitudes of the stress coefficient significantly exceed those of the displacement coefficient at angles below about $\theta = 60°$, suggesting that for this experimental system, anyway, stress is the dominant cause of surface smoothing near normal incidence and the transition to parallel-mode ripples around 45°. At higher angles, however, (presumably erosive) displacements seemed to remain a critical contributor to perpendicular-mode instabilities near grazing incidence.

In addition, this approach produced the first experimental estimate of the ion-enhanced fluidity $\eta^{-1}(\theta)$ and found it to be higher than estimated previously by scaled MD results assuming a linear dependence on flux. This result may resolve a puzzle described above—the apparent ability of both stress and redistribution to effectively predict observed pattern wavelengths for $250\,\text{eV Ar}^+ \to \text{Si}$. Whereas the wavelength predictions in Ref. 111 due to mass redistribution relied directly on the estimated values of η^{-1}, the predictions in Ref. 148 based on anisotropic plastic flow were able to avoid estimating η^{-1} because the predicted wavelength ended up depending only on the magnitude of the steady stress $|\mathbf{T}_0|$, which can be obtained directly via wafer curvature measurements. Hence, while the accuracy of the redistribution model is vulnerable to changes in estimates of η^{-1}—which is poorly understood—the accuracy of the anisotropic plastic flow model is not. The results here suggest that the former results were, indeed, based on an under-estimate of η^{-1} and that the latter are therefore more robust.

VII. CONCLUSIONS AND OUTSTANDING QUESTIONS

The past 10 years have seen a remarkable advance in understanding the ion-induced nanopatterning process. The classic results of pattern formation theory have been integrated into the field, in particular the relationship between the type of linear instability and the anticipated degree of order in the resulting pattern. The mechanistic Sigmund model of erosion has been extended into the full-spectrum linear regime. The Crater Function Framework has emerged as a robust means of
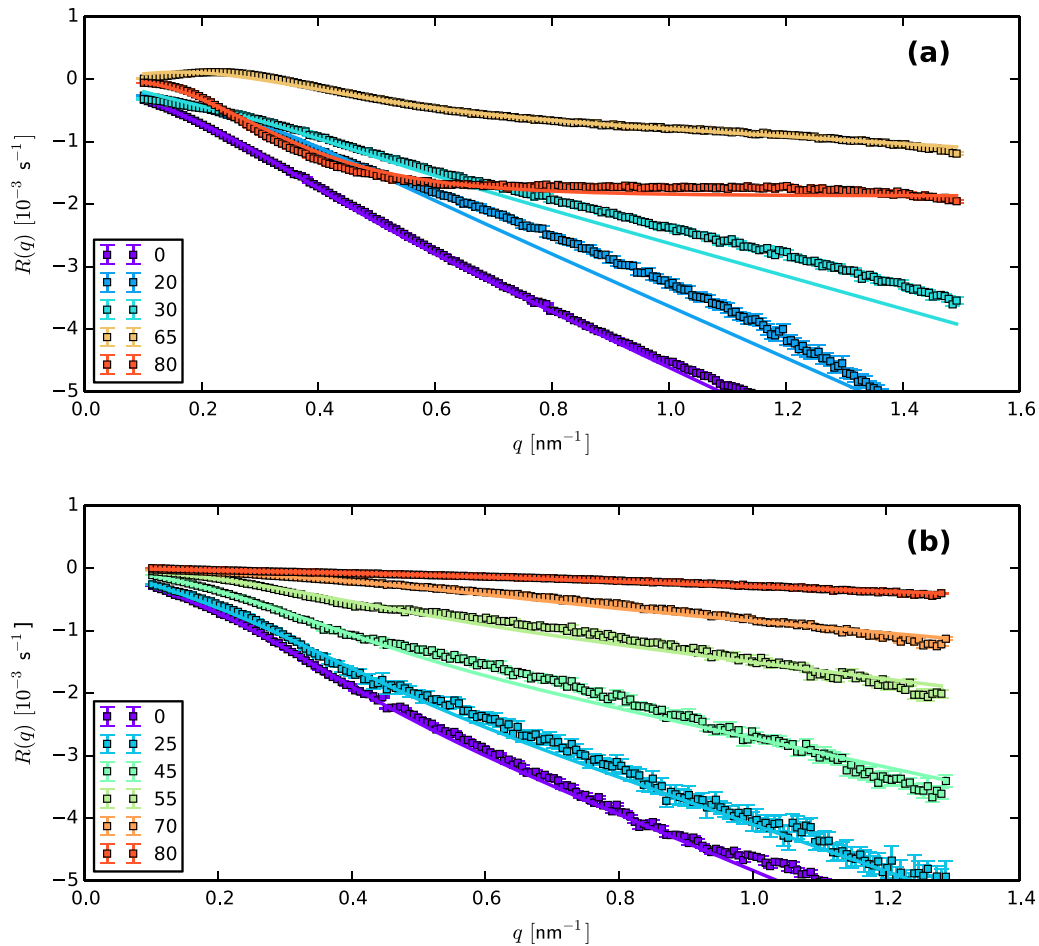
FIG. 10. Amplification rates inferred from GISAXS for the system 1 keV $Ar^+ \rightarrow$ Si. (a) Extracted values of $R(q_x, 0)$ for irradiation angles $\theta \in$ [0, 20, 35, 45, 65, and 80] over the range $q \in$ [–1.5, 1.5] nm^{-1} and (b) extracted values of $R(0, q_y)$ for irradiation angles of $\theta \in$ [0, 25, 45, 55, 70, 80, and 85] over the range $q \in$ [–1.25, 1.25] nm^{-1} are shown. In each case, the extracted values of R are indicated by squares, whereas fits of those data to the composite model in Eqs. (28) and (29) are shown by solid lines. Reproduced with permission from S. A. Norris *et al.*, Sci. Rep. **7**, 2016 (2017). Copyright 2017 Scientific Reports.

estimating the combined contribution of sputter erosion and mass redistribution in the longwave limit. The commonly observed buildup of stress in irradiated films has been revealed as an important, and perhaps even dominant, contributor to ion-induced nanopatterning. Finally, GISAXS has been shown to be a powerful tool to distinguish the relative contributions of different physical mechanisms by comparing the predictions of full-spectrum linear models with observed spectra at early times. Nevertheless, due to the complexity of multiple active mechanisms with multiple parametric dependencies, the goal of a robust predictive model remains frustratingly out of reach. We therefore conclude this paper with several concerns that transcend the individual approaches described above.

A. Simple fitted models are dangerous

In light of the unusually large number of physical mechanisms that are important to the nanopatterning process, we believe that a critical need going forward is a continued focus

on the methods of evaluating model parameters as a means to eliminate fit coefficients. This concern is illustrated by two recent series of studies that report success in accounting for the presence and wavelength of patterns across a range of ion energies and target materials. One is based solely on the Crater Function Framework,[120,123,128] whereas the other is based solely on Effective Body Force models of stress.[155,156] It is at first puzzling that such different approaches could each claim such success. However, upon deeper examination, one observes that the latter studies implicitly used the ion-enhanced fluidity η^{-1} as a free parameter, and the former included several ad-hoc modifications to the standard crater function framework. In each case, these decisions enabled better fits to data and, in a simpler, better understood system, would be *very reasonable*. However, in a system which has been repeatedly frustrated by the emergence of new relevant physical mechanisms, such an approach becomes dangerous, offering a false sense of confidence and frustrating the interpretation of competing explanations for

observed behavior. [As a cautionary tale, we note that there are several examples in the literature of studies that "succeeded" in explaining a regime of dots in the low-angle, low-energy corner of Fig. 2, which ultimately turned out to be an experimental artifact.] A careful study on the values of parameters such as η^{-1} in common experimental systems, although potentially difficult, would do much to clear up this confusion.

B. Needed: Efficient continuum parameter evaluation

Although, for some time, the evaluation of parameters associated with the collision cascade has been enabled by simulation methods, the evaluation of parameters associated with stress and viscous effects has been more problematic. In Sec. VI, we reviewed work showing that with an appropriate strategy for interpreting data at high wavenumbers, early-time GISAXS spectra could be a powerful tool for evaluating these parameters and distinguishing between physical mechanisms for which full-spectrum linear models exist. However, GISAXS experiments are expensive and time-consuming. A different strategy to experimentally distinguish mechanisms such as displacements and stress—which differ only at high wavenumbers—would be highly desirable. Similarly, a straightforward method to evaluate parameters associated with stress generation and ion-enhanced fluidity is needed. The method of Ishii et al.,[141] of fitting a "flow defect" model to wafer curvature measurements, seems highly promising in this light (indeed, the values extracted from GISAXS spectra in Ref. 137 exhibit very good agreement with values inferred from Ref. 141). It would also be desirable to extract parameters for this model from MD or BCA simulation, in the spirit of Refs. 28 and 164, and to systematize the variety of MD studies of stress.[156,160–163] Ideally, these would be packaged with something like the PyCraters library,[130] allowing a user to write a simple script that automates the performance of all simulations needed for a certain experimental system and performs any subsequent fitting and calculation required to obtain the needed parameter values.

C. Nonlinear terms: Bridging the gap

The work described in this review has focused primarily on the linear regime and is therefore applicable only to the early stages of the pattern formation process. As we have seen, the linear regime is entirely capable of providing much insight in the quest to identify the presence and strength of various relevant physical mechanisms, especially when compared with GISAXS spectra. However, to address questions on late-stage pattern formation (e.g., on saturation, coarsening, and defect interaction), nonlinear effects are of course required. The task here seems to be to "bridge the gap" between the top-down longwave models such as those in Refs. 64–67 on the one hand and bottom-up full-spectrum models such as Eqs. (28) and (29) on the other hand. The former, when fitted to experimental results, can readily produce astonishingly good visual agreement with experiment, and it is therefore quite likely that the terms used therein are indeed the dominant nonlinear terms. However, the work reviewed here collectively suggests that each mechanism in the latter is necessary to produce good agreement with GISAXS spectra in the

linear regime. This suggests that the remaining task is now to rigorously extend models of *each one of the relevant mechanisms* into the nonlinear regime to build a physical basis for the values of the coefficients of nonlinear terms.

APPENDIX: DERIVATION OF THE GENERIC CRATER FUNCTION FRAMEWORK

Equation (16) is a slight variation on the main result in Ref. 107, in which we have abandoned the use of "effective" moments, and the result is therefore only valid in the linear regime. In justifying this minor change, we take the opportunity to reproduce the derivation of Eq. (16) in more detail than appeared in Ref. 107. Specifically, limiting our attention to one spatial dimension, we show that

$$
\begin{aligned}
\frac{\partial h}{\partial t} &= \int I(\varepsilon u)\Delta h(-u; \mathcal{S}(\varepsilon u))\, du \\
&= \int R(u, -\varepsilon u)\, du \\
&= \int \left\{ [R]_{U=0} - (\varepsilon u)\left[\frac{\partial R}{\partial U}\right]_{U=0} + \cdots \right\} du \\
&= \int \left\{ [R]_{U=0} - \varepsilon \frac{\partial}{\partial U}[uR]_{U=0} + \cdots \right\} du \\
&= \left[\int R\, du\right]_{U=0} - \varepsilon \frac{\partial}{\partial U}\left[\int uR\, du\right]_{U=0} + \cdots \\
&= I(U)\int \Delta h(u; \mathcal{S}(U))\, du - \varepsilon \frac{\partial}{\partial U}\left[I(U)\int u\Delta h\, du\right] + \cdots \\
&= IM^{(0)} - \varepsilon \frac{\partial}{\partial U}\left[IM^{(1)}\right] + \cdots \\
&= IM^{(0)} - \frac{\partial}{\partial u}\left[IM^{(1)}\right] + \cdots.
\end{aligned}
$$

The first line simply repeats Eq. (15). In the second line, we have simplified the integrand to the generic multi-scale form $\int R(u, -\varepsilon u)du$ and made the substitution $u \rightarrow -u$. In the third line, we have performed a Taylor series in the small variable ε and used U to refer to the second argument of R. The key conceptual step is then to notice that on the right-hand side of that line, the symbols u and U are *formally independent*—the derivatives are taken with respect to the second argument of R, while the subsequent integral is taken only in the first argument. Therefore, u and U may be treated *as if* they were unrelated variables, allowing us to pull first the multiplication by u (fourth line) and then the integral in u (fifth line) into the derivatives with respect to U. In the sixth line, we recall the definition of R, and in the seventh line, note that each term contains a moment of the crater function Δh. The eighth line,

not previously published, reveals that ε was all along a tool used to manipulate the integral and does not appear explicitly in the final result.

REFERENCES

[1] I. Yamada, J. Matsuo, N. Toyoda, and A. Kirkpatrick, "Materials processing by gas cluster ion beams," Mater. Sci. Eng. R **34**, 231–295 (2001).

[2] B. Ziberi, F. Frost, T. Höche, and B. Rauschenbach, "Ripple pattern formation on silicon surfaces by low-energy ion-beam erosion: Experiment and theory," Phys. Rev. B **72**, 235310 (2005).

[3] S. Facsko, T. Dekorsy, C. Koerdt, C. Trappe, H. Kurz, A. Vogt, and H. L. Hartnagel, "Formation of ordered nanoscale semiconductor dots by ion sputtering," Science **285**, 1551–1553 (1999).

[4] M. J. Baldwin and R. P. Doerner, "Helium induced nanoscopic morphology on tungsten under fusion relevant plasma conditions," Nucl. Fusion **48**, 035001 (2008).

[5] H. H. Chen, O. A. Urquidez, S. Ichim, L. H. Rodriquez, M. P. Brenner, and M. J. Aziz, "Shocks in ion sputtering sharpen steep surface features," Science **310**, 294–297 (2005).

[6] J. Li, D. Stein, C. McMullan, D. Branton, M. J. Aziz, and J. A. Golovchenko, "Ion-beam sculpting at nanometre length scales," Nature **412**, 166 (2001).

[7] M. Moseler, P. Gumbsch, C. Casiraghi, A. C. Ferrari, and J. Robertson, "The ultrasmoothness of diamond-like carbon surfaces," Science **309**, 1545–1548 (2005).

[8] W. L. Chan and E. Chason, "Making waves: Kinetic processes controlling surface evolution during low energy ion sputtering," J. Appl. Phys. **101**, 121301 (2007).

[9] B. Ziberi, F. Frost, and B. Rauschenbach, "Formation of large-area nanostructures on Si and Ge surfaces during low energy ion beam erosion," J. Vac. Sci. Technol. A **24**, 1344 (2006).

[10] B. Ziberi, F. Frost, M. Tartz, H. Neumann, and B. Rauschenbach, "Ripple rotation, pattern transitions, and long range ordered dots on silicon by ion beam erosion," Appl. Phys. Lett. **92**, 063102 (2008).

[11] Q. Wei, J. Lian, S. Zhu, W. Li, K. Sun, and L. Wang, "Ordered nanocrystals on argon ion sputtered polymer film," Chem. Phys. Lett. **452**, 124–128 (2008).

[12] A. Cuenat, H. B. George, K.-C. Chang, J. M. Blakely, and M. J. Aziz, "Lateral templating for guided self-organization of sputter morphologies," Adv. Mater. **17**, 2845–2849 (2005).

[13] B. N. Singh and S. J. Zinkle, "Defect accumulation in pure fcc metals in the transient regime: A review," J. Nucl. Mater. **206**, 212–229 (1993).

[14] P. Sigmund, "Theory of sputtering. I. Sputtering yield of amorphous and polycrystalline targets," Phys. Rev. **184**, 383–416 (1969).

[15] P. Sigmund, "A mechanism of surface micro-roughening by ion bombardment," J. Mater. Sci. **8**, 1545–1553 (1973).

[16] R. M. Bradley and J. M. E. Harper, "Theory of ripple topography induced by ion bombardment," J. Vac. Sci. Technol. **6**, 2390–2395 (1988).

[17] G. Carter and V. Vishnyakov, "Roughening and ripple instabilities on ion-bombarded Si," Phys. Rev. B **54**, 17647–17653 (1996).

[18] B. P. Davidovitch, M. J. Aziz, and M. P. Brenner, "On the stabilization of ion sputtered surfaces," Phys. Rev. B **76**, 205420 (2007).

[19] C. Herring, "Effect of change of scale on sintering phenomena," J. Appl. Phys. **21**, 301 (1950).

[20] W. W. Mullins, "Theory of thermal grooving," J. Appl. Phys. **28**, 333–339 (1957).

[21] C. A. Volkert, "Stress and plastic flow in silicon during amorphization by ion bombardment," J. Appl. Phys. **70**, 3521 (1991).

[22] E. Snoeks, T. Weber, A. Cacciato, and A. Polman, "MeV ion irradiation-induced creation and relaxation of mechanical stress in silica," J. Appl. Phys. **78**, 4723 (1995).

[23] M. L. Brongersma, E. Snoeks, T. van Dillen, and A. Polman, "Origin of MeV ion irradiation-induced stress changes in SiO$_2$," J. Appl. Phys. **88**, 59–64 (2000).

[24] S. G. Mayr and R. S. Averback, "Ion-irradiation-induced stresses and swelling in amorphous Ge thin films," Phys. Rev. B **71**, 134102 (2005).

[25] W. L. Chan and E. Chason, "Stress evolution and defect diffusion in Cu during low energy ion irradiation: Experiments and modeling," J. Vac. Sci. Technol. A **26**, 44 (2008).

[26] S. E. Orchard, "On surface levelling in viscous liquids and gels," Appl. Sci. Res. A **11**, 451 (1963).

[27] C. C. Umbach, R. L. Headrick, and K.-C. Chang, "Spontaneous nanoscale corrugation of ion-eroded SiO$_2$: The role of ion-irradiation-enhanced viscous flow," Phys. Rev. Lett. **87**, 246104 (2001).

[28] S. Vauth and S. G. Mayr, "Relevance of surface viscous flow, surface diffusion, and ballistic effects in kev ion smoothing of amorphous surfaces," Phys. Rev. B **75**, 224107 (2007).

[29] C. S. Madi, B. P. Davidovitch, H. B. George, S. A. Norris, M. P. Brenner, and M. J. Aziz, "Multiple bifurcation types and the linear dynamics of ion sputtered surfaces," Phys. Rev. Lett. **101**, 246102 (2008).

[30] C. S. Madi, H. B. George, and M. J. Aziz, "Linear stability and instability patterns in ion-sputtered silicon," J. Phys. Condens. Matter **21**, 224010 (2009).

[31] C. S. Madi and M. J. Aziz, "Multiple scattering causes the low energy-low angle constant wavelength bifurcation of argon ion bombarded silicon surfaces," Appl. Surf. Sci. **258**, 4112–4115 (2012).

[32] J. Muñoz-García, L. Vázquez, M. Castro, R. Gago, A. Redondo-Cubero, A. Moreno-Barrado, and R. Cuerno, "Self-organized nanopatterning of silicon surfaces by ion beam sputtering," Mater. Sci. Eng. R **86**, 1–44 (2014).

[33] T. K. Chini, F. Okuyama, M. Tanemura, and K. Nordlund, "Structural investigation of keV Ar-ion-induced surface ripples in Si by cross-sectional transmission electron microscopy," Phys. Rev. B **67**, 205403 (2003).

[34] M. C. Cross and P. C. Hohenberg, "Pattern formation outside of equilibrium," Rev. Mod. Phys. **65**, 851–1123 (1993).

[35] M. Cross and H. Greenside, *Pattern Formation and Dynamics in Nonequilibrium Systems* (Cambridge University Press, 2009), ISBN: 0521770505.

[36] A. Oron, S. H. Davis, and S. G. Bankoff, "Long-scale evolution of thin liquid films," Rev. Mod. Phys. **69**, 931–980 (1997).

[37] J. N. Wilking, T. E. Angelini, A. Seminara, M. P. Brenner, and D. A. Weitz, "Biofilms as complex fluids," MRS Bull. **36**, 385–391 (2011).

[38] M. Kardar, G. Parisi, and Y. C. Zhang, "Dynamic scaling of growing interfaces," Phys. Rev. Lett. **56**, 889–892 (1986).

[39] T. Tsuzuki and Y. Kuramoto, "Persistent propagation of concentration waves in dissipative media far from thermal equilibrium," Prog. Theor. Phys. **55**, 356–369 (1976).

[40] G. I. Sivashinski, "Nonlinear analysis of hydrodynamic instability in laminar flames I. Derivation of basic equations," Acta Astronaut. **4**, 1177–1206 (1977).

[41] F. Frost, B. Ziberi, T. Höche, and B. Rauschenbach, "The shape and ordering of self-organized nanostructures by ion sputtering," Nucl. Instrum. Methods Phys. Res. B **216**, 9–19 (2004).

[42] R. Gago, L. Vásquez, R. Cuerno, M. Varela, C. Ballesteros, and J. María Albella, "Production of ordered silicon nanocrystals by low-energy ion sputtering," Appl. Phys. Lett. **78**, 3316 (2001).

[43] R. Gago, L. Vásquez, O. Plantevin, T. H. Metzger, J. Muñoz-García, R. Cuerno, and M. Castro, "Order enhancement and coarsening of self-organized silicon nanodot patterns induced by ion-beam sputtering," Appl. Phys. Lett. **89**, 233101 (2006).

[44] F. Frost, B. Ziberi, A. Schindler, and B. Rauschenbach, "Surface engineering with ion beams: From self-organized nanostructures to ultra-smooth surfaces," Appl. Phys. A **91**, 551–559 (2008).

[45] S. Facsko, T. Bobek, A. Stahl, and H. Kurz, "Dissipative continuum model for self-organized pattern formation during ion-beam erosion," Phys. Rev. B **69**, 153412 (2004).

[46] S. A. Norris, "Stability analysis of a viscoelastic model for ion-irradiated silicon," Phys. Rev. B **85**, 155325 (2012).

[47] G. Ozaydin, A. S. Ozcan, Y. Y. Wang, K. F. Ludwig, H. Zhou, R. L. Headrick, and D. P. Siddons, "Real-time x-ray studies of Mo-seeded Si nanodot formation during ion bombardment," Appl. Phys. Lett. **87**, 163104 (2005).

[48] G. Ozaydin, A. S. Özcan, Y. Wang, K. F. Ludwig, Jr., H. Zhou, and R. L. Headrick, "Real-time X-ray studies of the growth of Mo-seeded Si

nanodots by low-energy ion bombardment," Nucl. Instrum. Methods Phys. Res. B **264**, 47–54 (2007).

49G. Ozaydin, Jr., K. F. Ludwig, H. Zhou, and R. L. Headrick, "Effects of Mo seeding on the formation of Si nanodots during low-energy ion bombardment," J. Vac. Sci. Technol. B **26**, 551 (2008).

50G. Ozaydin-Ince and K. F. Ludwig, Jr., "*In situ* x-ray studies of native and Mo-seeded surface nanostructuring during ion bombardment of Si(100)," J. Phys. Condens. Matter **21**, 224008 (2009).

51S. Macko, F. Frost, B. Ziberi, D. F. Forster, and T. Michely, "Is keV ion-induced pattern formation on Si(001) caused by metal impurities?," Nanotechnology **21**, 085301 (2010).

52P. Sigmund, A. Oliva, and G. Falcone, "Sputtering of multicomponent materials: Elements of a theory," Nucl. Instrum. Methods Phys. Res. B **194**, 541–548 (1982).

53V. B. Shenoy, W. L. Chan, and E. Chason, "Compositionally modulated ripples induced by sputtering of alloy surfaces," Phys. Rev. Lett. **98**, 256101 (2007).

54S. L. Roy, E. Søndergård, I. S. Nerbø, M. Kildemo, and M. Plapp, "Diffuse-interface model for nanopatterning induced by self-sustained ion-etch masking," Phys. Rev. B **81**, 161401(R) (2010).

55R. Mark Bradley and P. D. Shipman, "Spontaneous pattern formation induced by ion bombardment of binary compounds," Phys. Rev. Lett. **105**, 145501 (2010).

56R. Mark Bradley, "Theory of nanodot and sputter cone arrays produced by ion sputtering with concurrent deposition of impurities," Phys. Rev. B **83**, 195410 (2011).

57G. Abrasonis and K. Morawetz, "Instability types at ion-assisted alloy deposition: From two-dimensional to three-dimensional nanopattern growth," Phys. Rev. B **86**, 085452 (2012).

58H. Hofsäss, K. Zhang, A. Pape, O. Bobes, and M. Brötzmann, "The role of phase separation for self-organized surface pattern formation by ion beam erosion and metal atom co-deposition," Appl. Phys. A **111**, 653–664 (2013).

59S. A. Norris, "Ion-assisted phase separation in compound films: An alternate route to ordered nanostructures," J. Appl. Phys. **114**, 204303 (2013).

60O. El-Atwani, S. A. Norris, K. Ludwig, S. Gonderman, and J. P. Allain, "Ion beam nanopatterning of III-V semiconductors: Consistency of chemistry-driven theory with experimental and simulation parametric trends," Sci. Rep. **5**, 18207 (2015).

61O. El-Atwani, S. Ortoleva, A. Cimaroli, and J. P. Allain, "Formation of silicon nanodots via ion beam sputtering of ultrathin gold thin film coatings on Si," Nanoscale Res. Lett. **6**, 403 (2011).

62O. El-Atwani, A. Suslova, A. DeMasi, S. Gonderman, J. Fowler *et al.*, "Real time x-ray studies during nanostructure formation on silicon via low energy ion beam irradiation using ultrathin iron films," Appl. Phys. Lett. **101**, 263104 (2012).

63M. P. Harrison and R. M. Bradley, "Producing virtually defect-free nanoscale ripples by ion bombardment of rocked solid surfaces," Phys. Rev. E **93**, 040802(R) (2016).

64M. Castro, R. Cuerno, L. Vásquez, and R. Gago, "Self-organized ordering of nanostructures produced by ion-beam sputtering," Phys. Rev. Lett. **94**, 016102 (2005).

65M. Castro, J. Muñoz-García, R. Cuerno, M. d. M. G. Hernández, and L. Vázquez, "Generic equations for pattern formation in evolving interfaces," New J. Phys. **9**, 102 (2007).

66J. Muñoz-García, R. Gago, L. Vázquez, J. A. Sánchez-García, and R. Cuerno, "Observation and modeling of interrupted pattern coarsening: Surface nanostructuring by ion erosion," Phys. Rev. Lett. **104**, 026101 (2010).

67M. Castro, R. Cuerno, M. M. García-Hernández, and L. Vázquez, "Pattern-wavelength coarsening from topological dynamics in silicon nanofoams," Phys. Rev. Lett. **112**, 094103 (2014).

68M. Navez, D. Chaperot, and C. Sella, "Microscopie electronique-etude de lattaque du verre par bombardement ionique," C. R. Hebd. Seances Acad. Sci. **254**, 240 (1962).

69W. W. Mullins, "Solid surface morphologies governed by capillarity," in *Metal Surfaces: Structure, Energetics, and Kinetics*, edited by W. D. Robertson and N. A. Gjostein (American Society for Metals, Metals Park, OH, 1963), Chap. 2, pp. 17–66.

70G. Carter, "The effects of surface ripples on sputtering erosion rates and secondary ion emission yields," J. Appl. Phys. **85**, 455–459 (1999).

71R. Cuerno and A.-L. Barabási, "Dynamic scaling of ion-sputtered surfaces," Phys. Rev. Lett. **74**, 4746–4749 (1995).

72M. A. Makeev, R. Cuerno, and A.-L. Barabási, "Morphology of ion-sputtered surfaces," Nucl. Instrum. Methods Phys. Res. B **197**, 185–227 (2002).

73M. Holmes-Cerfon, W. Zhou, A. L. Bertozzi, M. P. Brenner, and M. J. Aziz, "Development of knife-edge ridges on ion bombarded surfaces," Appl. Phys. Lett. **101**, 143109 (2012).

74M. Holmes-Cerfon, M. J. Aziz, and M. P. Brenner, "Creating sharp features by colliding shocks on uniformly irradiated surfaces," Phys. Rev. B **85**, 165441 (2012).

75J. C. Perkinson, M. J. Aziz, M. P. Brenner, and M. Holmes-Cerfon, "Designing steep, sharp patterns on uniformly ion-bombarded surfaces," Proc. Natl. Acad. Sci. U. S. A. **113**, 11425–11430 (2016).

76D. A. Pearson and R. M. Bradley, "Theory of terraced topographies produced by oblique-incidence ion bombardment of solid surfaces," J. Phys.: Condens. Matter **27**, 015010 (2015).

77M. P. Harrison, D. A. Pearson, and R. M. Bradley, "Emergence and detailed structure of terraced surfaces produced by oblique-incidence ion sputtering," Phys. Rev. E **96**, 032804 (2017).

78M. P. Harrison and R. M. Bradley, "Blazed diffraction gratings produced by ion bombardment of pre-patterned solid surfaces," J. Appl. Phys. **121**, 054308 (2017).

79J. C. Perkinson, J. M. Swenson, A. DeMasi, C. Wagenbach, K. F. Ludwig, Jr., S. A. Norris, and M. J. Aziz, "Sawtooth structure formation under nonlinear-regime ion bombardment," J. Phys.: Condens. Matter **30**, 294004 (2018).

80M. A. Makeev and A.-L. Barabási, "Ion-induced effective surface diffusion in ion sputtering," Appl. Phys. Lett. **71**, 2800–2802 (1997).

81R. M. Bradley, "Exact linear dispersion relation for the Sigmund model of ion sputtering," Phys. Rev. B **84**, 075413 (2011).

82S. N. More and R. Kree, "Non-local linear stability of ion beam eroded surfaces," Appl. Surf. Sci. **258**, 4179–4185 (2012).

83B. J. Alder and T. E. Wainwright, "Studies in molecular dynamics. I. General method," J. Chem. Phys. **31**(2), 459 (1959).

84J. B. Gibson, A. N. Goland, M. Milgram, and G. H. Vineyard, "Dynamics of radiation damage," Phys. Rev. **120**(4), 1229–1253 (1960).

85R. P. Webb and D. E. Harrison, Jr., "Computer simulation of pit formation in metals by ion bombardment," Phys. Rev. Lett. **50**, 1478–1481 (1983).

86M. Ghaly and R. S. Averback, "Effect of viscous flow on ion damage near solid surfaces," Phys. Rev. Lett. **72**(3), 364–367 (1994).

87O. S. Oen, D. K. Holmes, and M. T. Robinson, "Ranges of energetic atoms in solids," J. Appl. Phys. **34**, 302 (1963).

88M. T. Robinson and I. M. Torrens, "Computer simulation of atomic-displacement cascades in solids in the binary-collision approximation," Phys. Rev. B **9**(12), 5008–5024 (1974).

89J. P. Biersack and L. G. Haggmark, "A Monte Carlo computer program for the transport of energetic ions in amorphous targets," Nucl. Instrum. Methods **174**, 257–269 (1980).

90J. F. Ziegler, J. P. Biersack, and U. Littmark, *The Stopping and Range of Ions in Matter* (Pergamon Press, New York, 1985).

91A. Mutzke, R. Schneider, W. Eckstein, and R. Dohmen, SDTrimSP version 5.0, IPP Report 12/8, Garching, 2011.

92M. L. Nietiadi, L. Sandoval, H. M. Urbassek, and W. Möller, "Sputtering of Si nanospheres," Phys. Rev. B **90**, 045417 (2014).

93H. Zhou, L. Zhou, G. Özaydin, K. F. Ludwig, Jr., and R. L. Headrick, "Mechanisms of pattern formation and smoothing induced by ion-beam erosion," Phys. Rev. B **78**, 165404 (2008).

94K. B. Winterbon, P. Sigmund, and J. B. Sanders, "Spatial distribution of energy deposited by atomic particles in elastic collisions," Mat. Fys. Medd. **37**, 14 (1970).

95M. Z. Hossain, J. B. Freund, and H. T. Johnson, "Ion impact energy distribution and sputtering of Si and Ge," J. Appl. Phys. **111**, 103513 (2012).

[96]G. Hobler, R. M. Bradley, and H. M. Urbassek, "Probing the limitations of Sigmund's model of spatially resolved sputtering using Monte Carlo simulations," Phys. Rev. B **93**, 205443 (2016).

[97]P. Sigmund, "Sputtering by ion bombardment: Theoretical concepts," in *Sputtering by Particle Bombardment I*, edited by R. Behrisch, Topics in Applied Physics Vol. 47 (Springer-Verlag, Berlin, Heidelberg, 1981), pp. 9–72.

[98]Y. Yamamura, Y. Itikawa, and N. Itoh, "*Angular dependence of sputtering yields of monatomic solids*," *Technical Report*, Institute of Plasma Physics, Nagoya University, Nagoya, Japan, 1983.

[99]M. J. Vasile, Z. Niu, R. Nassar, W. Zhang, and S. Liu, "Focused ion beam milling: Depth control for three-dimensional microfabrication," J. Vac. Sci. Technol. B **15**, 2350 (1997).

[100]R. M. Bradley and H. Hofsäss, "A modification to the Sigmund model of ion sputtering," J. Appl. Phys. **116**, 234304 (2014).

[101]H. Hofsäss and R. M. Bradley, "Tests of the modified Sigmund model of ion sputtering using Monte Carlo simulations," J. Appl. Phys. **117**, 174310 (2015).

[102]G. Costantini, F. Buatier de Mongeot, C. Boragno, and U. Valbusa, "Is ion sputtering always a "negative homoepitaxial deposition"?," Phys. Rev. Lett. **86**, 838–841 (2001).

[103]E. M. Bringa, K. Nordlund, and J. Keinonen, "Cratering energy regimes: From linear collision cascades to heat spikes to macroscopic impacts," Phys. Rev. B **64**, 235426 (2001).

[104]N. Kalyanasundaram, M. Ghazisaeidi, J. B. Freund, and H. T. Johnson, "Single impact crater functions for ion bombardment of silicon," Appl. Phys. Lett. **92**, 131909 (2008).

[105]H. M. Urbassek and K. T. Waldeer, "Spikes in condensed rare gases induced by keV-atom bombardment," Phys. Rev. Lett. **67**, 105 (1991).

[106]M. J. Aziz, "Nanoscale morphology control using ion beams," Mat. Fys. Medd. **52**, 187–206 (2006).

[107]S. A. Norris, M. P. Brenner, and M. J. Aziz, "From crater functions to partial differential equations: A new approach to ion bombardment induced nonequilibrium pattern formation," J. Phys. Condens. Matter **21**, 224017 (2009).

[108]B. Davidovitch, M. J. Aziz, and M. P. Brenner, "Linear dynamics of ion sputtered surfaces: Instability, stability and bifurcations," J. Phys. Condens. Matter **21**, 224019 (2009).

[109]N. Kalyanasundaram, J. B. Freund, and H. T. Johnson, "A multiscale crater function model for ion-induced pattern formation in silicon," J. Phys. Condens. Matter **21**, 224018 (2009).

[110]M. P. Harrison and R. Mark Bradley, "Crater function approach to ion-induced nanoscale pattern formation: Craters for flat surfaces are insufficient," Phys. Rev. B **89**, 245401 (2014).

[111]S. A. Norris, J. Samela, L. Bukonte, M. Backman, D. F. K. Nordlund, C. S. Madi, M. P. Brenner, and M. J. Aziz, "Molecular dynamics of single-particle impacts predicts phase diagrams for large scale pattern formation," Nat. Commun. **2**, 276 (2011).

[112]J. C. Perkinson, E. Anzenberg, M. J. Aziz, and K. F. Ludwig, Jr., "Model-independent test of the truncated crater function theory of surface morphology evolution during ion bombardment," Phys. Rev. B **89**(11), 115433 (2014).

[113]K. Nordlund, M. Ghaly, R. S. Averback, M. Caturla, T. Diaz de la Rubia, and J. Tarus, "Defect production in collision cascades in elemental semiconductors and fcc metals," Phys. Rev. B **57**, 7556–7570 (1998).

[114]M. Ghaly, K. Nordlund, and R. S. Averback, "Molecular dynamics investigations of surface damage produced by kiloelectronvolt self-bombardment of solids," Philos. Mag. A **79**, 795–820 (1999).

[115]K. Nordlund, "Molecular dynamics simulation of ion ranges in the 1–100 keV energy range," Comput. Mater. Sci. **3**, 448 (1995).

[116]J. Samela, S. A. Norris, K. Nordlund, and M. J. Aziz, "Optimization of large amorphous silicon and silica structures for molecular dynamics simulations of energetic impacts," Nucl. Instrum. Methods Phys. Res. B **269**, 1568–1571 (2011).

[117]M. Z. Hossain, K. Das, J. B. Freund, and H. T. Johnson, "Ion impact crater asymmetry determines surface ripple orientation," Appl. Phys. Lett. **99**, 151913 (2011).

[118]G. Hobler, D. Maciążek, and Z. Postawa, "Crater function moments: The role of implanted noble gas atoms," Phys. Rev. B **97**, 155307 (2018).

[119]M. L. Nietiadi and H. M. Urbassek, "Influence of local curvature on sputtering," Appl. Phys. Lett. **103**, 113108 (2013).

[120]H. Hofsäss, "Surface instability and pattern formation by ion-induced erosion and mass redistribution," Appl. Phys. A **114**, 401–422 (2014).

[121]H. Hofsäss, O. Bobes, and K. Zhang, "Argon ion beam induced surface pattern formation on Si," J. Appl. Phys. **119**, 035302 (2016).

[122]R. M. Bradley and H. Hofsäss, "Nanoscale patterns produced by self-sputtering of solid surfaces: The effect of ion implantation," J. Appl. Phys. **120**, 074302 (2016).

[123]H. Hofsäss, K. Zhang, and O. Bobes, "Self-organized surface ripple pattern formation by ion implantation," J. Appl. Phys. **120**, 135308 (2016).

[124]S. A. Norris, J. Samela, K. Nordlund, M. Vestberg, and M. Aziz, "Crater functions for compound materials: A route to parameter estimation in coupled-PDE models of ion bombardment," Nucl. Instrum. Methods Phys. Res. B **318**, 245–252 (2014).

[125]Scott A. Norris, J. Samela, M. Vestberg, K. Nordlund, and M. J. Aziz, "Corrigendum to "Crater functions for compound materials: A route to parameter estimation in coupled-PDE models of ion bombardment" [Nucl. Instr. Meth. Phys. Res. B 318B (2014) 245–252]," Nucl. Instrum. Methods Phys. Res. B **349**, 192 (2015).

[126]B. Liedke, "Ion beam processing of surfaces and interfaces: Modeling and atomistic simulations," Ph.D. thesis (Helmholtz Zentrum Dresden Rossendorf, 2011).

[127]L. Bukonte, F. Djurabekova, J. Samela, K. Nordlund, S. A. Norris, and M. J. Aziz, "Comparison of molecular dynamics and binary collision approximation simulations for atom displacement analysis," Nucl. Instrum. Methods Phys. Res. B **297**, 23–28 (2013).

[128]H. Hofsäss, "Model for roughening and ripple instability due to ion-induced mass redistribution [addendum to H. Hofsäss, Appl. Phys. A 114 (2014) 401, "Surface instability and pattern formation by ion-induced erosion and mass redistribution"]," Appl. Phys. A **119**, 687–695 (2015).

[129]O. Bobes, H. Hofsäss, and K. Zhang, "Neon ion beam induced pattern formation on amorphous carbon surfaces," AIP Adv. **8**, 025205 (2018).

[130]Scott A. Norris, Pycraters: A python library for the crater function framework, 2015, https://github.com/scott-norris-math/pycraters.

[131]F. Spaepen, *Physics of Defects*, edited by R. Balian, M. Kleman, and J. P. Poirier, Vol. XXXV of *Les Houches Lectures* (North Holland, 1981).

[132]S. Roorda, W. C. Sinke, J. M. Poate, D. C. Jacobson, S. Dierker, B. S. Dennis, D. J. Eaglesham, F. Spaepen, and P. Fuoss, "Structural relaxation and defect annihilation in pure amorphous silicon," Phys. Rev. B **44**, 3702 (1991).

[133]S. T. Pantelides, "Defects in amorphous silicon: A new perspective," Phys. Rev. Lett. **57**, 2979 (1986).

[134]A. Witvrouw and F. Spaepen, "Viscosity and elastic constants of amorphous Si and Ge," J. Appl. Phys. **74**, 7154 (1993).

[135]W. W. Mullins, "Flattening of a nearly plane solid surface due to capillarity," J. Appl. Phys. **30**, 77–83 (1959).

[136]T. M. Mayer, E. Chason, and A. J. Howard, "Roughening instability and ion-induced viscous relaxation of SiO_2 surfaces," J. Appl. Phys. **76**, 1633 (1994).

[137]S. A. Norris, J. C. Perkinson, M. Mokhtarzadeh, E. Anzenberg, M. J. Aziz, and K. F. Ludwig, Jr., "Distinguishing physical mechanisms using GISAXS experiments and linear theory: The importance of high wavenumbers," Sci. Rep. **7**, 2016 (2017).

[138]K. Otani, X. Chen, J. W. Hutchinson, J. F. Chervinsky, and M. J. Aziz, "Three-dimensional morphology evolution of SiO_2 patterned films under MeV ion irradiation," J. Appl. Phys. **100**, 023535 (2006).

[139]T. van Dillen, A. Polman, C. M. van Kats, and A. van Blaaderen, "Ion beam-induced anisotropic plastic deformation at 300 keV," Appl. Phys. Lett. **83**, 4315–4317 (2003).

[140]C. S. Madi, "Linear stability and instability patterns in ion bombarded silicon surfaces," Ph.D. thesis (Harvard University, 2011).

[141]Y. Ishii, C. Madi, M. J. Aziz, and E. Chason, "Stress evolution in Si during low-energy ion bombardment," J. Mater. Res. **29**, 2942–2948 (2014).

142P. F. A. Alkemade, "Propulsion of ripples on glass by ion bombardment," Phys. Rev. Lett. **96**, 107602 (2006).

143R. J. Asaro and W. A. Tiller, "Interface morphology development during stress corrosion cracking: Part I via surface diffusion," Metall. Trans. **3**, 1789–1796 (1972).

144M. A. Grinfeld, "Instability of the separation boundary between a non-hydrostatically stressed body and a melt," Sov. Phys. Dokl. **31**, 831–834 (1986).

145D. J. Srolovitz, "On the stability of surfaces of stressed solids," Acta Metall. **37**, 621–625 (1989).

146R. Cuerno, M. Castro, J. Muñoz-García, R. Gago, and L. Vásquez, "Nanoscale pattern formation at surfaces under ion-beam sputtering: A perspective from continuum models," Nucl. Instrum. Methods Phys. Res. B **269**, 894–900 (2011).

147M. Castro and R. Cuerno, "Hydrodynamic approach to surface pattern formation by ion beams," Appl. Surf. Sci. **258**, 4171–4178 (2012).

148S. A. Norris, "Stress-induced patterns in ion-irradiated silicon: Model based on anisotropic plastic flow," Phys. Rev. B **86**, 235405 (2012).

149H. Trinkaus and A. I. Ryazanov, "Viscoelastic model for the plastic flow of amorphous solids under energetic ion bombardment," Phys. Rev. Lett. **75**, 5072–5075 (1995).

150H. Trinkaus, "Dynamics of viscoelastic flow in ion tracks: Origin of plastic deformation of amorphous materials," Nucl. Instrum. Methods Phys. Res. B **146**, 204–216 (1998).

151T. van Dillen, A. Polman, P. R. Onck, and E. van der Giessen, "Anisotropic plastic deformation by viscous flow in ion tracks," Phys. Rev. B **71**, 024103 (2005).

152Y.-R. Kim, P. Chen, M. J. Aziz, D. Branton, and J. J. Vlassak, "Focused ion beam induced deflections of freestanding thin films," J. Appl. Phys. **100**, 104322 (2006).

153H. B. George, Y. Tang, X. Chen, J. Li, J. W. Hutchinson, J. A. Golovchenko, and M. J. Aziz, "Nanopore fabrication in amorphous Si: Viscous flow model and comparison to experiment," J. Appl. Phys. **108**, 014310 (2010).

154C. A. Davis, "A simple model for the formation of compressive stress in thin films by ion bombardment," Thin Solid Films **226**, 30 (1993).

155M. Castro, R. Gago, L. Vázquez, J. Muñoz-García, and R. Cuerno, "Stress-induced solid flow drives surface nanopatterning of silicon by ion-beam irradiation," Phys. Rev. B **86**, 214107 (2012).

156A. Moreno-Barrado, M. Castro, R. Gago, L. Vázquez, J. Muñoz-García, R. Redondo-Cubero, B. Galiana, C. Ballesteros, and R. Cuerno, "Nonuniversality due to inhomogeneous stress in semiconductor surface nanopatterning by low-energy ion-beam irradiation," Phys. Rev. B **91**, 155303 (2015).

157A. Witvrouw, C. A. Volkert, and F. Spaepen, "The viscosity of amorphous metallic thin films," Mater. Sci. Eng. A **134**, 1274 (1991).

158A. Witvrouw and F. Spaepen, "Determination of the plane stress elastic constants of thin films from substrate curvature measurements: Applications to amorphous metals," J. Appl. Phys. **73**, 7344–7350 (1993).

159J. C. Perkinson, "Nanoscale pattern self-organization under ion bombardment," Ph.D. thesis (Harvard university, 2017).

160N. Kalyanasundaram, J. B. Freund, and H. T. Johnson, "Atomistic determination of continuum mechanical properties of ion-bombarded silicon," J. Eng. Mater. Technol. **127**, 457–461 (2005).

161N. Kalyanasundaram, M. C. Moore, J. B. Freund, and H. T. Johnson, "Stress evolution due to medium-energy ion bombardment of silicon," Acta Mater. **54**, 483–491 (2006).

162N. Kalyanasundaram, "Structure, stress, and surface evolution in silicon due to ion bombardment," Ph.D. thesis (University of Illinois at Urbana-Champaign, 2007).

163N. Kalyanasundaram, M. Wood, J. B. Freund, and H. T. Johnson, "Stress evolution to steady state in ion bombardment of silicon," Mech. Res. Commun. **35**, 50–56 (2008).

164S. G. Mayr, Y. Ashkenazy, K. Albe, and R. S. Averback, "Mechanisms of radiation-induced viscous flow: Role of point defects," Phys. Rev. Lett. **90**, 055505 (2003).

165C. S. Madi, E. Anzenberg, K. F. Ludwig, Jr., and M. J. Aziz, "Mass redistribution causes the structural richness of ion-irradiated surfaces," Phys. Rev. Lett. **106**, 066101 (2011).

166J. M. Swenson and S. A. Norris, "Swelling as a stabilizing mechanism in irradiated thin films," J. Phys.: Condens. Matter **30**, 304003 (2018).

167N. V. Medhekar, W. L. Chan, V. B. Shenoy, and E. Chason, "Stress-enhanced pattern formation on surfaces during low energy ion bombardment," J. Phys. Condens. Matter **21**, 224021 (2009).

168E. Anzenberg, C. S. Madi, M. J. Aziz, and K. F. Ludwig, Jr., "Time-resolved measurements of nanoscale surface pattern formation kinetics in two dimensions on ion-irradiated Si," Phys. Rev. B **84**, 214108 (2011).

169Scott A. Norris, PyGLIDRE: A python library for GISAXS LInear Dispersion Relation Extraction, 2017, https://github.com/scott-norris-math/pyglidre.