

**Distributional fixed-point equations for island nucleation in one dimension: The inverse problem**Hrvojka Krcelic,<sup>1,\*</sup> Michael Grinfeld,<sup>2,†</sup> and Paul A. Mulheran<sup>1,‡</sup><sup>1</sup>*Department of Chemical and Process Engineering, University of Strathclyde, Glasgow G1 1XJ, United Kingdom*<sup>2</sup>*Department of Mathematics and Statistics, University of Strathclyde, Glasgow G1 1XH, United Kingdom*

(Received 12 June 2018; published 12 November 2018)

The self-consistency of the distributional fixed-point equation (DFPE) approach to understanding the statistical properties of island nucleation and growth during submonolayer deposition is explored. We perform kinetic Monte Carlo simulations, in which point islands nucleate on a one-dimensional lattice during submonolayer deposition with critical island size  $i$ , and we examine the evolution of the inter-island gaps as they are fragmented by new island nucleation. The DFPE couples the fragmentation probability distribution within the gaps to the consequent gap size distribution (GSD), and we find a good fit between the DFPE solutions and the observed GSDs for  $i = 0, 1, 2$ , and  $3$ . Furthermore, we develop numerical methods to address the inverse problem, namely the problem of obtaining the gap fragmentation probability from the observed GSD, and again find good self-consistency in the approach. This has consequences for its application to experimental situations where only the GSD is observed and where the growth rules embodied in the fragmentation process must be deduced.

DOI: [10.1103/PhysRevE.98.052801](https://doi.org/10.1103/PhysRevE.98.052801)**I. INTRODUCTION**

Island nucleation and growth during submonolayer deposition is a topic of continuing research, with ongoing development of theoretical models to describe the scaling properties of the island sizes and spatial distribution [1]. Over the past two decades, the focus has tended to move from the problem of obtaining the correct form of the island size distribution to finding the capture zone distribution (CZD) [2–6]. An island's capture zone is defined as the region on the substrate closer to that island than to any other. It represents the growth rate of the island, since the deposited monomers that are inside the capture zone are most likely to be trapped by the parent island; the CZD is therefore a consequence of the spatial arrangement of the islands.

One common theoretical approach utilizes rate equations, with capture numbers reflecting the capture zones [2–4,7,8], although often this requires some empirically determined parameter(s). An alternative approach, of the type we adopt here, treats the process using fragmentation equations. An island nucleation means the creation of a new capture zone and the size reduction of the zones that were previously occupying that region of the substrate; so the parent capture zones are fragmented to create the daughter capture zone [5,6,9–13].

Aside from these analytical models, when looking only at the functional form of a simulated or experimentally obtained CZD in the scaling regime, the semiempirical Gamma distribution function has been used frequently as a fitting model, both in two dimensions [6,14–17] and in one dimension [18]. Similarly, Pimpinelli and Einstein proposed a generalized Wigner surmise [19] as a model function for the CZD, relating it back to a fragmentation process. An advantage is that

functional form can be used to deduce the island nucleation mechanism from a measured CZD [20], through the critical island size  $i$  for nucleation (the critical size is defined as the size above which an island will not dissociate into monomers). This distribution has also been applied to some experimental data [17,21,22]; however, there are also some controversies about the validity of this model [5,23,24].

In the present work, we adopt and explore a nucleation model on a one-dimensional substrate (modeling nucleation along a step edge, for example) in which island nucleation is seen as a fragmentation of an inter-island gap. The evolution of the gap size distribution (GSD) and the CZD is tracked by considering the parent gaps (capture zones) that were fragmented by a new island's nucleation.

In previous work it was proposed that the GSD and the CZD can be modeled with distributional fixed-point equations (DFPEs) on one-dimensional substrates [25]. The model equation for the GSD consists entirely of physical, easily measurable quantities, so in this paper we focus solely on the gaps.

The DFPE for the GSD reads

$$x \stackrel{\triangle}{=} a(1+x), \quad (1)$$

where  $x$  is a gap size scaled to the average at a given time (coverage) and  $a$  is a position in the gap where a new island nucleates, scaled to the size of the gap ( $a \in [0, 1]$ ). The symbol  $\stackrel{\triangle}{=}$  means that the left and the right hand side of the equation have the same distribution.

Equation (1) then says that the distribution of scaled gap sizes  $x$  is equal to the distribution of gap sizes that are created when a larger, parent gap of size  $x+y$  (and, by employing the mean-field assumption, we have set the scaled size  $y=1$ ) fragments into two gaps, of proportions  $a$  and  $1-a$ . DFPE (1) has an integral equation form:

$$\phi(x) = \int_0^{\min(x,1)} \phi\left(\frac{x}{a}-1\right) \frac{f(a)}{a} da, \quad (2)$$

\*hrvojka.krcelic@strath.ac.uk

†m.grinfeld@strath.ac.uk

‡paul.mulheran@strath.ac.uk

where  $\phi(x)$  is the probability distribution function for scaled gap sizes  $x$  and  $f(a)$  is the probability of breaking a gap into proportions  $a$  and  $(1-a)$ . A version of the DFPE (1) that does not involve a mean-field approximation can also be found in Ref. [25]; however, its corresponding integral equation does not offer the possibility of calculating  $f(a)$  from a known  $\phi(x)$ , so we will not use it in the present work.

Blackman and Mulheran [9] proposed an analytical form for  $f(a)$ :

$$f(a) = \frac{(2\alpha + 1)!}{(\alpha!)^2} a^\alpha (1-a)^\alpha. \quad (3)$$

Here  $\alpha$  reflects the mechanism of island nucleation: for  $\alpha = i$  nucleation is deposition driven and for  $\alpha = i + 1$  it is diffusion driven [10]. In a diffusion-driven nucleation event, an island is formed by the coming together of  $(i + 1)$  diffusing monomers and we assume that they all independently follow the steady-state monomer density distribution within the gap, yielding  $\alpha = i + 1$ . In a deposition-driven nucleation event, a smaller, unstable cluster of  $i$  monomers (each following the steady-state monomer distribution) is increased by the arrival of a newly nearby deposited monomer; the latter event is uniformly distributed across the gap, yielding  $\alpha = i$ . We assume that the real (experimental) nucleation process can be modelled as a combination of these two idealized cases. We also note the distinction of the deposition-driven nucleation mechanism from attachment-limited events; in our work we assume the attachment is irreversible.

Equation (3) is derived from the monomer density solutions  $n_1(x)$  of a long-time (steady-state;  $dn_1(x)/dt \simeq 0$ ) diffusion equation with a constant monolayer deposition rate within a gap [5,9]. The nucleation probability is then assumed to be  $\sim n_1(x)^\alpha$ , which gives Eq. (3); therefore it is only valid after the system has had time to reach steady-state conditions in which monomer density and, by extension,  $f(a)$  within a gap are time independent. Since the deposition rate is constant, provided there is no desorption we have coverage  $\theta = Ft$ . Then Eq. (3) is only valid in a scaling regime where the GSD, scaled to the average size, and  $f(a)$  are independent of  $\theta$ .

In this paper, we look further at the applicability of the DFPE approach for island nucleation and growth in one dimension. We are particularly interested in whether the DFPE provides a self-consistent approach to understanding the statistics of gaps. The following two questions are addressed.

(i) We can measure  $f(a)$  during a kinetic Monte Carlo (kMC) simulation. How does the measured form compare to that of Eq. (3), and how does the solution of Eq. (2), using the observed  $f(a)$ , compare to the kMC GSD?

(ii) Can we invert the argument of Eq. (2)? Can we find  $f(a)$  from a given GSD, and if so how does this recovered  $f(a)$  compare to that observed in the kMC?

## II. KINETIC MONTE CARLO SIMULATION

We use a standard kMC simulation model where monomers are deposited onto a one-dimensional lattice with a constant monolayer deposition rate  $F$  and are free to diffuse by nearest-neighbor hopping with the diffusion constant  $D$ .

Immobile point islands nucleate according to values of the critical island size  $i$  and subsequently grow by capturing either diffusing or deposited monomers. Island nucleation and growth are irreversible and re-evaporation of monomers from the surface is forbidden.

We start with an initially empty lattice with  $N = 10^6$  sites and the diffusion to deposition ratio  $R = D/F = 10^7$ . We allow monomers to hop on average 20 times before the next deposition event ( $R = 0.5 \times 20 \times N$ ). In total we deposit  $n = 10^6$  monomers to get  $\theta = 100\%$  coverage (not all of the  $n$  monomers will get incorporated into islands, typically at the end of a simulation there are up to a hundred free monomers in the  $i = 1$  case, and more for higher  $i$ ).

At each diffusion step a monomer is selected at random and moved by a unit length on the lattice, in a random direction. If it arrives to a position adjacent to another monomer or cluster of monomers, and the resulting number of monomers is larger than  $i$ , they will be fixed in a single lattice site and the newly nucleated island's size and position will be recorded. Islands capture monomers that diffuse to adjacent sites and monomers that are deposited on top or on an adjacent position. Increments in island sizes are recorded while the islands are kept as single points on the lattice; this way the islands do not coalesce for large coverage, which allows us to collect a lot of data while the system has still got a long way to go before the scaling breaks down [26].

In the  $i = 0$  case we set the probability that a monomer will stick to the site onto which it hopped or was deposited to be  $p = 10^{-7}$ .

To get the GSDs [ $\phi_{\text{kMC}}(x)$ ] we used outputs at coverage  $\theta = 100\%$  and averaged the data over 100 runs. Every time a new island nucleated, we recorded its position within the gap and used that data to create  $f_{\text{kMC}}(a)$  (as a histogram). Since we want to understand scaling properties when the steady-state conditions have been achieved, we need to find the coverage at which  $f_{\text{kMC}}$  stabilizes. It was previously shown in Ref. [10] that the monomer density  $n_1$  behaves in a manner that would yield Eq. (3) for small gaps, but not large ones. Those findings corresponds to large coverages (where we expect to find mainly smaller gaps) versus small coverages (large gaps). Figure 1 shows  $f_{\text{kMC}}$  for  $i = 1$ , reaching the steady-state condition above the coverage of approximately  $\theta = 1\%$ .

## III. DFPE METHODOLOGY

For a given  $f(a)$ , Eq. (2) is solved iteratively for  $\phi(x)$ . Following the procedure described in Ref. [25], we perform numerical integration on a mesh of 500 equally spaced points for  $x \in [0, 5]$ . With an initial guess of a rectangular  $\phi$ , we iterate Eq. (2) until the solution stabilizes to at least its third decimal place.

To solve the inverse problem of obtaining  $f$  from a given  $\phi$ , we use two different strategies.

### A. Tikhonov regularization for the inverse problem

Equation (2) belongs to the well-known class of Fredholm integral equations of the first kind,  $\phi(x) = \int k(x, a)f(a)da$ , which are ill posed. We also have an additional complication

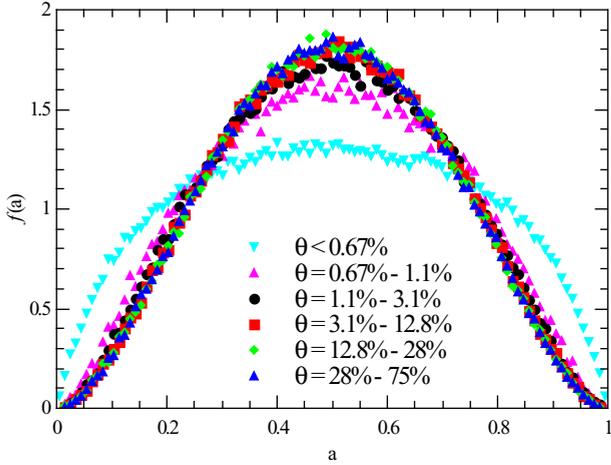


FIG. 1.  $f_{\text{kMC}}$  for various coverage intervals, from  $\theta = 0\%$  to  $\theta = 75\%$ , for  $i = 1$ . Between  $\theta = 1\%$  and  $3\%$  the system has reached steady-state conditions so subsequent  $f_{\text{kMC}}$  curves overlap.

of having the left-hand side  $\phi(x)$  appearing in the kernel function  $k(x, a)$ . This means that any noise in the input data will propagate in the kernel. Hence this is not a standard inverse problem and, to the best of our knowledge, there is no established way of solving this particular type of problem. We proceed to treat Eq. (2) as we would treat a standard Fredholm equation.

One of the most common ways to deal with ill-posed equations is the Tikhonov regularization procedure, in which a regularization term is added onto the original equation. This is a standard method found in many textbooks (see, for example, Ref. [27]); we describe it briefly.

The problem of finding an  $f$  that satisfies the matrix equation  $Kf = \phi$  [the discretized form of Eq. (2), where the operator  $K$  stands for the kernel function and the integral operator] can be treated as a minimization problem:  $\min_f \{\|Kf - \phi\|_2^2\}$ . By adding a regularization term, this problem is substituted with  $\min_f \{\|Kf - \phi\|_2^2 + \lambda \|Lf\|_2^2\}$ . Here,  $f$  is the sought solution,  $L$  is the regularization operator, usually chosen to be the identity operator or a differential operator, and the regularization parameter  $\lambda > 0$  controls how much weight is given to the minimization of  $\|Kf - \phi\|_2^2$  relative to the minimization of the added term  $\|Lf\|_2^2$ . Solving the inverse problem then includes choosing the appropriate operator  $L$  and optimizing for  $\lambda$ . For a particular value of  $\lambda$ , the matrix equation to be solved for  $f$  is

$$(K^T K + \lambda L^T L)f_\lambda = K^T \phi, \quad (4)$$

where  $K^T$  is the transpose of the matrix operator  $K$ . It is straightforward to solve Eq. (4) numerically.

If  $K$  is ill conditioned, with an ill-determined rank, the addition of the regularization operator  $L$  has a function of making Eq. (4) well posed; then Eq. (4) will have a unique solution  $f_\lambda$  for all  $\lambda$ .

The procedure then involves solving Eq. (4) while varying  $\lambda$  to find an optimal value of  $\lambda$  which stabilizes the solution without oversmoothing it. Good values of  $\lambda$  are typically taken

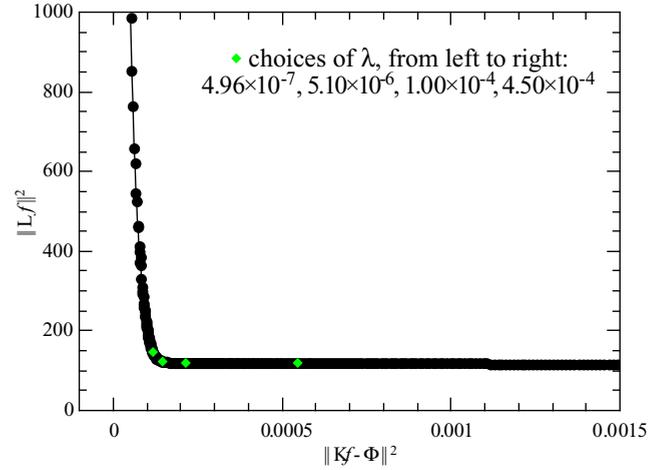


FIG. 2. L curve for inverting Eq. (2) with  $\alpha = 1$  and  $f$  defined by Eq. (3). Four chosen values of  $\lambda$  are marked with green diamond symbols; corresponding solutions  $P_\lambda$  and their integrals  $\phi_\lambda$  are shown in Fig. 3.

to be within the corner of the  $\|Lf\|_2^2$  vs  $\|Kf - \phi\|_2^2$  plot, the so-called L curve.

We tested the identity and the second derivative operator as candidates for the regularization operator  $L$  and, despite the fact that the second derivative should be the first choice for damping oscillatory behavior in unstable solutions, we found that we get better results when using the identity operator. In our calculations, we used routines from Ref. [28] (see chapter therein on Linear Regularisation Methods).

In Figs. 2 and 3 (where  $L$  is the identity operator), we use  $f(a) = 6a(1 - a)$  [Eq. (3) with  $\alpha = 1$ , corresponding to the deposition case for the critical island size  $i = 1$  or the diffusion case for  $i = 0$ ] to show the results of the Tikhonov regularization method for a known function.

With this  $f(a)$ , we integrated Eq. (2) iteratively to obtain  $\phi(x)$  and solved Eq. (4) for  $f_\lambda$  (solved the inverse problem). Figure 2 shows the L curve, where each point of the curve corresponds to a different  $\lambda$  ( $\lambda$  increases from left to right). Note that  $\|Kf_\lambda - \phi\|_2$  is the root mean error between the input  $\phi$  and  $\phi_\lambda$  [where  $\phi_\lambda = Kf_\lambda$  is the result of integrating Eq. (2) with  $f = f_\lambda$ ].

We chose four values of  $\lambda$ , highlighted on the L-curve plot, and show the four solutions  $f_\lambda$  in Fig. 3 (bottom panel), as well as the corresponding  $\phi_\lambda$  (upper panel), alongside the original  $\phi$  and  $f$ . While different  $\phi_\lambda$  lie almost perfectly on top of each other and on top of input  $\phi$ , the solutions  $f_\lambda$  show how strongly this problem is ill posed. For the two smaller values of  $\lambda$ , the solutions  $f_\lambda$  exhibit high oscillations and the largest  $\lambda$  begins to show signs of oversmoothing in the interval  $a \in [0.7, 1]$ . The best solution still has some noise, it is not symmetric and needs normalization; the area under the curve is  $\|f_\lambda\|_1 \approx 0.99$ . We found that a general trend is a decreasing  $\|\cdot\|_1$  norm with growing  $\lambda$  (moving away from the corner of the L curve to the right). The same problems are amplified when applying the method to  $\phi_{\text{kMC}}$ , obtained from (noisy) kMC data, with an additional problem that the solutions  $f_\lambda$  sometimes dropped slightly below zero near

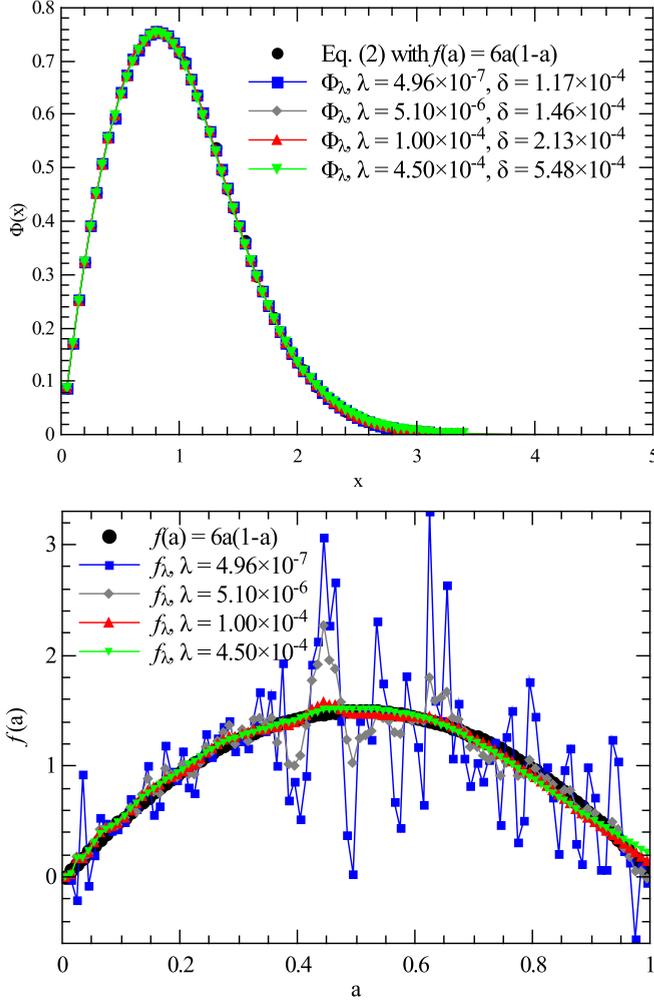


FIG. 3. Top: Original  $\phi$  (black circles) to be inverted: the solution of integrating Eq. (2) [where  $P$  was taken to be Eq. (3) with  $\alpha = 1$ ]. After obtaining the Tikhonov results of inversion,  $f_\lambda$ , we integrated them again according to Eq. (2) to get the shown  $\phi_\lambda$ ; all of the curves overlap. Bottom: Tikhonov results  $f_\lambda$  ( $\lambda = 4.96 \times 10^{-7}$ ,  $5.1 \times 10^{-6}$ ,  $1 \times 10^{-4}$ , and  $4.5 \times 10^{-4}$ ), shown with the true  $f(a) = 6a(1-a)$  (black circles).

$a = 1$ , although seemingly within the noise error we would expect when solving for kMC data input.

Since this implementation of Tikhonov regularization does not give entirely satisfactory results (loss of symmetry,  $\|\cdot\|_1$  norm or positivity), we would need to modify it. Normalization can be always done by hand, but adding extra symmetry and positivity constraints on the regularization, while theoretically possible, would turn the L curve of our minimization problem into a three-dimensional hypersurface in  $\mathbb{R}^4$ . This would extend the scope of the work enormously, so instead we look for an alternative approach.

### B. Fourier representation for the inverse problem

To complement the Tikhonov regularization results, we develop an alternative method of solving the inverse problem. In this method, we represent  $f(a)$  as a finite Fourier series whose corresponding  $\phi$  matches the true, kMC-obtained  $\phi_{\text{kMC}}$ .

Whether we take Eq. (3) to be an accurate model of physical systems or not,  $f(a)$  at least has to be equal to zero at  $a = 0$  and  $1$ , and it is physically reasonable to assume it is symmetrical. Therefore we only use sine waves and odd wave numbers to enforce symmetry around  $a = 1/2$  and the requirement  $f_F(0) = f_F(1) = 0$ . We start with a single normalized sine;  $f_F(a) = N \sin(\pi a)$ ,  $N$  being the normalization constant. We integrate this  $f_F$  according to Eq. (2) to obtain  $\phi_F$  and calculate the error:

$$\delta = \sum_i [\phi_F(x_i) - \phi_{\text{kMC}}(x_i)]^2. \quad (5)$$

Then we proceed to build  $f_F$  by adding higher random harmonics  $A \sin(k\pi a)$ , where in each step we randomly choose the wave number (from the allowed values  $k = 3, 5, \dots, k_{\text{max}}$ ) and the amplitude  $A$  ( $A \in [-A_{\text{max}}, A_{\text{max}}]$ ), normalize the new  $f_F$ , and recalculate  $\phi_F$  and  $\delta$ . Then, provided that the resulting  $f_F$  is everywhere positive, we keep the newly added harmonic with the Boltzmann probability  $\exp[-(\delta_{\text{new}} - \delta_{\text{old}})\beta]$ . We repeat this cycle with a fixed  $\beta$  (initially set to 1)  $m$  times before increasing  $\beta$  by a factor of 2 (i.e., perform a simulated anneal). After increasing  $\beta$  in such a way  $M$  times, we narrow in on the solution by a search in which we only keep the newly added harmonics if  $\delta_{\text{new}} < \delta_{\text{old}}$ .

Since there are five search parameters (values of  $k_{\text{max}}$  and  $A_{\text{max}}$ , number of cycles  $M$  and  $m$ , and the number of search attempts while only accepting moves with  $\delta_{\text{new}} < \delta_{\text{old}}$ ), we needed to find the optimal parameters on a known problem before proceeding to calculate  $f_F$  for  $\phi_{\text{kMC}}$ .

Therefore we first integrated Eq. (2) with  $f(a)$  given by Eq. (3) and then used the resulting  $\phi$  in place of  $\phi_{\text{kMC}}$  in Eq. (5) to see how could we correctly reconstruct  $f_F$ . Because Eq. (2) is ill posed, adding higher harmonics actually leads to a worse, less stable solution  $f_F$  with high-frequency noise, as shown in Fig. 4. At the same time the error  $\delta$  [Eq. (5)] can decrease (here with the rest of the search parameters fixed, although in general, when increasing  $k_{\text{max}}$ , a higher number of search cycles is needed to reach a stable solution). This happens regardless of the amount (or absence) of noise in the input and cannot be avoided. It is a consequence of the following property of the equation  $Kf = \phi$ : the inverse  $K^{-1}$  of the operator  $K : U \rightarrow V$  is unbounded, and the equation is ill posed, if  $U$  is an infinite dimensional space [29]. Hence decreasing the dimension of space, spanned with the harmonics, in which we build  $f$ , is a form of regularization.

Because of that, we limited the maximum allowed wave number to 11. With  $k_{\text{max}} = 11$  and allowed maximum amplitude  $A_{\text{max}} = 0.05$  per one search attempt, we ran the simulated anneal with  $m = 30$  and  $M = 500$  cycles ( $30 \times 500$  random harmonic choices) and then ran through another 300 attempts, accepting only  $\delta_{\text{new}} < \delta_{\text{old}}$ . These are the parameters we then used to calculate  $f_F$  for  $\phi_{\text{kMC}}$ , for all the values of  $i$  (for  $i = 0$  we also used  $k_{\text{max}} = 5$  as explained below).

## IV. RESULTS

We show the diffusion and deposition  $\phi$  [Eq. (2)] with  $f$  given by two cases of Eq. (3), the GSD obtained from kMC ( $\phi_{\text{kMC}}$ ), and  $\phi_F$  and  $\phi_\lambda$  plotted together in the upper panels in Figs. 5–8, for critical island sizes  $i = 1, 2, 3$ , and 0,

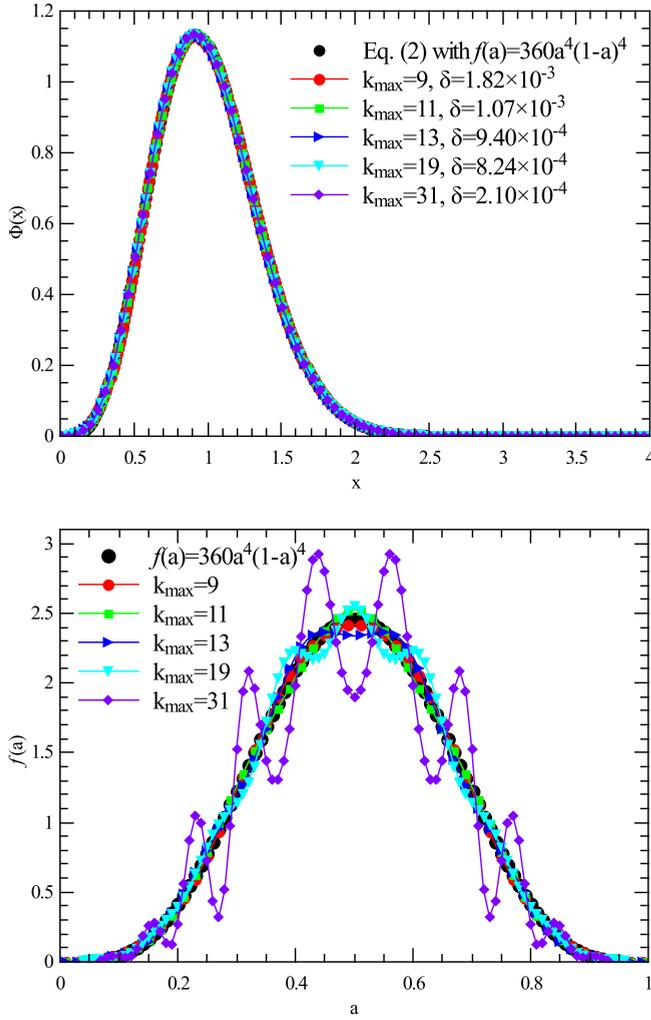


FIG. 4. Searching for the optimal parameter  $k_{\max}$ . Top: Original  $\phi$  to be inverted [Eq. (2) in which  $f$  was taken to be Eq. (3) with  $\alpha = 4$ ]. After obtaining the results of inversion built with the Fourier series,  $f_F$ , for different values of  $k_{\max}$ , we integrated them again according to Eq. (2) to get the  $\phi_F$  plotted on top of each other; all of the curves overlap. Bottom: Inversion results  $f_F$  ( $k_{\max} = 9, 11, 13, 19,$  and  $31$ ), shown with the true  $f(a) = 360a^4(1-a)^4$ .

respectively. The solutions of integrating Eq. (2) with  $f(a) = f_{\text{kMC}}(a)$  are plotted with empty square symbols.

The bottom panels of Figs. 5–8 show the deposition and diffusion  $f(a)$  given by Eq. (3),  $f_{\text{kMC}}(a)$  obtained from kMC simulations, and the solutions of the inverse problems  $f_F$  and  $f_\lambda$ .

Errors  $\delta$  listed in the legends are the sum of squares differences between kMC-obtained  $\phi_{\text{kMC}}$  and  $\phi_F$  and  $\phi_\lambda$  obtained by integrating the solutions  $f_F$  and  $f_\lambda$  according to Eq. (2) [for  $\phi_F$  the error is given with Eq. (5) and for  $\phi_\lambda$  it is given with  $\|Kf_\lambda - \phi_{\text{kMC}}\|_2^2$ ]. The solutions  $f_F$  are always normalized during the procedure of adding new harmonics, but the Tikhonov procedure only deals with the  $\|\cdot\|_2$  norm, so none of the  $f_\lambda$  solutions shown have the  $\|\cdot\|_1$  norm equal to 1. We have found, however, that all  $i$  solutions with optimal choices of  $\lambda$  have a norm close to 1, and it only significantly

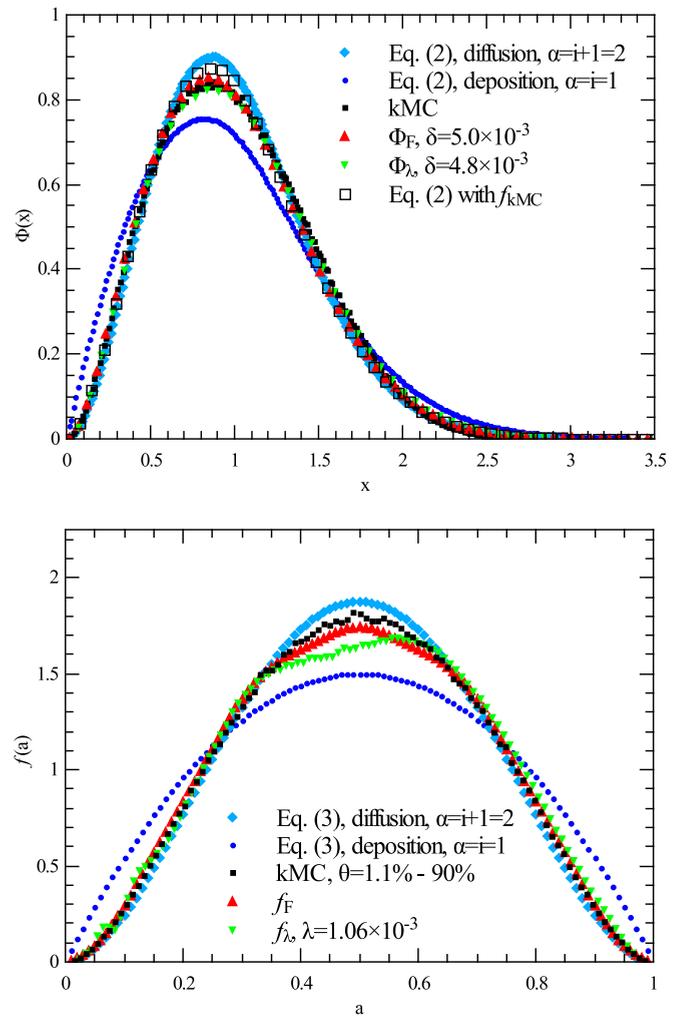


FIG. 5. Critical island size  $i = 1$ . Top: Solutions of integrating Eq. (2) with  $f$  given by Eq. (3) ( $\alpha = i + 1$  case is shown in light blue diamonds and  $\alpha = i$  in dark blue circles). The kMC-obtained GSD  $\phi_{\text{kMC}}$  (solid black squares) is inverted according to Eq. (2); when the resulting  $f_{F,\lambda}$  are used to integrate Eq. (2) we get  $\phi_{F,\lambda}$  (shown in red up and green down-facing triangles, respectively). The empty squares show the result of integrating Eq. (2) with  $f = f_{\text{kMC}}$ . Bottom:  $\alpha = i + 1$  and  $\alpha = i$  cases of Eq. (3) (light blue diamonds and dark blue circles), the kMC result  $f_{\text{kMC}}$  (solid black squares), and the results of inverting  $\phi_{\text{kMC}}$ :  $f_F$  (red up triangles) and  $f_\lambda$  (green down triangles).

drops (below 0.95) for too high  $\lambda$  which also gave a large error  $\delta$ .

We note here that our  $f_{\text{kMC}}$  results are similar to the nucleation probabilities for  $i = 1, 2,$  and  $3$  shown in a recent publication by González, Pimpinelli, and Einstein [6].

When we use  $f_{\text{kMC}}$  to integrate Eq. (2), the resulting GSD (empty squares in the upper panels of Figures 5–8) fits the kMC-obtained GSD ( $\phi_{\text{kMC}}$ ) quite well for all the  $i$  cases, but it does not match it perfectly. We remind the reader here that the DFPE model we are using involves a mean-field approximation; a non-mean-field version suggested in Ref. [25] gives more accurate results.

Returning to the inverse problem, for the  $i = 1$  and  $2$  cases (Figs. 5 and 6) both the Fourier and the Tikhonov method gave

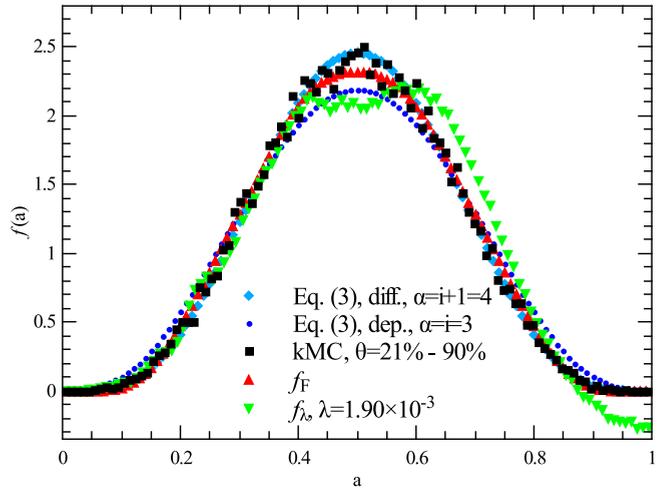
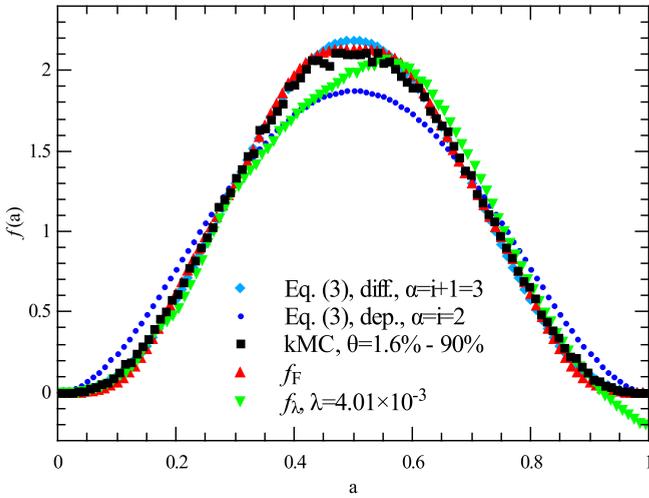
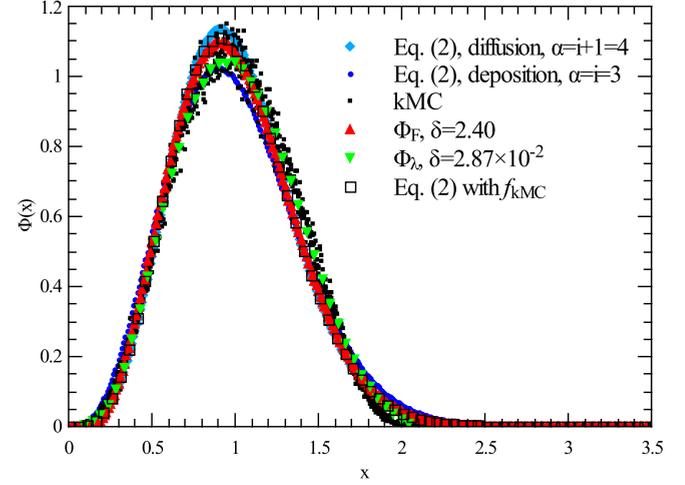
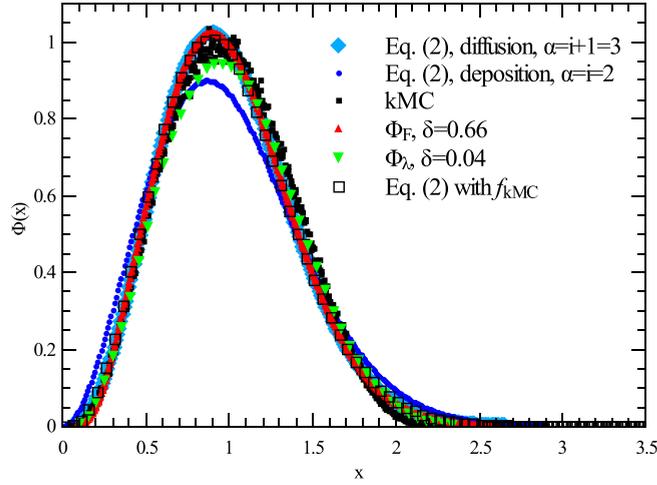


FIG. 6. Critical island size  $i = 2$ . The symbols used in this figure have the same meaning as those in Fig. 5.

FIG. 7. Critical island size  $i = 3$ . The symbols used in this figure have the same meaning as those in Fig. 5.

good  $f_F$  and  $f_\lambda$  results, but in the  $i = 2$  case we start to see the effect of increased noise in the input  $\phi_{\text{kMC}}$  relative to the  $i = 1$  case:  $f_\lambda$  is noticeably negative near  $a = 1$ . In the  $i = 3$  case the situation is even worse (see Fig. 7), so here the Tikhonov solution is more of a guideline for the behavior of the true  $f(a)$ . On the other hand, the Fourier construction was successful in all the cases so we can conclude that, by using both methods for assurance, we can find reliable solutions in problems where  $f(a)$  is not directly measurable (e.g., many experiments to create nanostructures).

In the  $i = 0$  case (Fig. 8), only the diffusion limit ( $\alpha = i + 1$ ) of Eq. (3), as introduced in Ref. [9], has physical meaning. In addition, the Fourier result for  $\phi_{\text{kMC}}$  inversion with  $k_{\text{max}} = 11$  is problematic. Its high oscillations around  $a = 0.5$  suggest a higher degree of regularization is needed, so even though the previously established cutoff  $k_{\text{max}} = 11$  gave excellent results when inverting Eq. (2) for all  $\alpha$  values in Eq. (3) (including the here relevant  $\alpha = i + 1 = 1$ ), we additionally show the inverse  $f_F$  where we used  $k_{\text{max}} = 5$ . This result is backed by the Tikhonov solution ( $\lambda$  is taken from the corner area of the L curve).

The measured  $f_{\text{kMC}}(a)$  for  $i = 0$  lies almost perfectly on top of the diffusion curve. However, the solution of Eq. (2)

with  $f = f_{\text{kMC}}$  is (in this case most noticeably) not matching  $\phi_{\text{kMC}}$ , which shows the limitations of the mean-field approximation used to formulate this approach.

For critical island sizes higher than 0, the measured  $f_{\text{kMC}}(a)$  (and, consequently,  $\phi_{\text{kMC}}$ ) is at least a little below the diffusion prediction, allowing for a small contribution of the deposition-driven nucleation. We finish our analysis by quantifying the level of this contribution for different  $i$ . Table I shows the result of fitting  $f_{\text{kMC}}$  on a convex combination of analytic expressions for diffusion and deposition  $f(a)$  from Eq. (3):

$$f_{\text{kMC}} = \beta f_{\alpha=i+1}^{\text{diffusion}} + (1 - \beta) f_{\alpha=i}^{\text{deposition}} \quad (6)$$

with the least squares method. We also show the fit of  $\phi_{\text{kMC}}$  on the convex combination of the diffusion and deposition cases,

$$\phi_{\text{kMC}} = \gamma \phi_{\alpha=i+1}^{\text{diffusion}} + (1 - \gamma) \phi_{\alpha=i}^{\text{deposition}}, \quad (7)$$

where we fitted kMC curves  $\phi_{\text{kMC}}$  on the results of the numerical integration of Eq. (2). From the results, we can safely conclude that diffusion is the dominant mechanism of island nucleation. (Note that the result for  $\gamma$  in the  $i = 2$  case is larger than 1, but not if its allowed error is subtracted.)

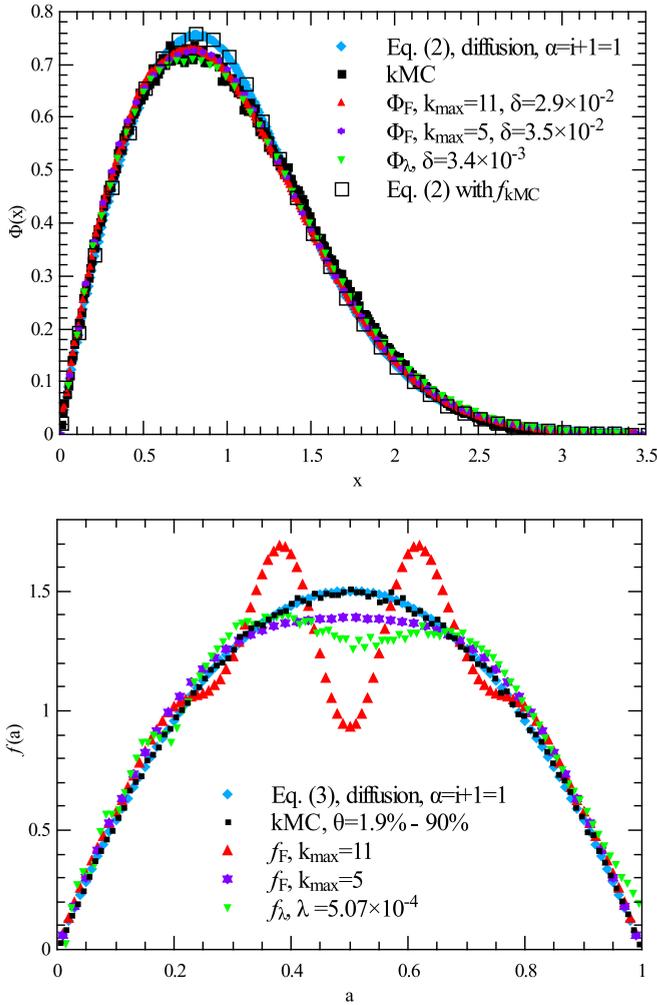


FIG. 8. Critical island size  $i = 0$ . The symbols used in this figure have the same meaning as those in Fig. 5.

## V. SUMMARY

In this paper, we have revisited the mean-field DFPE (1) model of gap fragmentation on a one-dimensional substrate from Ref. [25]. Using the Tikhonov regularization method, from the kMC-obtained GSD and the integral equation form of the DFPE for the GSD [Eq. (2)], we were able to calculate the gap fragmentation probability, that is, the probability of a new island nucleation occurring at a position  $a$  inside a gap

TABLE I. Results of fitting kMC results according to Eqs. (6) and (7).

$i$	$\beta$	$\gamma$
1	$0.821 \pm 0.007$	$0.728 \pm 0.010$
2	$0.819 \pm 0.015$	$1.015 \pm 0.018$
3	$0.844 \pm 0.004$	$0.714 \pm 0.020$

[ $f(a)$ ]. The results show fair agreement with the probability  $f_{\text{kMC}}(a)$  that we measured directly from kMC simulations, although they lack the expected symmetry and strict positivity. Growing amounts of numerical noise (in cases of higher  $i$ ) aggravate this problem.

We developed an alternative method of inverting Eq. (2) to obtain  $f$ , in which we represent  $f$  as a finite Fourier series and use the series properties. This allows us to impose symmetry and positivity, however a downside is a more time-consuming procedure due to the large amount of search parameters. The results of this method are in better agreement with the measured  $f_{\text{kMC}}(a)$ , so this method, especially when backed by the well-known Tikhonov method, makes for a good tool in solving problems where it is not possible to measure  $f(a)$  directly.

The DFPE model we use involves two limiting cases of island nucleation: diffusion (via colliding adatoms) and deposition driven. As expected, within this framework our results (both the kMC-obtained GSD and  $f_{\text{kMC}}$ ) favor the diffusion-driven nucleation as the dominant mechanism. We found no correlation between  $i$  and one mechanism's contribution amount relative to the other; however, if there were a trend, a model with a built-in mean-field approximation would most likely be too crude for it to be observed, especially from noisy data.

Finally, we emphasize that, while the DFPE we employ here may not offer a perfect fit [as seen with the solutions of Eq. (2) with  $f_{\text{kMC}}$  values which have slightly higher peaks than  $\phi_{\text{kMC}}$  values], its strength lies in the unique possibility of calculating  $f(a)$  from a given GSD, without the need for additional information. This provides the opportunity for further interpretation of experimental data and greater insight into the physics of nucleation and growth by helping to identify the dominant mechanisms that drive the gap fragmentation probability distribution  $f(a)$ . Our approach to tackling the inverse solutions to the DFPE might also be applied and adapted to the analysis of allied problems [30].

[1] M. Einax, W. Dieterich, and P. Maass, *Rev. Mod. Phys.* **85**, 921 (2013).  
 [2] M. N. Popescu, J. G. Amar, and F. Family, *Phys. Rev. B* **64**, 205404 (2001).  
 [3] P. Mulheran and D. Robbie, *Europhys. Lett.* **49**, 617 (2000).  
 [4] J. W. Evans and M. C. Bartelt, *Phys. Rev. B* **66**, 235410 (2002).  
 [5] M. Grinfeld, W. Lamb, K. O'Neill, and P. Mulheran, *J. Phys. A: Math. Theor.* **45**, 015002 (2011).

[6] D. L. González, A. Pimpinelli, and T. L. Einstein, *Phys. Rev. E* **96**, 012804 (2017).  
 [7] M. Körner, M. Einax, and P. Maass, *Phys. Rev. B* **86**, 085403 (2012).  
 [8] F. Gibou, C. Ratsch, and R. Caflisch, *Phys. Rev. B* **67**, 155403 (2003).  
 [9] J. A. Blackman and P. A. Mulheran, *Phys. Rev. B* **54**, 11681 (1996).

- [10] K. P. O'Neill, M. Grinfeld, W. Lamb, and P. A. Mulheran, *Phys. Rev. E* **85**, 021601 (2012).
- [11] J. Blackman, M. Grinfeld, and P. Mulheran, *Phys. Lett. A* **379**, 3146 (2015).
- [12] V. Tokar and H. Dreyssé, *J. Phys. A: Math. Theor.* **50**, 375002 (2017).
- [13] D. L. González, A. Pimpinelli, and T. L. Einstein, *Phys. Rev. E* **84**, 011601 (2011).
- [14] P. A. Mulheran and J. A. Blackman, *Phys. Rev. B* **53**, 10261 (1996).
- [15] M. Fanfoni, E. Placidi, F. Arciprete, E. Orsini, F. Patella, and A. Balzarotti, *Phys. Rev. B* **75**, 245312 (2007).
- [16] C. P. Joshi, Y. Shim, T. P. Bigioni, and J. G. Amar, *Phys. Rev. E* **90**, 032406 (2014).
- [17] S. Miyamoto, O. Moutanabbir, E. E. Haller, and K. M. Itoh, *Phys. Rev. B* **79**, 165415 (2009).
- [18] V. I. Tokar and H. Dreyssé, *Phys. Rev. E* **92**, 062407 (2015).
- [19] A. Pimpinelli and T. L. Einstein, *Phys. Rev. Lett.* **99**, 226102 (2007).
- [20] T. L. Einstein, A. Pimpinelli, and D. L. González, *J. Cryst. Growth* **401**, 67 (2014).
- [21] M. Groce, B. Conrad, W. Cullen, A. Pimpinelli, E. Williams, and T. Einstein, *Surf. Sci.* **606**, 53 (2012).
- [22] T. Potocar, S. Lorbek, D. Nabok, Q. Shen, L. Tumbek, G. Hlawacek, P. Puschnig, C. Ambrosch-Draxl, C. Teichert, and A. Winkler, *Phys. Rev. B* **83**, 075423 (2011).
- [23] F. Shi, Y. Shim, and J. G. Amar, *Phys. Rev. E* **79**, 011602 (2009).
- [24] M. Li, Y. Han, and J. W. Evans, *Phys. Rev. Lett.* **104**, 149601 (2010).
- [25] P. A. Mulheran, K. P. O'Neill, M. Grinfeld, and W. Lamb, *Phys. Rev. E* **86**, 051606 (2012).
- [26] C. Ratsch, Y. Landa, and R. Vardavas, *Surf. Sci.* **578**, 196 (2005).
- [27] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*, Mathematical Modeling and Computation (SIAM, Philadelphia, 1998).
- [28] W. H. Press, W. T. Vetterling, S. A. Teukolsky, and B. P. Flannery, *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, Cambridge, 1992).
- [29] R. Kress, in *Linear Integral Equations*, Applied Mathematical Sciences (Springer, New York, 2014), p. 300.
- [30] P. Seba, *Acta Phys. Pol. A* **112**, 681 (2007).