Model discovery for studies of surface morphological modifications based on Kuramoto-Sivashinsky dynamics

D. Reiser

Forschungszentrum Jülich GmbH, Institut für Energie- und Klimaforschung – Plasmaphysik, Partner of the Trilateral Euregio Cluster (TEC), 52425 Jülich, Germany

(Received 21 May 2019; published 23 September 2019)

A wide range of observations in studies of surfaces exposed to ion beams can be explained and analyzed successfully by continuum models of the Kuramoto-Sivashinsky type. Despite certain progress in the theoretical understanding of the model parameters on the basis of atomistic models, much of the applications are based on phenomenological determination of several unknown quantities. In this work a numerical tool is discussed and investigated, which allows us to determine model coefficients and complex model structures from experimental findings. The method resembles known approaches in machine learning and data-driven reconstruction techniques. To keep the discussion on a fundamental level, numerical simulations are conducted by employing a scaled test model. The reconstruction technique is demonstrated for this model system and shows a high accuracy in recovering input parameters for situations without beam noise. As an application to an unknown system to be explored, the algorithm is then applied to a system with lognormal distributed ion bombardment. The impact of the beam fluctuations in the proposed model are discussed. Perspectives of the numerical algorithm for an analysis of experimental data are addressed.

DOI: 10.1103/PhysRevE.100.033312

I. INTRODUCTION

This paper deals with the interplay of fluctuating plasma or ion beams and material surfaces where erosion takes place due to the according bombardment with high energetic ions. As in previous studies, a continuum model is employed to get insight into the particular changes in morphology and plasma-wall interaction for macroscopic dimensions of the target material [1–21]. The implementation of beam fluctuations is usually taken into account by inclusion of a random term describing additive noise. However, most of the coefficients in the continuum models depend on the material properties and on the ion-beam characteristics as well. Therefore, in this work a modified Kuramoto-Sivashinsky (KS) equation is derived which reflects the beam fluctuations via fluctuating coefficients describing the morphology-dependent erosion processes. A second and perhaps even more important point of this paper is the detailed description of a method to identify surface models of the KS type. The ultimate goal of this numerical tool is the analysis of future experimental data to complement *ab initio* theories for atomistic processes by half-empirical methods. In this work the basic algorithm is introduced and its suitability is demonstrated for the analysis of surface dynamics with and without ion-beam fluctuations. Currently missing experimental data are replaced by simulation results in this initial study. For this preparatory work a detailed statistical analysis of the reconstruction and data-driven model discovery is presented. In Sec. II an introduction to the Kuramoto-Sivashinsky dynamics is given and, after that, in Sec. III, the inclusion of beam fluctuations by multiplicative noise is elucidated. The method to reconstruct or discover a surface model from the given data is detailed in Sec. IV and in Sec. V numerical results are presented for systems with and without beam fluctuations. The beam fluctuations are chosen to be lognormally distributed with prescribed temporal correlation. Simulations taking into account different choices for the noise are compared with cases of constant model coefficients. The characterization of the results is done by means of pattern inspection, Fourier analysis, and probability distribution functions for the surface height. It is found that the results can show a strong change in the surface morphology as a function of the fluctuation parameters, i.e., fluctuation amplitude and correlation time. With respect to the model discovery method discussed here, the important result is the possibility of a quite accurate reconstruction of model parameters for cases with and without beam fluctuations. Statistical analysis of the reconstructed model parameters gives some hints on possible optimization of the numerical approach based on a careful adjustment of sampling time with respect to characteristic timescales of the physical system considered.

II. CONTINUUM MODEL FOR SURFACE EVOLUTION

In this section a quite general continuum model is discussed, which contains several models discussed in the literature as limiting cases. The surface structure of a material exposed to the plasma beam is described by the surface height function $h = h(x, y)$, where x and y are the Cartesian coordinates in the reference plane defined by the unmodified flat surface. The evolution of the surface height is assumed to be given by an equation of the form

$$\frac{\partial h}{\partial t} = v - a(h - \bar{h}) - b\bar{h} + \sum_{i=x,y} \gamma_i \partial_i h + \sum_{i=x,y} \nu_i \partial_{ii} h + \sum_{i,j=x,y} K_{ij} \partial_{ii} \partial_{jj} h + \sum_{i=x,y} \lambda_i (\partial_i h)^2 + \eta. \quad (1)$$

Here $\overline{h}$ denotes the x-y-averaged profile. The coefficient v is a constant erosion rate. The rate coefficient a represents a flattening effect, pushing the surface height to the average height profile $\overline{h}$ and the rate coefficient b introduces a tendency of the system to keep the averaged height at zero. The coefficients ν_i denote the surface tension and the terms proportional to λ_i describe a slope-dependent erosion. The coefficients K_{ij} are components of a symmetric diffusion matrix and the related terms introduce surface diffusion. The coefficients γ_i represent an advective motion of surface structures and, in many model variants, an anisotropic advection with $\gamma_y = 0$ and a finite γ_x is introduced to describe grazing incidence of the ion beam. For normal incidence the ion beam induced advection disappears, giving $\gamma_i = 0$ [4–6]. An additional random contribution is denoted by η and represents an additive noise introduced to mimic the fluctuations of the ion beam. One finds for a periodic domain that the average height $\overline{h}$ obeys the evolution equation

$$\frac{\partial \overline{h}}{\partial t} = \overline{v} - b\overline{h} + \sum_{i=x,y} \lambda_i \overline{(\partial_i h)^2} + \overline{\eta}. \tag{2}$$

For the structured piece $\widetilde{h} = h - \overline{h}$ one obtains the evolution equation

$$\frac{\partial \widetilde{h}}{\partial t} = \widetilde{v} - a\widetilde{h} + \sum_{i=x,y} \gamma_i \partial_i \widetilde{h} + \sum_{i=x,y} \nu_i \partial_{ii} \widetilde{h} + \sum_{i,j=x,y} K_{ij} \partial_{ii} \partial_{jj} \widetilde{h} + \sum_{i=x,y} \lambda_i (\partial_i \widetilde{h})^2 - \sum_{i=x,y} \lambda_i \overline{(\partial_i \widetilde{h})^2} + \widetilde{\eta}. \tag{3}$$

It can be concluded from Eq. (3) that the homogeneous component $\overline{h}$ does not affect the dynamics of $\widetilde{h}$ because $\overline{h}$ does not appear on the right-hand side. On the contrary the spatially averaged profile $\overline{h}$ is affected by the inhomogeneous piece $\widetilde{h}$ via the slope-dependent terms proportional to the coefficients λ_i, as can be seen from the evolution equation Eq. (2). Therefore, the damping coefficient a changes the dynamics of the morphology represented by $\widetilde{h}$, whereas the damping term $b\overline{h}$ affects the average height $\overline{h}$ only. This is of importance for the numerical studies presented in the subsequent sections. There a damping term with finite b is chosen, but a damping via the term $a\widetilde{h}$ is excluded because it is known that this would have a strong impact on the dynamics [20]. In general it can be said that the model defined by Eq. (1) is suitable for a variety of applications. Equation (1) includes a large number of effects considered to be relevant for morphological changes due to erosion, surface diffusion, additive noise of the ion-beam flux, and other effects. Actually it represents a damped Kuramoto-Sivashinsky equation, and continuum models like Eq. (1) and variants of it have been discussed extensively in the literature (see Refs. [1–14] and references therein). There were many attempts to relate the model parameters appearing in KS continuum models with microscopic processes or to extract their particular values from experiments. According to Makeev *et al.* [4] the coefficients γ_i, ν_i, and λ_i in the KS model (1) can be expressed as functions of the incident angle of the ions and atomistic parameters such as penetration depth, distribution width, etc. Another example of a set of coefficients based on microscopic theory has been published by Cuerno and Barabási [5]. Derived from Sigmund's sputtering theory, they also obtained formulas for v, ν_i, and λ_i. Lauritsen *et al.* considered a microscopic model for ion sputtering and derived expressions for v, ν_i, λ_i, K_{ij}, and the noise term γ in Ref. [9]. Muñoz-García *et al.* [10] derived model coefficients from a two-field model and Sigmund's sputtering theory. A semi-empirical approach has been discussed by Muñoz-García *et al.* [11]. They derived model coefficients ν_i, λ_i, K_{ij} (and in addition the coefficient σ_{ij} introduced in their model) from analytical estimates and experimental data. In summary, Eq. (1) comprises several models used in the literature to link experimental data and fundamental microscopic processes on the plasma-beam-exposed surface. The need for an accurate determination of model coefficients was also the motivation for the development of model discovery techniques such as the one presented here. However, despite the success of models like Eq. (1), in many cases it has become necessary to include further effects in the continuum description of surface dynamics. This is still an ongoing discussion, and a very comprehensive overview of such extensions, current issues, and more refined theoretical approaches can be found in Ref. [12]. A few possible extensions of Eq. (1) are discussed in Appendix B where the candidate model for the model discovery method is elucidated.

III. INCLUSION OF BEAM FLUCTUATIONS

In this section the KS model (1) will be simplified to obtain a model with convenient scaling properties while still being relevant for realistic experimental conditions; namely, normal incidence of the ion beam. In this work—and in contrast to previous studies—ion-beam fluctuations are not considered by incorporation of an additive noise, i.e., a particular choice for the stochastic term η in Eq. (1). Instead of that it is assumed that $\eta = 0$ and that the coefficients ν_i and λ_i representing the erosion process are strictly proportional to the beam flux. This introduces multiplicative noise to the model system. Moreover, it is assumed that K_{ij} is constant, i.e., surface diffusion is independent of the impinging ions. To simplify matters it is also assumed that the ions hit the surface under normal incidence and that the surface has no anisotropies. Due to the symmetry of the assumptions made the number of coefficients in the model is strongly reduced: $\gamma_x = \gamma_y = 0$, $\nu_x = \nu_y = \nu$, $\lambda_x = \lambda_y = \lambda$, $K_{xx} = K_{xy} = K_{yx} = K_{yy} = K$. For normal incidence the coefficients γ_x and γ_y are expected to vanish due to their dependence on the angle of incidence of the ion beam [4–6]. It should be noted that these simplifications do not restrict the flexibility of the model reconstruction introduced in the next section. The simplifications are only introduced for the purpose of making the discussions clearer while still maintaining a realistic physical picture. Next, the fluctuating coefficients ν and λ and the plasma beam flux J are split according to $\nu = \langle \nu \rangle + \widetilde{\nu}$, $\lambda = \langle \lambda \rangle + \widetilde{\lambda}$, and $J = \langle J \rangle + \widetilde{J}$, respectively. The bracket $\langle \cdots \rangle$ denotes a temporal average. Finally, a damping term proportional to $\overline{h}$ is included to control the average height, which is pushed to zero in the simulations. As mentioned above and shown by Eqs. (2) and (3), this kind of damping has an impact only on the evolution of the averaged height $\overline{h}$ but it does not change the dynamics of the surface morphology, i.e., the structured piece $\widetilde{h}$, which

is actually in the focus of this work. Note that this is very different to a damping via the term proportional to a in Eq. (1). This term would have a strong impact on the dynamics as discussed, e.g., in Ref. [20]. Here such an effect is excluded from the model. Thus, the basic test model considered here reads

$$\begin{aligned}\frac{\partial h}{\partial t} &= -\nu\nabla^2 h + \frac{\lambda}{2}(\nabla h)^2 - K\nabla^2\nabla^2 h - b\bar{h} \\ &= -\left[\nu_0\nabla^2 h - \frac{\lambda_0}{2}(\nabla h)^2\right]\frac{\langle J\rangle}{J_0} - b\bar{h} \\ &\quad -\left[\nu_0\nabla^2 h - \frac{\lambda_0}{2}(\nabla h)^2\right]\frac{\widetilde{J}}{J_0} - K_0\nabla^2\nabla^2 h. \end{aligned} \tag{4}$$

The reference flux J_0 has been introduced and ν_0 and λ_0 are the corresponding constant model coefficients. The diffusion coefficient $K_0 = K$ is a constant, too, and, as mentioned above, it does not depend on the beam flux. Using K_0 instead of K only serves to make the notation uniform. It is to be stressed again that the reduced model of Eq. (4) misses several effects which are considered to be of importance in many experiments. Besides additional effects mentioned in Appendix B the additive noise η is often considered as crucial to explain experimental observations (see, e.g., Ref. [21]). However, the model considered here is an attempt to focus on the effect of multiplicative noise with temporal correlation and to test the model discovery method elucidated in the next section. In future applications the model discovery method allows us to analyze data in the framework of the full candidate model listed in Appendix B. Therefore, it would also be able to identify effects which have been neglected in the preparatory tests of this work. After this remark, the derivation of the reduced KS model will be finalized by rewriting Eq. (4) using the convenient scaling $h \to \nu_0/\lambda_0 h$, $x \to \sqrt{K_0/\nu_0}x$, $t \to K_0/\nu_0^2 t$ to obtain the dimensionless equation

$$\begin{aligned}\frac{\partial h}{\partial t} &= -\left[\nabla^2 h - \frac{1}{2}(\nabla h)^2\right]\frac{\langle J\rangle}{J_0} - \frac{bK_0}{\nu_0^2}\bar{h} \\ &\quad -\left[\nabla^2 h - \frac{1}{2}(\nabla h)^2\right]\frac{\widetilde{J}}{J_0} - \nabla^2\nabla^2 h. \end{aligned} \tag{5}$$

Due to this scaling, a large variety of model systems can be studied by considering Eq. (5) with varying statistics of the scaled beam flux J/J_0. To give an impression of the dimensions hidden behind the scaling, typical values for the characterization of the $\mathrm{Si} - \mathrm{Ar}^+$ system from Ref. [11] are quoted: For a flux of $J_0 = 6.00 \times 10^{18}$ ions $\mathrm{m}^{-2}\,\mathrm{s}^{-1}$ the model coefficients have been found to be $\nu_0 = 1.67 \times 10^{-1}$ nm^2/s, $\lambda_0 = 1.00 \times 10^{-1}$ nm/s and $K_0 = 2.60$ nm^4/s. This gives for the scaling length and timescales of the system: $\nu_0/\lambda_0 = 1.67$ nm, $\sqrt{K_0/\nu_0} = 3.95$ nm, and $\tau_0 = K_0/\nu_0^2 = 93.60$ s. At this point the reference timescale τ_0 is introduced for later considerations. Note that the linear analysis of instabilities for small perturbations in the height h, with $\lambda = 0$ and negligible beam noise η gives for Eq. (4) a most unstable mode with wave number k_* and corresponding growth rate γ_* and length scale l_*. These are given by

$$k_* = \sqrt{\frac{\nu}{4K}}, \quad \gamma_* = \frac{\nu^2}{4K}, \quad l_* = 4\pi\sqrt{\frac{K}{\nu}}. \tag{6}$$

Therefore, fluctuations in the beam will have an impact on the linear scales of the system due to the variation of surface tension ν. In this work and according to previous findings for plasma beam fluctuations [22], the scaled flux $z = J/J_0$ is assumed to obey a lognormal distribution $f(z)$ specified by

$$f(z) = \frac{1}{z\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(\ln z - \mu)^2}{2\sigma^2}\right]. \tag{7}$$

Its first and second moments are

$$\langle z\rangle = \left\langle\frac{J}{J_0}\right\rangle = e^{\mu+\sigma^2/2}, \tag{8}$$

$$\langle (z - \langle z\rangle)^2\rangle = \left\langle\frac{\widetilde{J}^2}{J_0^2}\right\rangle = e^{2\mu+2\sigma^2} - e^{2\mu+\sigma^2}. \tag{9}$$

The parameters μ and σ are determined by prescribing the mean value of the scaled flux $\langle J/J_0\rangle$ and its scaled variance $\langle\widetilde{J}^2/J_0^2\rangle$, respectively:

$$\mu = \ln\left\langle\frac{J}{J_0}\right\rangle - \frac{1}{2}\ln\left(1 + \frac{\langle\widetilde{J}^2/J_0^2\rangle}{\langle J/J_0\rangle^2}\right), \tag{10}$$

$$\sigma^2 = \ln\left(1 + \frac{\langle\widetilde{J}^2/J_0^2\rangle}{\langle J/J_0\rangle^2}\right). \tag{11}$$

The parameters μ and σ are used to generate a Gaussian random variable Y with mean μ and variance σ^2 numerically. Finally, the assignment $Z = e^Y$ gives the desired random variable Z having a lognormal distribution. Details on the numerical generation of lognormal noise ψ with temporal correlation are given in Appendix A. This generated noise is used to describe the fluctuations of the ion flux, $J/J_0 = \psi$.

IV. MODEL RECONSTRUCTION AND MODEL DISCOVERY

The following section is devoted to a topic that could be described as the "inverse simulation approach." This term distinguishes the method from a "forward simulation" in which equations such as Eq. (1), (4), or (5) are solved in order to compare the results with the experimental findings. The forward method then tries to optimize the simulations as systematically as possible by selecting suitable model parameters and thus to improve the agreement with the measurements. An inverse approach would be to look at a large number of experimental findings and use them to adapt a mathematical model that would best describe the data. This resembles the standard regression problem of curve fitting, but here it is meant that a model defined by a partial differential equation similar to Eq. (1) is subject to a fit procedure. Such approaches have been discussed in the context of machine learning and data-driven statistical methods (see Refs. [23–28] and references therein). Also in studies of interfaces, a least-squares method very similar to the one presented here has already been discussed in Refs. [16,17]. However, although the method of this work has many similarities with ideas discussed in connection with Gaussian process regression in Ref. [24] and essential points of the approach have already been presented in Refs. [25,27] and especially in Refs. [16,17], it will be outlined here in detail for reasons of clarity. In addition, the

method applied here differs in a few but significant numerical details from previous studies, and this may allow a higher accuracy than similar approaches would suggest. Since the focus is on the analysis of surface structures in ion-beam experiments by the use of continuum models, the method is introduced by starting with an evolution equation similar to Eq. (1). The discretization is done with respect to a spatial grid consisting of points $\boldsymbol{x}_k$, $k = 1, \ldots, N$, in the x-y plane:

$$\frac{\partial h(\boldsymbol{x}_k)}{\partial t} = \sum_{i=1}^{P} \alpha_i f_i(\boldsymbol{x}_k). \tag{12}$$

Here P is the number of terms constituting the right-hand side of the assumed mathematical model description and the α_i are constant model parameters. The functions f_i are assumed to be polynomials of the height function h and its derivatives, i.e., $\partial_i h$, $\partial_{ij} h$, $(\partial_i h)^2$, etc. An example of such a candidate model is given Eq. (B1) in Appendix B, which defines the discovery model [actually the right-hand side of Eq. (12)] used in this work. It is important to note that the stochastic term η in Eq. (1) is excluded from the set of functions f_i. Here, only deterministic terms are taken into account. Now, the discretized model of Eq. (12) is rewritten by defining

$$T_k = \frac{\partial h(\boldsymbol{x}_k)}{\partial t}, \quad H_{ki} = f_i(\boldsymbol{x}_k). \tag{13}$$

This gives for Eq. (12) a matrix equation

$$\boldsymbol{H} \cdot \boldsymbol{\alpha} = \boldsymbol{T}, \tag{14}$$

where $\boldsymbol{H} \in \mathbb{R}^{N\times P}$, $\boldsymbol{\alpha} \in \mathbb{R}^P$, $\boldsymbol{T} \in \mathbb{R}^N$, and $P < N$. The number of grid points N is assumed to be larger than the number of model parameters P. Therefore, an overdetermined set of linear equations for the parameters α_i is considered. One can proceed using a singular value decomposition (SVD) [29,30] to decompose the matrix $\boldsymbol{H}$ as

$$\boldsymbol{H} = \boldsymbol{U} \cdot \boldsymbol{S} \cdot \boldsymbol{V}^{\mathrm{T}}, \tag{15}$$

where $\boldsymbol{U} \in \mathbb{R}^{N\times N}$ and $\boldsymbol{V} \in \mathbb{R}^{P\times P}$ are orthonormal matrices. The matrix $\boldsymbol{S} \in \mathbb{R}^{N\times P}$ has r nonzero components on the diagonal only. These define the singular values σ_1, σ_2, ..., σ_r, $r \leqslant P$. The other diagonal elements are $\sigma_i = 0$ for $i > r$. The squared singular values σ_i^2 are the eigenvalues of the products $\boldsymbol{H}^{\mathrm{T}} \cdot \boldsymbol{H}$ and $\boldsymbol{H} \cdot \boldsymbol{H}^{\mathrm{T}}$, as can be seen from the relations

$$\boldsymbol{H} \cdot \boldsymbol{H}^{\mathrm{T}} = \boldsymbol{U} \cdot \boldsymbol{S} \cdot \boldsymbol{S}^{\mathrm{T}} \cdot \boldsymbol{U}^{\mathrm{T}}, \quad \boldsymbol{H}^{\mathrm{T}} \cdot \boldsymbol{H} = \boldsymbol{V} \cdot \boldsymbol{S}^{\mathrm{T}} \cdot \boldsymbol{S} \cdot \boldsymbol{V}^{\mathrm{T}}, \tag{16}$$

where $\boldsymbol{S} \cdot \boldsymbol{S}^{\mathrm{T}} \in \mathbb{R}^{N\times N}$ and $\boldsymbol{S}^{\mathrm{T}} \cdot \boldsymbol{S} \in \mathbb{R}^{P\times P}$ are diagonal square matrices with diagonal elements $\sigma_1^2, \sigma_2^2, \ldots, \sigma_r^2$, $r \leqslant P$. For the purpose of solving Eq. (14), the pseudoinverse $\boldsymbol{H}^+$ is defined by

$$\boldsymbol{H}^+ = \boldsymbol{V} \cdot \boldsymbol{S}^+ \cdot \boldsymbol{U}^{\mathrm{T}}, \tag{17}$$

where $\boldsymbol{S}^+$ has nonzero components on the diagonal only and these are given by the inverse of the singular values σ_i of $\boldsymbol{S}$, except where the singular values are zero, there also $\boldsymbol{S}^+$ has a zero entry. The solution $\boldsymbol{\alpha}$ that minimizes the functional D defined by

$$D = |\boldsymbol{H} \cdot \boldsymbol{\alpha} - \boldsymbol{T}|^2 \tag{18}$$

is called the optimal solution $\boldsymbol{\alpha}^+$, and it can be shown that this is given by [29]

$$\boldsymbol{\alpha}^+ = \boldsymbol{H}^+ \cdot \boldsymbol{T}. \tag{19}$$

The practical implementation of this approach is in the evaluation of the matrix components (13) and the subsequent calculation of the least-squares parameters α_i^+ by Eq. (19). A single reconstruction requires the evaluation of T_k and H_{ki} by the use of "snapshots" of the discrete height profiles $h^t(\boldsymbol{x}_k)$ and $h^{t+\tau_\Delta}(\boldsymbol{x}_k)$ at two subsequent time points t and $t + \tau_\Delta$. This allows us to approximate the matrix components of time derivatives $T_k^t = (h_k^{t+\tau_\Delta} - h_k^t)/\tau_\Delta$. The corresponding components of H_{ki}^t are found by finite differences on the x-y grid using the profile $h^t(\boldsymbol{x}_k)$. Note that the parameter τ_Δ denotes the time interval between the snapshots, which might be very different from the time step Δt used in the numerical solution of the model equation Eq. (5) to generate the snapshots in the following sections. In future applications this parameter will be dictated by the experimental conditions. In this work, however, it is possible to select it arbitrarily and compare it with the typical system times. Inserting the results for the optimal parameters α_i^+ into Eq. (12) (in its continuous version) then represents the regression model. This is exemplified in the next sections. At this point, it is necessary to comment on the relationship between this method and the method presented in Ref. [16]. In fact, despite differences in application and numerical details, both methods can be regarded as the result of the same basic idea; namely, discretizing a model equation like that defined in Eq. (12) and minimizing an error functional, which in the notation used here is given by Eq. (18). Minimizing D with respect to the parameters α_i gives the matrix equation

$$\boldsymbol{H}^{\mathrm{T}} \cdot \boldsymbol{H} \cdot \boldsymbol{\alpha} = \boldsymbol{H}^{\mathrm{T}} \cdot \boldsymbol{T}. \tag{20}$$

This is the basic equation solved in Ref. [16], but no further details on the numerical scheme are given there. It is reasonable to assume, and this interpretation has also been discussed in Ref. [17], that the least-squares solution in Ref. [16] was calculated by inversion of the matrix $\boldsymbol{H}^{\mathrm{T}} \cdot \boldsymbol{H}$ via

$$\boldsymbol{\alpha}^* = (\boldsymbol{H}^{\mathrm{T}} \cdot \boldsymbol{H})^{-1} \cdot \boldsymbol{H}^{\mathrm{T}} \cdot \boldsymbol{T}. \tag{21}$$

Here, this solution is denoted by $\boldsymbol{\alpha}^*$ to distinguish it from $\boldsymbol{\alpha}^+$. As long as the matrix $\boldsymbol{H}^{\mathrm{T}} \cdot \boldsymbol{H}$ is nonsingular, i.e., all singular values σ_i are nonzero, this is exactly the same result as given by Eq. (19). Although this looks like only a minor limitation at first, in practice it is found that the matrix $\boldsymbol{H}$, and consequently the matrix $\boldsymbol{H}^{\mathrm{T}} \cdot \boldsymbol{H}$, is often ill-conditioned, i.e., a few singular values σ_i are orders of magnitude smaller than the rest and the matrix $\boldsymbol{H}^{\mathrm{T}} \cdot \boldsymbol{H}$ is "almost singular." This leads quite often to significant numerical errors in the computation of the optimal parameters α_i when using Eq. (21) (see, e.g., Ref. [30]). In the cases considered here these errors were of the order of at least 10% or even led to extreme numerical outliers. This problem is avoided in the approach elucidated in this section by the use of the pseudo-inverse $\boldsymbol{H}^+$ to find the least-squares minimizing solution $\boldsymbol{\alpha}^+$. Note that an alternative method to find optimal model parameters has been proposed in Ref. [17], which has certain advantages over least-squares methods like the one presented here and in Ref. [16]. However, this approach has been derived for the case of Kardar-Parisi-Zhang (KPZ)

models. For the particular application in model discovery intended in this work, a least-squares method seems to be the natural and simplest choice.

V. SIMULATION OF SURFACE MORPHOLOGY AND RECONSTRUCTION ANALYSIS

In this section the surface dynamics in the framework of model (5) is discussed. First, cases without beam fluctuations are considered for reference. In later sections the impact of beam fluctuations is discussed as well. In both cases the simulation results for the two-dimensional profiles of the surface height h are used as an input for the model regression method presented in Sec. IV. Time-dependent profiles of the surface height h for cases with zero beam fluctuations, $\widetilde{J} = 0$, serve as a benchmark test to prove the quality of the regression. Since these data originate from a well-defined dynamical system, the algorithm should reconstruct the input parameters of Eq. (5) without errors. In further applications of the model regression analysis the profiles for finite ion-beam fluctuations are considered. These will serve as noisy input and the assumption of a deterministic model in the regression provides approximate and time-dependent coefficients which mimic the stochastic evolution of the surface morphology. Therefore, it is also an example of extracting effective model coefficients, even if the staked frame of the assumed model is approximate. For the numerical analysis, a surface of size $L \times L$ is considered, where $L = 100\sqrt{K_0/\nu_0}$. The reference plane is discretized by using 128×128 grid points. The time integration of Eq. (5) was done by the explicit second-order Ralston's method [Eq. (3.5) in Ref. [31]] with time step $\Delta t = 10^{-3}\tau_0$. For the integration of the evolution equation (5) with finite beam fluctuations included no particular stochastic analysis was used. Although the fluctuating coefficients were evaluated by using the relations from Appendix A, i.e., $J/J_0 = \psi$, obtained from known stochastic integrals of the Itô calculus, the coefficients thus determined were regarded as smooth functions of time and the resulting integrals to advance the height function h in time were evaluated in the Riemann sense. This is similar to approaches in numerical studies of passive scalars where a stochastic velocity field is generated according to stochastic calculus and the resulting advection equations are integrated over time with an explicit Euler method [32–34]. The damping coefficient is chosen as $b = 10^2\tau_0^{-1}$ to keep the averaged height $\overline{h}$ effectively at zero. This serves just for numerical convenience by avoiding a continuously increasing mean value $\overline{h}$ due to the terms proportional to λ_i in Eq. (2). A cross-check with simulations where $b = 0$ showed that the dynamics of the structured piece $\widetilde{h}$ is indeed unaffected.

A. Cases without beam fluctuations

In the first part of this numerical study of surface dynamics and model reconstruction three cases with constant beam amplitude are considered. These are characterized by $\langle J/J_0\rangle = 0.5$, 1.0, 1.5 and $\widetilde{J} = 0$. The choice is guided by the amplitude of beam fluctuations studied in the next section. The values $\langle J/J_0\rangle = 0.5$ and $\langle J/J_0\rangle = 1.5$ can considered as approximate limiting values for the coefficients fluctuating around the reference case with $\langle J/J_0\rangle = 1.0$.

1. Analysis of surface morphology

The simulations are initialized with a noisy height profile of very small amplitude. These perturbations lead to a growth of unstable structures and at a certain level nonlinear processes provide a saturation in the local fluctuations of the surface height. To get some insight into the dynamics of the model systems the layer thickness W is used to characterize the temporal evolution of the surface morphology. It is defined by by the variance of the height fluctuations,

$$W = \sqrt{\overline{h^2} - \overline{h}^2}. \tag{22}$$

As can be seen in Fig. 1, in all three cases, statistically stationary states result after a growth phase with exponential increase of W, which is guided by linear physics. This exponential growth for small times t corresponds to the linear growth rate $\gamma_* = \langle J/J_0\rangle^2/(4\tau_0)$, which is a function of the beam flux due to $\gamma_* \sim \nu^2$. Straight line fit functions in the semilogarithmic plot of Fig. 1 clearly demonstrate the expected time dependence. In the nonlinear phase the layer thickness W fluctuates around a mean value $\langle W\rangle \approx 4\nu_0/\lambda_0$, whereby all three cases show similar behavior. If one considers the probability distribution PDF(h) in Fig. 2 for the fluctuations of the height h in the statistically stationary state, similar distributions appear in all three cases. The PDFs can be approximated very accurately by Gaussian distributions, as shown by the regression curves. Looking at the morphology of the respective surfaces in Fig. 3, however, clear differences can be seen in the structure of the surface patterns. The variation of $\langle J/J_0\rangle$ obviously leads to an enlargement of pattern size for $\langle J/J_0\rangle < 1$ and a reduction for $\langle J/J_0\rangle > 1$ when compared with the reference case with $\langle J/J_0\rangle = 1$. This is in accordance with the scaling of the most unstable mode found in the linear analysis: $l_* \sim \nu^{-1/2} \sim \langle J/J_0\rangle^{-1/2}$. Since the mean deviation of the height is almost the same in all cases, the quantity W cannot be used to distinguish the different morphological features. However, a clear distinction can be achieved by looking at the mean squared gradient $\overline{|\nabla h|^2}$. The squared Fourier components $|c_{ij}|^2$ of the gradient are displayed in Fig. 4. They are defined by

$$\overline{|\nabla h|^2} = \sum_{m,n} \left(m^2k_x^2 + n^2k_y^2\right)|h_{m,n}|^2 \equiv \sum_{m,n} |c_{m,n}|^2, \tag{23}$$

where $h_{m,n}$ denotes the Fourier components of the height function h and $k_x = k_y = 2\pi/L$. This representation clearly reflects the change in the structures. The enlarged pattern size for $\langle J/J_0\rangle < 1$ corresponds to a circular pattern in Fourier space peaked at smaller wave numbers than the reference case with $\langle J/J_0\rangle = 1$. For $\langle J/J_0\rangle > 1$ the situation is reversed. To underline the usefulness of this quantity for the analysis here and in subsequent sections, Fig. 5 shows the evolution of $\overline{|\nabla h|^2}$ over time analogous to W in Fig. 1. The different saturation levels are well separated at $\overline{|\nabla h|^2} = 2.5$, 4.9, and 7.2 $\nu_0^3K_0^{-1}\lambda_0^{-2}$. This corresponds to the change of pattern size by a factor of two when comparing the figures in Fig. 3. Note that, for a single dominant mode, Eq. (23) gives immediately a relation for the average pattern size $\bar{l}^2 \sim W^2/\overline{|\nabla h|^2}$. Now,

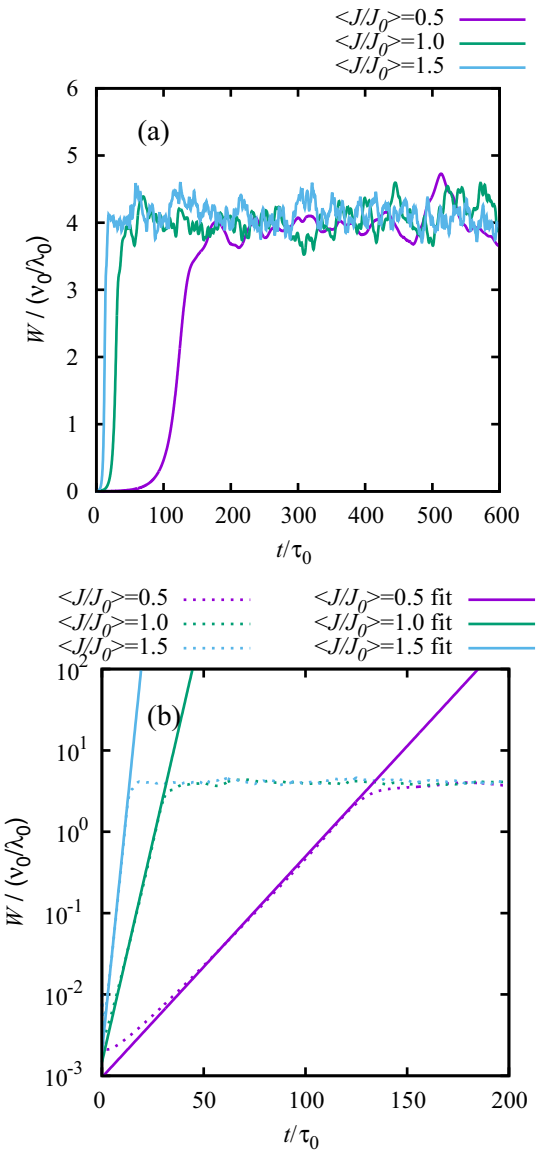

FIG. 1. Panel (a) shows the temporal evolution of the layer thickness W for the cases with constant beam flux characterized by $\langle J/J_0\rangle = 0.5$, 1.0, 1.5 and $\langle \widetilde{J}^2\rangle = 0$. The reference time is defined by $\tau_0 = K_0/\nu_0^2$. The saturation level $\langle W\rangle \approx 4\nu_0/\lambda_0$ is almost the same for all cases. The exponential growth for small times t corresponds to the beam flux dependence of the linear growth rate γ_* of Eq. (6). To illustrate this panel (b) shows a semilogarithmic plot of the same time traces as in the top figure with fit functions (giving the straight lines) included. According to linear theory the fit functions were chosen as $f(t) = b\exp(\gamma_* t)$, where $\gamma_* = \langle J/J_0\rangle^2/(4\tau_0)$. Fit parameters found are $b = 9.6\times10^{-4}$, $b = 1.5\times10^{-3}$, and $b = 2.0\times10^{-3}$ for $\langle J/J_0\rangle = 0.5$, 1.0, 1.5, respectively. Therefore, the straight lines in the semilogarithmic plot become steeper as $\langle J/J_0\rangle$ increases.

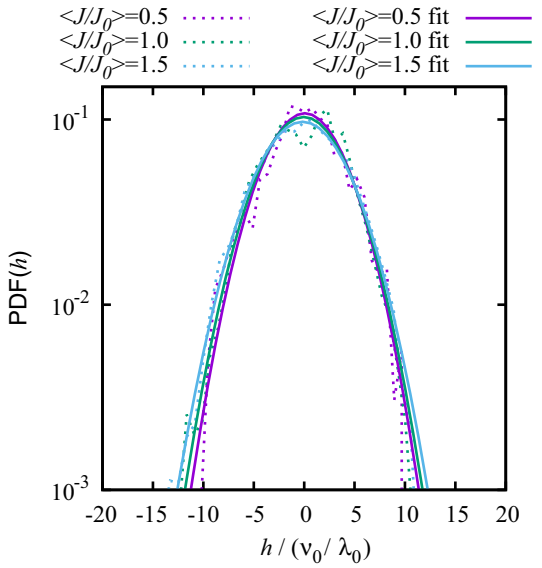

FIG. 2. Distribution functions PDF(h) for the cases with constant beam flux characterized by $\langle J/J_0\rangle = 0.5$, 1.0, 1.5 and $\langle \widetilde{J}^2\rangle = 0$ (dashed lines). Gaussian fits are also shown in the figure for all three cases based on the first and second moments computed for the time traces of h (solid lines). The statistical analysis was performed for the saturated and statistically stationary state.

the results of Figs. 1 and 3 might be compared with the experimental data of Ref. [11], where normal beam incidence has also been considered. In Ref. [11], model parameters have been extracted from experimental data illustrated by Figs. 1 and 2 there. Pattern sizes $\ell_l \sim 35$ nm and timescales $t_l \sim 6.25$ min were found, where the timescale t_l represents a growth rate, similar to $t_* = \gamma_*^{-1}$ in this work. Using the numbers of Sec. III of this work for ν_0, λ_0, and K_0 and the results for the temporal evolution of W for $\langle J/J_0\rangle = 1$, as illustrated by Fig. 1, one obtains $\tau_* \sim 6.25$ min. This is identical to t_l but, of course, this is not too surprising, because the numbers for ν_0, λ_0, and K_0 have been taken from Muñoz-García's work. However, the model discussed by Muñoz-García differs from the one used here and the results prove the accuracy of the simulation results at least for the initial phase, where linear effects dominate. But, also for the nonlinearly saturated phase, one finds by simple dot pattern inspection of Fig. 3 of this work an average pattern size of $\bar{l} \sim 34$ nm. This is fairly close to the value of ℓ_l given in Ref. [11]. Finally, note that the hexagonal nanodot patterns of Fig. 1 in Ref. [11] and of Fig. 1 of this work look quite similar.

2. Analysis of reconstructed model coefficients ν, λ, K

The simulation results for the temporal development of the surfaces for the three examples without ion-beam fluctuations are now used for the model reconstruction described in Sec. IV. For the reconstruction an extended evolution equation for the height h has been used. The detailed *ansatz* is given by Eq. (B1) in Appendix B which defines the "discovery model," where far more terms are considered than were included in the simulations based on Eq. (5) to provide the snapshots. As mentioned in Appendix B, a few of these additional terms have been discussed in the literature as possible and important extensions of KS models. But, for the present study based on the simulations with the model defined by Eq. (5), these terms also serve to test the accuracy of the method in such cases when complicated but irrelevant effects are included in the model identification. Indeed, it has been proven that the reconstruction algorithm is able to determine the relevant coefficients with high accuracy and to identify irrelevant terms.

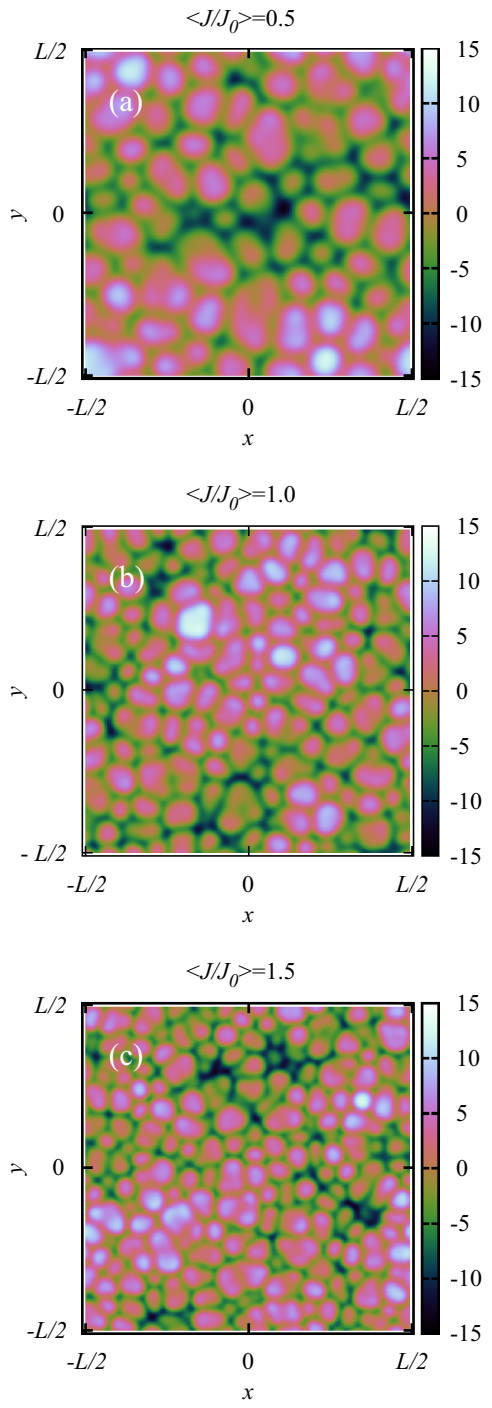

FIG. 3. Snapshots of the morphological structure for the cases with constant beam flux characterized by (a) $\langle J/J_0\rangle = 0.5$, (b) 1.0, (c) 1.5 and $\langle \widetilde{J}^2\rangle = 0$. The domain size is chosen as $L = 100K_0^{1/2}/\nu_0^{1/2}$. The change in the size of the patterns follows the scaling $l_* \sim \nu^{-1/2} \sim \langle J/J_0\rangle^{-1/2}$.

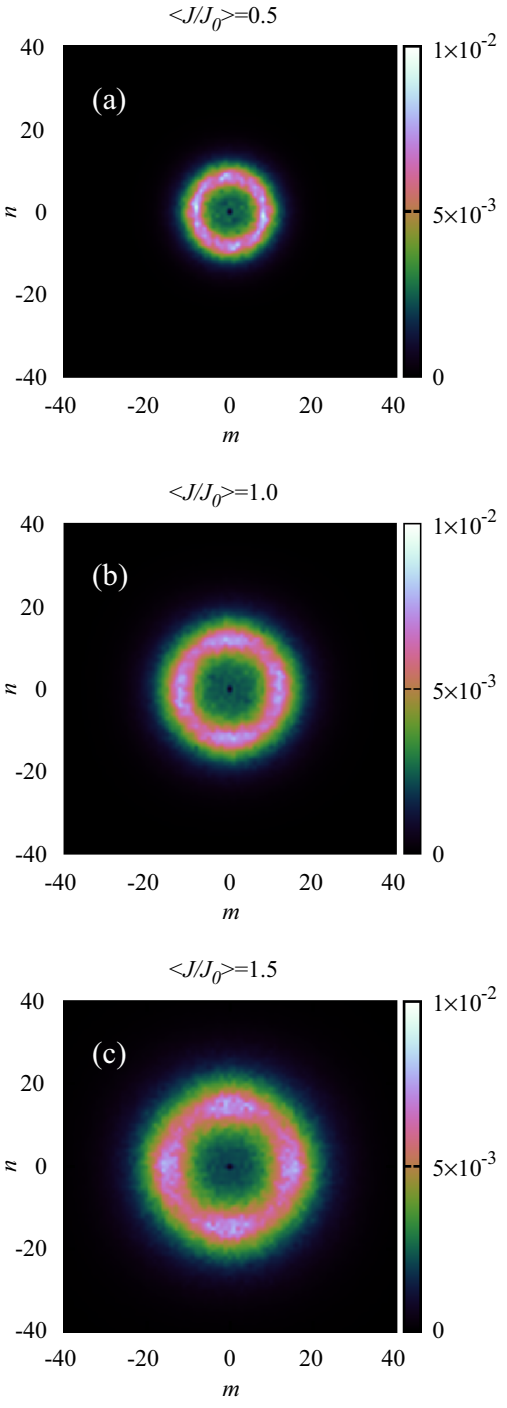

FIG. 4. Shown are the squared components $|c_{ij}|^2$ of the time-averaged squared gradient $\overline{|\nabla h|^2}$ in m-n wave-number plane for the cases with constant beam flux characterized by (a) $\langle J/J_0\rangle = 0.5$, (b) 1.0, (c) 1.5 and $\langle \widetilde{J}^2\rangle = 0$. The Fourier decomposition is given by Eq. (23). The change in structural size displayed in Fig. 3 is reflected by circular patterns of different radii.

This is demonstrated by an example of numerical results listed in Table I in Appendix B. For the reconstruction analysis in this section 200 snapshots from the simulations were used. According to the elucidation of the reconstruction technique in Sec. IV two consecutive snapshots with a time interval of τ_Δ were used for a single evaluation of the model coefficients. By repeating the evaluation for the entire series of 200 snapshots available a series of model parameters is obtained. The results for the 199 reconstructions of the coefficient λ, obtained for the reference case with $\langle J/J_0\rangle = 1$, are shown in Fig. 6. It can be seen that the choice of the temporal increment τ_Δ

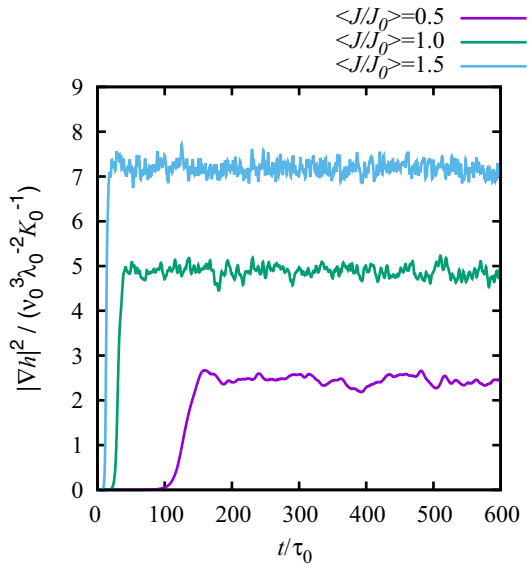

FIG. 5. Temporal evolution of the squared gradient $\overline{|\nabla h|^2}$ for the cases with constant beam flux characterized by $\langle J/J_0\rangle = 0.5$, 1.0, 1.5 and $\langle \widetilde{J}^2\rangle = 0$. The different saturation levels are well separated at $\overline{|\nabla h|^2} = 2.5$, 4.9 and $7.2\,\nu_0^3 K_0^{-1}\lambda_0^{-2}$ for $\langle J/J_0\rangle = 0.5$, 1.0, 1.5, respectively.

plays an important role for the statistics of the 199 individual results. Obviously a very strong reduction of the scatter can be achieved by a suitable choice of τ_Δ. A similar result is obtained for the input parameters ν and K. The coefficients not included in the model simulations lead to regression coefficients of the discovery model Eq. (B1) being smaller

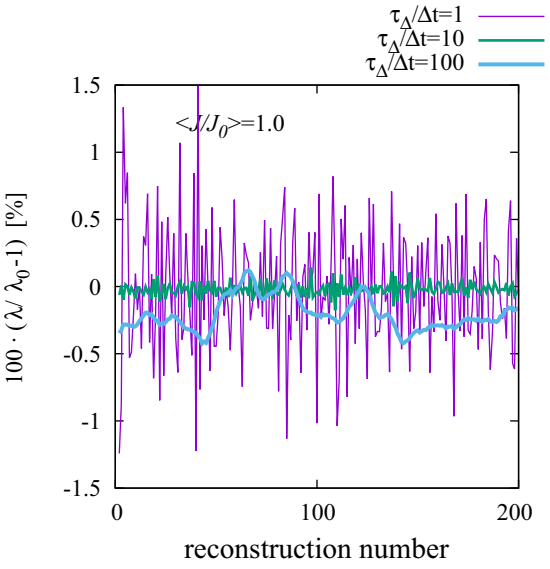

FIG. 6. Shown are the 199 different results for the reconstruction of the coefficient λ obtained by evaluation of subsequent pairs of 200 snapshots of the simulations all of them with $\langle J/J_0\rangle = 1$ but with different increments $\tau_\Delta/\Delta t = 1$, 10, 100 between the two respective snapshots used for the reconstruction method. The calculated values are represented by their relative deviation from the known input parameter λ_0 in percent.

than ν, λ, and K by several orders of magnitude. Therefore, the extra terms included in the discovery model are found to be negligible—as they should be in this numerical test. However, note in this context that the damping term plays a special role, as is discussed in Appendix B. The distribution of the relevant coefficients found in the regression procedure have clear Gaussian signatures, as illustrated by the plots in Fig. 7. There the distribution functions of the coefficients ν, λ, and K for differently selected increments are displayed. In addition, Gaussian fits $G(s)$ are drawn, which were determined from the mean values μ_ν, μ_λ, and μ_K and the variances σ_ν^2, σ_λ^2, and σ_K^2 of each series of results using Gaussian functions for $s = \nu, \lambda, K$:

$$G(s) = \frac{1}{\sqrt{2\pi\sigma_s^2}} \exp\left[-\frac{(s-\mu_s)^2}{2\sigma_s^2}\right], \tag{24}$$

and for the computation of averages the relations

$$\mu_s = \frac{1}{R}\sum_{k=1}^{R} s_k, \quad \sigma_s^2 = \frac{1}{R}\sum_{k=1}^{R}(s_k-\mu_s)^2, \tag{25}$$

with R denoting the number of reconstructions. It can be stated that the reconstructed coefficients are clearly Gaussian distributed, and the variance and the mean value are influenced by the choice of the increment τ_Δ. An explanation for this obvious Gaussian signature in Fig. 7 is pending. To shed more light on this connection between reconstruction accuracy and time increment τ_Δ, Fig. 8 shows the deviations of the numerical mean values μ_ν, μ_λ, and μ_K for the three example cases as a function of the ratio τ_Δ/τ_*. The timescale τ_* is the inverse of the growth rate γ_* It takes into account the changes of the characteristic timescale of the system with varying beam flux:

$$\tau_* = \frac{4K_0}{\nu_0^2}\left\langle\frac{J}{J_0}\right\rangle^{-2} = 4\tau_0\left\langle\frac{J}{J_0}\right\rangle^{-2}. \tag{26}$$

In the plots of Fig. 8 it can be seen that the deviations of ν and λ from the correct value are less than 0.1% for appropriately chosen increments τ_Δ. The qualitative picture is similar for μ_K, but there the error for the case with $\langle J/J_0\rangle = 0.5$ is 1% at best. A certain optimal choice of the increment seems to be possible, at which the coefficients can be determined particularly well, but the particular value for the ratio τ_Δ/τ_* seems to be different for the different choices of $\langle J/J_0\rangle$. In Fig. 9 the standard deviations σ_ν, σ_λ, and σ_K are displayed, which are scaled by the respective true values of the model coefficients. Again a plot versus the scaled increment τ_Δ/τ_* is chosen and, surprisingly, the functional form of σ_ν, σ_λ, and σ_K are all close to each other. The scatter in the reconstructed coefficients has a pronounced minimum close to $5\times10^{-3}\tau_\Delta/\tau_*$. For both cases, reduction and increase of τ_Δ, the noise in the determination of the model coefficients increases. Careful inspection shows that the standard deviations can be approximated reasonably well by a functional form

$$\frac{\sigma_s}{s} = c_1\frac{\tau_*}{\tau_\Delta} + c_2\left(\frac{\tau_\Delta}{\tau_*}\right)^{3/4}\left(\frac{\tau_\Delta}{\Delta t}\right)^{1/4}, \tag{27}$$

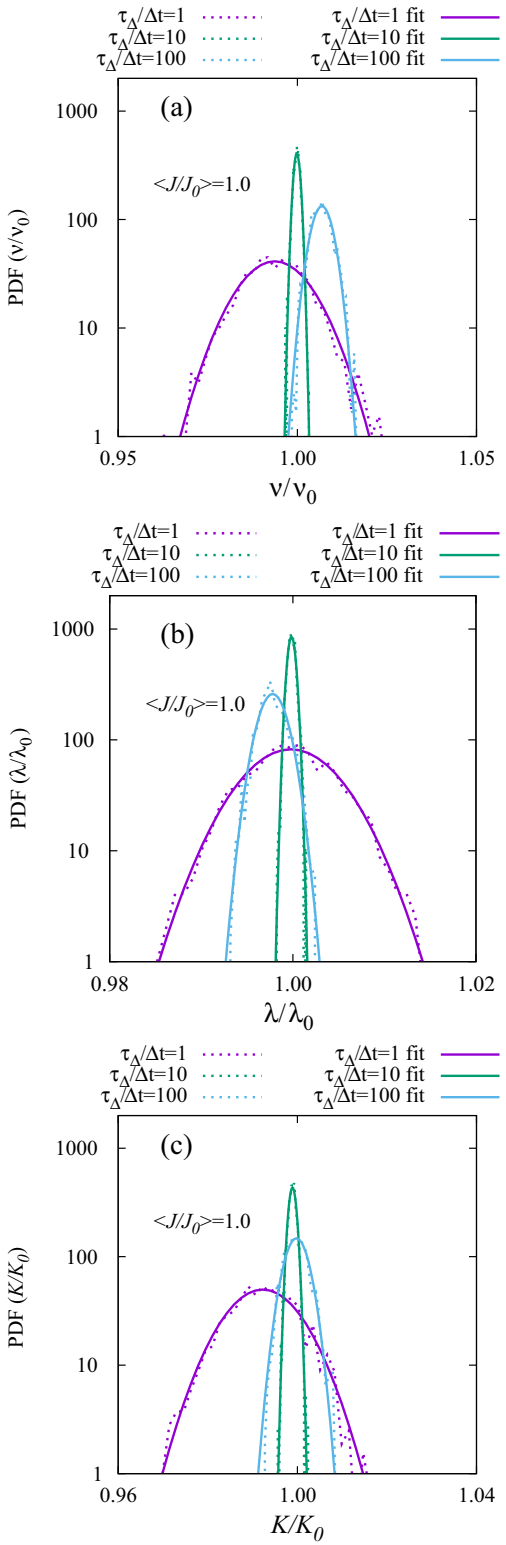

FIG. 7. Shown are the distribution functions (a) PDF(ν/ν_0), (b) PDF(λ/λ_0), and (c) PDF(K/K_0) for the reconstruction using different increments $\tau_\Delta/\Delta t = 1, 10, 100$ and based on 200 snapshots of the simulation with $\langle J/J_0\rangle = 1$. Gaussian fits $G(\nu/\nu_0)$, $G(\lambda/\lambda_0)$, and $G(K/K_0)$ are also shown in the figure based on moments of ν, λ, and K computed for the series of reconstruction results. The narrowest distribution occurs in all cases (a), (b) and (c) for the parameter $\tau_\Delta/\Delta t = 10$, the widest for $\tau_\Delta/\Delta t = 1$.

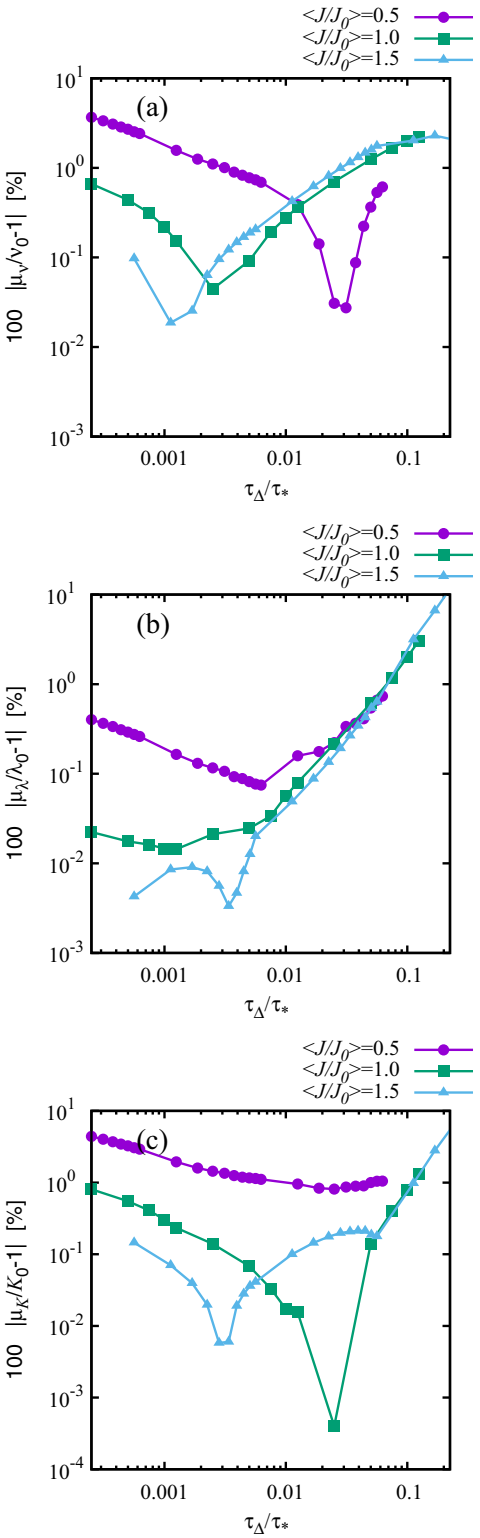

FIG. 8. Plots of the relative error of the averaged values (a) μ_ν, (b) μ_λ, and (c) μ_K obtained by averaging the reconstructed coefficients over all 199 reconstructions. For all cases $\langle J/J_0\rangle = 0.5$, 1.0, 1.5 the relative error can be less than 1% as long as the incremental time τ_Δ is smaller than $5 \times 10^{-3}\tau_*$.

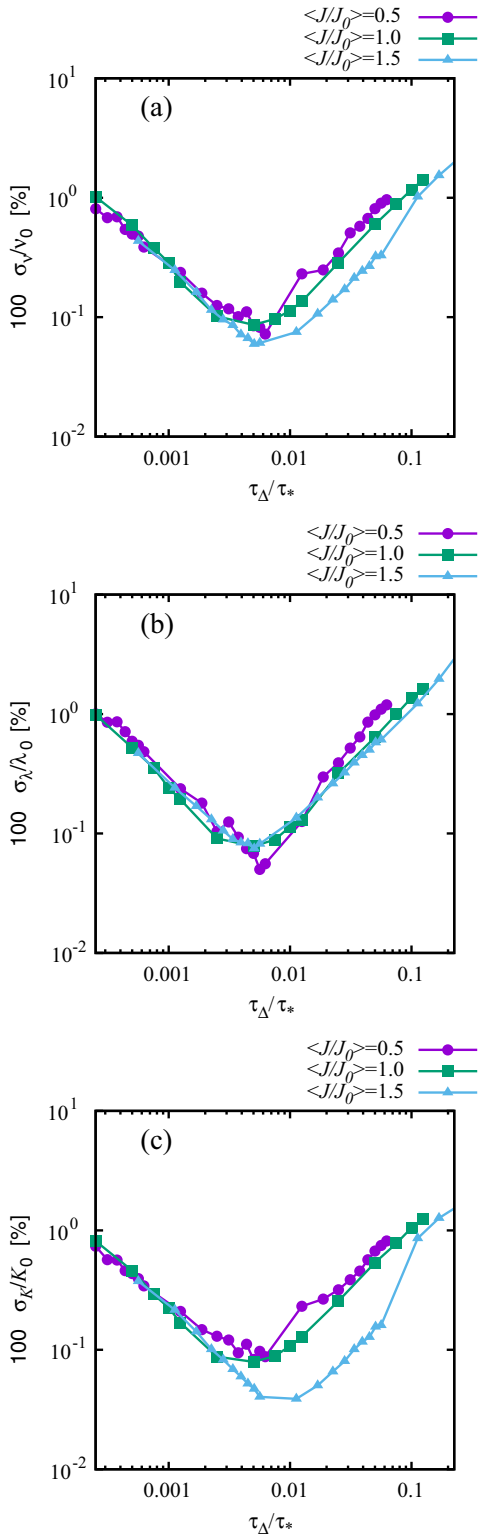

FIG. 9. Plots of the relative standard deviation (a) σ_v, (b) σ_λ, and (c) σ_K obtained by averaging the reconstructed coefficients over all 199 reconstructions. For all cases $\langle J/J_0\rangle = 0.5$, 1.0, 1.5 the curves are almost identical and a pronounced minimum occurs at around $\tau_\Delta \sim 5 \times 10^{-3}\tau_*$.

where c_1 and c_2 are constant. However, the formula (27) has only been verified for the present cases and cannot yet be derived from basic principles. Nevertheless, the qualitative trends become plausible if one considers that a small value for the ratio τ_Δ/τ_* is synonymous with only a slight change of h between consecutive snapshots. The corresponding discrete approximation of the time derivative $\partial h/\partial t$ tends to zero with decreasing τ_Δ/τ_*. Then the reconstruction model given by Eqs. (12) and (13) and leading to the solution (19) is not well defined anymore. On the other hand, an increase of τ_Δ/τ_* can be considered as an averaging procedure, where responses of the model system are not resolved, but rather a time-averaged reaction is considered. Both types of inaccuracies are caused by some kind of information loss. This impairs the usefulness of the method for both limits $\tau_\Delta/\tau_* \to 0$ and $\tau_\Delta/\tau_* \to \infty$. Thus, an optimal situation is found somewhere in an intermediate range of τ_Δ/τ_* and, according to the simple approximation of Eq. (27), this is found close to $\tau_\Delta = \sqrt{c_1/c_2}\tau_*^{7/8}\Delta t^{1/8}$. Therefore, in order to evaluate unknown data as accurately as possible, e.g., from an experiment, information about typical timescales of the system to be studied is required, i.e., information about the linear instabilities which also determine the timescale τ_* in these studies.

Before the influence of ion-beam fluctuations is studied in the next section, the results of this section shall be briefly summarized.

(1) When initialized with small amplitude noise in the height, the simulations evolve along well-known linear instabilities with exponential growth. Then, even without beam fluctuations, the model system shows fluctuating surface structures and nonlinear saturation (Figs. 1 and 5).

(2) The surface patterns strongly depend on a reduction or an increase of the ion flow (Figs. 3 and 4). The changes can be explained by rescaling the characteristic temporal and spatial scales, γ_* and l_*, respectively.

(3) Although the structures in the three cases are different, the distribution functions of the local fluctuations of the surface height h are almost identical and show a clear Gaussian characteristic (Fig. 2).

(4) The scattering of the reconstructed model coefficients determined from a large number of snapshots has a clear Gaussian characteristic. This Gaussian noise in the reconstruction introduces inherent uncertainties in the model discovery (Figs. 6 and 7).

(5) The reconstruction of the model coefficients can be optimized to an accuracy of 0.1% or less if the temporal increment is chosen appropriately. Then the mean value of the coefficients from a large number of evaluated snapshots deviates from the true value by a maximum of 0.1% and the standard deviation is of the same order (Figs. 8 and 9).

(6) The choice of the temporal increment τ_Δ represents a critical part of the method, for which no general explanation and optimization method can be given at the moment. Despite these uncertainties, the analysis offers a good approach to optimizing the method—at least for the moment and for the cases considered. This calibration might be guided by the formula given in Eq. (27) which gives an empirical estimate of the standard deviation of reconstructed parameters as a

function of the characteristic timescale τ_* and the sampling time τ_Δ.

B. Cases with beam fluctuations

In the following sections, fluctuations of the ion beam are taken into account, whereby for the temporal variation of the ion-beam amplitude a lognormal distribution is considered. The lognormal distribution was chosen because of its relevance in turbulent fluctuations of plasma beams [22]. The beam fluctuations are prescribed by a mean value $\langle J/J_0\rangle = 1$ and a variance $\langle \widetilde{J}^2/J_0^2\rangle = 0.09$. A scan is performed for the correlation time τ of the lognormally distributed beam noise along the series $\tau/\tau_0 = 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^{0}, 10^{1}$.

1. Analysis of surface morphology

As in the previous section the layer thickness W and the mean-squared gradient $\overline{|\nabla h|^2}$ are used to follow the evolution of the surface morphology. The results for the series of correlation times τ/τ_0 are displayed in Fig. 10. Again, statistically stationary states with $\langle W\rangle \approx 4$ are found after an exponential growth dominated by the most unstable surface mode. The plot of $\overline{|\nabla h|^2}$ shows an increasing scatter when the scaled correlation time τ/τ_0 is increased. Also it can be seen that for the larger values of τ/τ_0 considered the mean-squared gradient is found to be roughly in the range between 3 and 7 $\nu_0^3 K_0^{-1}\lambda_0^{-2}$, which are close to the saturation levels $\langle\overline{|\nabla h|^2}\rangle \approx 2.5\nu_0^3 K_0^{-1}\lambda_0^{-2}$ and $\langle\overline{|\nabla h|^2}\rangle \approx 7.2\nu_0^3 K_0^{-1}\lambda_0^{-2}$ of the constant-beam results of Fig. 5. Of course, this agreement is not accidental, but is due to the fact that the parameters of the constant beam simulations have been chosen so that the model parameters limit approximately the range covered by the fluctuations. Consequently, it can be concluded that the size of the structures oscillates between the two boundary cases with $\langle J/J_0\rangle = 0.5$ and $\langle J/J_0\rangle = 1.5$ considered in Sec. V A.

2. Analysis of reconstructed model coefficients ν, λ, K

Due to the oscillations in the surface properties and the results from the above investigations for constant ion beams, it can be expected that the reconstructed model coefficients also have a distribution reflecting the fluctuations of the ion beam. To analyze details, the PDFs of ν/ν_0 and λ/λ_0 are displayed in Fig. 11. It is found that the numerical PDFs can be approximated very well with lognormal fits $H(s)$:

$$H(s) = \frac{1}{s\sqrt{2\pi\ln\left(\sigma_s^2/\mu_s^2\right)}}\exp\left[-\frac{\left(\ln s - \ln\mu_s^2/\sigma_s\right)^2}{2\ln\left(\sigma_s^2/\mu_s^2\right)}\right], \tag{28}$$

where $s = \nu$, λ and μ_s and σ_s^2 are defined by Eq. (25). The values of ν/ν_0 and λ/λ_0 show a pronounced scatter which is about 10% for the smaller values of τ/τ_0 but is increased up to 50% for $\tau/\tau_0 > 10^{-1}$. The comparison with the beam flux distribution also shown in Fig. 11 proves that the increase of the correlation time τ/τ_0 pushes the distribution of the scaled model parameters closer to the distribution of the scaled beam flux. Again, a qualitative explanation can be found for this

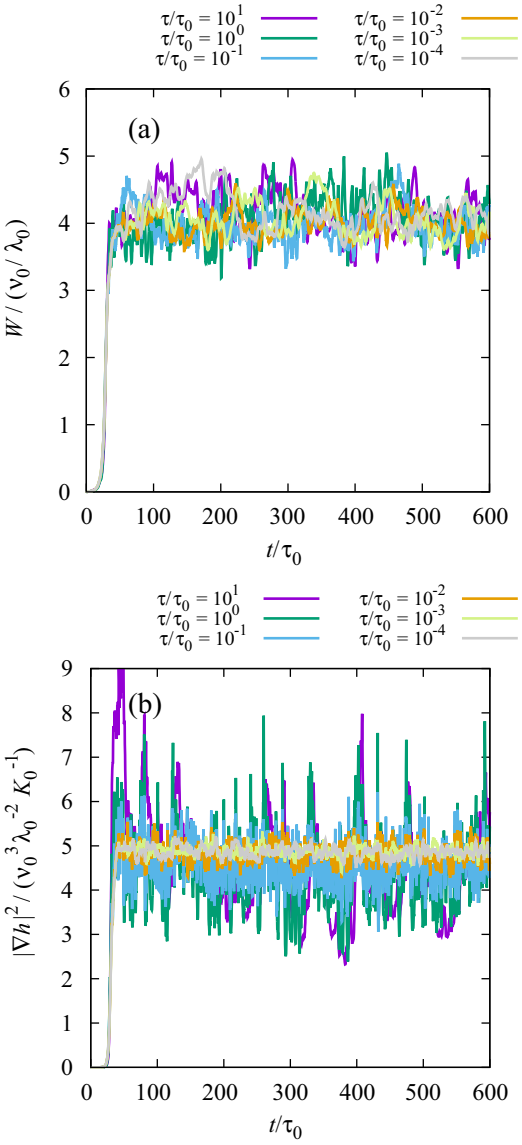

FIG. 10. Temporal evolution of (a) the layer thickness W and (b) the squared gradient $\overline{|\nabla h|^2}$ for the cases with a fluctuating beam flux characterized by $\langle J/J_0\rangle = 1.0$, $\langle \widetilde{J}^2/J_0^2\rangle = 0.09$, and $\tau/\tau_0 = 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^{0}, 10^{1}$. For all values of τ/τ_0 in the statistically stationary phase the temporal average of the layer thickness is $\langle W\rangle \approx 4\,\nu_0/\lambda_0$. The respective mean values of the squared gradient $\langle\overline{|\nabla h|^2}\rangle$ lie in a range between 4.3 and 4.9 $\nu_0^3 K_0^{-1}\lambda_0^{-2}$. The scattering of $\overline{|\nabla h|^2}$ decreases when the scaled correlation time τ/τ_0 is reduced.

trend: The reference time τ_0 characterizes the model system as far as it describes the linear growth rate on the one hand and stands for a response time on the other hand. For small values $\tau/\tau_0 \ll 1$ the system cannot follow the fluctuations of the beam on the timescale τ. The model coefficients stay effectively close to the average value and the beam fluctuations are canceled out to a certain extent. Some noise level is left and the standard deviation is clearly affected by this. On the contrary, a large value of τ/τ_0 represents slow beam fluctuations with respect to the response time of the system. Then the dynamics can easily follow through the entire range of parameter variations. In this case the lognormal beam

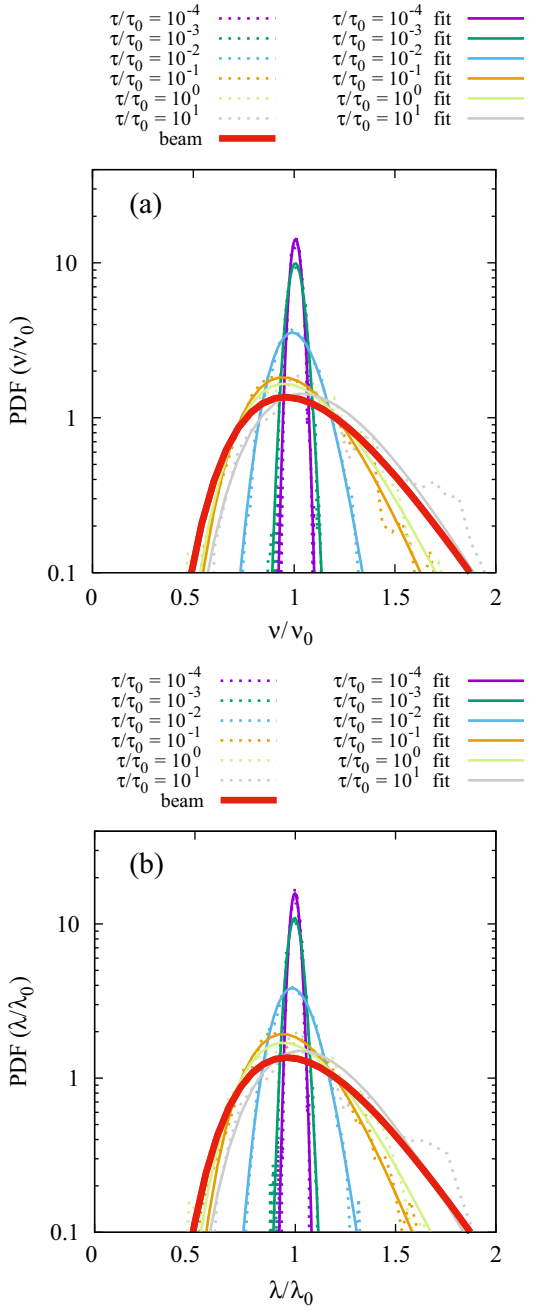

FIG. 11. Reconstruction results for the cases with fluctuating beam flux characterized by $\langle J/J_0\rangle = 1.0$, $\langle \widetilde{J}^2/J_0^2\rangle = 0.09$, and $\tau/\tau_0 = 10^{-4}$, 10^{-3}, 10^{-2}, 10^{-1}, 10^{0}, 10^{1}. Shown are the distribution functions (a) PDF(ν/ν_0) and (b) PDF(λ/λ_0) for an increment $\tau_\Delta/\Delta t = 100$ and different correlation lengths $\tau/\Delta t = 10^{-1}$, 10^{0}, 10^{1}, 10^{2}, 10^{3}, 10^{4}. The dotted lines represent the raw numerical data and the solid line the lognormal fits $H(\nu/\nu_0)$ and $H(\lambda/\lambda_0)$. The evaluation is done for 2000 snapshots. The lognormal distribution of the beam flux is also indicated by the thick red curve.

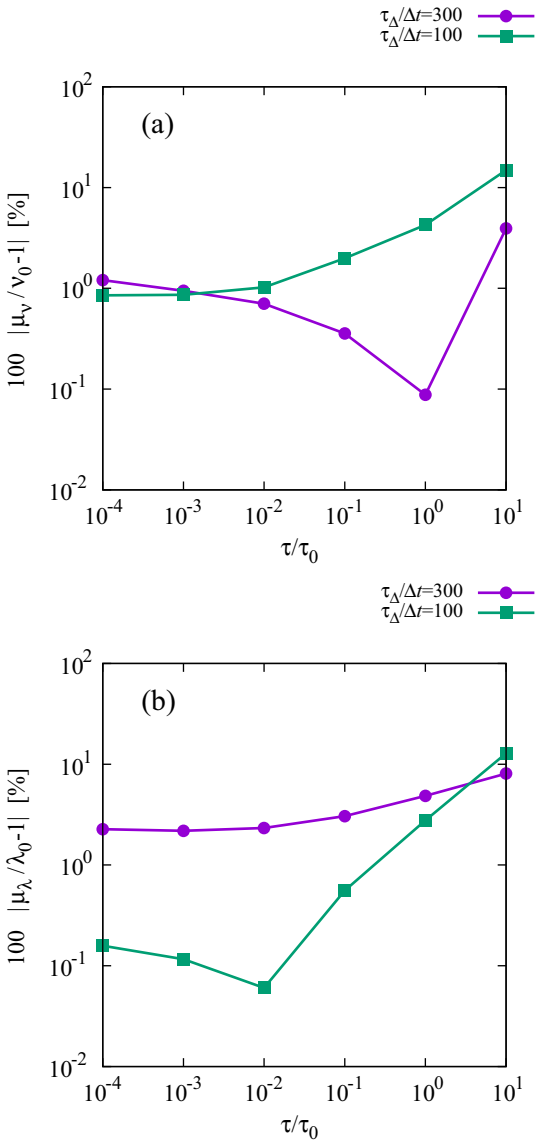

FIG. 12. Plots of the relative error of the averaged values (a) μ_ν and (b) μ_λ obtained by averaging the reconstructed coefficients over 1999 reconstructions. Fluctuating beams have been considered with $\langle J/J_0\rangle = 1.0$, $\langle \widetilde{J}^2/J_0^2\rangle = 0.09$, and $\tau/\tau_0 = 10^{-4}$, 10^{-3}, 10^{-2}, 10^{-1}, 10^{0}, 10^{1}.

distribution is recovered by the distribution of reconstruction results. To quantify this effect even more, Figs. 12 and 13 show the numerical results for the moments of the PDFs of Fig. 11. The deviation of the averaged mean values and standard deviations of ν and λ are given with respect to the mean values ν_0 and λ_0. Now the mean values ν_0 and λ_0 represent the model parameters corresponding to the mean value $\langle J/J_0\rangle = 1$ of the beam lognormal distribution. As can be seen in Fig. 12 the reconstructed values of ν/ν_0 and λ/λ_0 are close to 1 with a few percent deviation if $\tau/\tau_0 \ll 1$. For larger values of τ/τ_0 the results are far from this mean and also show a strong dependence on the choice of increment τ_Δ, indicated by additional results obtained with $\tau_\Delta/\Delta t = 300$. Actually, this is in accordance with the arguments given above: For $\tau/\tau_0 \ll 1$ the beam fluctuations are canceled out to a certain extent and one is left with a noisy but fairly accurate reconstruction of the mean values ν_0 and λ_0. For values $\tau/\tau_0 > 10^{-2}$ the results become completely random and unpredictable. Then the average values of ν and λ are far from the mean value ν_0 and λ_0. Also the choice of the sampling time τ_Δ changes the results drastically. This is explained by the bad resolution

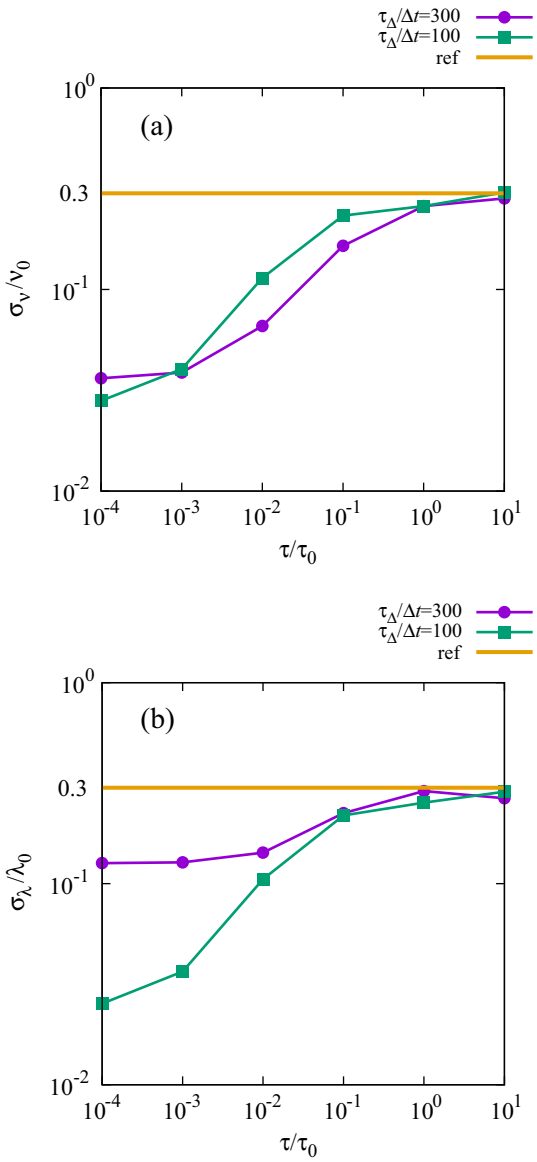

FIG. 13. Plots of the relative standard deviation (a) σ_ν and (b) σ_λ obtained by averaging the reconstructed coefficients over 1999 reconstructions. Fluctuating beams have been considered with $\langle J/J_0\rangle = 1.0$, $\langle \tilde{J}^2/J_0^2\rangle = 0.09$ and $\tau/\tau_0 = 10^{-4}$, 10^{-3}, 10^{-2}, 10^{-1}, 10^0, 10^1. The reference value denoted by "ref" in the plots marks the standard deviation of the imposed beam fluctuations $\langle \tilde{J}^2/J_0^2\rangle^{1/2} = 0.3$, which represents the limit of long correlation times $\tau/\tau_0 \gg 1$, i.e., for cases when the surface dynamics is fast enough to adapt the slow beam fluctuation dynamics.

of the beam correlation time, due to the fact that the total period covered by the snapshots of the reconstructions does not give sufficient statistics of the slow beam fluctuations. These conclusions are supported by the standard deviations displayed in Fig. 13. For small τ/τ_0 the standard deviation is in a range of a few percent, but it increases for increasing τ/τ_0 to the value $\sigma_\nu/\nu_0 = \sigma_\lambda/\lambda_0 = 0.3$. This value is imposed by the beam statistics with $\langle \tilde{J}^2/J_0^2\rangle^{1/2} = 0.3$. If one considers the results of Figs. 12 and 13 together, it is to be expected that an improvement of statistics by increasing the number of snapshots would permit a more accurate calculation of the mean values ν_0 and λ_0. In spite of all this, in a typical experimental arrangement, the requirement $\tau/\tau_0 \ll 10^{-2}$ is usually met, and as a result a reconstruction can be expected with an accuracy of a few percent.

VI. SUMMARY

As preparatory work for future analysis of experimental data a model discovery technique for the identification of Kuramoto-Sivashinsky model parameters in ion-beam-induced surface morphological dynamics has been discussed. The reconstruction method has been applied to a paradigmatic model described by a scaled evolution equation for the surface height including ion-beam fluctuations. The reconstruction works very well for examples with constant beam flux. Prescribed model coefficients could be extracted with an error less than 1%. The coefficients found have statistical variations that have been studied in detail. It is shown that an inherent Gaussian noise occurs in the reconstruction of the model parameters, which varies strongly with the physical timescales of the model system and the numerical parameters of the method, especially the sampling rate. Parameter scans show the possibility and the need for an optimal choice of sampling intervals and time steps for the data evaluation. In addition to these calculations with a constant ion beam, further studies were carried out to examine the influence of ion beam fluctuations on the surface morphology. There the reconstruction method has been applied to the case of nonconstant model coefficients. To cope with typical features of plasma beam statistics a lognormal multiplicative noise was included in the surface model. In this case of fluctuating coefficients, it could be seen that the quality of the reconstruction depends essentially on the interplay of the correlation time of the beam fluctuations and the sampling time. The distribution of the coefficients in the reconstructions follows the statistics of the beam and an improvement in the quality of the results is achieved by a sampling time that is greater than or at least equal to the correlation time of the beam.

In summary, the presented method provides a very accurate tool for identifying model parameters for continuum models of surface dynamics. This can prove to be a useful method for deriving empirical models if there is a sufficient amount of experimental data available. Another application would be to derive effective macroscopic models for numerical simulations of microscopic processes, such as a continuum approximation for atomistic Monte Carlo simulations. However, the inevitable statistical variations of the reconstruction results require a careful analysis of the numerical parameters to account for the characteristic times of the system under consideration (growth rates of instabilities, correlation times of fluctuations, etc.). For a further assessment of the method, it is particularly necessary to examine Eq. (27) in more detail by statistically analyzing a large number of further simulations. This equation gives an empirical estimate of the quality (the variance) of reconstructed parameters as a function of the characteristic timescale τ_* of the physical system and the sampling time τ_Δ.

ACKNOWLEDGMENTS

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under Grant Agreement No. 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission. Special thanks go to the unknown reviewer who contributed much to the improvement of the paper.

APPENDIX A: LOGNORMAL DISTRIBUTED VARIABLE WITH TEMPORAL CORRELATION

The random quantity ϕ is assumed to obey the stochastic Langevin equation

$$\tau \frac{\partial \phi}{\partial t} = -\phi + \epsilon \zeta. \tag{A1}$$

Here ζ is Gaussian white noise with $\langle \zeta \rangle$ and correlation

$$\langle \zeta(t_1)\zeta(t_2)\rangle = \delta(t_1 - t_2). \tag{A2}$$

The constant ϵ is the amplitude of the noise and τ is the correlation time. This defines an Ornstein-Uhlenbeck process whose solution is

$$\phi(t + \Delta t) = \phi(t) e^{-\Delta t/\tau} + \omega, \tag{A3}$$

where

$$\omega = \frac{\epsilon}{\tau} \int_t^{t+\Delta t} \exp\left(-\frac{t + \Delta t - t'}{\tau}\right) \zeta \, dt'. \tag{A4}$$

It follows that ω is a Gaussian distributed random number with mean $\langle \omega \rangle = 0$ and variance

$$\langle \omega^2 \rangle = \sigma^2 (1 - e^{-2\Delta t/\tau}), \tag{A5}$$

where

$$\sigma^2 = \frac{\epsilon^2}{2\tau}. \tag{A6}$$

Thus, in the statistically stationary phase, the random variable ϕ is characterized by

$$\langle \phi(t + \Delta t)\rangle = \langle \phi(t)\rangle = 0, \tag{A7}$$

$$\langle \phi^2(t + \Delta t)\rangle = \langle \phi^2(t)\rangle = \sigma^2. \tag{A8}$$

Details on this can be found in Refs. [35,36]. For the covariance C_ϕ one finds

$$C_\phi(\Delta t) = \langle \phi(t + \Delta t)\phi(t)\rangle = \sigma^2 e^{-\Delta t/\tau}, \tag{A9}$$

and its correlation coefficient R_ϕ reads

$$R_\phi(\Delta t) = \frac{\langle \phi(t + \Delta t)\phi(t)\rangle}{[\langle \phi(t + \Delta t)^2\rangle \langle \phi(t)^2\rangle]^{1/2}} = e^{-\Delta t/\tau}. \tag{A10}$$

This in turn gives the correlation time τ according to

$$\int_0^\infty R_\phi(t') dt' = \tau. \tag{A11}$$

Using these results, it follows for the random quantity $\psi = e^{\mu+\phi}$ that

$$\psi(t + \Delta t) = e^{\omega + \mu - \mu e^{-\Delta t/\tau}} \psi(t)^{e^{-\Delta t/\tau}}. \tag{A12}$$

One obtains for the first two moments

$$\langle \psi(t + \Delta t)\rangle = \langle \psi(t)\rangle = e^{\mu + \sigma^2/2}, \tag{A13}$$

$$\langle \psi^2(t + \Delta t)\rangle = \langle \psi^2(t)\rangle = e^{2\mu + 2\sigma^2}. \tag{A14}$$

Thus, the covariance and correlation coefficient result as

$$C_\psi(\Delta t) = \exp[2\mu + \sigma^2(1 + e^{-\Delta t/\tau})], \tag{A15}$$

$$R_\psi(\Delta t) = \exp[\sigma^2(e^{-\Delta t/\tau} - 1)]. \tag{A16}$$

Note that the correlation coefficient of ψ does not vanish for $\Delta t \to \infty$ and a lower limit exists: $R_\psi \geqslant e^{-\sigma^2}$.

APPENDIX B: DISCOVERY MODEL

The particular *ansatz* for the discretized Eq. (12) used to conduct the reconstruction and discovery analysis of Secs. V A 2 and V B 2 is chosen as follows:

$$\begin{aligned}
\frac{\partial h}{\partial t} = {} & \alpha_1 + \alpha_2 h + \alpha_3 h^2 + \alpha_4 \partial_x h + \alpha_5 \partial_y h \\
& + \alpha_6 (\partial_x h)^2 + \alpha_7 (\partial_y h)^2 + \alpha_8 (\partial_x h)(\partial_y h) \\
& + \alpha_9 \partial_{xx} h + \alpha_{10} \partial_{yy} h + \alpha_{11} \partial_{xy} h \\
& + \alpha_{12} \partial_{xxxx} h + \alpha_{13} \partial_{yyyy} h + \alpha_{14} \partial_{xxyy} h \\
& + \alpha_{15} (\partial_{xx} h + \partial_{yy} h)[(\partial_x h)^2 + (\partial_y h)^2] \\
& + \alpha_{16} (\partial_x h)^3 + \alpha_{17} (\partial_y h)^3 + \alpha_{18} \partial_{xxx} h \\
& + \alpha_{19} \partial_{xxy} h + \alpha_{20} \partial_{xyy} h + \alpha_{21} \partial_{yyy} h \\
& + \alpha_{22} (\partial_x h)(\partial_{xx} h) + \alpha_{23} (\partial_x h)(\partial_{yy} h) \\
& + \alpha_{24} (\partial_y h)(\partial_{xx} h) + \alpha_{25} (\partial_y h)(\partial_{yy} h) \\
& + \alpha_{26} \partial_{xx} (\partial_x h)^2 + \alpha_{27} \partial_{xx} (\partial_y h)^2 \\
& + \alpha_{28} \partial_{yy} (\partial_x h)^2 + \alpha_{29} \partial_{yy} (\partial_y h)^2 + \alpha_{30} \overline{h}.
\end{aligned} \tag{B1}$$

Even though in the present study many of the terms of this candidate model for data analysis are just used to demonstrate the ability of the model discovery method to identify correctly also zero coefficients for noncontributing processes, many of the terms listed in Eq. (B1) have a particular meaning and have been discussed in detail in the literature. The coefficient α_1 represents a constant erosion and terms proportional to α_2 and α_{30} represent damping effects. In particular the damping term $\alpha_2 h$ has been proven to be very important in the formation of hexagonal patterns [20]. The terms with α_4 and α_5 represent constant velocities of the surface structures. These are usually considered as a consequence of grazing incidence of the ion beam [4–6]. The coefficients α_6 and α_7 describe the basic process of slope-dependent erosion. The terms with α_9, α_{10}, and α_{11} describe the effect of surface tension in the erosion and α_{12}, α_{13}, and α_{14} represent surface diffusion. An extension of the KS model with a term like the one proportional to α_{15} has been discussed in the context of step morphology and Cahn-Hilliard models [18,19]. Also, an extension by the terms with coefficients α_{26}, α_{27}, α_{28}, and α_{29} has been found important to take into account the coupling between erosion and surface transport to lowest order [12]. It is often called the conserved Kardar-Parisi-Zhang term (CKPZ term) and its isotropic form ($\alpha_{26} = \alpha_{27} = \alpha_{28} = \alpha_{29}$)

TABLE I. Reconstruction results for the model parameters $\alpha_1, \ldots, \alpha_{30}$ obtained for a simulation run with $\langle J/J_0 \rangle = 1$ as described in Sec. V A 2. For the averaging, 400 snapshots have been used.

α_1	α_2	α_3	α_4	α_5	α_6
-2.5	7.9×10^{-4}	-1.3×10^{-5}	5.2×10^{-4}	2.5×10^{-4}	0.4999
α_7	α_8	α_9	α_{10}	α_{11}	α_{12}
0.4999	2.2×10^{-5}	-0.993	-0.993	-5.8×10^{-5}	-0.992
α_{13}	α_{14}	α_{15}	α_{16}	α_{17}	α_{18}
-0.992	-1.990	2.9×10^{-4}	-9.0×10^{-6}	-4.5×10^{-5}	8.9×10^{-5}
α_{19}	α_{20}	α_{21}	α_{22}	α_{23}	α_{24}
2.8×10^{-4}	1.4×10^{-5}	-2.5×10^{-5}	3.3×10^{-5}	-2.1×10^{-6}	1.7×10^{-5}
α_{25}	α_{26}	α_{27}	α_{28}	α_{29}	α_{30}
2.9×10^{-5}	1.1×10^{-3}	7.4×10^{-4}	6.6×10^{-4}	1.1×10^{-3}	-1.34

has been discussed, e.g., in Refs. [10,37–39]. A derivation and discussion of an anisotropic form is reported in Ref. [40]. Further information on the significance of these model terms and theoretical derivations from microscopic theories can be found in Refs. [1–21] and the references listed there. The other terms of Eq. (B1) not mentioned in this section do not have an obvious meaning and have been included just to keep the model discovery flexible for future applications using experimental data. Now the comparison of Eq. (B1) with the particular simulation model (5) and the parameters used in Sec. V A results in the following assignment:

$$\alpha_6 = \alpha_7 = \frac{1}{2}\left\langle \frac{J}{J_0} \right\rangle, \quad \alpha_9 = \alpha_{10} = -\left\langle \frac{J}{J_0} \right\rangle, \tag{B2}$$

$$\alpha_{12} = \alpha_{13} = -1, \quad \alpha_{14} = -2, \tag{B3}$$

$$\alpha_{30} = -100. \tag{B4}$$

The other coefficients should be zero for the cases considered in Sec. V. This is indeed obtained to a high level of accuracy as illustrated by the numbers in the Table I containing the averaged coefficients α_i for a reconstruction done for 400 snapshots based on the simulations of Sec. V A 2 with $\langle J/J_0 \rangle = 1$. The coefficients α_6, α_7, α_9, α_{10}, α_{12}, α_{13}, and α_{14} are found with an error less than than 1%. One important point should be emphasized: At first glance, the damping coefficient α_{30} seems to be completely wrong. However, a closer look reveals that the coefficient α_1 supports the damping effect. If one takes into account that the average height $\bar{h}$ in the simulations was not absolutely zero, but was found as $\bar{h} \approx 0.025\,07$ an effective coefficient given by the sum $\alpha_1/\bar{h} + \alpha_{30} \approx -99.87$ provides almost the right value for the damping. This kind of mismatch in the constant damping term α_1 and the damping of the average height $\bar{h}$ via α_{30} is found as a systematic bias in the algorithm. It is recommended that these two shares are always considered together, because these two terms cannot be separated cleanly in the reconstruction. On the other hand, all other terms, despite of their complexity and nonlinearity can be reproduced with quite high accuracy.

[1] R. M. Bradley and J. M. E. Harper, J. Vac. Sci. Technol., A **6**, 2390 (1988).
[2] G. I. Sivashinsky, Acta Astronaut. **4**, 1177 (1977).
[3] Y. Kuramoto, T. Tsuzuki, Prog. Theor. Phys. **54**, 687 (1975).
[4] M. A. Makeev *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. B **197**, 185 (2002).
[5] R. Cuerno and A. L. Barabási, Phys. Rev. Lett. **74**, 4746 (1995).
[6] A. Keller, S. Facsko, and R. Cuerno, Numerical Integrator for Continuum Equations of Surface Growth and Erosion, in *Computational Nanotechnology*, edited by S. M. Musa (CRC Press, Boca Raton, 2012).
[7] V. O. Kharchenko and D. O. Kharchenko, Condens. Matter Phys. **14**, 23602 (2011).
[8] D. O. Kharchenko, V. O. Kharchenko, I. O. Lysenko, and S. V. Kokhan, Phys. Rev. E **82**, 061108 (2010).
[9] K. B. Lauritsen, R. Cuerno, and H. A. Makse, Phys. Rev. E **54**, 3577 (1996).
[10] J. Muñoz-García, R. Cuerno, and M. Castro, J. Phys.: Condens. Matter **21**, 224020 (2009).
[11] J. Muñoz-García, R. Gago, L. Vázquez, J. A. Sánchez-García, and R. Cuerno, Phys. Rev. Lett. **104**, 026101 (2010).
[12] J. Muñoz-García *et al.*, Mater. Sci. Eng., R **86**, 1 (2014).
[13] M. Kardar, G. Parisi, and Y.-C. Zhang, Phys. Rev. Lett. **56**, 889 (1986).
[14] S. Park, B. Kahng, H. Jeong, and A. L. Barabasi, Phys. Rev. Lett. **83**, 3486 (1999).
[15] D. Reiser, in *Proceedings 45st EPS Conference on Plasma Physics, Prague, Czech Republic*, 2–6 July (2018), Europhys. Conf. Abstracts Vol. 42A; http://ocs.ciemat.es/EPS2018PAP/pdf/P4.1024.pdf
[16] B. M. Boghosian, C. C. Chow, and T. Hwa, Phys. Rev. Lett. **83**, 5262 (1999).
[17] A. Giacometti and M. Rossi, Phys. Rev. E **62**, 1716 (2000).
[18] S. Saito and M. Uwaha, J. Phys. Soc. Jpn. **65**, 3576 (1996).

[19] A. A. Golovin, A. A. Nepomnyashchy, S. H. Davis, and M. A. Zaks, Phys. Rev. Lett. **86**, 1550 (2001).
[20] S. Facsko, T. Bobek, A. Stahl, H. Kurz, and T. Dekorsy, Phys. Rev. B **69**, 153412 (2004).
[21] C. Madi, E. Anzenberg, K. F. Ludwig, and M. J. Aziz, Phys. Rev. Lett. **106**, 066101 (2011).
[22] D. Reiser *et al.*, Phys. Scr. **2017**, 014039 (2017).
[23] H. Schaeffer, Proc. R. Soc. London, Ser. A **473**, 20160446 (2017).
[24] M. Raissi and G. M. Karniadakis, J. Comput. Phys. **357**, 125 (2018).
[25] S. H. Rudy, S. L. Brunton, J. L. Proctor, and J. N. Kutz, Sci. Adv. **3** (2017).
[26] S. Roberts *et al.*, Philos. Trans. R. Soc., A **371**, 20110550 (2018).
[27] S. L. Brunton, J. L. Proctor, and J. N. Kutz, Proc. Natl. Acad. Sci. USA **113**, 3932 (2016).
[28] E. Schultz, M. Speekenbrink, and A. Krause, J. Math. Psychol. **85**, 1 (2018).
[29] G. H. Golub and C. Reinsch, Numer. Math. **14**, 403 (1970).
[30] J. Mandel, Am. Stat. **36**, 15 (1982).
[31] A. Ralston, Math. Comput. **16**, 431 (1962).
[32] A. Careta *et al.*, J. Stat. Phys. **71**, 235 (1993).
[33] A. Careta *et al.*, Phys. Fluids **6**, 349 (1994).
[34] A. C. Marti *et al.*, Phys. Fluids **9**, 1078 (1997).
[35] R. Mannella and V. Palleschi, Phys. Rev. A **40**, 3381 (1989).
[36] J. D. Bao *et al.*, J. Comput. Phys. **197**, 241 (2004).
[37] F. Caballero, C. Nardini, F. van Wijland, M. E. Cates, Phys. Rev. Lett. **121**, 020601 (2018).
[38] T. J. Kim *et al.*, Phys. Rev. Lett. **92**, 246104 (2004).
[39] T. Sun, H. Guo, and M. Grant, Phys. Rev. A **40**, 6763 (1989).
[40] J. Muñoz-García, R. Cuerno, and M. Castro, Phys. Rev. B **78**, 205408 (2008).