

# PURE and APPLIED ANALYSIS

# PAM

YUAN GAO, ANYA E. KATSEVICH, JIAN-GUO LIU,  
JIANFENG LU AND JEREMY L. MARZUOLA

**ANALYSIS OF A FOURTH-ORDER EXPONENTIAL PDE  
ARISING FROM A CRYSTAL SURFACE JUMP PROCESS  
WITH METROPOLIS-TYPE TRANSITION RATES**



# ANALYSIS OF A FOURTH-ORDER EXPONENTIAL PDE ARISING FROM A CRYSTAL SURFACE JUMP PROCESS WITH METROPOLIS-TYPE TRANSITION RATES

YUAN GAO, ANYA E. KATSEVICH, JIAN-GUO LIU, JIANFENG LU AND JEREMY L. MARZUOLA

We analytically and numerically study a fourth-order PDE modeling rough crystal surface diffusion on the macroscopic level. We discuss existence of solutions globally in time and long-time dynamics for the PDE model. The PDE, originally derived by Katsevich is the continuum limit of a microscopic model of the surface dynamics, given by a Markov jump process with Metropolis-type transition rates. We outline the convergence argument, which depends on a simplifying assumption on the local equilibrium measure that is valid in the high-temperature regime. We provide numerical evidence for the convergence of the microscopic model to the PDE in this regime.

## 1. Introduction

We analyze a fourth-order exponential PDE which models the macroscopic dynamics of crystal surface relaxation. The PDE arises as the time and space scaling limit of a microscopic Markov jump process, which evolves via discrete surface hopping events. These events occur at specified transition rates. The transition rates, along with the form of the interaction potential, determine the microscopic dynamics and therefore also shape the macroscopic PDE. (The choice of interaction potential affects the Hamiltonian, through which the equilibrium probability distribution is defined). Here we focus on a quadratic interaction potential and transition rates of Metropolis type, meaning that they are a function only of the difference in energy between the pre- and postjump crystal states. Below we will restrict ourselves to one spatial dimension with periodic boundary conditions for simplicity of exposition, but higher-dimensional generalizations and other types of boundary condition are indeed possible and treatable using the presented techniques.

In the continuum limit, the crystal surface is represented as a height profile  $h(t, x)$ ,  $t \geq 0$ ,  $x \in \mathbb{T}$ , where  $\mathbb{T}$  is the unit interval with periodic boundary conditions. The resulting PDE limit is of the form

$$\partial_t h(t, x) = \partial_x (\exp[-\beta \partial_x^3 h(t, x)] - \exp[\beta \partial_x^3 h(t, x)]), \quad h(0, x) = h_0(x), \quad (1)$$

where  $\beta$  denotes inverse temperature. Conditions on the initial data required for the existence theory will be discussed below. In this paper, we develop analytic tools for the PDE (1). In particular, we prove the existence of global solutions to (1) in Section 3. We also explore properties of (1) numerically in Section 4.

This PDE has several key features that distinguish it from analogous exponential PDEs formally derived as the macroscopic limit of Arrhenius rate dynamics in [Marzuola and Weare 2013; Krug et al. 1994],

MSC2010: 35Q70, 82D25.

Keywords: crystal surface relaxation, gradient flows, metropolis rates.

and recently revisited in [Gao et al. 2020]. The Arrhenius rates are another set of rates for microscopic dynamics which are used in the solid-on-solid (SOS) model, a well-known (and well-studied) model for crystal surface relaxation [Binh 1983]. The Arrhenius rate PDE in [Krug et al. 1994] assumes an absolute value potential, while the PDE in [Marzuola and Weare 2013] is arrived at using a more general interaction potential, including in particular the quadratic one. For the quadratic interaction potential, it is given by an equation of the form (taking  $\beta = 1$  for simplicity)

$$\partial_t h = \partial_{xx} \exp(-\partial_{xx} h), \quad h(0, x) = h_0(x). \quad (2)$$

This PDE can be interpreted as a weighted  $H^{-1}$  gradient flow in a similar fashion to the framework laid out in the related works [Liu et al. 2019; Craig et al. 2020]. Recent analytic progress has been made relating to existence of weak solutions, characterization of dynamics, construction of strong solutions, and classification of the breakdown of regularity for this equation. See for instance the works [Liu and Xu 2016; 2017; Liu et al. 2019; Gao et al. 2019; Xu 2018].

One symmetry property of (1) not present in (2) demonstrates an important intrinsic distinction between the microscopic Metropolis and Arrhenius rate dynamics. Namely, if  $h$  is a solution to (1) with initial profile  $h_0$ , then  $-h$  is another solution with corresponding initial data  $-h_0$ . This is not the case for (2). In fact, solutions to (2) form singularities in convex regions but not in concave regions [Marzuola and Weare 2013]. Using the structure of the Metropolis and Arrhenius rates, one can show this same symmetry property (or lack thereof) holds on a microscopic level.

We note that the exponential PDEs (1) and (2) arise by applying a nonstandard but potentially more informative scaling regime to the microscopic dynamics (i.e., relative scaling of time, space, and height in the large crystal limit). In a more standard scaling regime, the exponentials become linearized.

In [Marzuola and Weare 2013], the limiting macroscopic PDEs in both scaling regimes are formally derived using a probabilistic framework. These PDEs can also be formally derived using physical arguments. In [Krug et al. 1994], for example, the authors use physical principles to derive a PDE for the Arrhenius rate broken bond model with an absolute value potential. The resulting PDE does not have exponential dependence on the derivatives of  $h$ , and corresponds to the standard scaling regime. A separate argument in the last section of that work suggests an exponential PDE similar to (2) as an alternative. In [Krug et al. 1994], as well as in [Gao et al. 2020], the exponential PDE is obtained by applying the so-called Gibbs–Thomson relation as an underlying approximation for how the density varies with respect to the chemical potential; see [Liu et al. 2019, Section 1.1] for details on this approach.

The formal derivation of (1) uses the same probabilistic framework as [Marzuola and Weare 2013]. The argument is exactly the same, except of course that one should replace the Arrhenius rates with the Metropolis rates. For this reason, and because the analysis of exponential PDE (1) is the primary focus of this paper, we only give a high-level overview of the argument. The argument relies on the assumption that in local equilibrium the process's distribution is given by a local Gibbs measure. This assumption is correct for the Arrhenius rates but is not entirely accurate for the Metropolis rates due to their jump asymmetry. As a consequence, the PDE (1) is not necessarily accurate in general; see the end of Section 2A for further discussion. Nevertheless, the equation is an excellent fit for the microscopic dynamics when  $\beta$  is small.

We add that the equation retains its nonlinearity for small  $\beta$  (i.e., one cannot linearize the exponentials), as evidenced by the evolution of numerical solutions to (1) shown in Section 2A. We note that a fully rigorous derivation of the convergence of the microscopic models to the macroscopic PDEs does not exist for any of the models, neither the Arrhenius cases considered in [Marzuola and Weare 2013; Krug et al. 1994; Liu et al. 2019; Gao et al. 2020], nor the Metropolis settings considered here or a related version in [Krug et al. 1994].

The paper will proceed as follows. In Section 2, we describe the microscopic process and scaling regime leading to the PDE (1). We then compare the evolution of the PDE solution to that of the microscopic process. The section concludes with a discussion of the slight discrepancy between the PDE dynamics and the true large crystal microscopic dynamics. In Section 3, we prove the global existence and long-time behavior of the solutions to PDE (2). In Section 4, we explore properties of the PDE numerically. In Appendix A we compute the Metropolis rate expectation with respect to the local Gibbs measure. This is the main computation needed to formally derive the PDE from the microscopic dynamics. In Appendix B, we give an alternative approach to the analysis of long-time dynamics for (1).

## 2. Motivation: PDE as scaling limit

The PDE (1) arises as the continuum limit of a discrete microscopic Markov jump process modeling the relaxation of a crystal surface. We briefly describe the state space and dynamics of this microscopic process. The process is represented by a height profile

$$h_N(t) = (h_N^0(t), \dots, h_N^{N-1}(t)) \in \mathbb{Z}^N, \quad t \geq 0.$$

For fixed  $t$ , we may think of  $h_N(t)$  as a step function in space on the torus  $\mathbb{T} = [0, 1)$  (with endpoints identified) which takes the value  $h_N^j(t)$  on  $[j/N, (j+1)/N)$ . We call the interval  $[j/N, (j+1)/N)$  “site  $j$ ”. The integer values  $h_N^j$  represent the number of particles at site  $j$  stacked in a column, above or below a fixed level surface representing zero height.

The process evolves by means of particles jumping between neighboring sites. Each such jump occurs instantaneously at a certain transition or jump rate. The jumps and their corresponding jump rates fully determine the dynamics and must be specified in advance.

We denote the event in which a particle jumps from site  $i$  to site  $j$  (with  $|i - j| = 1$ ) by  $h \mapsto J_i^j h$ . That is, if the profile was given by  $h_N(t) = h$  before the jump, then after the jump it is given by  $J_i^j h$ , where

$$\begin{aligned} (J_i^j h)^i &= h^i - 1, & (J_i^j h)^j &= h^j + 1, \\ (J_i^j h)^k &= h^k, & k &\neq i, j. \end{aligned} \tag{3}$$

The rate  $r^{i,j}(h)$  at which this transition occurs is defined through

$$\mathbb{P}(h_N(t + \delta) = J_i^j h \mid h_N(t) = h) = r^{i,j}(h)\delta + o(\delta).$$

In particular, we assume  $r^{i,j}(h)$  is independent of time. The jump rates determine the expected instantaneous change in height induced by a jump. Thus while a jump from site  $i$  to  $j$  increases  $h_N^j$  by 1, the expected increase in  $h_N^j$  is given by  $r^{i,j}(h_N(t)) dt$ .

Let us write down the evolution of  $h_N^i$  at a given moment in time. It will decrease if a jump occurs from  $i$  to  $i + 1$  or from  $i$  to  $i - 1$  and increase if a jump occurs from  $i + 1$  or  $i - 1$  to  $i$ . We therefore obtain

$$dh_N^i(t) = [(r^{i-1,i} - r^{i,i-1})(h_N(t)) - (r^{i,i+1} - r^{i+1,i})(h_N(t))] dt + d\xi_N^i(t),$$

where  $\xi_N^i$  represents random fluctuations.

If we define  $\mathcal{J}_N^i = r^{i,i+1} - r^{i+1,i}$  as the “current” from  $i$  to  $i + 1$ , then we can write the above as a microscopic conservation law:

$$dh_N^i(t) = -(\mathcal{J}_N^i - \mathcal{J}_N^{i-1})(h_N(t)) dt + d\xi_N^i(t).$$

By scaling time and height with  $N$ , the random fluctuations vanish and we obtain the deterministic limit

$$h(t, x) = \lim_{N \rightarrow \infty} N^{-3} h_N^{Nx}(N^4 t),$$

with

$$\partial_t h = -\partial_x \mathcal{J}^x(h). \quad (4)$$

Here  $\mathcal{J}^x(h) = \lim_{N \rightarrow \infty} \langle \mathcal{J}_N^{Nx}(h_N) \rangle$ , where  $\langle \cdot \rangle$  denotes expectation with respect to the local Gibbs measure (see [Appendix A](#) for a definition of the measure).

We note that the above argument is only heuristic. In particular, we do not necessarily expect pointwise convergence of the  $h_N^i$  to  $h(t, x)$ . One can only expect the convergence of averages of  $h_N^i$  in mesoscopic-sized intervals to the macroscopic profile.

Note that (1) is of the form (4), with

$$\mathcal{J}^x(h) = e^{-\frac{3}{2}\beta} \sinh(\beta h_{xxx}(t, x)).$$

It is the scaling limit of the microscopic process under a certain choice of transition rates, which we now describe. Namely, we set

$$r^{i,j}(h) = \exp\left(-\frac{\beta}{2}[H(J_i^j h) - H(h)]\right), \quad |i - j| = 1, \quad (5)$$

where  $\beta$  is an inverse temperature and  $H(h)$  is the Hamiltonian, i.e., surface energy, of a configuration  $h$ , defined by

$$H(h) = \sum_{i=0}^{N-1} (h^{i+1} - h^i)^2.$$

Note that the energy is independent of how the zero height level surface is chosen; i.e., a uniform shift  $h^i \mapsto h^i + c$ ,  $i = 0, \dots, N - 1$ , does not affect the energy.

These rates give preference to atomistic motion that lowers the surface energy  $H$ . Importantly, the dynamics is not symmetric with respect to jumping left or right, instead favoring the direction yielding lower energy.

The rates are in detailed balance with respect to the Gibbs distribution

$$p_\infty(h) \propto e^{-\beta H(h)},$$

so that the Gibbs distribution is the global equilibrium ( $t \rightarrow \infty$ ) distribution of the process. These rates belong to a class of rates of what we call “Metropolis type”, in that they are in detailed balance with

the Gibbs measure and depend only on the energy difference  $\Delta H$  between the pre- and postjump states. Indeed, any rates of the form  $\phi(\Delta H)$  which satisfy

$$\phi(-\Delta H) = \phi(\Delta H)e^{\beta\Delta H}$$

are in detailed balance with the Gibbs measure. The more well-known set of rates of Metropolis type are  $\phi(\Delta H) = e^{-\beta\Delta H} \wedge 1$ . We note that while this is also the acceptance probability in standard Metropolis–Hastings algorithms, the goal of that algorithm is to sample the invariant Gibbs measure, whereas we are interested in the Markov process dynamics itself. We choose rates of the form  $e^{-\beta\Delta H/2}$  due to their analytic tractability, since it is more straightforward to compute the rate expectations  $\langle r^{i,j}(h) \rangle$  for these rates than for the Metropolis rates involving a minimum.

See [Appendix A](#) for the computation of  $\langle r^{i,i+1} - r^{i+1,i} \rangle$ , from which the PDE follows.

**2A. Scaling limit: numerics.** In [Figures 1](#) and [2](#), we numerically compare the microscopic evolution with that of the solution to the PDE [\(1\)](#).

To solve the PDE, we discretized spatially using centered difference schemes and applied a numerical ODE timestepper designed for stiff ODEs. To simulate the microscopic process, we used the kinetic Monte Carlo (KMC) method. In this method, one iteratively updates the state of the process,  $\mathbf{h} \mapsto \mathbf{h}'$ , and the physical time,  $t \mapsto t + \Delta$ , until the desired final time is reached. The new state  $\mathbf{h}'$  is randomly chosen with probability proportional to the transition rate  $r(\mathbf{h} \mapsto \mathbf{h}')$ , and  $\Delta$ , which represents the amount of time the process spent in state  $\mathbf{h}$ , is chosen from an exponential distribution  $\exp(\lambda)$ , with  $\lambda = \sum_{\mathbf{h}'} r(\mathbf{h} \mapsto \mathbf{h}')$ .

To compare the macroscopic and rescaled microscopic processes, we fix an initial nontrivial (out of equilibrium) macroscopic profile  $h(x, 0) = h_0(x)$  for which  $\|h\|_\infty > 0$  in order to ensure nontrivial dynamics in both the microscopic and macroscopic flows. We evolve the PDE forward from  $h_0$  to some macroscopic times of interest  $T$ . Then, for various  $N$ , we run KMC from the initial microscopic profile  $\mathbf{h}_N(0) = (N^3 h_0(j/N))_{j=0}^{N-1}$  up to the microscopic time  $N^4 T$ . We should then expect to see that  $h(j/N, T) \approx N^{-3} h_N^j(N^4 T)$  for large  $N$ .

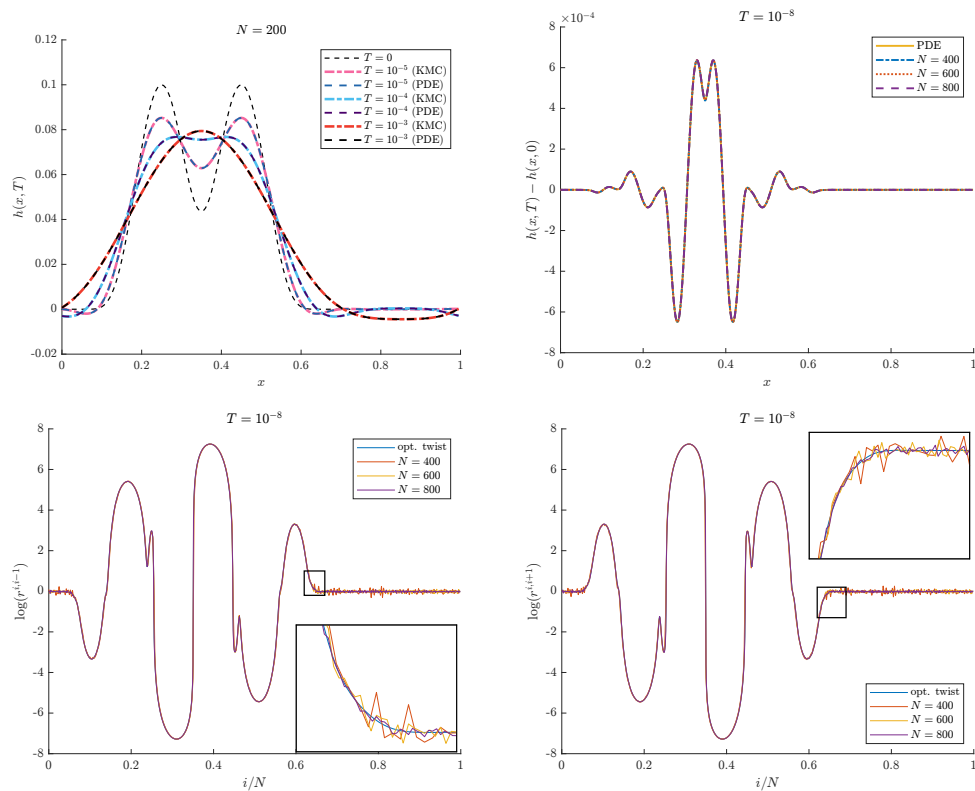
In addition to comparing the evolution of the KMC and PDE height profiles, we also check whether the Metropolis rate expectation computed using the local Gibbs measure converges to the true rate expectation. If this is the case, then one should have

$$\langle r^{i,i+1}(h_N(t)) \rangle \approx \frac{1}{2N^4\delta} \int_{N^4(T-\delta)}^{N^4(T+\delta)} r^{i,i+1}(h_N(s)) \, ds, \quad (6)$$

where the left-hand side is the rate expectation with respect to the local Gibbs measure; see [Appendix A](#) for this expression. The expectation depends on the macroscopic profile  $h(t, x)$ , which we estimate using the PDE solution. The right-hand side is the integral of a step function, since the microscopic process is a Markov jump process. It can therefore be simply computed from the KMC simulation by keeping track of the rate values and time between jumps.

We consider two initial profiles and values of  $\beta$ : in [Figure 1](#), we take  $h_0(x) = g(x) + g(x + 0.2 \bmod 1)$  with  $\beta = 0.01$ , where

$$g(x) = \frac{1}{10} \exp\left(8 - \frac{1}{x} - \frac{1}{\frac{1}{2} - x}\right) \mathbb{1}_{(0, \frac{1}{2})}(x)$$

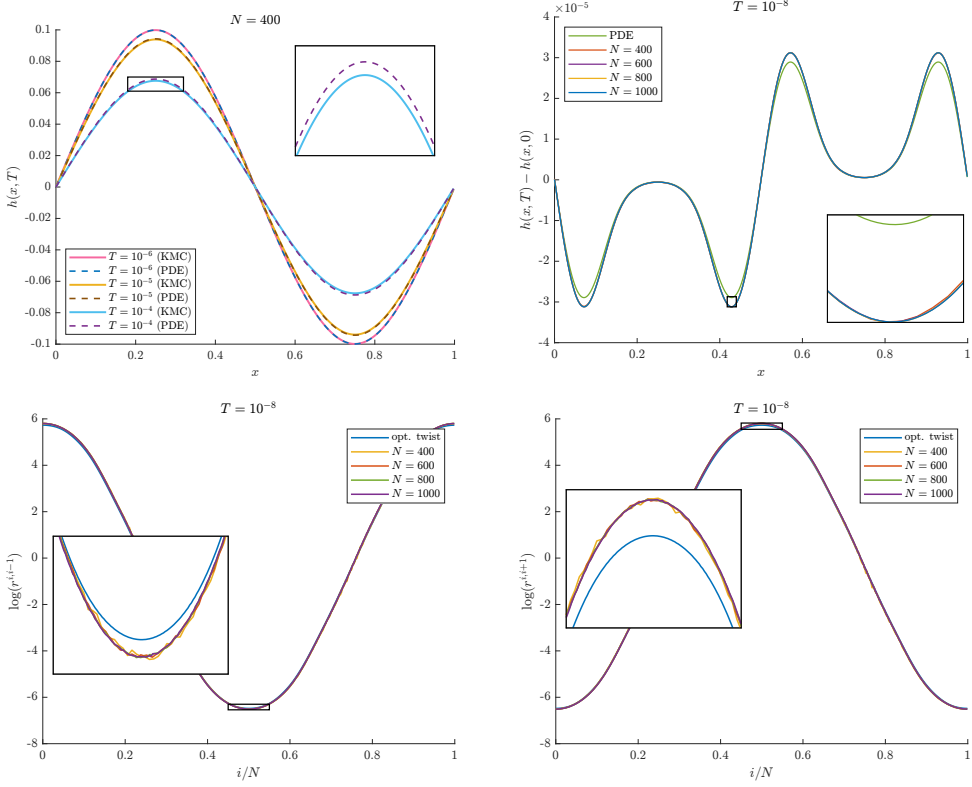


**Figure 1.** Metropolis rate dynamics with temperature  $\beta = 0.01$ , from initial profile  $h(x, 0)$  given by a sum of a bump function  $g(x)$  and its translation,  $g(x + 0.2 \bmod 1)$ . Top left: PDE vs. KMC evolution with fixed  $N$  at various times  $T$ . Top right: PDE vs. KMC short time evolution,  $h(x, T) - h(x, 0)$ , for fixed  $T$  and various  $N$ . Bottom left (bottom right): log of local Gibbs average vs. log of empirical average of rates to jump left (jump right) for fixed  $T$  and various  $N$ .

is a smooth bump function supported on  $(0, \frac{1}{2})$ . In Figure 2,  $h_0(x) = \frac{1}{10} \sin(2\pi x)$ , with  $\beta = 0.25$ . The reason for choosing small amplitude for the initial profile and small  $\beta$  is to limit how large the rates can be, since they depend exponentially on  $\beta$  and the curvature of  $h_N$ .

Figure 1 shows results from the experiment with a compactly supported two-bump initial profile, and  $\beta = 0.01$ . We observe an excellent fit of the PDE solution to the microscopic profile obtained from KMC. The top left figure shows that on the scale of the initial height amplitude, the PDE dynamics exactly coincides with the microscopic dynamics for  $N = 200$ . The top right figure shows that the PDE also fits the microscopic dynamics on the scale of small shifts in amplitude. Moreover, the microscopic dynamics (after rescaling) has already converged for  $N = 400$ , since increasing  $N$  does not affect the dynamics.

The bottom left (bottom right) figure compares the time average of the right-jump rates (left-jump rates) with its expectation with respect to the local Gibbs measure. The zoomed-in part of the plot shows that the KMC rate time average oscillates more closely around the local Gibbs rate expectation as  $N$  increases.



**Figure 2.** Metropolis rate dynamics with temperature  $\beta = 0.25$ , from initial profile  $h(x, 0) = 0.1 \sin(2\pi x)$ . Top left: PDE vs. KMC evolution with fixed  $N$  at various times  $T$ . Top right: PDE vs. KMC short time evolution,  $h(x, T) - h(x, 0)$  for fixed  $T$  and various  $N$ . Bottom left (bottom right): optimal twist average vs. empirical average of rates to jump left (jump right) for fixed  $T$  and various  $N$ .

Figure 2 shows results from the experiment with a sinusoidal initial profile and  $\beta = 0.25$ . Qualitatively, the PDE fully captures the microscopic dynamics. However, compared to the two-bump initial profile, we observed a slightly poorer fit between the PDE and microscopic dynamics. The top right and bottom figures, in which results from different values of  $N$  are presented, show that the microscopic process has nearly converged. Thus the discrepancy between the PDE and microscopic profiles is not attributable to  $N$  being insufficiently large. Instead, it must be related to the assumption of a local Gibbs measure. Indeed, the bottom plots show that the expected value of the rate observable is not given exactly by the expectation with respect to the local Gibbs measure, though of course it is clearly a very close approximation to the actual measure.

**Discussion of Discrepancy in Figure 2.** The PDE (1) is derived using the assumption that once the microscopic process equilibrates locally, its probability distribution is well-approximated by a local Gibbs measure. In actuality, this measure requires a correction which involves the sum of the local currents,  $\sum_{i=0}^{N-1} \mathcal{J}^i$ . This sum (and the correction) vanishes for jump rates which satisfy the gradient condition, under which the current can be written as a discrete spatial derivative of another observable. As an



example, symmetric jump rates (i.e., those for which jumping left from site  $i$  has the same rate as jumping right) satisfy this condition. Indeed, we have  $\mathcal{J}^i = r^i - r^{i+1}$ , where  $r^i := r^{i,i+1} = r^{i,i-1}$ .

The correction to the measure affects the expectation of the current. It should manifest in the expectation as a nonconstant multiplicative factor, called a mobility by physicists. For a discussion of this, and for the form of the multiplicative factor, see, e.g., [Krug et al. 1994; Spohn 1991]. As we see in the numerical simulations, the correction has a negligible (albeit nonzero) effect on the dynamics. We expect the PDE (1) to be an accurate approximation to the microscopic dynamics when  $\beta$  is small. Indeed, for  $\beta$  small the jump rates are close to constant, so that  $\sum_{i=0}^{N-1} \mathcal{J}^i$  is small.

For higher  $\beta$  (lower temperature), the PDE requires a nontrivial correction. The form of this correction is investigated in a forthcoming paper.

3. Global solution and long-time dynamics

In this section, we interpret the fourth-order exponential PDE for Metropolis rate dynamics as a gradient flow of a proper convex functional with a  $L^2$  dissipation. Then using the minimizing movement and the convergence analysis in [Ambrosio et al. 2008] (henceforth abbreviated [AGS]), we construct a global strong solution and prove there is no singularity formation. In Appendix B, we present ideas for another approach using the bivariational structures.

After setting most physical constants to be 1, we obtain the continuous equation for surface growth with Metropolis-type rates

$$\partial_t u = \frac{1}{2} \partial_x [e^{-\partial_x^3 u} - e^{\partial_x^3 u}]. \tag{7}$$

Set  $h := \partial_x u$ . We obtain formally the equation for  $h$

$$\partial_t h = \frac{1}{2} \partial_{xx} [e^{-\partial_{xx} h} - e^{\partial_{xx} h}] = -\partial_{xx} (\sinh(\partial_{xx} h)). \tag{8}$$

We will validate this equation by proving the global existence and the long-time behavior of solutions to (8) with periodic boundary condition; see Theorem 3.6.

**3A. Gradient flow in  $L^2(\mathbb{T})$ .** Let us first define formally a convex functional with some formal observations and recast (8) into a  $L^2(\mathbb{T})$  gradient flow. Let  $\phi$  be

$$\phi(h) := \int_{\mathbb{T}} \cosh(h_{xx}) \, dx. \tag{9}$$

The first variation of  $\phi$  is

$$\frac{\delta \phi}{\delta h} = \partial_{xx} (\sinh(h_{xx})),$$

and then formally we have

$$h_t = -\frac{\delta \phi}{\delta h}. \tag{10}$$

To study the global strong solution to (8), we plan to apply the gradient flow theory in the metric space  $L^2(\mathbb{T})$ . Let us first make some inspiring observations, which will be made rigorous in the proof later.

**Observation 1** (conservation laws). Thanks to the periodic assumption, we have

$$\frac{d}{dt} \int_{\mathbb{T}} h \, dx = 0, \quad (11)$$

which implies  $\int_{\mathbb{T}} h \, dx = \int_{\mathbb{T}} h_0 \, dx$ . Moreover from  $\int_{\mathbb{T}} h_{xx} \, dx = 0$ , we know

$$\int_{\mathbb{T}} (h_{xx})^+ \, dx = \int_{\mathbb{T}} (h_{xx})^- \, dx. \quad (12)$$

Here  $(h_{xx})^-$  is the negative part of  $h_{xx}$  and  $(h_{xx})^+$  is the positive part of  $h_{xx}$ .

**Observation 2** (dissipation inequalities). From the gradient flow structure (10),

$$\frac{d\phi}{dt} = \int_{\mathbb{T}} \frac{\delta\phi}{\delta h} h_t \, dx = - \int_{\mathbb{T}} \left| \frac{\delta\phi}{\delta h} \right|^2 \, dx = - \int_{\mathbb{T}} h_t^2 \, dx \leq 0, \quad (13)$$

which gives the observation

$$\phi(h(t)) \leq \phi(h(0)) \quad \text{for any } t \geq 0.$$

One shall notice the boundedness of the functional  $\phi$  gives us good estimates to prevent both the positive and negative parts of  $h_{xx}$  from becoming singular. Indeed, we have the uniform estimate

$$\begin{aligned} \frac{1}{p} \int_{\mathbb{T}} (h_{xx}^+)^p \, dx &\leq \int_{\mathbb{T} \cap (h_{xx}^+ > 0)} e^{(h_{xx})^+} \, dx \leq \int_{\mathbb{T} \cap (h_{xx}^+ > 0)} (e^{(h_{xx})^+} + e^{-(h_{xx})^+}) \, dx \\ &\leq \int_{\mathbb{T}} (e^{h_{xx}} + e^{-h_{xx}}) \, dx = 2\phi(h(t)) \leq 2\phi(h(0)). \end{aligned}$$

Similarly, we have the same estimate for the negative part  $h_{xx}^-$ . Thus for any  $p \in \mathbb{N}^+$ ,

$$\frac{1}{p} \int_{\mathbb{T}} |h_{xx}|^p \, dx \leq 4\phi(h(0)). \quad (14)$$

For simplicity we choose the working space

$$H := \{u \in L^2(\mathbb{T}) : \int_{\mathbb{T}} u \, dx = 0\}, \quad V := \{u \in H^2(\mathbb{T}) : \int_{\mathbb{T}} u \, dx = 0\} \quad (15)$$

with standard  $L^2$ -norm, denoted by  $\|\cdot\|$ , and  $H^2$ -norm, denoted by  $\|\cdot\|_V$ . Denote the best constant for Poincaré's inequality by  $\kappa > 0$ , which depends only on the size of the domain  $\mathbb{T}$ . Specifically, if the period of the domain is  $L$ , then  $\kappa = (L/(2\pi))^2$ .

**3B. Variational inequality solution.** Let  $\phi$  be a functional:

$$\phi : H \rightarrow [0, +\infty], \quad \phi(h) := \begin{cases} \int_{\mathbb{T}} \cosh(h_{xx}) \, dx & \text{if } h \in V, \\ +\infty & \text{otherwise.} \end{cases} \quad (16)$$

**3B1. Euler scheme.** First let us establish the gradient flow evolution in the metric space  $(H, \text{dist})$ , with distance  $\text{dist}(u, v) := \|u - v\|$ . Let  $h_0(x) \in H$  be a given initial datum and  $0 < \tau \ll 1$  be a given parameter. We consider a sequence  $\{x_n^\tau\}$  which satisfies the unconditional-stable backward Euler scheme

$$\begin{cases} x_n^{(\tau)} \in \operatorname{argmin}_{x' \in H} \left\{ \phi(x') + \frac{1}{2\tau} \|x' - x_{n-1}^{(\tau)}\|^2 \right\}, & n \geq 1, \\ x_0^{(\tau)} := h_0 \in H. \end{cases} \quad (17)$$

The existence and uniqueness of the sequence  $\{x_n^\tau\}$  can be proved by direct methods in the calculus of variations after establishing the convexity and lower semicontinuity of  $\phi$  in [Proposition 3.2](#). Thus we consider the gradient descent with respect to  $\phi$  in the space  $(H, \text{dist})$ .

Now for any  $0 < \tau \ll 1$  we define the resolvent operator, also known as proximal mapping of  $\phi$  (see [\[AGS, p. 40\]](#)),

$$\mathcal{J}_\tau[h] := \operatorname{argmin}_{v \in H} \left\{ \phi(v) + \frac{1}{2\tau} \|v - h\|^2 \right\};$$

then the variational approximation of  $h$  at  $t$  is obtained by Euler scheme [\(17\)](#) as

$$h_n(t) := (\mathcal{J}_{t/n})^n[h_0]. \quad (18)$$

In [Proposition 3.5](#), we will use the theory for gradient flow in metric space [\[AGS, Theorem 4.0.4\]](#) to establish the convergence of the variational approximation  $h_n(t)$  to variational inequality solution to [\(8\)](#), which is defined below.

**Definition 3.1.** Given initial data  $h_0 \in H$ , we call  $h : [0, +\infty) \rightarrow H$  a variational inequality solution to [\(8\)](#) if  $h(t)$  is a locally absolutely continuous curve such that  $\lim_{t \rightarrow 0} h(t) = h_0$  in  $H$  and

$$\langle h_t(t), h(t) - v \rangle_{H', H} \leq \phi(v) - \phi(h(t)) \quad \text{for a.e. } t > 0, \text{ for all } v \in D(\phi). \quad (19)$$

Next we study some properties, including convexity and lower semicontinuity in  $H$ , of the functional  $\phi$ .

**3C. Convexity and lower semicontinuity of function  $\phi$  in  $H$ .** We will prove the  $\lambda$ -convexity and lower semicontinuity of function  $\phi$  in  $H$ . We note  $\lambda > 0$  is important for the long-time behavior of the global solution.

**Proposition 3.2.** *The functional  $\phi : H \rightarrow [0, +\infty]$  is proper,  $\lambda$ -convex, lower semicontinuous in  $H$  and satisfies coercivity defined in [\[AGS, \(2.4.10\)\]](#).*

*Proof.* Clearly since the typical function satisfies  $h = 0 \in D(\phi)$ , we have  $D(\phi) = \{\phi < +\infty\}$  is nonempty and  $\phi$  is proper. Due to the positivity of  $\phi$ , coercivity [\[AGS, \(2.4.10\)\]](#), i.e., there exists  $u_* \in D(\phi)$ ,  $r_* > 0$  such that  $\inf\{\phi(v) : v \in H, \text{dist}(v, u_*) \leq r_*\} > -\infty$ , is obvious.

*$\lambda$ -convexity.* Given  $u, v \in H$ ,  $t \in (0, 1)$ , without loss of generality we assume  $u, v \in D(\phi)$ ; otherwise the convexity inequality is trivial. Therefore from the definition of  $\lambda$ -convexity, we only need to prove for any  $t \in [0, 1]$  and any  $u, v \in H$  we have

$$\phi((1-t)u + tv) \leq (1-t)\phi(u) + t\phi(v) - \frac{1}{2}\lambda t(1-t)\|u - v\|_{L^2}^2. \quad (20)$$

Define

$$I(t) := \int_{\mathbb{T}} (1-t) \cosh(u_{xx}) + t \cosh(v_{xx}) - \frac{\lambda}{2} t(1-t) \|u - v\|_{L^2}^2 - \cosh((1-t)u_{xx} + tv_{xx}) \, dx$$

and notice  $I(1) = I(0) = 0$ . Thus we only need to prove  $I''(t) \leq 0$ . It is easy to calculate that

$$I''(t) = \int_{\mathbb{T}} -\cosh[(1-t)u_{xx} + tv_{xx}](v_{xx} - u_{xx})^2 + \lambda(u - v)^2 \, dx \leq \int_{\mathbb{T}} -(u_{xx} - v_{xx})^2 + \lambda(u - v)^2 \, dx \leq 0,$$

due to Poincaré's inequality with  $\lambda = 1/\kappa^2$ . Hence  $\phi$  is  $\lambda$ -convex for  $\lambda = 1/\kappa^2 = (2\pi/L)^4 > 0$ .

*Lower semicontinuity.* Consider a sequence  $h_n \rightarrow h$  in  $H$ . We need to check

$$\phi(h) \leq \liminf_n \phi(h_n).$$

If  $h_n \in D(\phi)$  does not hold for all large  $n$ , then lower semicontinuity holds. Without loss of generality, we can assume  $h_n \in D(\phi)$  for all  $n$ , and also

$$\liminf_n \phi(h_n) = \lim_n \phi(h_n).$$

First notice  $h_n \in D(\phi)$  for any  $n$  implies the uniform estimate

$$\begin{aligned} \frac{1}{p} \int_{\mathbb{T}} |(\partial_{xx} h_n)^+|^p dx &\leq \int_{\mathbb{T} \cap (\partial_{xx} h_n^+ > 0)} e^{(\partial_{xx} h_n)^+} dx \leq \int_{\mathbb{T} \cap (\partial_{xx} h_n^+ > 0)} (e^{(\partial_{xx} h_n)^+} + e^{-(\partial_{xx} h_n)^+}) dx \\ &\leq \int_{\mathbb{T}} (e^{\partial_{xx} h_n} + e^{-\partial_{xx} h_n}) dx = 2\phi(h_n) < \infty. \end{aligned}$$

Similarly, we have the same estimate for the negative part  $h_{xx}^-$ . Thus for  $p = 2$ ,

$$\int_{\mathbb{T}} |h_{xx}|^2 dx \leq C, \quad (21)$$

which yields that there exists  $h^* \in V$  such that  $h_n \rightharpoonup h^*$  in  $V$ . From the strong convergence  $h_n \rightarrow h$  in  $H$  we know  $h_n \rightharpoonup h$  in  $V$ . Therefore from the convexity of cosh function, we know  $\phi$  is also convex in  $V$  and lower semicontinuous with respect to the weak topology of  $V$

$$\liminf_n \phi(h_n) \geq \phi(h).$$

Thus the lower semicontinuity in  $H$  is proved.  $\square$

As long as we have the convexity of  $\phi$ , the  $(\tau^{-1} + \lambda)$ -convexity is standard, and the proof can be found in [AGS, Section 2.4].

**Proposition 3.3** ( $(\tau^{-1} + \lambda)$ -convexity). *For any  $h, v_0, v_1 \in D(\phi)$ , there exists a curve  $v : [0, 1] \rightarrow D(\phi)$  such that  $v(0) = v_0$ ,  $v(1) = v_1$  and the functional*

$$\Phi(\tau, h; v) := \phi(v) + \frac{1}{2\tau} \|h - v\|_H^2 \quad (22)$$

*satisfies  $\tau^{-1}$ -convexity, i.e.,*

$$\Phi(\tau, h; v(t)) \leq (1-t)\Phi(\tau, h; v_0) + t\Phi(\tau, h; v_1) - \frac{1}{2} \left( \frac{1}{\tau} + \lambda \right) t(1-t) \|v_0 - v_1\|_H^2 \quad (23)$$

*for all  $\tau > 0, t \in [0, 1]$ .*

**3D. Existence and long-time behavior of the global solution.** After studying convexity and lower semicontinuity in the last section, we shall apply the convergence result in [AGS, Theorem 4.0.4] to derive that the discrete solution  $h_n$  obtained by Euler scheme (17) converges to the variational inequality solution defined in Definition 3.1. For  $v \in D(\phi)$ , denote the local slope by

$$|\partial\phi|(v) := \limsup_{w \rightarrow v} \frac{\max\{\phi(v) - \phi(w), 0\}}{\text{dist}(v, w)}. \quad (24)$$



**Remark 3.4.** In particular, by [AGS, Proposition 1.4.4], we have the local slope is

$$|\partial\phi|(v) = \min\{\|\xi\| : \xi \in \partial\phi(v)\}$$

for any  $v \in H$ . Since  $\phi$  is a smooth functional, its subdifferential  $\partial\phi$  is single-valued and equals its Fréchet differential

$$\partial\phi = D\phi(h) := [\sinh(h_{xx})]_{xx}.$$

**Proposition 3.5.** *Given  $h_0 \in H$ , for any  $t > 0$ ,  $t = n\tau$ , let  $h_n(t)$  defined in (18) be the approximation solution obtained by Euler scheme (17), then:*

(1) *There exists a local Lipschitz curve  $h(t) : [0, +\infty) \rightarrow H \in MM(\Phi; h_0)$  (i.e., minimizing movement for  $\Phi$ ) such that*

$$h_n(t) \rightarrow h(t) \quad \text{in } L^2(\mathbb{T}) \quad (25)$$

*and  $h : [0, +\infty) \rightarrow H$  is the unique EVI solution in the sense that  $h$  is unique among all the locally absolutely continuous curves such that  $\lim_{t \rightarrow 0} h(t) = h_0$  in  $H$  and*

$$\frac{1}{2} \frac{d}{dt} \|h(t) - v\|^2 + \frac{\lambda}{2} \|h(t) - v\|^2 \leq \phi(v) - \phi(h(t)) \quad \text{a.e. } t > 0, \text{ for all } v \in D(\phi). \quad (26)$$

(2) *We have the regularities*

$$\phi(h(t)) \leq \phi(v) + \frac{1}{2t} \|v - h_0\|_H^2 \quad \text{for all } v \in D(\phi), \quad (27)$$

$$|\partial\phi|^2(h(t)) \leq |\partial\phi|^2(v) + \frac{1}{t^2} \|v - h_0\|_H^2 \quad \text{for all } v \in D(|\partial\phi|). \quad (28)$$

(3) *There exist  $t_0 > 0$  and we have the exponential decay of  $h(t)$*

$$\frac{\lambda}{2} \|h(t) - h^*\| \leq (\phi(h_0) - \phi(h^*)) e^{-2\lambda(t-t_0)}, \quad \text{for any } t \geq t_0 > 0, \quad (29)$$

where  $h^* = 0$  is the unique minimizer of  $\phi$ .

This proposition is a direct result of combining Theorems 4.0.4 and 2.4.14 in [AGS] with Propositions 3.2 and 3.3. We remark the exponential convergence rate is  $\lambda = 1/\kappa^2 = (2\pi/L)^4$ . So the convergence speed is the square of the Poincaré–Wirtinger constant.

Next by Proposition 3.2 and [AGS, Theorem 2.4.15], we claim that given a better initial data  $h_0$ , the EVI solution obtained above is a global strong solution to (8) with better properties as follows.

**Theorem 3.6.** *Given any  $T > 0$  and initial datum  $h_0 \in D(\partial\phi)$  such that  $\phi(h_0) < +\infty$ , the solution obtained in Proposition 3.5 is a global strong solution in the sense that  $\partial_t h = -\partial\phi = -\partial_{xx}(\sinh(h_{xx}))$  holds for all  $t \geq 0$  with the regularities*

$$h \in C([0, T]; D(\partial\phi)) \cap C^1([0, T]; H)$$

*and the decay estimate*

$$\|h(t)\| \leq \|h_0\| \quad \text{for all } t \geq 0, \quad \|\partial_t h(t)\| = \|\partial\phi(h(t))\| \leq e^{-\lambda t} \|\partial\phi(h_0)\|_{L^2} \quad \text{for all } t \geq 0. \quad (30)$$

**Remark 3.7.** In general  $D(\partial\phi)$  is given by all  $v \in D(\phi)$  such that  $\partial\phi(v) \neq \emptyset$ . In our case,  $D(\partial\phi) = \{h \in V : (\sinh(h_{xx}))_{xx} \in H\}$ .

*Proof of Theorem 3.6.* Since Proposition 3.2 satisfies all the assumptions in [AGS, Theorem 2.4.15], we directly apply that theorem for initial datum  $h_0 \in D(\partial\phi)$  (see [AGS, Remark 2.4.16]). First  $\partial_t h = -\partial\phi = -\partial_{xx}(\sinh(h_{xx}))$  follows from [AGS, Theorem 2.3.3 and Corollary 1.4.2]. Then the statements (i) and (iii) in [Ambrosio et al. 2008, Theorem 2.4.15] show that  $h \in C([0, T]; D(\partial\phi))$  and

$$|\partial_t h|(t) := \lim_{s \rightarrow t^+} \frac{\|h(s) - h(t)\|}{s - t} = |\partial\phi|(h(t)) = \|\sinh(h_{xx})_{xx}\| \quad \text{for any } t \geq 0. \quad (31)$$

Finally the nonincreasing property of  $\|h(t)\|$  follows from (26), while the nonincreasing property of  $e^{\lambda t} \|\partial\phi(h(t))\|$  follows from the statement (ii) in [AGS, Theorem 2.4.15].  $\square$

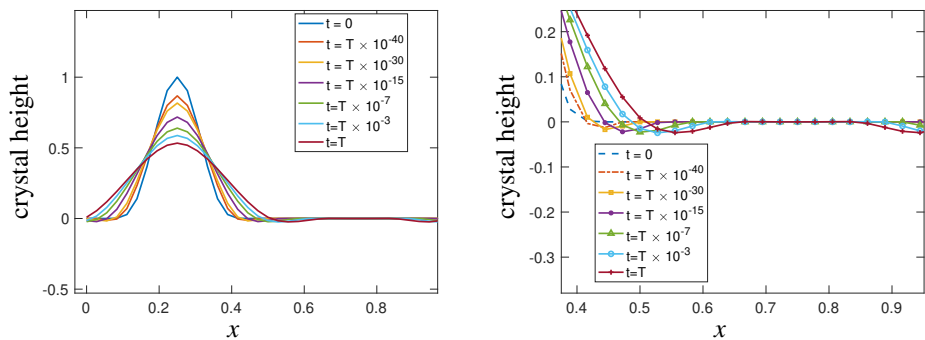
## 4. Numerics

In this section we numerically explore some properties of the PDE (1). Due to the exponential dependence on  $\beta \partial_{xxx} h$  in (1), we are limited to the setting of either relatively high temperature (small  $\beta$ ) or small curvature, in which case the PDE can be solved numerically. As in [Marzuola and Weare 2013], we focus on two key phenomena: wetting, or how compactly supported solutions evolve to fill the domain, and self-similar structures in the collapse to equilibrium. Similar numerical studies were undertaken in [Marzuola and Weare 2013] for the Arrhenius rate PDE (2). While certainly not an exhaustive study of phenomena in these models, they are key features of the dynamics one would like to understand for the evolution of crystal surfaces. The numerical method implemented below involved using a symmetric finite difference stencil to discretize spatial derivatives, along with a stiff ODE solver such as *ode15s* in Matlab in time. Similar methods were employed in [Marzuola and Weare 2013; Liu et al. 2019] and are described in some more detail there.

**4A. Wetting.** Motivated by properties of (2) shown in [Marzuola and Weare 2013], one phenomenon we investigate for (1) is how quickly mass spreads from regions of nonzero height into regions with zero height. This process is known as wetting in the study of thin films. In order for facets (macroscopic flat regions on the crystal surface) to be stable features of a surface, the wetting rate should be finite. In Figure 3, we study this phenomenon for the PDE (1). Similarly to (2), it appears numerically that the solution can wet at finite rate. One interesting difference between (1) and (2) is that for an initial nonnegative compactly supported profile, the numerical solution to (2) remains positive, while the numerical solution to (1) dips below zero before leveling off. The wetting rate was investigated for the initial profile

$$h(0, x) = \begin{cases} e^{8-|x|^{-1}-(0.5-|x|)^{-1}} & \text{for } 0 < |x| < \frac{1}{2}, \\ 0 & \text{otherwise.} \end{cases} \quad (32)$$

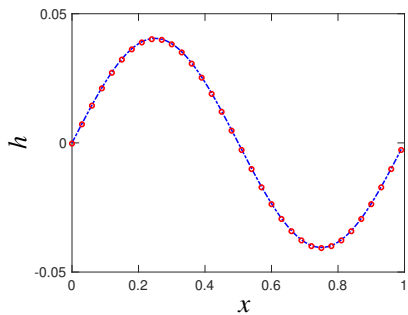
Given the high curvatures present in this model, the discretization in space is taken to be somewhat coarse when using the discrete approximation we have implemented here to the flow of (1) given by using symmetric finite difference operators in space and the variable time step stiff solver *ode15s* in Matlab in time. Our methods are comparable to the methods used in [Marzuola and Weare 2013; Liu et al. 2019]. The results are consistent as we refine the spatial grid to the extent possible, but the grid can only be refined so much for these extremely stiff approximations using the PDE solvers implemented here due to



**Figure 3.** Snapshots of solution of PDE (1) with  $\beta = .05$ , from the initial profile in (32), at various times in an interval of length  $T = 5 \times 10^{-10}$  (left) and a blowup in the region of zero initial height (right).

the high curvatures present. For related algorithms that can be implemented with much finer grids but that require longer running times, see the recent works of [Liu and Xu 2016; Craig et al. 2020], for which similar and consistent dynamics have been observed for PDEs of the form (2).

**4B. Self-similarity.** Again following the analysis in [Marzuola and Weare 2013], we study the behavior of the surfaces as they near equilibrium ( $h \equiv 0$ ). In Figure 4 we show that the surfaces appear to approximately factor as  $h(t, x) = \phi(t)g(x)$  for very large  $t$ . The results in that figure are generated via a fixed-point iteration in which the surface is evolved for some length of time and then rescaled so that the surface’s maximal (in absolute value) height is 1, and then evolved and rescaled repeatedly until convergence. The plot shows the last two fixed-point iterations (before rescaling). The fact that they nearly coincide indicates that the iterations have converged to  $g(x)$ . We note that the function  $g(x)$  will typically have some dependence on the particular initial profile. In this simulation we took  $h(0, x) = \sin(2\pi x)$  and  $\beta = .25$ . It appears that the self-similar solution is quite regular. This is in contrast to (2), in which a singularity forms in the self-similar profile at its minimum.



**Figure 4.** Results of fixed-point iteration in which PDE (1) is evolved for some interval of time, then rescaled to have maximum height (in absolute value) equal to 1 and then evolved and rescaled repeatedly until convergence. The plot shows the last two fixed point iterations (before rescaling). The solution appears to be approximately of the form  $h(t, x) = g(x)\phi(t)$ . Here, we renormalize after intervals of length  $T = 5e - 4$ .

### Appendix A: Expectation of jump rates

Recall the definition  $z_N^i = h_N^{i+1} - h_N^i$  and the expression for the rates

$$r^{i,j}(h) = \exp\left(-\frac{\beta}{2}[H(J_i^j h) - H(h)]\right), \quad |i - j| = 1, \quad (33)$$

where

$$H(h) = \sum_{i=0}^{N-1} (z^i)^2.$$

A simple computation gives the explicit expressions

$$r^{i,i+1}(z_N) = e^{-3\beta} \exp(-\beta(z^{i-1} - 2z^i + z^{i+1})), \quad (34)$$

$$r^{i+1,i}(z_N) = e^{-3\beta} \exp(\beta(z^{i-1} - 2z^i + z^{i+1})). \quad (35)$$

We will compute the expectation  $\langle \mathcal{J}^i \rangle = \langle r^{i,i+1} \rangle - \langle r^{i+1,i} \rangle$  with respect to the local Gibbs measure, assumed to be a good approximation to the true measure at macroscopic time  $t > 0$  (microscopic time  $N^4 t$ ) and for large enough  $N$ . The local Gibbs approximation is one in which we simply shift the mean of the distribution to that of the most probable state. This measure is given by

$$p_\lambda(z_N) \propto \exp\left(-\beta \sum_{i=0}^{N-1} (z^i)^2 + \sum_{i=0}^{N-1} \lambda^i z^i\right), \quad z_N \in \mathbb{Z}^N.$$

The  $\lambda^i$  are chosen so that under  $\rho_\lambda$ , the expectation of  $z_N^i$  is  $N^2 h_x(t, i/N)$ . One can show that this implies

$$N^{-2} \lambda^{Nx} \rightarrow 2\beta h_x(t, x) \quad (36)$$

as  $N \rightarrow \infty$ . To see this, observe that the equations for  $\lambda$  arise in the assumption that the local Gibbs measure captures the out of equilibrium dynamics. Namely,  $\lambda_N(t)$  is determined by the implicit equations

$$\langle z^i \rangle_{\lambda_N(t)} = N^2 \partial_x h\left(\frac{i}{N}, t\right), \quad i = 0, \dots, N-1, \quad (37)$$

where  $\langle \cdot \rangle_\lambda$  denotes expectation with respect to  $p_\lambda$  and  $h$  is the assumed macroscopic limit of  $\bar{h}_N$ . Thus, the equations for  $\lambda$  break up into  $N$  independent equations, and we have

$$\lambda_i(t) = \lambda\left(N^2 \partial_x h\left(\frac{i}{N}, t\right)\right),$$

where  $\lambda(u)$  is defined implicitly via

$$u = \frac{\sum_{n \in \mathbb{Z}} n e^{-\beta n^2 + \lambda(u)n}}{\sum_{n \in \mathbb{Z}} e^{-\beta n^2 + \lambda(u)n}} =: \langle z \rangle_\lambda,$$

where  $\langle \cdot \rangle_\lambda$  denotes expectation with respect to  $\rho_\lambda(n) \propto e^{-\beta n^2 + \lambda n}$ ,  $n \in \mathbb{Z}$ . In [Marzuola and Weare 2013, Section 6] it is shown that  $\lambda(N^\alpha u) = O(N^\alpha)$ , and if  $\alpha > 0$  then

$$\lambda(N^\alpha u)/N^\alpha \rightarrow 2\beta u, \quad N \rightarrow \infty.$$

Note that  $p_\lambda$  is a product of one-dimensional measures, which we denote by  $p_{\lambda_i}$ , where  $p_\lambda(n) \propto e^{-\beta n^2 + \lambda n}$ . We let  $\langle \cdot \rangle_\lambda$  denote expectation with respect to  $p_\lambda$  and  $\langle \cdot \rangle_{\lambda_i}$  denote expectation with respect to  $p_{\lambda_i}$ . Using the expression (34), we have

$$\langle r^{i,i+1}(z_N) \rangle_\lambda = e^{-3\beta} \langle e^{\beta z} \rangle_{\lambda_{i-1}} \langle e^{-2\beta z} \rangle_{\lambda_i} \langle e^{\beta z} \rangle_{\lambda_{i+1}}. \quad (38)$$

**Lemma A.1.** *Let  $m \in \mathbb{Z}$  and  $z \in \mathbb{Z}$  be a random variable distributed according to  $p_\lambda$ , where  $p_\lambda(n) \propto e^{-\beta n^2 + \lambda n}$ . Then*

$$\langle e^{m\beta z} \rangle_\lambda = \begin{cases} e^{\frac{1}{4}\beta m^2 + \frac{1}{2}\lambda m} & \text{if } m \text{ even,} \\ Z\left(\beta, \frac{\lambda}{2\beta}\right) e^{\frac{1}{4}\beta m^2 + \frac{1}{2}\lambda m} & \text{if } m \text{ odd,} \end{cases} \quad (39)$$

where

$$Z(\beta, \alpha) = \frac{\sum_{n=-\infty}^{\infty} e^{-\beta(n-(\alpha+\frac{1}{2}))^2}}{\sum_{n=-\infty}^{\infty} e^{-\beta(n-\alpha)^2}}.$$

If  $\beta$  is small, then  $Z(\beta, \alpha) = 1 + o(\beta)$ .

*Proof.* We have

$$\langle e^{m\beta z} \rangle_\lambda = \frac{\sum_{n=-\infty}^{\infty} e^{m\beta n - \beta n^2 + \lambda n}}{\sum_{n=-\infty}^{\infty} e^{-\beta n^2 + \lambda n}} = \exp\left(\left(\beta\left(\frac{m}{2} + \frac{\lambda}{2\beta}\right)\right)^2 - \left(\frac{\lambda}{2\beta}\right)^2\right) \frac{\sum_{n=-\infty}^{\infty} e^{-\beta(n-(\frac{m}{2} + \frac{\lambda}{2\beta}))^2}}{\sum_{n=-\infty}^{\infty} e^{-\beta(n-\frac{\lambda}{2\beta})^2}}. \quad (40)$$

The factor in front of the ratio of sums simplifies to

$$e^{\frac{1}{4}\beta m^2 + \frac{1}{2}\lambda m}.$$

If  $m$  is even then, by summing over  $n - \frac{m}{2}$ , we see that the numerator of the sum ratio equals the denominator. If  $n$  is odd, we can sum over  $n - \frac{m-1}{2}$  in the numerator to obtain

$$\frac{\sum_{n=-\infty}^{\infty} e^{-\beta(n-\frac{1}{2}-\frac{\lambda}{2\beta})^2}}{\sum_{n=-\infty}^{\infty} e^{-\beta(n-\frac{\lambda}{2\beta})^2}}.$$

To see that  $Z(\beta, \alpha) = 1 + o(\beta)$  one can express the sums in terms of the Jacobi theta function  $\vartheta_3$  and use properties of this function.  $\square$

Using Lemma A.1, we can finish computing the expectation  $\langle r^{i,i+1} \rangle_\lambda$ . Substituting (39) into (38), we obtain

$$\begin{aligned} \langle r^{i,i+1} \rangle_\lambda &= e^{-3\beta} (e^{\frac{1}{4}\beta + \frac{1}{2}\lambda^{i-1}}) (e^{\beta - \lambda^i}) (e^{\frac{1}{4}\beta + \frac{1}{2}\lambda^{i+1}}) Z(\beta, \lambda^{i-1}) Z(\beta, \lambda^{i+1}) \\ &= e^{-\frac{3}{2}\beta} \exp\left(\frac{1}{2}(\lambda^{i-1} - 2\lambda^i + \lambda^{i+1})\right) Z(\beta, \lambda^{i-1}) Z(\beta, \lambda^{i+1}) \\ &= e^{-\frac{3}{2}\beta} \exp\left(\frac{1}{2}(\lambda^{i-1} - 2\lambda^i + \lambda^{i+1})\right) + o(\beta), \end{aligned} \quad (41)$$

where the last line is for  $\beta$  small. Similarly, we have

$$\langle r^{i+1,i} \rangle_\lambda = e^{-\frac{3}{2}\beta} \exp\left(-\frac{1}{2}(\lambda^{i-1} - 2\lambda^i + \lambda^{i+1})\right) + o(\beta).$$



Since  $N^{-2}\lambda^i \approx 2\beta h_x(t, i/N)$  for  $N$  large using (36), we have

$$\lambda^{i-1} - 2\lambda^i + \lambda^{i+1} \approx 2\beta h_{xxx}\left(t, \frac{i}{N}\right),$$

so that

$$\langle \mathcal{J}^i \rangle_\lambda = \langle r^{i,i+1} - r^{i+1,i} \rangle_\lambda \approx e^{-\frac{3}{2}\beta} \sinh\left(\beta h_{xxx}\left(t, \frac{i}{N}\right)\right).$$

## Appendix B: Another method for long-time behavior using bivariational structures

Since  $\phi$  is a smooth functional, its subdifferential  $\partial\phi$  is single-valued and equals its Fréchet differential  $\partial\phi(h) = [\sinh(h_{xx})]_{xx}$ . We give another simple proof for the exponential decay to 0 of the global classical solution  $h$ .

On one hand, it is easy to obtain the energy dissipation

$$\frac{d\phi}{dt} = \int \partial\phi(h) h_t \, dx = \int -|\partial\phi(h)|^2 \, dx =: -D. \quad (42)$$

On the other hand, we use the  $\lambda$ -convexity of  $\phi$  to establish the connection between  $D$  and  $\phi$ . First, from the  $\lambda$ -convexity (20) with some  $\lambda > 0$ , we know  $\phi(h) - \frac{\lambda}{2}\|h\|^2$  is convex. Thus we have

$$\phi(h(t)) - \frac{\lambda}{2}h^2(t) - \phi(v) + \frac{\lambda}{2}v^2 \leq \langle \partial\phi(h(t)) - \lambda h(t), h(t) - v \rangle \quad (43)$$

for any  $v \in L^2$ . Then we obtain

$$\phi(h(t)) - \phi(v) \leq -\frac{\lambda}{2}\|h(t) - v\|^2 + \langle \partial\phi(h(t)), h(t) - v \rangle \quad (44)$$

for any  $v \in L^2$ . Using Young's inequality, we have

$$\langle \partial\phi(h(t)), h(t) - v \rangle \leq \frac{\lambda}{2}\|h(t) - v\|^2 + \frac{1}{2\lambda}\|\partial\phi(h(t))\|^2. \quad (45)$$

From this, the estimate (44) becomes

$$\phi(h(t)) - \phi(v) \leq -\frac{\lambda}{2}\|h(t) - v\|^2 + \langle \partial\phi(h(t)), h(t) - v \rangle \leq \frac{1}{2\lambda}\|\partial\phi(h(t))\|^2. \quad (46)$$

Combining (46) with (42), we obtain

$$\frac{d}{dt}(\phi(h(t)) - \inf \phi) = -\|\partial\phi(h(t))\|^2 \leq -2\lambda(\phi(h(t)) - \inf \phi), \quad (47)$$

which gives the exponential decay to  $\inf \phi = \phi(h^*)$  with  $h^* = 0$ ; i.e.,

$$\frac{\lambda}{2}\|h(t) - h^*\|^2 \leq \phi(h(t)) - \phi(h^*) \leq (\phi(h(0)) - \phi(h^*))e^{-2\lambda t}. \quad (48)$$

## Acknowledgements

This project was started while Marzuola was on sabbatical at Duke University in the Spring of 2019. Marzuola thanks Bob Kohn, Dio Margetis and Jonathan Weare for many valuable conversations regarding modeling of kinetic Monte Carlo. Katsevich is supported by the DOE Computational Science Graduate Fellowship. Liu was supported by the National Science Foundation (NSF) grants DMS-1812573 and DMS-2106988. Lu was supported by the National Science Foundation via grant DMS-1454939. Marzuola

acknowledges support from the NSF through NSF CAREER grant DMS-1352353 and NSF grant DMS-1909035. The authors also thank the anonymous referees for very valuable suggestions to improving the final version of the paper.

## References

- [Ambrosio et al. 2008] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, 2nd ed., Birkhäuser, Basel, 2008. [MR](#) [Zbl](#)
- [Binh 1983] V. T. Binh (editor), *Surface mobilities on solid materials*, NATO Advanced Science Institutes **86**, Springer, 1983.
- [Craig et al. 2020] K. Craig, J.-G. Liu, J. Lu, J. L. Marzuola, and L. Wang, “A proximal-gradient algorithm for crystal surface evolution”, preprint, 2020. [arXiv](#)
- [Gao et al. 2019] Y. Gao, J.-G. Liu, and X. Y. Lu, “Gradient flow approach to an exponential thin film equation: global existence and latent singularity”, *ESAIM Control Optim. Calc. Var.* **25** (2019), art. id. 49. [MR](#) [Zbl](#)
- [Gao et al. 2020] Y. Gao, J.-G. Liu, J. Lu, and J. L. Marzuola, “Analysis of a continuum theory for broken bond crystal surface models with evaporation and deposition effects”, *Nonlinearity* **33**:8 (2020), 3816–3845. [MR](#) [Zbl](#)
- [Krug et al. 1994] J. Krug, H. Dobbs, and S. Majaniemi, “Adatom mobility for the solid-on-solid model”, *Z. Phys. B* **97**:2 (1994), 281–291.
- [Liu and Xu 2016] J.-G. Liu and X. Xu, “Existence theorems for a multidimensional crystal surface model”, *SIAM J. Math. Anal.* **48**:6 (2016), 3667–3687. [MR](#) [Zbl](#)
- [Liu and Xu 2017] J.-G. Liu and X. Xu, “Analytical validation of a continuum model for the evolution of a crystal surface in multiple space dimensions”, *SIAM J. Math. Anal.* **49**:3 (2017), 2220–2245. [MR](#) [Zbl](#)
- [Liu et al. 2019] J.-G. Liu, J. Lu, D. Margetis, and J. L. Marzuola, “Asymmetry in crystal facet dynamics of homoepitaxy by a continuum model”, *Phys. D* **393** (2019), 54–67. [MR](#) [Zbl](#)
- [Marzuola and Weare 2013] J. L. Marzuola and J. Weare, “Relaxation of a family of broken-bond crystal-surface models”, *Phys. Rev. E* **8**:3 (2013), art. id. 032403.
- [Spohn 1991] H. Spohn, *Large scale dynamics of interacting particles*, Springer, 1991. [Zbl](#)
- [Xu 2018] X. Xu, “Existence theorems for a crystal surface model involving the  $p$ -Laplace operator”, *SIAM J. Math. Anal.* **50**:4 (2018), 4261–4281. [MR](#) [Zbl](#)

Received 16 Mar 2020. Revised 19 Nov 2020. Accepted 29 Apr 2021.

YUAN GAO: [gao662@purdue.edu](mailto:gao662@purdue.edu)

Department of Mathematics, Purdue University, West Lafayette, IN, United States

ANYA E. KATSEVICH: [katsevich@cims.nyu.edu](mailto:katsevich@cims.nyu.edu)

Courant Institute of Mathematical Sciences, New York University, New York City, NY, United States

JIAN-GUO LIU: [jliu@math.duke.edu](mailto:jliu@math.duke.edu)

Departments of Mathematics and Physics, Duke University, Durham, NC, United States

JIANFENG LU: [jianfeng@math.duke.edu](mailto:jianfeng@math.duke.edu)

Departments of Mathematics, Physics and Chemistry, Duke University, Durham, NC, United States

JEREMY L. MARZUOLA: [marzuola@math.unc.edu](mailto:marzuola@math.unc.edu)

Department of Mathematics, University of North Carolina, Chapel Hill, NC, United States