

REVIEW ARTICLE OPEN

Recent developments in computational modelling of nucleation in phase transformations

Lei Zhang¹, Weiqing Ren^{2,3}, Amit Samanta⁴ and Qiang Du⁵

Nucleation is one of the most common physical phenomena in physical, chemical, biological and materials sciences. Owing to the complex multiscale nature of various nucleation events and the difficulties in their direct experimental observation, development of effective computational methods and modeling approaches has become very important and is bringing new light to the study of this challenging subject. Our discussions in this manuscript provide a sampler of some newly developed numerical algorithms that are widely applicable to many nucleation and phase transformation problems. We first describe some recent progress on the design of efficient numerical methods for computing saddle points and minimum energy paths, and then illustrate their applications to the study of nucleation events associated with several different physical systems.

npj Computational Materials (2016) **2**, 16003; doi:10.1038/npjcompumats.2016.3; published online 8 April 2016

INTRODUCTION

The recent call of Materials Genome Initiative (MGI) exemplifies the use of computational modelling in new materials design.¹ One of the most effective ingredients to design materials with certain desired properties is through the control of their phase transformations and microstructure evolution. These processes often start with the nucleation of nanoscale new-phase particles, followed by growth and particle impingement or coarsening.

Generically, nucleation of a new phase requires overcoming a minimum thermodynamic barrier, which leads to a saddle point configuration along the minimum energy path on the energy landscape. Being a rare event, nucleation is difficult to observe directly in physical experiments as a critical nucleus typically appears transiently at very fast time scales. Therefore, there have been many theoretical studies of the nucleation event. Early classical nucleation theories mainly study phase changes in fluids, e.g., nucleation of a liquid droplet from a vapour phase. The thermodynamic properties of a nucleus are assumed to be the same as in the corresponding bulk. Consequently, the size of a critical nucleus is determined as a result of bulk free-energy reduction and interfacial energy increase, $r^* = -2\gamma/\Delta G_v$, where γ is the interfacial energy per unit area between a nucleus and the parent matrix and ΔG_v is the free-energy-driving force per unit volume. The nucleation rate then depends on the height of the critical energy barrier ΔE^* , i.e., $I = I_0 \exp(-\Delta E^*/k_B T)$ with the pre-exponential factor I_0 calculated from fundamental statistical approaches, k_B and T being the Boltzmann's constant and the absolute temperature, respectively. However, nucleation process in general phase transformation problems is much more challenging to characterise due to the possible complex geometry and structure of critical nuclei. Therefore, much computational effort has been called for to study the nucleation events in various applications.^{2–9} Being a saddle point of the free energy, the critical nucleus satisfies the Euler–Lagrange equation of the energy.

However, solving the Euler–Lagrange equation directly or classical optimisation methods are inefficient for this problem due to the unstable nature of the saddle point. Our goal here is to provide a review of some relatively new progress in this direction to the computational materials science community.

This review is by no means a comprehensive treatment of nucleation modelling and transition state search. We refer to^{10–12} for other excellent reviews on the subject. In preparing for this article, over a hundred papers on the topic were researched, about half of them are included as references. The main objective of this paper is to provide a sampler of some relatively new progress on the development of numerical algorithms that are applicable to general nucleation and phase transformation problems. Our discussions are focused on a few approaches developed in the past decade, and our reviews of the literature are narrowed to those highly relevant works. This serves as our particular search criterion.

We organise the review by describing some developments of algorithmic works first, followed by illustrations of their applications to nucleation events in various material systems. On the algorithmic side, we first introduce some recent developments of methods for finding saddle points and minimum energy paths. One of the popular approaches is the class of surface walking methods. These methods locate the saddle point starting from an initial state. Here we mainly consider methods based on the dimer method,^{13,14} the gentlest ascent method such as the recent works on the gentlest ascent dynamics^{15–18} and the shrinking dimer dynamics.¹⁹ Another approach is categorised as path-finding methods that are to compute the minimum energy path (MEP). Some of the representative methods include the nudged elastic band (NEB) method^{20,21} and the string method^{22–27} with the latter being a focus of our discussion for this type of approaches. For complex and rough energy landscapes, we review some methods that can be used to compute either the transition tubes^{28–30} or the mean free-energy path in the space of collective variables.^{31,32}

¹Beijing International Center for Mathematical Research, Center for Quantitative Biology, Peking University, Beijing, China; ²Department of Mathematics, National University of Singapore, Singapore; ³Institute of High Performance Computing, Agency for Science, Technology and Research, Singapore; ⁴Physics Division, Lawrence Livermore National Laboratory, Livermore, CA, USA and ⁵Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY, USA.

Correspondence: Q Du (qd2125@columbia.edu)

Received 28 November 2015; revised 14 January 2016; accepted 15 January 2016

To illustrate how the numerical methods can be utilised in the nucleation studies, we present their applications to the nucleation processes in three different systems. In the first application, we consider solid-state phase transformation, and present predicted morphologies of critical nuclei with long-range elastic interactions.^{8,33,34} In the second example, we consider the search of transition pathways in micromagnetics, and show the application of the string method in the study of thermally activated switching and the energy landscape of submicron-sized magnetic elements.³ The last example is on the solid melting problem, as an illustration of nucleation in solid-fluid phase transition, the multiple barrier-crossing events within the solid basin are discussed, which reveal the importance of nonlocal behaviour.⁵

COMPUTING SADDLE POINTS AND TRANSITION PATHS

The classical transition state theory gives a sufficiently accurate description of the transition process for systems with smooth energy landscapes.^{35,36} For such systems, the transition state is a saddle point with the lowest energy that connects two neighbouring local minima. Here we focus on the numerical algorithms for computing saddle points and MEPs in systems with smooth potential energy. Generally speaking, there are two distinct classes of numerical methods: (1) surface walking methods for finding saddle points starting from a single state; and (2) path-finding methods for computing MEPs, which involve two end states. In this subsection, we illustrate the ideas behind these approaches via some recently developed methods.

Finding saddle points/surface walking methods

Several surface walking methods have been developed to locate saddle points. An important character of such methods is to perform a systematic search for a saddle point starting from a given initial state, without knowing the final states. Here we focus on a special class of surface walking methods, the so-called minimum mode following methods, where only the lowest eigenvalue and the corresponding eigenvector of the Hessian are calculated and subsequently used together with the energy gradient (often referred as the force) to locate the saddle point. The representative methods include the dimer method,^{13,14} the gentlest ascent method/dynamics,^{15–17} and the shrinking dimer dynamics.^{19,37} Besides the minimum mode following methods, there are some other surface walking methods as well, such as the eigenvector-following method,³⁸ the minimax method,^{39,40} the activation-relaxation technique,^{41–43} the step and slide method,⁴⁴ to name a few.

The gentlest ascent method/dynamics. The gentlest ascent method was first developed by Crippen and Scheraga¹⁵ as a numerical algorithm to search the transition path from a local minimum to a neighbouring minimum via an intervening saddle point on an energy surface. Later, E and Zhou¹⁶ reformulated the gentlest ascent method as a dynamical system and proposed the gentlest ascent dynamics (GAD).

Consider a system with N degrees of freedom (DOF) contained in a vector $x \in \mathbb{R}^N$, the GAD refers to the following dynamic system

$$\begin{cases} \dot{x} = -\nabla V(x) + 2\frac{(\nabla V, v)}{(v, v)}v, \\ \dot{v} = -\nabla^2 V(x)v + \frac{(v, \nabla^2 V v)}{(v, v)}v, \end{cases} \quad (1)$$

where ∇V is the gradient of the potential energy V , $\nabla^2 V$ is the Hessian of V , and (\cdot, \cdot) denotes the standard inner product. The second equation in Equation (1) determines the orientation vector v to be the eigenvector that corresponds to the smallest eigenvalue of $\nabla^2 V$, which is used in the first equation in Equation

(1) as an ascent direction to find the saddle point. The stable fixed points of this dynamical system were proved to be the index-1 saddle points in ref. 16. The GAD has also been applied to non-gradient systems.

The dimer method. A dimer system consists of two nearby points x_1 and x_2 separated by a small distance, that is, with a small dimer length $l = \|x_1 - x_2\|$. The dimer orientation is given by a unit vector v so that $x_1 - x_2 = lv$. The (rotating) center of the dimer is usually taken as the midpoint of the dimer, i.e., $x_c = \frac{x_1}{2} + \frac{x_2}{2}$. The dimer method developed by Henkelman and Jonsson¹³ uses only first-order derivatives of the energy to calculate the forces F_1 and F_2 on the two end points of the dimer.

The dimer method proceeds by alternately performing the rotation and translation steps. The rotation step is to find the lowest eigenmode at the center of the dimer. This is done by a single rotation towards the configuration which minimises the dimer energy with its center fixed. In practice, a conjugate gradient algorithm is used to choose the plane of rotation, whereas the minimisation of the force on the dimer within the plane is carried out by using the modified Newton method. In the translation step, the potential force is first modified so that its component along the dimer is reversed. Then the dimer is translated using the modified force with either the steepest descent algorithm or the conjugate gradient method.

Much effort has been made to improve the dimer method. In ref. 14, it was argued that the method is more effective when the rotation step is proceeded until convergence, compared with using a single rotation after every translation step. More recently, in the work of Gould, Ortner and Packwood,⁴⁵ a preconditioner technique is used in the dimer iteration and a line search technique is applied for finding the step size to achieve better efficiency and convergence. A recent attempt was made in ref. 46 to unify various techniques for accelerating the dimer methods though most of the approaches discussed there do not share superlinear convergence property. In contrast, an interesting development was given by Kastner and Sherwood in ref. 47, which used the Limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm for the dimer translation to greatly improve the convergence of the dimer method. Moreover, the numbers of gradient calculations per dimer iteration are reduced through an extrapolation of the gradients during repeated dimer rotations.

The shrinking dimer dynamics. The shrinking dimer dynamics (SDD) was proposed by Zhang and Du in ref. 19 to find the index-1 saddle points based on the original dimer method.¹³ It follows a dynamic system as

$$\begin{cases} \mu_1 \dot{x}_a = (I - 2vv^T)((1 - \alpha)F_1 + \alpha F_2), \\ \mu_2 \dot{v} = (I - vv^T)\frac{(F_1 - F_2)}{l}, \\ \mu_3 \dot{l} = -\nabla E_{\text{dimer}}(l), \end{cases} \quad (2)$$

where μ_1, μ_2, μ_3 are nonnegative relaxation constants, and α is a constant parameter between 0 and 1, which determines the rotating center of the dimer.

The first two equations of SDD in Equation (2) represent the translation step and the rotation step, respectively, which are essentially same with the classical dimer method. The operator $(I - 2vv^T)$ is the Householder mirror reflection that reverses the component of the force along v . The operator $(I - vv^T)$ is a projection that makes v of unit length. Instead of using a fixed small dimer length in the classical dimer method, the third equation of SDD in Equation (2) follows a gradient flow of the energy function $E_{\text{dimer}}(l)$, which is generally taken as a monotonically increasing function in l such that it allows the shrinking of the dimer length over time by forcing it to approach zero at the steady state.

The dynamic system (Equation (2)) is similar to Equation (1), but avoids the calculation of the second-order derivatives by requiring only the evaluation of the natural forces. Rigorous analysis for both the continuous dynamic system and its time discretisation were carried out in ref. 19 to demonstrate the linear stability and convergence, in particular, the importance of shrinking the dimer length for the guaranteed convergence.

In terms of numerical schemes, the SDD employs either explicit or modified Euler method to obtain the linear convergence in ref. 19. In ref. 37, a constrained SDD has been proposed as an variant of SDD to handle the saddle point search on a constrained manifold. The use of preconditioner has also been alluded to in refs 37,48 but without implementation. To accelerate the convergence and further improve the efficient of the dimer method, the optimisation-based shrinking dimer (OSD) method was recently proposed by Zhang *et al.* in ref. 49 as the generalised formulation of SDD. The OSD method translates the rotation and translation steps of the dimer in Equation (2) to the corresponding optimisation problems such that the efficient optimisation methods can be naturally employed to substantially speed up the computation of saddle points. In ref. 49, the Barzilai–Borwein gradient method was used as an implementation of OSD and showed a superlinear convergence.

Finding minimum energy path

In the case when the object of interest is the most probable transition path between metastable states of the smooth potential energy, it is known that for overdamped Langevin dynamics the most probable path for the transition is the MEP, which is the path in the configuration space such that the potential force is parallel to the tangents along the path, i.e.,

$$(\nabla V)^\perp(\varphi) := \nabla V - (\nabla V, \hat{\tau})\hat{\tau} = 0, \quad (3)$$

where $\hat{\tau}$ is the unit tangent vector to the curve. Several numerical methods have been developed for finding MEPs. Below we review two popular ones, the (zero temperature) string method^{24,25,50} and the nudged elastic band method,^{20,21,51} for computing the MEPs. Once the MEP is found, the transition states can be identified from the maxima of the energy along the MEP.

The (zero-temperature) string method. The string method was first proposed by E *et al.* in ref. 24, and it proceeds by evolving a string in the configuration space by using the steepest descent dynamics. Let $\varphi(\alpha, t)$ denote the string at the time t with parameterisation $\alpha \in [0, 1]$, then the string evolves according to

$$\dot{\varphi} = -\nabla V(\varphi) + \lambda \hat{\tau}, \quad 0 < \alpha < 1, \quad (4)$$

where λ is the Lagrange multiplier to impose the equal arc-length constraint, and $\tau = \varphi' / |\varphi'|$ is the unit tangent vector to the string. Here we use $\dot{\varphi}$ and φ' to denote the temporal and spatial derivatives, respectively. The above evolution equation is supplemented with the boundary conditions:

$$\varphi(0, t) = x_a, \quad \varphi(1, t) = x_b, \quad (5)$$

where x_a and x_b are the two minima of the potential energy $V(x)$.

In the numerical implementation, the discretised string is composed of a number of images $\{\varphi_i(t), i = 0, 1, \dots, N\}$, where $\varphi_i(t) = \varphi(i/N, t)$. Equation (4) is solved using a time-splitting scheme, and the string method iterates the following two steps:

String evolution updates the images on the string over some time interval Δt according to the potential force:

$$\dot{\varphi}_i = -\nabla V(\varphi_i), \quad i = 1, 2, \dots, N-1. \quad (6)$$

This equation can be integrated by any ODE solver, e.g., the Euler method, or Runge–Kutta methods.

String reparametrisation is applied to redistribute the images along the string using linear or cubic spline interpolation according to equal arc-length parameterisation.

The string method is a simple but an effective technique for finding MEPs. It only requires an ODE solver and an interpolation scheme, thus it is easy to implement and can be readily incorporated into any existing code as long as the force evaluation is provided.

In the case that the exact locations of the minima x_a and/or x_b are unknown beforehand, the two end points can be computed on-the-fly by following the relaxation dynamics as the string evolves towards the MEP, i.e.,

$$\dot{\varphi}_0 = -\nabla V(\varphi_0), \quad \dot{\varphi}_N = -\nabla V(\varphi_N). \quad (7)$$

At the steady state, the end points converge to the minima x_a and x_b , as long as they initially lie in the basins of attraction of these minima, respectively.

In ref. 26, the dynamics of the final point φ_N is modified so that it converges to a saddle point. This can be used for saddle point search. Specifically, the dynamics of φ_N follows the modified potential force

$$\dot{\varphi}_N = -\nabla V(\varphi_N) + 2(\nabla V(\varphi_N), \hat{\tau}_N)\hat{\tau}_N, \quad (8)$$

where $\hat{\tau}_N$ is the unit tangent vector to the string at φ_N . In Equation (8), the component of the potential force in the direction along the string is reversed. This makes φ_N climb uphill towards a saddle point.

In practice, an initial string is constructed in the basin of the initial state, e.g., the linear interpolation of the initial state and its small perturbation. Then following the dynamics in Equations (4) and (8), the final state converges to a saddle point, and the string converges to the MEP connecting the initial state and the saddle point.

The string method can also be generalised to compute the mean free-energy path (MFEP) in collective variable spaces. This will be discussed later in this review.

The nudged elastic band method. The NEB method is another approach to compute MEPs. It connects two minima x_a and x_b by a chain of states, then evolves this chain by using a combination of the potential force and a spring force:

$$\dot{x}_i = -(\nabla V(x_i))^\perp + (F_i^s \cdot \hat{\tau}_i)\hat{\tau}_i, \quad i = 1, \dots, N-1, \quad (9)$$

where $-(\nabla V)^\perp = -\nabla V + (\nabla V, \hat{\tau})\hat{\tau}$ denotes the projection of the potential force in the hyperplane perpendicular to the chain, $F_i^s = k(x_{i+1} - 2x_i + x_{i-1})/\Delta\alpha$ is the spring force with $\Delta\alpha = 1/N$ and $k > 0$ being a parameter, and $\hat{\tau}_i$ denotes the unit tangent vector along the chain at x_i .

The NEB is an improvement upon the elastic band method.^{52–54} The elastic band method, which evolves the chain using the total spring forces, fails to converge to the MEP as the spring force tends to make the chain straight, which leads to corner-cutting. The NEB overcomes this difficulty by using only the normal component of the potential force and the tangential component of the spring force.

Compared with the string method, which evolves a curve with intrinsic arc-length parameterisation, the NEB uses a spring force to connect the states along the chain. This requires the prescription of an additional parameter k . The value of k controls the strength of the spring force: if it is too large, then the dynamics in Equation (9) becomes stiff, which may limit the time step; on the other hand, a small k leads to weak spring force then the states may not be evenly distributed along the chain.

COMPLEX ENERGY LANDSCAPES

The concept of transition states and MEPs become inappropriate when the energy landscape is rough with many saddle points, and most of these saddle points have potential-energy barriers that are less than or comparable to the strength of the noise, thus do not act as barriers for the transition. In this case, one approach is to compute the so-called transition tubes. The transition tube carries most of the current of the transition in the configuration space. The other way is to consider the free-energy surface (FES) instead of potential energy surface. One first selects a set of collective variables, then computes the transition state or the MFEP in the space of these collective variables. The FES is typically much more smooth than the potential energy surface. Below we review these two approaches.

Computing transition tubes

The transition tube is defined by considering the committor function and the current of reaction trajectories. The committor function specifies, at each point of the configuration space, the probability that the reaction or transition will succeed if the system is initiated at that point. The isosurfaces of the committor function, called the iso-committor surfaces, foliate the space between the metastable states under consideration. Assume the reaction current is localised in the configuration space, then it intersects with the isocommittor surfaces in one or a few isolated regions. The collection of these regions form one or a few tubes, which are called transition tubes (Figure 1).^{28–30}

Under the assumptions that the transition tube is thin compared with the local radius of the curvature of the centerline and the isocommittor surfaces can be approximated by hyperplanes within the transition tube, it was shown that the centerline of the transition tube, which is defined as

$$\varphi(\alpha) = \langle x \rangle_{p_\alpha} \equiv \frac{\int_{p_\alpha} x e^{-\beta V} d\sigma(x)}{\int_{p_\alpha} e^{-\beta V} d\sigma(x)}, \quad (10)$$

is normal to the isocommittor surfaces, i.e.,

$$\hat{n}(\alpha) \parallel \varphi'(\alpha), \quad (11)$$

where the centerline φ is parameterised by α (e.g., the arc length), $\varphi'(\alpha)$ is the tangent vector to the centerline, and \hat{n} is the normal vector of the hyperplanes.

The finite-temperature string method is an iterative procedure to solve Equations (10) and (11).^{30,55,56} Let φ^n denote the centerline of the tube at the n th step, the new configuration of the centerline is computed following two steps:

Expectation step: sample on the hyperplanes perpendicular to the current configuration of the centerline and compute the center of mass $\langle x \rangle_{p_\alpha}$ on each hyperplane.

Relaxation step: evolve the centerline to the new configuration according to

$$\varphi^{n+1} = \varphi^n + \Delta t \left(\langle x \rangle_{p_\alpha} - \varphi^n \right). \quad (12)$$

In practice, each image is evolved according to Equation (12), followed by a reparameterisation step using interpolation as in the zero-temperature string method. The centerline of mass $\langle x \rangle_{p_\alpha}$ can be estimated using constrained simulations in the hyperplanes^{30,55} or Voronoi cells.⁵⁶ At convergence, the Voronoi cells also have the remarkable property that they form a centroidal Voronoi tessellation or CVT—a concept introduced in ref. 57.

Exploring the free-energy surface

We consider the case when the transition of interest can be described by the collective variables $\xi(x) = (\xi_1(x), \dots, \xi_d(x))$. These collective variables correspond to the slow manifold along which the transition occurs, therefore, the dimension of the collective variable space is usually much smaller than the total DOF in the system. The free energy associated with $\xi(x)$ is given by

$$F(z) = -k_B T \ln \left(\frac{1}{Z} \int_{R^N} e^{-\beta V(x)} \prod_{i=1}^d \delta(\xi_i(x) - z_i) dx \right), \quad (13)$$

where $z = (z_1, \dots, z_d)$ are the coordinates in the collective variable space, $\beta = 1/k_B T$, and Z is the partition function. On the FES, the metastable and transition states are given by the local minima of $F(z)$ and the saddle points between them, respectively. The path of maximum likelihood for the transition follows the MFEP,³¹ which satisfies

$$(M(\varphi) \nabla F(\varphi))^\perp = 0, \quad (14)$$

where $(\dots)^\perp$ denotes the projection on the hyperplane perpendicular to the path φ , and M is an $d \times d$ tensor whose entries are given by

$$M_{ij}(z) = \frac{1}{Z} e^{\beta F(z)} \int_{R^N} \sum_{k=1}^N \frac{\partial \xi_i(x)}{\partial x_k} \frac{\partial \xi_j(x)}{\partial x_k} e^{-\beta V(x)} \quad (15)$$

$$\times \delta(z_1 - \xi_1(x)) \cdots \delta(z_d - \xi_d(x)) dx. \quad (16)$$

If one is interested in the transition between two metastable states z_a and z_b , then Equation (14) is supplemented with the boundary conditions such that the path φ connects the two minima.

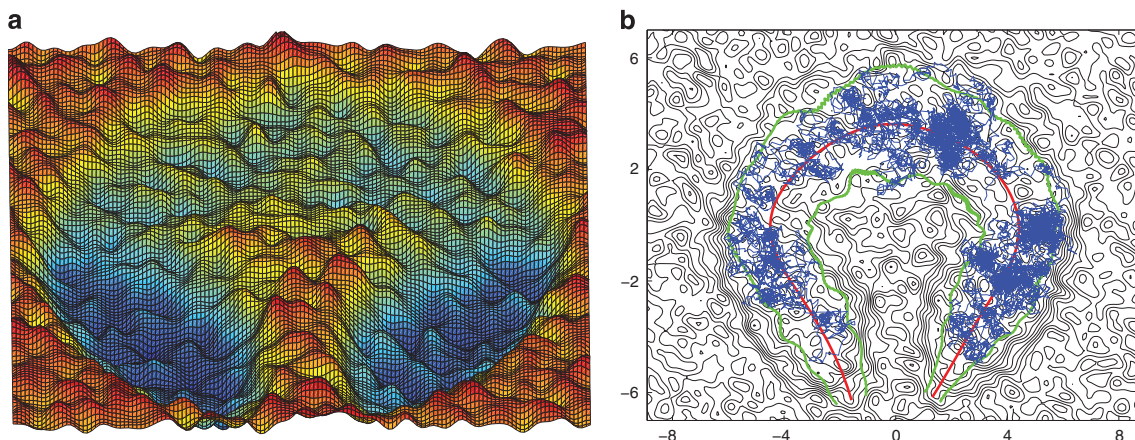


Figure 1. Example of the rough energy landscape (a) and the transition tube (b) identified by the string method. Figure modified from ref. 55. Permission has been granted by the publisher to use the original figures.

The numerical techniques in ‘Finding saddle points/surface walking methods’ and ‘Finding minimum energy path’ can be readily used to compute the saddle points and the MFEP on the FES. For example, the MFEP can be computed using the zero-temperature string method by evolving a string $\varphi(\alpha)$ according to

$$\gamma\dot{\varphi}(\alpha, t) = -M(\varphi(\alpha, t))\nabla F(\varphi(\alpha, t)) + \lambda(\alpha, t)\hat{t}, \quad (17)$$

where the boundary conditions $\varphi(0, t) = z_a$ and $\varphi(1, t) = z_b$. At the steady state, the string converges to the MFEP with the prescribed parameterisation.

In the above algorithm or the methods for computing the saddle points, one needs to compute the mean force $-\nabla F$. When the dimension of the free-energy space is low, e.g., $d=2$ or 3 , one can first construct the free-energy function, for example, using a variational reconstruction scheme as in the single sweep method.^{5,58} The method follows two steps: (a) first a sufficiently long temperature-accelerated molecular dynamics (TAMD)/adiabatic free-energy dynamics (AFED) trajectory is obtained and a set of centers in the space of the collective variables and; (b) the free energy is expressed as a superposition of Gaussian radial basis functions (RBFs) placed at these centers. The optimal values of the coefficients of the RBFs are obtained by optimising the cost function, which is the standard mean-squared error in the gradient of the free energy.

The reconstruction of the FES is only possible when the number of collective variables is small. When the number is large, the mean force can be computed on the fly without the requirement of a globally explicit form of the free-energy function *a priori*. This can be done by averaging over long trajectories with the collective variables fixed at their current locations.³¹ Alternatively, one may couple the microscopic dynamics with the evolution of the string simultaneously.³² Specifically, each image along the string is coupled with two independent microscopic systems, denoted by $x^{(1)}$ and $x^{(2)}$, respectively, where the first system is for the computation of M and the second system is for the computation of the mean force $-\nabla F$. These systems evolve by molecular dynamics or the Langevin dynamics at the temperature T in the extended potential

$$U_k(x; \varphi(\alpha, t)) = V(x) + \frac{1}{2}\kappa|\xi(x) - \varphi(\alpha, t)|^2. \quad (18)$$

The second term in the extended potential is to constrain the system at the current location of the string, so that dynamics of the microscopic systems are slaved to the evolution of the string. The tensor M and the mean force $-\nabla F$ are computed from the instantaneous configurations of $x^{(1)}$ and $x^{(2)}$:

$$M_{ij} \approx \sum_{k=1}^N \frac{\partial \xi_i}{\partial x_k}(x^{(1)}) \frac{\partial \xi_j}{\partial x_k}(x^{(1)}), \quad (19)$$

$$\nabla F \approx -\kappa(\xi(x^{(2)}) - \varphi(\alpha, t)). \quad (20)$$

To ensure the microscopic systems to have enough time to equilibrate before the string move significantly, the dynamics of the string in Equation (17) is slowed down by taking $\gamma 1$ in Equation (17).³²

APPLICATIONS IN MATERIALS SCIENCE

Nucleation in solid-state phase transformations

Following the seminal work by Cahn and Hilliard,⁵⁹ the phase field (diffuse interface) model has been successfully employed to investigate nucleation and microstructure evolution in phase transformations.⁶⁰ In the diffuse interface description, a set of conserved field variables c_1, c_2, \dots, c_m and non-conserved field

variables $\eta_1, \eta_2, \dots, \eta_n$ are often used to describe the compositional/structural domains and the interfaces, and the total free energy of an inhomogeneous microstructure system is formulated as

$$E_{total} = \int \left[\sum_{i=1}^m \alpha_i (\nabla c_i)^2 + \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^n \beta_{ij} \nabla_i \eta_k \nabla_j \eta_k \right] dx \quad (21)$$

$$+ f(c_1, c_2, \dots, c_m, \eta_1, \eta_2, \dots, \eta_n) dx \\ + \int \int G(x-x', \vec{c}, \vec{\eta}) dx dx', \quad (22)$$

where the gradient coefficient α_i and β_{ij} can be used to reflect the interfacial energy anisotropy and the function f corresponds to the local free-energy density. The last integral in the above equation represents a nonlocal term that includes a general long-range interaction such as elastic interactions in solids.

In solid-state phase transformations, the local free-energy function f is often described by a polynomial of order parameters with a conventional Landau-type of expansion, for instance, a simple double-well potential,

$$f(\eta) = \frac{\eta^4}{4} - \frac{\eta^2}{2} + \lambda \frac{\eta^3 - 3\eta}{4}, \quad (23)$$

with two energy wells at $\eta = \pm 1$ and λ determines the well depth difference.

Furthermore, the lattice mismatch between solid phases and domains are accommodated by elastic displacements, so the computation of the long-range elastic energy is needed. For the case that the elastic modulus is anisotropic but homogeneous, the microscopic elasticity theory of Khachaturyan⁶¹ is often used in phase field simulation and the total elastic energy of a microstructure can be given by

$$E_{elastic} = \frac{1}{2} \int_{\Omega} C_{ijkl} e_{ij}^{el} e_{kl}^{el} dx, \quad (24)$$

where the elastic strain e^{el} is the difference between the total strain and stress-free strain as stress-free strain does not contribute to the total elastic energy.

Examples of predicted critical profiles in the presence of long-range elastic energy for a cubic crystal are shown in Figure 2a. By combining the diffuse-interface approach with the minimax technique,³⁹ it demonstrated that the elastic interactions can markedly change the critical nucleus morphology, thus revealing the fascinating possibility of nuclei with non-convex shapes, together with the phenomenon of shape-bifurcation and the formation of critical nuclei whose symmetry is lower than both the new phase and the original parent matrix.^{8,33,62}

For the conserved solid field with profile $c = c(x)$, a combination of diffuse-interface description and a constrained string method²³ is able to predict both the critical nucleus and equilibrium precipitate morphologies simultaneously (Figure 2b).³⁴ Using the cubic to cubic or cubic to tetragonal transformation as examples, simulations showed that the morphology of a critical nucleus can be markedly different from the equilibrium one due to the elastic energy contributions.^{34,63}

The general framework of diffuse interface model has been greatly extended to investigate complex nucleation phenomena, such as heterogenous nucleation,^{48,64,65} homogeneous/heterogenous crystal nucleation^{66–68} and nucleation dynamics.^{69–72} In particular, to investigate the nucleation and growth kinetics in real alloys, the thermodynamic properties of critical nuclei with the chemical free energy and interfacial energy can be assessed from thermodynamic calculation and atomistic simulation, and then efficient numerical method of computing saddle points can be applied to quantitatively predict the critical nucleus, nucleation energy barriers and growth kinetics.⁷³

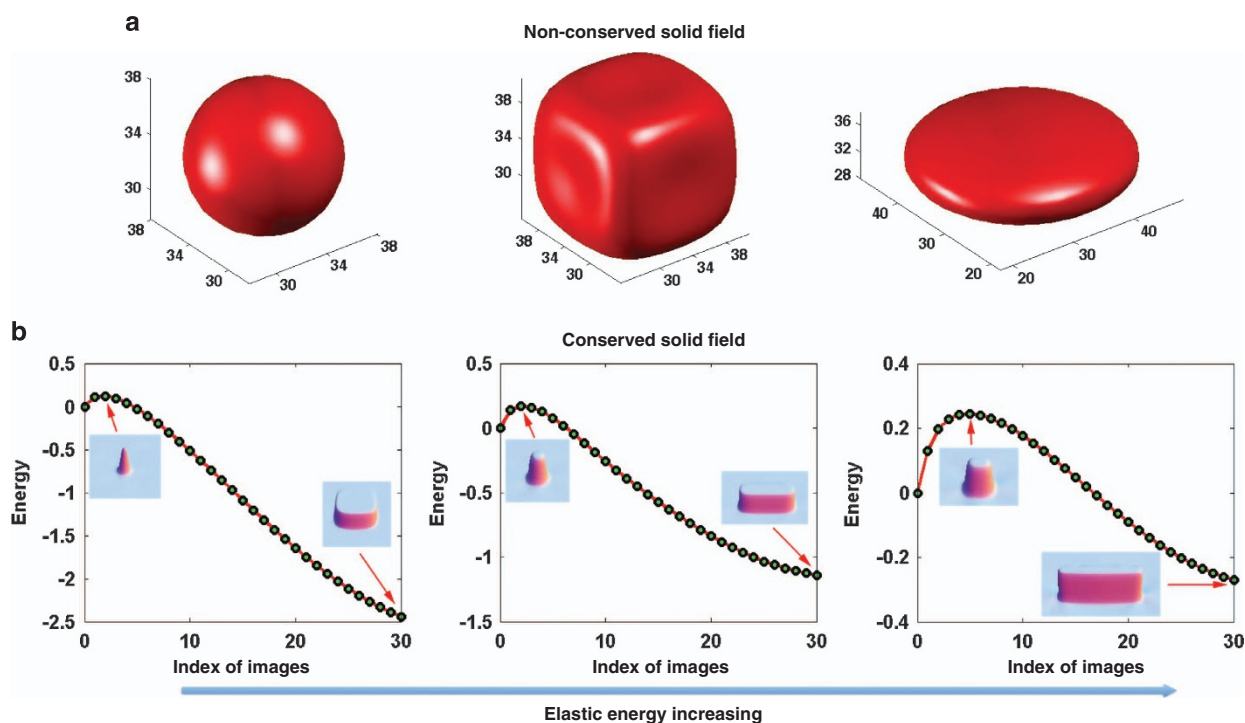


Figure 2. Critical nuclei with increasing of elastic energy in non-conserved solid field (a) and conserved solid field (b). Figure reproduced from refs 33,34. Permission has been granted by the publisher to use the original figures.

Transition pathways in micromagnetics

Submicron-sized magnetic elements have found a wide range of applications in science and technology, particularly as storage devices. As the elements get smaller, the effect of thermal noise and the issue of data retention time become a concern. In ref. 3, the string method was applied to study thermally activated switching and the energy landscape of submicron-sized magnetic elements. The free energy of the system is modelled by the Landau–Lifshitz functional:

$$F[\mathbf{M}] = \frac{1}{2} \int_{\Omega} \left\{ \frac{C_{\text{ex}}}{M_s^2} |\nabla \mathbf{M}|^2 + \phi \left(\frac{\mathbf{M}}{M_s} \right) \right\} dx \quad (25)$$

$$- \frac{1}{2} \int_{\Omega} 2\mu_0 \mathbf{H}_e \cdot \mathbf{M} dx + \frac{\mu_0}{2} \int_{R^3} |\nabla U|^2 dx, \quad (26)$$

where \mathbf{M} is the magnetisation distribution normalised so that $|\mathbf{M}| = 1$, Ω is the domain occupied by the element, $C_{\text{ex}} |\nabla \mathbf{M}|^2 / M_s^2$ is the exchange interaction energy the the spins, $\phi(\mathbf{M}/M_s)$ is the energy due to material anisotropy, $-2\mu_0 \mathbf{H}_e \cdot \mathbf{M}$ is the energy due to the external applied field. The last term is the energy due to the field induced by the magnetisation distribution inside the material. This induced field $-\nabla U$ can be computed by solving

$$\Delta U = \begin{cases} \nabla \cdot \mathbf{M}, & \text{in } \Omega, \\ 0, & \text{outside } \Omega, \end{cases} \quad (27)$$

with the jump conditions

$$[U] = 0, \quad \left[\frac{\partial U}{\partial \nu} \right] = -\mathbf{M} \cdot \nu, \quad (28)$$

at the boundary of the element. In the above conditions, $[\cdot]$ denotes the jump across the boundary, and the vector ν is the outward unit normal on the boundary of Ω .

The string method was used to compute the MEPs and analyse the energy landscape of the Landau–Lifshitz free energy. It is found that the switching proceeds by two generic scenarios: Domain-wall propagation and reconnection followed by edge

domain switching, or vortex nucleation at the boundary followed by vortex propagation through the element. These two pathways are shown in Figure 3. The second pathway is preferred for thicker films, whereas the first is preferred for thin films.

Nucleation in solid–liquid phase transition

Atomic processes associated with complex systems often exhibit a variety of time scales. The origins of this disparity in time scales can often be traced to the spatial scale associated with different process. Localised processes involve small subsets of the total degrees of freedom (DOF) of the system while collective processes involve large number of DOF; the latter typically occur at a slower rate. However, when collective processes are initiated through a series of one or more localised events, then the processes are inherently multiscale in nature. An important example of type of physical process is the melting of a solid.^{5,74–78}

As a model system, copper is used to study the melting of a solid. The interatomic interactions were modelled using the embedded atom method potential for Cu developed by Mishin *et al.*⁷⁹ To efficiently explore the relevant parts of the configuration space and study the microscopic mechanisms of melting, the AFED,^{80–82} a recently developed exploration technique, was used. Techniques like the AFED and the TAMD^{82,83} exploit the adiabatic time scale separation between the evolution of the atomic DOF and the collective variables (CVs) by assigning fictitious masses to the CVs that are much larger than the atomic masses, and at the same time, maintaining the CVs at a temperature much higher than the physical temperature. We have used the volume (V) of the system and the Steinhardt orientation order parameters Q_4 , Q_6 as the collective variables. The volume of the system captures the changes in the density, whereas the orientation order parameters capture the changes in the symmetry of the atomic structure.

The dynamic sampling of the FES in AFED/TAMD schemes is obtained by using the equations of motions obtained from the potential-energy surface $U_k(x, z)$ that spans over an extended space of atomic and coarse-grained DOF. The TAMD scheme

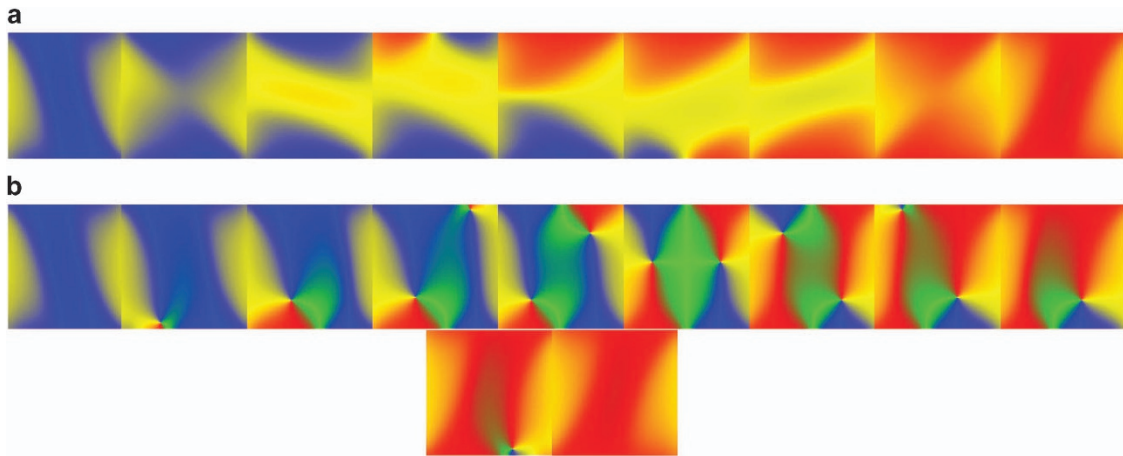


Figure 3. Two generic pathways (MEPs) (a) and (b) followed by the magnetisation during the switching of submicron-sized ferromagnetic elements. The figures show the successive minima and saddle points. The colour code represents the in-plane components of the magnetisation: blue = right, red = left, yellow = up, green = down. Figure modified from ref. 24. Permission has been granted by the publisher to use the original figures.

within the isobaric–isothermal conditions was used to sample the FES of Cu at different physical temperatures.⁸² From these sampling trajectories, a set of CVs $\{\mathbf{z}_k\}_{k=0}^m$ was selected to reconstruct the Gibbs FES. The mean forces $\{\mathbf{f}_k\}_{k=0}^m$ on these CVs were obtained from an average of the instantaneous forces on these CVs along the sampling trajectories.

The reconstruction of the FES, using the knowledge of the mean forces at the set of CVs selected above, is essentially an inverse problem. The free-energy was expressed as a superposition of Gaussian RBFs placed at the chosen centers.⁵⁸

$$\mathcal{F}(\mathbf{z}) = \sum_{k=1}^m a_k \phi_\sigma(|\mathbf{z} - \mathbf{z}_k|^2), \quad (29)$$

where $|\cdot|$ is the L_2 norm, $\phi_\sigma(r) = e^{-r/2\sigma^2}$ and σ is the width of the Gaussian RBFs. The optimal values of the coefficients a_k were obtained by optimising the following cost function

$$\mathcal{E}_r = \frac{1}{m} \sum_{k=1}^m |\nabla \mathcal{F}(\mathbf{z}_k) + \mathbf{f}_k|^2 + \lambda |\mathbf{a}|^2. \quad (30)$$

Here, λ is a regularisation parameter and $\lambda |\mathbf{a}|^2$ is a smoothness constraint introduced to stabilise the solution. From a statistical point of view, a proper handling of noisy data entails proper training as well as testing of the fitting model. To this end, multifold cross validation is used for model validation.⁸⁵

The FES of Copper calculated close to the melting temperature, shown in Figure 4, illustrates the multitude of barriers that a system has to overcome to travel from the solid super-basin to the liquid basin. Some of the barriers inside the solid basin are associated with the formation of isolated point defects and are on the order of 1–5 eV. In contrast, the barriers to cross from these isolated point defect states to metastable states corresponding to extended defects like defect cluster, are on the order of 20 eV. The barrier to melting, on the other hand, is on the order of 120 eV. This disparity in the Gibbs free-energy barriers results in time scale disparities between the processes. In addition, the presence of a large number of metastable states inside the solid basin suggests that contrary to the tenets of the classical nucleation theory melting is not a simple, single thermally activated barrier crossing event but involves multiple thermally activated transition events.⁵

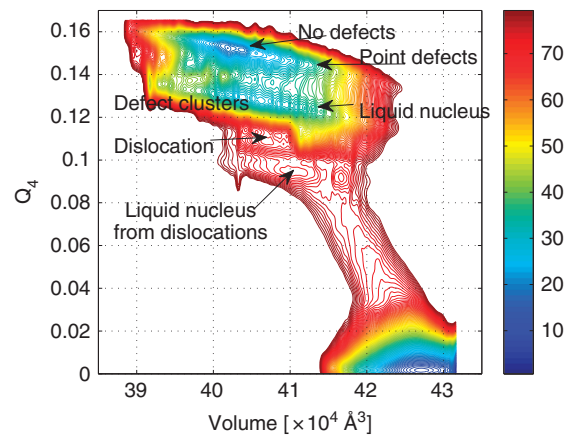


Figure 4. The free-energy surface of copper close to the melting temperature as a function of the collective variables—volume and Steinhardt orientation order parameter. The presence of multiple locally stable states illustrates that melting is a complex thermally activated transition event. Figure reproduced from ref. 5. Permission has been granted by the publisher to use the original figures.

SUMMARY AND OUTLOOK

Nucleation is a complex multiscale problem. Recent development of new numerical algorithms and modelling approaches on MEP calculation and saddle point search as well as transition path theory have brought new light to this challenging subject. The computational methods reviewed in ‘Computing saddle points and transition paths’ and ‘Complex energy landscapes’ are generic methods. We classify them according to the different purposes they serve, such as saddle point search, computing MEPs/MFEPs or transition tubes. Subject to the properties and structures of the problems, one is free to use different methods to study the same nucleation event. The examples in ‘Applications in materials science’ provided some illustrative applications of these new approaches.

There are naturally many other relevant and important issues that are not addressed here. For example, it is an interesting subject to see how global optimisation techniques may be adopted to efficiently identify saddle points with a low and/or minimum energy barrier. For complex models in high dimensions, it is worthwhile to mention the importance of data analysis techniques to our understanding of such systems. As it is difficult

to *a priori* predict the relevant order parameters, one may combine FES exploration techniques in conjunction with dimension reduction techniques like diffusion maps⁸⁴ to obtain quantitative information about the relevant low-dimensional manifolds.

Furthermore, metastable states, saddle points, obtained from GAD or SDD or dimer like methods, in the space of collective variables, can be used to construct a weighted graph that is representative of the transition events taking place in a system. Such a network can have its own trapping regions, critical saddles and so on. One may also explore how to utilise the information encoded in such a network to conduct multiscale dynamic simulation of microstructure evolution that involves fast nucleation and slow coarsening processes. Preliminary studies along this direction can be found in ref. 71 These are only a few examples of the wide spectrum of multiscale analysis that is possible from the huge trove of that is obtained from saddle-point/MEP search and FES sampling schemes presented here.

ACKNOWLEDGEMENTS

We thank Professors Long-Qing Chen, Weinan E and Eric Vanden-Eijnden for fruitful discussions and collaborations on the subject. The work by L.Z. was supported by China NSFC No. 11421110001 and 91430217. The work of W.R. was partially supported by AcRF Tier-1 grant R-146-000-216-112. The work by A.S. was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. The work of Q.D. was supported in part by NSF-DMS1318586.

COMPETING INTERESTS

The authors declare no conflict of interest.

REFERENCES

- Materials Genome Initiative for Global Competitiveness. *National Science and Technology Council, Office of Science and Technology Policy*. (Washington DC, 2011).
- Cheng, X., Lin, L., E, W., Zhang, P. & Shi, A.-C. Nucleation of ordered phases in block copolymers. *Phys. Rev. Lett.* **104**, 148301 (2010).
- E, W., Ren, W. & Vanden-Eijnden, E. Energy landscape and thermally activated switching of submicron-sized ferromagnetic elements. *J. Appl. Phys.* **93**, 2275–2282 (2003).
- Li, T., Zhang, P. & Zhang, W. Nucleation rate calculations for the phase transition of diblock copolymers under stochastic Cahn-Hilliard dynamics. *SIAM Multi. Model. Simul.* **11**, 385–409 (2013).
- Samanta, A., Tuckerman, M. E., Yu, T.-Q. & E, W. Microscopic mechanisms of equilibrium melting of a solid. *Science* **346**, 729–732 (2014).
- Schlegel, H. Exploring potential energy surfaces for chemical reactions: an overview of some practical methods. *J. Comput. Chem.* **24**, 1514–1527 (2003).
- Wales, D. Energy landscapes: calculating pathways and rates. *Int. Rev. Phys. Chem.* **25**, 237–282 (2006).
- Zhang, L., Chen, L.-Q. & Du, Q. Morphology of critical nuclei in solid state phase transformations. *Phys. Rev. Lett.* **98**, 265703 (2007).
- Zhang, W., Li, T. & Zhang, P. Numerical study for the nucleation of one-dimensional stochastic Cahn-Hilliard dynamics. *Commun. Math. Sci.* **10**, 1105–1132 (2012).
- W. E. & Vanden-Eijnden, E. Transition-path theory and path-finding algorithms for the study of rare events. *Annu. Rev. Phys. Chem.* **61**, 391–420 (2010).
- Laaksonen, A., Talanquer, V. & Oxtoby, D. W. Nucleation: measurements, theory, and atmospheric applications. *Annu. Rev. Phys. Chem.* **46**, 489–524 (1995).
- Xu, X., Ting, C. L., Kusaka, I. & Wang, Z. G. Nucleation in polymers and soft matter. *Annu. Rev. Phys. Chem.* **65**, 449–475 (2014).
- Henkelman, G. & Jónsson, H. A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives. *J. Chem. Phys.* **111**, 7010 (1999).
- Olsen, R., Kroes, G., Henkelman, G., Arnaldsson, A. & Jónsson, H. Comparison of methods for finding saddle points without knowledge of the final states. *J. Chem. Phys.* **121**, 9776–9792 (2004).
- Crippen, G. & Scheraga, H. Minimization of polypeptide energy XI. The method of gentlest ascent. *Arch. Biochem. Biophys.* **144**, 462–466 (1971).
- E, W. & Zhou, X. The gentlest ascent dynamics. *Nonlinearity* **24**, 18311842 (2011).
- Gao, W., Leng, J. & Zhou, X. An iterative minimization formulation for saddle-point search. *SIAM J. Numer. Anal.* **53**, 1786–1805 (2015).
- Samanta, A., Chen, M., Yu, T. Q., Tuckerman, M. & E, W. Sampling saddle points on a free energy surface. *J. Chem. Phys.* **140**, 164109 (2014).
- Zhang, J. Y. & Du, Q. Shrinking dimer dynamics and its applications to saddle point search. *SIAM J. Numer. Anal.* **50**, 1899–1921 (2012).
- Henkelman, G. & Jónsson, H. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *J. Chem. Phys.* **113**, 9978–9985 (2000).
- Henkelman, G., Uberuaga, B. P. & Jónsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **113**, 9901–9904 (2000).
- Carilli, M. F., Delaney, K. T. & Fredrickson, G. H. Truncation-based energy weighting string method for efficiently resolving small energy barriers. *J. Chem. Phys.* **143**, 054105 (2015).
- Du, Q. & Zhang, L. A constrained string method and its numerical analysis. *Commun. Math. Sci.* **7**, 1039–1051 (2009).
- E, W., Ren, W. & Vanden-Eijnden, E. String method for the study of rare events. *Phys. Rev. B.* **66**, 052301 (2002).
- E, W., Ren, W. & Vanden-Eijnden, E. Simplified and improved string method for computing the minimum energy paths in barrier-crossing events. *J. Chem. Phys.* **126**, 164103 (2007).
- Ren, W. & Vanden-Eijnden, E. A climbing string method for saddle point search. *J. Chem. Phys.* **138**, 134105 (2013).
- Samanta, A. & E, W. Optimization-based string method for finding minimum energy path. *Commun. Comput. Phys.* **14**, 265–275 (2013).
- E, W., Ren, W. & Vanden-Eijnden, E. Transition pathways in complex systems: reaction coordinates, isocommittor surface, and transition tubes. *Chem. Phys. Lett.* **413**, 242–247 (2005).
- E, W. & Vanden-Eijnden, E. Towards a theory of transition paths. *J. Stat. Phys.* **123**, 503–523 (2006).
- Ren, W., Vanden-Eijnden, E., Maragakis, P. & E, W. Transition pathways in complex systems: application of the finite-temperature string method to the alanine dipeptide. *J. Chem. Phys.* **123**, 134109 (2005).
- Maragliano, L., Fischer, A., Vanden-Eijnden, E. & Ciccotti, G. String method in collective variables: minimum free energy paths and isocommittor surfaces. *J. Chem. Phys.* **125**, 024106 (2006).
- Maragliano, L. & Vanden-Eijnden, E. On-the-fly string method for minimum free energy paths calculation. *Chem. Phys. Lett.* **446**, 182–190 (2007).
- Zhang, L., Chen, L.-Q. & Du, Q. Diffuse-interface description of strain-dominated morphology of critical nuclei in phase transformations. *Acta Mater.* **56**, 3568–3576 (2008).
- Zhang, L., Chen, L.-Q. & Du, Q. Simultaneous prediction of morphologies of a critical nucleus and an equilibrium precipitate in solids. *Commun. Comput. Phys.* **7**, 674–682 (2010).
- Eyring, H. The activated complex and the absolute rate of chemical reactions. *Chem. Rev.* **17**, 6577 (1935).
- Wigner, E. The transition state method. *Trans. Farad. Soc.* **34**, 29–41 (1938).
- Zhang, J. Y. & Du, Q. Constrained shrinking dimer dynamics for saddle point search with constraints. *J. Comput. Phys.* **231**, 4745–4758 (2012).
- Cerjan, C. J. & Miller, W. H. On finding transition states. *J. Chem. Phys.* **75**, 2800–2806 (1981).
- Li, Y. & Zhou, J. A minimax method for finding multiple critical points and its applications to semilinear PDEs. *SIAM J. Sci. Comput.* **23**, 840–865 (2001).
- Rabinowitz, P. *Minimax Methods in Critical Point Theory with Applications to Differential Equations*. (American Mathematical Society, 1986).
- Cances, E., Legoll, F., Marinica, M.-C., Minoukadeh, K. & Willaime, F. Some improvements of the activation-relaxation technique method for finding transition pathways on potential energy surfaces. *J. Chem. Phys.* **130**, 114711 (2009).
- Machado-Charry, E. *et al.* Optimized energy landscape exploration using the ab initio based activation-relaxation technique. *J. Chem. Phys.* **135**, 034102 (2011).
- Mousseau, N. & Barkema, G. T. Traveling through potential energy landscapes of disordered materials: The activation-relaxation technique. *Phys. Rev. E.* **57**, 2419–2424 (1998).
- Miron, R. & Fichtorn, K. The step and slide method for finding saddle points on multi-dimensional potential surfaces. *J. Chem. Phys.* **115**, 8742–8750 (2001).
- Gould, N., Ortner, C. & Packwood, D. An Efficient Dimer Method With Preconditioning And Linesearch. Preprint at <http://arxiv.org/abs/1407.2817> (2014).
- Zeng, Y., Xiao, P. & Henkelman, G. Unification of algorithms for minimum mode optimization. *J. Chem. Phys.* **140**, 044115 (2014).
- Kastner, J. & Sherwood, P. Superlinearly converging dimer method for transition state search. *J. Chem. Phys.* **128**, 014106 (2008).

48. Zhang, L., Zhang, J. Y. & Du, Q. Finding critical nuclei in phase transformations by shrinking dimer dynamics and its variants. *Commun. Comput. Phys.* **16**, 781–798 (2014).
49. Zhang, L., Du, Q. & Zheng, Z. Optimization-based shrinking dimer method for finding transition states. *SIAM J. Sci. Comput.* (in the press).
50. Ren, W. Higher order string method for finding minimum energy paths. *Comm. Math. Sci.* **1**, 377–384 (2003).
51. Jónsson, H., Mills, G. & Jacobsen, K. W. in *Classical and Quantum Dynamics in Condensed Phase Simulations*, (eds Berne B. J., Ciccotti G. & Coker D. F.) (World Scientific, 1998).
52. Elber, R. & Karplus, M. A method for determining reaction paths in large molecules: application to myoglobin. *Chem. Phys. Lett.* **139**, 375 (1987).
53. Gillilan, R. E. & Lilien, R. H. Optimization and dynamics of protein-protein complexes using b-splines. *J. Comput. Chem.* **25**, 1630 (2004).
54. Ulitsky, A. & Elber, R. A new technique to calculate steepest descent paths in flexible polyatomic systems. *J. Chem. Phys.* **92**, 1510 (1990).
55. W. E., Ren, W. & Vanden-Eijnden, E. Finite temperature string method for the study of rare events. *J. Phys. Chem. B.* **109**, 6688–6693 (2005).
56. Vanden-Eijnden, E. & Venturoli, M. Revisiting the finite-temperature string method for the calculation of reaction tubes and free energies. *J. Chem. Phys.* **130**, 194103 (2009).
57. Du, Q., Faber, V. & Gunzburger, M. Centroidal Voronoi tessellations: applications and algorithms. *SIAM Rev.* **41**, 637–676 (1999).
58. Maragliano, L. & Vanden-Eijnden, E. Single-sweep methods for free energy calculations. *J. Chem. Phys.* **128**, 184110 (2008).
59. Cahn, J. & Hilliard, J. Free energy of a nonuniform system. III. Nucleation in a two-component incompressible fluid. *J. Chem. Phys.* **31**, 688–699 (1959).
60. Chen, L.-Q. Phase-field models for microstructure evolution. *Annu. Rev. Mater. Res.* **32**, 113–140 (2002).
61. Khachatryan, A. G. *Theory of Structural Transformations in Solids*. (Wiley, 1983).
62. Zhang, L., Chen, L. Q. & Du, Q. Mathematical and numerical aspects of phase-field approach to critical morphology in solids. *J. Sci. Comput.* **37**, 89–102 (2008).
63. Zhang, L., Chen, L.-Q. & Du, Q. Diffuse-interface approach to predicting morphologies of critical nucleus and equilibrium structure for cubic to tetragonal transformations. *J. Comput. Phys.* **229**, 6574–6584 (2010).
64. Gránásy, L., Pusztai, T., Saylor, D. & Warren, J. A. Phase field theory of heterogeneous crystal nucleation. *Phys. Rev. Lett.* **98**, 035703 (2007).
65. Laurila, T., Carlson, A., Do-Quang, M., Ala-Nissila, T. & Amberg, G. Thermo-hydrodynamics of boiling in a van der Waals fluid. *Phys. Rev. E* **85**, 026320 (2012).
66. Backofen, R. & Voigt, A. A phase-field-crystal approach to critical nuclei. *J. Phys. Condens. Matter* **22**, 364104 (2010).
67. Backofen, R. & Voigt, A. A phase field crystal study of heterogeneous nucleation—application of the string method. *Eur. Phys. J. Special Topics* **223**, 497–509 (2014).
68. Gránásy, L., Podmaniczky, F., Tóth, G. I., Tegze, G. & Pusztai, T. Heterogeneous nucleation of/on nanoparticles: a density functional study using the phase-field crystal model. *Chem. Soc. Rev.* **43**, 2159–2173 (2014).
69. Elder, K. R., Drolet, F., Kosterlitz, J. M. & Grant, M. Stochastic eutectic growth. *Phys. Rev. Lett.* **72**, 677 (1994).
70. Gránásy, L. *et al.* Phase-field modeling of polycrystalline solidification: from needle crystals to spherulites: a review. *Metall. Mater. Trans. A* **45**, 1694–1719 (2014).
71. Heo, T., Zhang, L., Du, Q. & Chen, L.-Q. Incorporating diffuse-interface nuclei in phase-field simulations. *Scripta Mater.* **63**, 8–11 (2010).
72. Roy, A., Rickman, J. M., Gunton, J. D. & Elder, K. R. Simulation study of nucleation in a phase-field model with nonlocal interactions. *Phys. Rev. E* **57**, 2610–2617 (1998).
73. Li, Y., Hu, S., Zhang, L. & Sun, X. Non-classical nuclei and growth kinetics of Cr precipitates in FeCr alloys during aging. *Model. Simul. Mater. Sci. Eng.* **22**, 025002 (2014).
74. Brillouin, L. On thermal dependence of elasticity in solids. *Physical Review* **54**, 916–917 (1938).
75. Cahn, R. W. Crystal defects and melting. *Nature* **273**, 491–492 (1978).
76. Gorecki, T. Vacancies and changes of physical properties of metals at the melting point. *Z. Metallk.* **65**, 426–431 (1974).
77. Lindemann, F. A. The calculation of molecular natural frequencies. *Phys. Z.* **11**, 609–612 (1910).
78. Mott, N. F. Theories of the liquid state. *Nature* **145**, 801–802 (1940).
79. Mishin, Y., Mehl, M. J., Papaconstantopoulos, D. A., Voter, A. F. & Kress, J. D. Structural stability and lattice defects in copper: Ab initio, tight-binding, and embedded-atom calculations. *Phys. Rev. B* **63**, 224106 (2001).
80. Abrams, J. B. & Tuckerman, M. E. Efficient and direct generation of multi-dimensional free energy surfaces via adiabatic dynamics without coordinate transformations. *J. Phys. Chem. B* **112**, 15742 (2008).
81. Rosso, L., Mináry, P., Zhu, Z. & Tuckerman, M. On the use of the adiabatic molecular dynamics technique in the calculation of free energy profiles. *J. Chem. Phys.* **116**, 4389–4402 (2002).
82. Yu, T. Q. *et al.* Order-parameter-aided temperature-accelerated sampling for the exploration of crystal polymorphism and solid-liquid phase transitions. *J. Chem. Phys.* **140**, 214109 (2014).
83. Maragliano, L. & Vanden-Eijnden, E. A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations. *Chem. Phys. Lett.* **426**, 168–175 (2006).
84. Coifman, R. R. *et al.* Geometric diffusion as a tool for harmonic analysis and structure definition of data, part I: diffusion maps. *Proc. Natl Acad. Sci. USA* **102**, 7426–7431 (2005).
85. Stone, M. Cross-validatory choice and assessment of statistical predictions. *J. R. Stat. Soc. B* **36**, 111–147 (1974).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>