

Simplified gentlest ascent dynamics for saddle points in non-gradient systems

Shuting Gu, and Xiang Zhou

Citation: *Chaos* **28**, 123106 (2018); doi: 10.1063/1.5046819

View online: <https://doi.org/10.1063/1.5046819>

View Table of Contents: <http://aip.scitation.org/toc/cha/28/12>

Published by the [American Institute of Physics](#)



Don't let your writing
keep you from getting
published!

AIP | Author Services

Learn more today!

Simplified gentlest ascent dynamics for saddle points in non-gradient systems

Shuting Gu^{1,a)} and Xiang Zhou^{2,b)}

¹*School of Mathematical Sciences, South China Normal University, Guangzhou 510631, People's Republic of China*

²*School of Data Science and Department of Mathematics, City University of Hong Kong, Tat Chee Ave., Kowloon, Hong Kong*

(Received 2 July 2018; accepted 15 November 2018; published online 5 December 2018)

The gentlest ascent dynamics (GAD) [W. E and X. Zhou, *Nonlinearity* **24**, 1831 (2011)] is a time continuous dynamics to efficiently locate saddle points with a given index by coupling the position and direction variables together. These saddle points play important roles in the activated process of randomly perturbed dynamical systems. For index-1 saddle points in non-gradient systems, the GAD requires two direction variables to approximate, respectively, the eigenvectors of the Jacobian matrix and its transposed matrix. In the particular case of gradient systems, the two direction variables are equal to the single minimum mode of the Hessian matrix. In this note, we present a simplified GAD which only needs one direction variable even for non-gradient systems. This new method not only reduces the computational cost for the direction variable by half but also avoids inconvenient transpose operation of the Jacobian matrix. We prove the same convergence property for the simplified GAD as that for the original GAD. The motivation of our simplified GAD is the formal analogy with Hamilton's equation governing the noise-induced exit dynamics. Several non-gradient examples are presented to demonstrate our method, including a two dimensional model and the Allen-Cahn equation in the presence of shear flow. *Published by AIP Publishing.* <https://doi.org/10.1063/1.5046819>

Saddle point plays the role of “transition state” in the dynamical systems perturbed by small noise. The exit path driven by noise typically crosses a transition state and the dynamics of the path is governed by a Hamilton's equation which arises from the large deviation theory. The gentlest ascent dynamics, as a saddle point search dynamics, solely relies on the local Jacobian or Hessian mode. Our work discovers a new and simplified gentlest ascent dynamics when exploring the analogy between the existing method and Hamilton's equation. The new method reduces the computational costs by half as an economic tool for locating saddle points, which is helpful for the metastability study in non-gradient systems. The application to the Allen-Cahn equation in the presence of shear flow discovers a variety of transition states.

on the most probable transition path between different local wells. The steepest descent flow that minimizes the potential energy gives rise to a gradient dynamical system. For such gradient systems, a large amount of numerical methods have been developed to locate their saddle points, such as the eigenvector-following method,¹ the dimer method,¹⁰ the gentlest ascent dynamics (GAD),^{5,17} the iterative minimization algorithm,^{7,8} and others.^{13,21} To pinpoint the saddle point on the basin boundary of a prescribed local well, the climbing string method¹⁵ works robustly due to the synergy between the string method² in the path space and the single state dynamics of the endpoint in the configuration space.

However, most of these methods were designed for the gradient systems, and few of them are applicable to the non-gradient systems. There are many important non-gradient models arising from biology, fluid dynamics,^{19,20} and so on. One prominent example is the phase field model^{11,14} such as the Allen-Cahn equation associated with a double-well potential, but subject to the influence of shear flow. The extra forcing from the fluid certainly turns the system into a non-gradient model. The saddle points in such non-gradient systems are still of great importance since they are also relevant to the non-equilibrium processes in the randomly perturbed dynamical systems.⁶

Among many saddle search methods mentioned previously, only the GAD⁵ proposed by one of the authors in this note is capable to address the saddle point in general dynamical systems by using two eigenvectors and oblique projection. This result extends the saddle point search method to the non-gradient systems. In this note, we present a new form of the gentlest ascent dynamics associated with the following

I. INTRODUCTION

Locating the saddle points in a dynamical system has been of broad interest in many fields of scientific research, especially in the understanding of the exit process from linearly stable states when the system is randomly perturbed by small noise. In computational chemistry,¹⁸ one of the most important objects on the potential energy surface is the transition state, a special type of the saddle point with index 1 which is defined as the one with exactly one dimensional unstable manifold. Such transition states are the bottlenecks

^{a)}Electronic mail: shutingu@m.scnu.edu.cn

^{b)}Author to whom correspondence should be addressed: xiang.zhou@cityu.edu.hk

non-gradient system

$$\dot{x} = b(x), \quad (1)$$

where b is a smooth vector field in \mathbb{R}^d . We are interested in the index-1 saddle point of the vector field b . To achieve this, the original GAD⁵ generates a position variable x and two direction variables v and w so that the linearly stable states of this new dynamics are index-1 saddle points of b . The dynamics of v and w in the GAD need the multiplication of the Jacobian matrix $Db(x)$ with v and the multiplication of the transpose $Db(x)^\top$ with w , respectively. The matrix-vector multiplication $Db(x)v = \lim_{h \rightarrow 0} [b(x+hv) - b(x)]/h$ can be easily approximated by the finite difference method. The difficulty comes from the calculation of the transposed term $Db(x)^\top w$, which lacks the interpretation of the directional derivative. In our simplified GAD below, we shall show that it suffices to use the dynamics of *one* direction variable (either v or w), without affecting the convergence property of the original GAD.

Despite the simple form of our result, we find an interesting connection between the simplified GAD and the underlying Hamilton's equation. Hamilton's equation describes the optimal transition path in the randomly perturbed equation

$$dX = b(X)dt + \sqrt{\varepsilon}dW_t, \quad (2)$$

where W_t is the standard Brownian motion and ε is a small constant. In fact, the study of rare events in Eq. (2) is one important motivation to study the saddle points of the vector field b . By the Freidlin-Wentzell large deviation theory,⁶ the most probable transition path is a minimizer of the Freidlin-Wentzell action functional and this path, as a function of time, satisfies Hamilton's equation with the zero Hamiltonian. The position and momentum in Hamilton's equation might be thought as the counterpart of the position and direction in the GAD. Our original inspiration is to understand this formal analogy and the result obtained in the end is the simplified GAD.

The rest of the paper is organized as follows. In Sec. II, we propose the simplified GAD for non-gradient systems after a short review of the GAD. Then, we apply the simplified GAD to the multiscale model of non-gradient slow-fast systems. Section III is our numerical examples. In particular, we study the Allen-Cahn equation in the presence of shear flow and investigate how the shear rate affects the transition states in this system. The conclusions and discussions are given in Sec. IV. We show the relation between the simplified GAD and Hamilton's dynamics in Appendixes A and B.

II. METHOD

A. Review of gentlest ascent dynamics (GAD)

The GAD⁵ for the flow $\dot{x}(t) = b(x)$ involves a position variable x and two direction variables v and w as follows:

$$\begin{cases} \dot{x}(t) = b(x) - 2 \frac{\langle b(x), w \rangle}{\langle w, v \rangle} v, & (3a) \\ \gamma \dot{v}(t) = J(x)v - \alpha v, & (3b) \\ \gamma \dot{w}(t) = J(x)^\top w - \beta w, & (3c) \end{cases}$$

where $J(x) = Db(x)$ is the Jacobian matrix $(Db)_{ij} \doteq \frac{\partial b_i}{\partial x_j}$, which is generally asymmetric. α and β are the Lagrangian multipliers to impose certain normalization conditions for v and w . For instance, if the normalization condition is $\langle v, v \rangle \equiv \langle w, v \rangle \equiv 1$, then $\alpha = \langle v, J(x)v \rangle$ and $\beta = 2 \langle w, J(x)v \rangle - \alpha$. Equation (3) is a flow in \mathbb{R}^{3d} .

As a special case, the GAD for a gradient system $\dot{x}(t) = -\nabla V(x)$ only involves v :

$$\begin{cases} \dot{x}(t) = -\nabla V(x) + 2 \frac{\langle \nabla V(x), v \rangle}{\langle v, v \rangle} v, & (4a) \\ \gamma \dot{v}(t) = -\nabla^2 V(x)v + \langle v, \nabla^2 V(x)v \rangle v, & (4b) \end{cases}$$

where $\gamma > 0$ is the relaxation parameter. A large γ means a fast relaxation for the direction variable $v(t)$ toward the steady state. For a frozen x , this steady state v is the min mode of the Hessian matrix $\nabla^2 V(x)$: the eigenvector corresponding to the smallest eigenvalue of $\nabla^2 V(x)$.

It is proved that the above GAD [the general form (3) and the gradient form (4)] has the property that its stable critical point corresponds to an index-1 saddle point of the original dynamics, $\dot{x} = b(x)$ or $\dot{x} = -\nabla V(x)$. Our simplified GAD has the exactly same property, which will be given below in detail.

In the GAD (3) for non-gradient systems, both $J(x)v$ and $J(x)^\top w$ in (3b) and (3c) must be calculated. One can apply the finite difference scheme to compute the matrix-vector multiplication $J(x)v$. But this trick cannot be applied to the term $J(x)^\top w$. $J(x)^\top$ can only be computed by a numerical transpose operation in most cases. Then, the computation of $J(x)^\top w$ imposes a severe challenge for large scale problems.

B. Simplified GAD

Our new GAD takes *one* of the following two forms (not simultaneously)

$$\begin{cases} \dot{x} = b(x) - 2 \langle b(x), v(t) \rangle v(t) / \|v(t)\|^2, & (5a) \\ \dot{v} = J(x)v - \langle v, Jv \rangle v & (5b) \end{cases}$$

or

$$\begin{cases} \dot{x} = b(x) - 2 \langle b(x), w(t) \rangle w(t) / \|w(t)\|^2, & (6a) \\ \dot{w} = J^\top(x)w - \langle w, J^\top w \rangle w. & (6b) \end{cases}$$

Hence, the simplified GAD is a flow in \mathbb{R}^{2d} in both forms. Initially, $\|v_0\| = 1$ or $\|w_0\| = 1$ so that v and w are always unit vectors. The difference between (5) and (6) is the matrix-vector multiplication $J(x)v$ or $J(x)^\top w$. As discussed above, to avoid computing $J(x)^\top w$, one prefers Eq. (5) practice. It will be seen later that in theory, Eq. (6) may be of more interest. For the gradient system $\dot{x}(t) = -\nabla V(x)$, $J = -H$, where $H = \nabla^2 V = H^\top$ is the Hessian matrix, the above two forms are identical and they become the GAD (4) for the gradient system.

The simplified GAD (5) or (6) converges to the index-1 saddle point of the original dynamics $\dot{x} = b(x)$; see the following theorem.

Theorem 1. (a) If (x_*, v_*) is a fixed point of the simplified GAD (5) and v_* is the normalized vector, $\|v_*\| = 1$, then, v_* is the eigenvector of $J(x_*)$ corresponding to an eigenvalue λ_* ,

i.e.,

$$J(x_*)v_* = \lambda_*v_*$$

and x_* is a fixed point of the original dynamical system, i.e., $b(x_*) = 0$;

(b) Let x_s be a fixed point of the dynamical system $\dot{x} = b(x)$. If the Jacobian matrix $J(x_s)$ has n distinct real eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ corresponding to the n linearly independent eigenvectors v_i , i.e.,

$$J(x_s)v_i = \lambda_i v_i, i = 1, 2, \dots, n$$

and $\|v_i\| = 1, \forall i$. Then $(x_s, v_i), \forall i$, is a fixed point of the simplified GAD (5). Furthermore, there is one fixed point (x_s, v_i) among these n fixed points, which is linearly stable if and only if x_s is an index-1 saddle point of the original dynamical system $\dot{x} = b(x)$ and the eigenvalue λ_{v_i} corresponding to v_i is the only positive eigenvalue of $J(x_s)$.

The proof is quite similar to that for the original GAD⁵ and is presented in Appendix A. The method of proof in both cases is the linear analysis at the fixed point by analytically computing the eigenvalues and eigenvectors. The only difference is that the simplified GAD only needs to handle the $2d \times 2d$ matrix. In fact, the simplified GAD and the original GAD share the same eigenvalues, but the original GAD has some eigenvalues with the multiplicity 2.

Remark 1. A positive constant τ can be used in the simplified GAD as in the GAD (3) or (4): $\dot{v} \rightarrow \tau \dot{v}$ [or $\dot{w} \rightarrow \tau \dot{w}$ as in Eq. (3)] to represent the time scale ratio between x and v (or w). τ controls the convergence speed of the system. It provides the flexibility of possible numerical acceleration by adaptively adjusting its value in practice. Since we do not pursue this point further in this note, we drop this factor to ease the presentation.

Finally, we remark on the relation of the simplified GAD (6) with Hamilton’s equation arising from the Freidlin–Wentzell large deviation theory. The full discussion on this topic is placed in Appendix B for interested readers.

C. Application to multiscale model

The idea of GAD has been extended to the slow-fast stochastic system and is called MsGAD.⁹ As a corollary of our result, the simplified GAD here can be applied to the same multi-scale model straightforwardly. For the background and more details, the reader can refer to the MsGAD paper.⁹ We here directly present the scheme based on the above simplified GAD. The slow-fast system in consideration is

$$\begin{cases} \dot{X}^\varepsilon(t) = f(X^\varepsilon, Y^\varepsilon), & (7a) \\ \dot{Y}^\varepsilon(t) = \frac{1}{\varepsilon}b(X^\varepsilon, Y^\varepsilon) + \frac{1}{\sqrt{\varepsilon}}\sigma(X^\varepsilon, Y^\varepsilon)\eta(t), & (7b) \end{cases}$$

where ε is a small parameter and η is the noisy perturbation. X^ε is the slow variable and Y^ε is the fast variable. As $\varepsilon \rightarrow 0$, the effective dynamics of the slow variable is

$$\dot{\bar{X}} = F(\bar{X}), \text{ where } F(x) \doteq \int f(x, y)\mu_x(dy), \quad (8)$$

where $\mu_x(dy)$ is the invariant measure of the fast process with the density function denoted by $\rho(x, y)$. F usually does not

have an analytical form. The simplified MsGAD for the saddle point of Eq. (8) is

$$\begin{cases} \dot{x}^\varepsilon(t) = f(x^\varepsilon, y^\varepsilon) - 2\frac{\langle f(x^\varepsilon, y^\varepsilon), v^\varepsilon \rangle}{\langle v^\varepsilon, v^\varepsilon \rangle}v^\varepsilon, & (9a) \\ \dot{y}^\varepsilon(t) = \frac{1}{\varepsilon}b(x^\varepsilon, y^\varepsilon) + \frac{\sigma}{\sqrt{\varepsilon}}\eta(t), & (9b) \\ \dot{v}^\varepsilon(t) = [D_x f(x^\varepsilon, y^\varepsilon) + C(x^\varepsilon, y^\varepsilon)]v^\varepsilon - \alpha^\varepsilon v^\varepsilon & (9c) \end{cases}$$

or

$$\begin{cases} \dot{x}^\varepsilon(t) = f(x^\varepsilon, y^\varepsilon) - 2\frac{\langle f(x^\varepsilon, y^\varepsilon), w^\varepsilon \rangle}{\langle w^\varepsilon, w^\varepsilon \rangle}w^\varepsilon, & (10a) \\ \dot{y}^\varepsilon(t) = \frac{1}{\varepsilon}b(x^\varepsilon, y^\varepsilon) + \frac{\sigma}{\sqrt{\varepsilon}}\eta(t), & (10b) \\ \dot{w}^\varepsilon(t) = [D_x f(x^\varepsilon, y^\varepsilon) + C(x^\varepsilon, y^\varepsilon)]^T w^\varepsilon - \beta^\varepsilon w^\varepsilon, & (10c) \end{cases}$$

where $D_x f(x, y)$ is the Jacobian matrix of $f(x, y)$ with respect to x . $\alpha = \langle v, (D_x f + C)v \rangle$, $\beta = \langle w, (D_x f + C)^T w \rangle$, $C(x, y) = [f(x, y) - F(x)] \otimes [g(x, y) - G(x)]$, $g(x, y) = -\nabla_x U(x, y)$, $U(x, y) = -\log \rho(x, y)$, and $G(x) = \int g(x, y)\mu_x(dy)$. In general, the covariance function C has no closed form and needs to be computed on the fly by the heterogeneous multiscale method (HMM) or the seamless coupling method.⁹

We show an example where the analytical expression of the covariance C is available

$$\begin{cases} \dot{X}_i = -\sum_{j=1}^2 A_{ij}X_j + Y_i^2, \quad i = 1, 2, & (11a) \\ \dot{Y}_i = -\frac{1}{\varepsilon}\frac{Y_i}{\Gamma_i(X)} + \frac{\sigma}{\sqrt{\varepsilon}}\eta_i(t), & (11b) \end{cases}$$

where A is a 2 by 2 matrix and $\Gamma_i(x), i = 1, 2$, are positive-valued functions. σ is a constant. Each fast process Y_i is an Ornstein–Uhlenbeck process and thus has the invariant measure $\mu_x(dy_i) = \mathcal{N}[0, \sigma^2\Gamma_i(x)/2]$. On the one hand, as $\varepsilon \rightarrow 0$, the closed form for the effective dynamics of (11) is

$$\dot{\bar{X}}_i = -\sum_{j=1}^2 A_{ij}\bar{X}_j + \frac{\sigma^2}{2}\Gamma_i(\bar{X}) =: F_i(\bar{X}), i = 1, 2. \quad (12)$$

The Jacobian matrix J takes the form $J(x) = DF(x) = -A + \frac{1}{2}\sigma^2\text{diag}\{\Gamma'_i(x) : i = 1, 2\}$ in this case. On the other hand, the covariance function C in Eq. (9) takes the form $C_{ij}(x, y) = \frac{\Gamma'_j(x)}{\Gamma_j^2(x)}\left[\frac{y_i^2 y_j^2}{\sigma^2} - \frac{1}{2}y_i^2 \Gamma_j(x) - \frac{1}{2}y_j^2 \Gamma_i(x) + \frac{\sigma^2}{4}\Gamma_i(x)\Gamma_j(x)\right]$. By the averaging principle applied to Eq. (9), we find that $\bar{C} = \int C(x, y)\mu_x(dy) = \frac{1}{2}\sigma^2\text{diag}\{\Gamma'_i(x) : i = 1, 2\}$ and $\int D_x f(x, y)\mu_x(dy) = -A$. This shows the effective dynamics of the MsGAD Eq. (9) is indeed the GAD equation associated with the effective Eq. (12).

III. NUMERICAL EXAMPLES

A. A two-dimensional example

The first test is to locate the saddle point of the following two dimensional non-gradient system

$$\dot{x}_i = -\sum_{j=1}^2 A_{ij}x_j + \frac{\sigma^2}{2}\Gamma_i(x), \quad i = 1, 2, \quad (13)$$

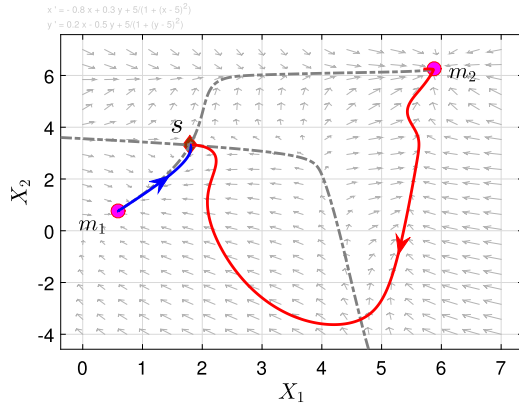


FIG. 1. GAD trajectories of the x component from two locally stable fixed points (m_1 and m_2) to the saddle point s . The blue and the red curves with arrows are the trajectories of the x component for the simplified GAD applied to the dynamics (13) with the initial vector $v = [1, 0]$ and $[0, 1]$, respectively. The flow indicated by the arrows corresponds to the non-gradient system (13). The dashed-dotted curves are the stable/unstable manifolds of the saddle point: the stable manifold is the basin boundary between m_1 and m_2 .

where $\sigma^2 = 10$, $A = \begin{bmatrix} 0.8 & -0.3 \\ -0.2 & 0.5 \end{bmatrix}$ and $\Gamma_i(x) = [1 + (x_i - 5)^2]^{-1}$, $i = 1, 2$. This dynamics has two stable fixed points $m_1 = (0.5931, 0.7655)$, $m_2 = (5.8770, 6.2507)$ and a unique saddle point $s = (1.7954, 3.3088)$. Figure 1 shows the simplified GAD trajectories of the x component (solid lines) starting from m_1 and m_2 , respectively.

Note that Eq. (13) is the effective dynamics of the slow-fast system (11) as shown above in Sec. II C. Our second test is to use the simplified MsGAD (9) to find the saddle point of (13) with the same setting of A , Γ_i , and σ . If we use the HMM, the trajectories will be almost identical to the result in Fig. 1. So, we show the result computed from the seamless coupling method, which requires to select a sufficiently small number ε . Figure 2 presents the trajectories of the x component computed from this method from m_1 or m_2 to the saddle point s .

B. Nucleation in the presence of shear flow

In the second example, we consider a more challenging problem: the nucleation in the reaction-diffusion equation in the presence of shear flow. Nucleation is a very important

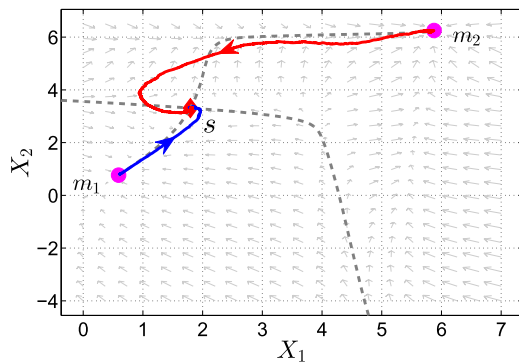


FIG. 2. GAD trajectories computed from the simplified MsGAD by the seamless coupling method. The small parameter ε is 10^{-5} . The initial direction vectors v are the same as in Fig. 1. The symbols and annotations have the same meaning as in Fig. 1.

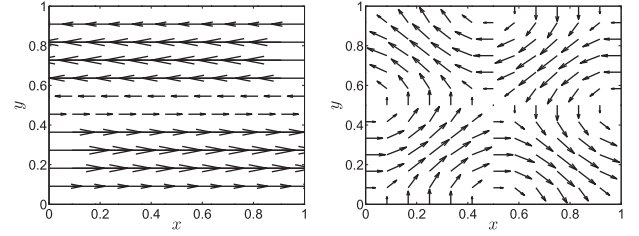


FIG. 3. The vector fields of the shear flows used in Eqs. (15) and (16), respectively (from left to right).

physical phenomenon^{11,12,14,16,22} and the nucleus is usually described by the saddle point of the Ginzburg-Landau free energy. In the case of gradient systems driven only by the free energy, the string method^{2,4} can be applied to calculate the minimum energy path. When the shear flow field is present, one is faced with a non-gradient system and in principle, one needs the minimum action method³ to compute the minimum action path and the minimal action.¹¹ The saddle point can be extracted after the whole path is computed. By our new method, the saddle point in the shear flow case can be calculated directly.

The Ginzburg-Landau free energy of the order parameter $\phi(x, y)$ is

$$E(\phi) = \int_{\Omega} \frac{\kappa}{2} |\nabla \phi|^2 + \frac{1}{4} (1 - \phi^2)^2 dx dy, \quad (14)$$

where $\kappa = 0.01$, the domain $\Omega = [0, 1] \times [0, 1]$. The periodic boundary condition is considered. We study two cases of the shear flow as illustrated in Fig. 3, then the corresponding dynamics of the Allen-Cahn equation in the presence of the shear flow are

$$\partial_t \phi = -\frac{\delta E}{\delta \phi} + \gamma \sin(2\pi y) \partial_x \phi \quad (15)$$

and

$$\partial_t \phi = -\frac{\delta E}{\delta \phi} + \gamma \sin(2\pi y) \partial_x \phi + \gamma \sin(2\pi x) \partial_y \phi, \quad (16)$$

respectively, where γ is the shear rate and the Fréchet derivative of E is $\delta_{\phi} E = -\kappa \Delta \phi - \phi + \phi^3$.

We want to locate the index-1 saddle point $\phi(x, y)$ in the dynamics (15) and (16) by the simplified GAD in Sec. II B. Denote the right hand side in (15) or (16) by $b(\phi)$, then, the simplified GAD (5) in this case is

$$\begin{cases} \partial_t \phi = b(\phi) - 2 \langle b(\phi), v \rangle v / \|v\|^2, & (17a) \\ \partial_t v = Db(\phi)v - \langle v, (Db)v \rangle v / \|v\|^2, & (17b) \end{cases}$$

where $v = v(x, t)$, $\langle \cdot, \cdot \rangle$, and $\|\cdot\|$ are the L^2 inner product and the L^2 norm in space.

Remark 2. Here, the dynamics is a PDE model and we can have the analytical expression for the Jacobian and its transpose. We take the case in Eq. (15), for example, to show Db and its adjoint $(Db)^T$. $b(\phi) = \kappa \Delta \phi + \phi - \phi^3 + \gamma \sin(2\pi y) \partial_x \phi$. $Db(\phi)v = \kappa \Delta v + v - 3\phi^2 v + \gamma \sin(2\pi y) \partial_x v$. Then, $[Db(\phi)]^T w = \kappa \Delta w + w - 3\phi^2 w - \gamma \sin(2\pi y) \partial_x w$ since the adjoint of ∂_x is $-\partial_x$. This example shows that when b is a differential operator, one may obtain the “transpose” (adjoint) of the Jacobian analytically. Then, the two forms of

the simplified GAD (5) and (6) are both applicable in such cases.

In numerical tests, we use the mesh points with $N_x = N_y = 128$ in the finite difference method for spatial discretization. The two metastable states are $\phi \equiv 1$ and $\phi \equiv -1$, regardless of the shear flow. By solving the simplified GAD (17), we shall get different index-1 saddle points for various shear rates γ . We are interested in the effect of the shear rate on the profiles of saddle points. It is noted that the steady states for any shear flow preserve the symmetry $\phi \rightarrow -\phi$ and Eq. (16) additionally preserves another symmetry $\phi(x, y) \rightarrow \phi(y, x)$. So, there are multiple symmetric images of one steady state. All of our plots below display only one of the symmetric images.

For the case in Eq. (15), the shear force is applied only in the x direction. As γ increases, the sequence of the profiles of the saddle point is shown in Fig. 4. We have the following interesting observation from this figure: the profiles of the saddle points get more and more stretched along the shear direction until a lamellar phase is attained for γ large enough. In fact, the lamellar phase shown in the last two subfigures [Figs. 4(e) and 4(f)] is always a saddle point for γ . It seems to have a critical γ_* between 0.05 and 0.065 such that for $\gamma < \gamma_*$, there are two saddle points: one is twisted and the other is lamellar, and for $\gamma > \gamma_*$, it seems only one index-1 saddle

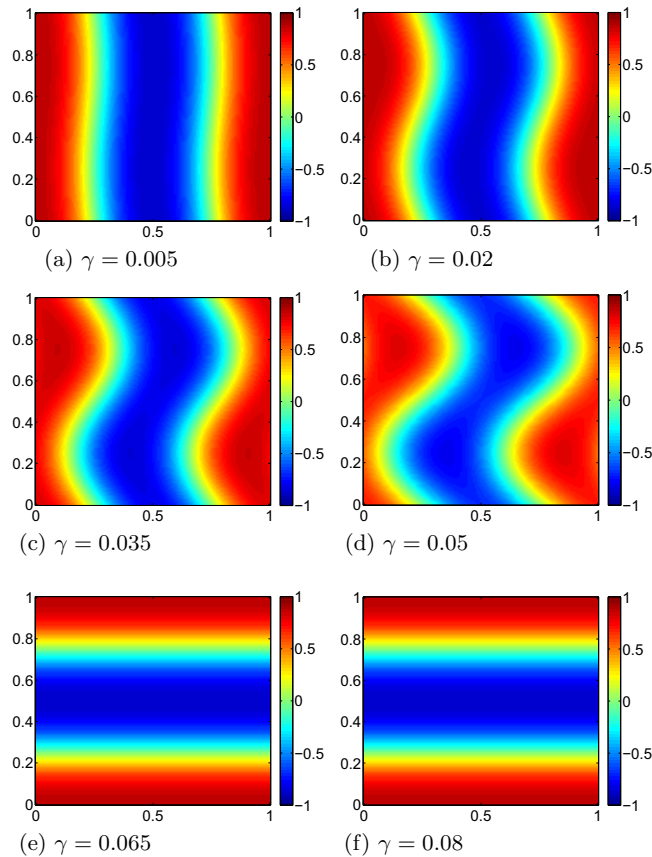


FIG. 4. The contour plots of the saddle points $\phi(x, y)$ of the model (15) at various shear rates γ . The profile of ϕ becomes stretched horizontally as γ gets larger. But it becomes lamellar phase and does not change any more when γ is sufficiently large. The color bar on the right of each figure shows the values of ϕ . (a) $\gamma = 0.005$ (b) $\gamma = 0.02$ (c) $\gamma = 0.035$ (d) $\gamma = 0.05$ (e) $\gamma = 0.065$, and (f) $\gamma = 0.08$.

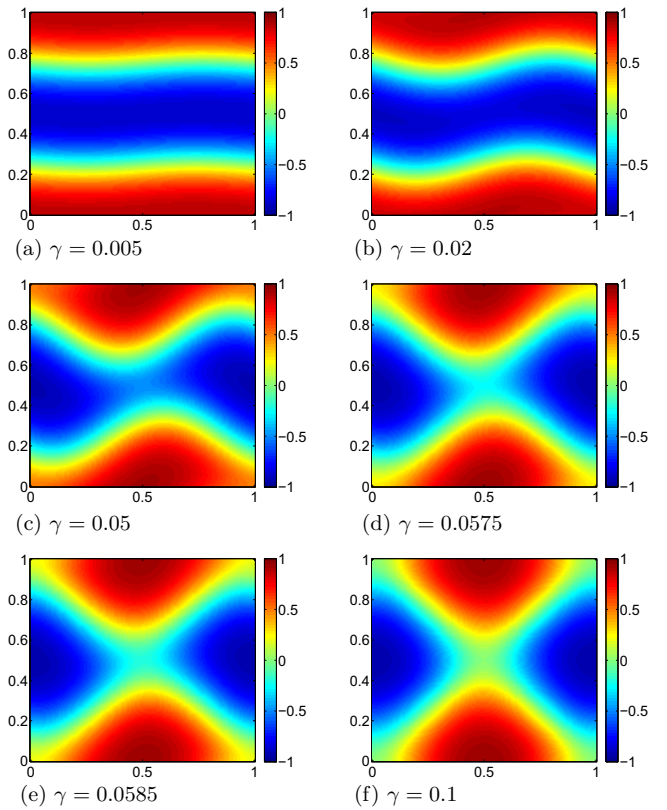


FIG. 5. The contour plots of the saddle points $\phi(x, y)$ of the model (16) at various shear rates γ . The profile of ϕ gets twisted more and more as γ increases and forms an “X” shape in the end. The color bar on the right of each figure also shows the values of ϕ . (a) $\gamma = 0.005$, (b) $\gamma = 0.02$, (c) $\gamma = 0.05$, (d) $\gamma = 0.0575$, (e) $\gamma = 0.0585$, and (f) $\gamma = 0.1$.

point, the lamellar phase. To determine which saddle point has the minimal action of escaping from a metastable state for a specific $\gamma < \gamma_*$, one needs to further run the minimum action method.¹¹

For the case in Eq. (16), the shear flow is no longer restricted in the x direction and is more general. The transition states corresponding to the various shear rates are shown in Fig. 5. The shear “twists” the profiles again but in different patterns. Similar to the first case, the saddle point is finally unchanged when γ is sufficiently large. And eventually, the saddle point forms an “X” shape. For the small shear rate, the “X” shaped saddle point in Fig. 5(f) does not exit, unlike the lamellar phase in the previous shear flow case. Thus, there seems to have only one index-1 saddle point at any γ , except for the symmetric images. In summary, the shear flow acting on the Ginzburg-Landau energy landscape induces a variety of different patterns of the saddle points and transition mechanisms. Our simplified GAD offers a useful tool to locate these saddle points with economic computational costs.

IV. CONCLUDING REMARKS

We present a simplified GAD for the non-gradient system in \mathbb{R}^d to search saddle points. It is a flow in \mathbb{R}^{2d} rather than in \mathbb{R}^{3d} . Only one direction variable and one position variable are required in this new GAD. So, it has the same computational costs as the GAD for the gradient system. Although we only show the result for index-1 saddle points in this paper, it is not

difficult to extend to index- k saddle points by following the approach in the original GAD paper.⁵

Our numerical tests include the Allen-Cahn equation in a periodic box with the presence of shear flow, and we have investigated the changes of saddle points when the system is subject to various shear flows. In the end, we need to point out that although index-1 saddle points seem important for the rare-event transitions in the non-gradient systems, the saddle point found by the GAD may not directly be the true transition state. To quantify the minimal action, the minimum action method may still be necessary to evaluate the action of the saddle points. One may combine the MAM for the path and the GAD for the final ending point on the path to construct a hybrid method, in a similar style to the climbing string method¹⁵ for the gradient system.

ACKNOWLEDGMENTS

The research of X.Z. was supported by the grants from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project Nos. CityU 11337216 and 11305318).

APPENDIX A: THE PROOF OF THEOREM THM1

Proof. (a) By the condition that (x_*, v_*) is a fixed point of the simplified GAD (5), we have

$$\begin{cases} b(x_*) - 2 \langle b(x_*), v_* \rangle v_* = 0, & (A1a) \\ J(x_*)v_* = \langle v_*, J(x_*)v_* \rangle v_*. & (A1b) \end{cases}$$

Equation (A1b) implies that v_* is the eigenvector of $J(x_*)$ corresponding to the eigenvalue $\lambda_* \doteq \langle v_*, J(x_*)v_* \rangle$. Making inner product with v_* on both sides of (A1a), we get

$$\langle b(x_*), v_* \rangle - 2 \langle b(x_*), v_* \rangle \langle v_*, v_* \rangle = 0.$$

Since $\|v_*\| = 1$, we have $\langle b(x_*), v_* \rangle = 0$. Thus, $b(x_*) = 0$ by (A1a).

(b) Since x_s is a fixed point of the system $\dot{x} = b(x)$, we have $b(x_s) = 0$, thus

$$b(x_s) - 2 \langle b(x_s), v_i \rangle v_i / \|v_i\|^2 = 0, \quad i = 1, 2, \dots, n. \quad (A2)$$

Since $J(x_s)v_i = \lambda_i v_i$, by taking the inner product with v_i on both sides of (A1a) and using the condition $\|v_i\| = 1$, we get $\lambda_i = \langle J(x_s)v_i, v_i \rangle$ and

$$J(x_s)v_i - \langle J(x_s)v_i, v_i \rangle v_i = 0, \quad i = 1, 2, \dots, n. \quad (A3)$$

Equations (A2) and (A3) imply that (x_s, v_i) is the fixed point of the simplified GAD (5) for all $i = 1, 2, \dots, n$.

Next, we write down the eigenvalues and corresponding eigenvectors of the Jacobian matrix of the simplified GAD at any fixed points (x_s, v_i) . First, the Jacobian matrix of the simplified GAD (5) has the expression

$$\tilde{\mathbb{J}}(x_s, v_i) = \begin{bmatrix} \mathbb{N}_1 & 0 \\ *, & \mathbb{M}_1 \end{bmatrix}, \quad (A4)$$

where

$$\mathbb{N}_1 := J - 2\lambda_i v_i v_i^T, \quad \mathbb{M}_1 := J - \lambda_i - v_i v_i^T (\lambda_i + J).$$

The eigenvalues of $\tilde{\mathbb{J}}$ can be obtained from the eigenvalues of its two diagonal blocks \mathbb{N}_1 and \mathbb{M}_1 . It is verified that

$$\begin{aligned} \mathbb{N}_1 v_i &= J v_i - 2\lambda_i v_i v_i^T v_i = -\lambda_i v_i, \\ \mathbb{N}_1 v_j &= J v_j - 2\lambda_i v_i v_i^T v_j = \lambda_j v_j, \\ \mathbb{M}_1 v_i &= J v_i - \lambda_i v_i - v_i v_i^T (\lambda_i + J) v_i = -2\lambda_i v_i \end{aligned}$$

and

$$\begin{aligned} \mathbb{M}_1 \left[v_j - \left(v_j^T v_i \right) v_i \right] &= \mathbb{M} v_j - \left(v_j^T v_i \right) \mathbb{M} v_i \\ &= [J - \lambda_i - v_i v_i^T (\lambda_i + J)] v_j + 2\lambda_i \left(v_j^T v_i \right) v_i \\ &= (\lambda_j - \lambda_i) v_j - v_i (\lambda_i + \lambda_j) v_i^T v_j + 2\lambda_i \left(v_j^T v_i \right) v_i \\ &= (\lambda_j - \lambda_i) v_j - v_i (\lambda_j - \lambda_i) v_i^T v_j \\ &= (\lambda_j - \lambda_i) \left[v_j - \left(v_i^T v_j \right) v_i \right]. \end{aligned}$$

Hence, the eigenvalues of the Jacobian matrix $\tilde{\mathbb{J}}$ at any fixed points $(x_s, v_i), i = 1, 2, \dots, n$ are

$$-2\lambda_i, -\lambda_i, \{\lambda_j : j \neq i\}, \quad \{\lambda_j - \lambda_i : j \neq i\}. \quad (A5)$$

The linear stability condition is that all the above eigenvalues of $\tilde{\mathbb{J}}$ are negative. Thus, one fixed point $(x_s, v_{i'})$ is linearly stable if and only if $\lambda_{i'} > 0$ and all other eigenvalues $\lambda_j < 0$ for $j \neq i'$. In this case, the fixed point x_s is an index-1 saddle point of the system $\dot{x} = b(x)$. \square

Remark 3. Theorem 1 also holds for the simplified GAD (6). In this case, the Jacobian matrix of the simplified GAD (6) becomes $\tilde{\mathbb{J}}(x_s, w_i) = \begin{bmatrix} \mathbb{N}_2 & 0 \\ *, & \mathbb{M}_2 \end{bmatrix}$, where $\mathbb{N}_2 := J - 2\lambda_i w_i w_i^T$ and $\mathbb{M}_2 := J^T - \lambda_i - w_i w_i^T (\lambda_i + J^T)$. This Jacobian matrix has the same eigenvalues (A5) as the Jacobian matrix (A4) of the simplified GAD (5).

APPENDIX B: RELATION WITH HAMILTON'S EQUATION

In this part, we discuss Hamilton's equation associated with the rare event study of Eq. (2). According to the Freidlin–Wentzell large deviation principle (LDP),⁶ as $\varepsilon \rightarrow 0$, the most probable transition path over the time interval $[0, T]$ of the system (2) is a minimizer of the following Freidlin–Wentzell action functional

$$S[\phi] = \int_0^T L(\phi, \dot{\phi}) dt, \quad (B1)$$

where the Lagrangian $L(x, y)$ is defined as

$$L(x, y) := \frac{1}{2} \langle y - b(x), y - b(x) \rangle. \quad (B2)$$

$\langle \cdot, \cdot \rangle$ is the inner product in \mathbb{R}^d . The Hamiltonian $H(x, p)$, as the conjugate of the Lagrangian $L(x, y)$, is

$$H(x, p) = \langle b(x), p \rangle + \langle p, p \rangle / 2. \quad (B3)$$

It is well-known that the minimizer of $S[\phi]$, denoted as $x(t)$, satisfies Hamilton's equations

$$\begin{cases} \dot{x} = H_p = b(x) + p(t), & (B4a) \\ \dot{p} = -H_x = -J(x)^T p(t), & (B4b) \end{cases}$$

where $p(t)$ is the (generalized) momentum. $J(x) = Db(x)$ is the Jacobian matrix we have used before in the GAD. The eigenvalues of $J(x)$ are denoted as $\{\lambda_i\}$. Equation (B4) looks superficially similar to Eq. (6) with two differences: (i) the signs before $J(x)^T$ are the opposite and (ii) the momentum p is not a unit vector as the direction variable w . In fact, the critical point of (B4) is (x_*, p_*) , where $b(x_*) = 0$ and $p_* = 0$ by assuming that $J(x_*)$ is non-degenerate. Assume that $J(x)$ has the right-eigenvectors v_i and the left-eigenvectors w_i

$$Jv_i = \lambda_i v_i \quad \text{and} \quad J^T w_i = \lambda_i w_i, \quad 1 \leq i \leq d,$$

where all eigenvalues are assumed distinctive and the left or right eigenvectors both form a basis of \mathbb{R}^d . We introduce the normalized unit vector u to represent the direction of p . Define the scalar $l \doteq \|p\|^2$, then $u = p/\sqrt{l}$ and $\dot{l} = 2\langle p, \dot{p} \rangle = -2\langle p, J^T p \rangle = -2l\langle u, J^T u \rangle$. So,

$$\dot{u} = \frac{d}{dt} \left(\frac{p}{\sqrt{l}} \right) = -J^T(x)u + \langle u, J^T u \rangle u. \quad (\text{B5})$$

By the important zero-Hamiltonian condition⁶ $H \equiv 0$, we have

$$l = \|p\|^2 = -2\langle b, p \rangle = -2\sqrt{l}\langle b, u \rangle,$$

that is,

$$l = 0, \quad \text{or} \quad \sqrt{l} = -2\langle b(x), u \rangle.$$

$l = 0$ means $p = 0$, which corresponds to the original dynamics $\dot{x} = b(x)$. l is not always zero for the exit dynamics, then $\sqrt{l} = -2\langle b(x), u \rangle$ and Eq. (B4a) becomes

$$\dot{x} = b(x) + \sqrt{l}u = b(x) - 2\langle b(x), u \rangle u. \quad (\text{B6})$$

So far, by (B5) and (B6), we get the momentum-normalized version for Hamilton's equations (B4) restricted on the zero- H hypersurface

$$\begin{cases} \dot{x} = b(x) - 2\langle b(x), u(t) \rangle u(t) / \|u(t)\|^2, & (\text{B7a}) \\ \dot{u} = -J^T(x)u + \langle u, J^T u \rangle u, & (\text{B7b}) \end{cases}$$

where $\|u_0\| = 1$ is assumed. Note that this dynamics (B7) is not exactly identical to the original Hamilton's equation (B4) since the branch of $p \equiv 0$ has been discarded.

Now, the only difference between Hamilton's equation (B7) and the simplified GAD (6) is the opposite sign on the right hand sides of (B7a) and (6b). By Remark 3, the Jacobian matrix of (B7) is $\begin{bmatrix} \mathbb{N}_2 & 0 \\ * & -\mathbb{M}_2 \end{bmatrix}$, whose eigenvalues are $-\lambda_i, 2\lambda_i, \{\lambda_j, j \neq i\}, \{\lambda_i - \lambda_j, j \neq i\}$. The position dynamics in (6) and (B7) have the same form of applying the projection matrix $I - 2ww^T$ or $I - 2uu^T$ in front of the original force $b(x)$. The difference is which direction they select. If x were frozen, the w dynamics in Eq. (6b) picks up the least stable direction, while the momentum direction in Hamilton's equations uses the most stable direction. Thus, the GAD (6) can converge to the saddle point of the vector field $b(x)$, while the

Hamiltonian dynamics (B7) has no stable steady state. So, one may view the simplified GAD as a modification of Hamilton's equation by flipping the sign of the (normalized) momentum to stabilize the saddle point. Note that although we can introduce a factor γ for (6b) as shown in Remark 1 to speed up the clock for the direction dynamics, there is no such a freedom for Hamilton's equation (B7b).

¹G. M. Crippen and H. A. Scheraga, "Minimization of polypeptide energy: XI. The method of gentlest ascent," *Arch. Biochem. Biophys.* **144**, 462–466 (1971).

²W. E, W. Ren, and E. Vanden-Eijnden, "String method for the study of rare events," *Phys. Rev. B* **66**, 052301 (2002).

³W. E, W. Ren, and E. Vanden-Eijnden, "Minimal action method for rare events," *Commun. Pure Appl. Math.* **57**, 637–656 (2004).

⁴W. E, W. Ren, and E. Vanden-Eijnden, "Simplified and improved string method for computing the minimum energy paths in barrier-crossing events," *J. Chem. Phys.* **126**, 164103 (2007).

⁵W. E and X. Zhou, "The gentlest ascent dynamics," *Nonlinearity* **24**, 1831 (2011).

⁶M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, 3rd ed., Grundlehren der mathematischen Wissenschaften (Springer-Verlag, New York, 2012).

⁷W. Gao, J. Leng, and X. Zhou, "An iterative minimization formulation for saddle point search," *SIAM J. Numer. Anal.* **53**, 1786–1805 (2015).

⁸W. Gao, J. Leng, and X. Zhou, "Iterative minimization algorithm for efficient calculations of transition states," *J. Comput. Phys.* **309**, 69–87 (2016).

⁹S. Gu and X. Zhou, "Multiscale gentlest ascent dynamics for saddle point in effective dynamics of slow-fast system," *Commun. Math. Sci.* **15**, 2279–2302 (2017).

¹⁰G. Henkelman and H. Jónsson, "A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives," *J. Chem. Phys.* **111**, 7010–7022 (1999).

¹¹M. Heymann and E. Vanden-Eijnden, "Pathways of maximum likelihood for rare events in nonequilibrium systems: Application to nucleation in the presence of shear," *Phys. Rev. Lett.* **100**, 40601 (2008).

¹²L. Lin, X. Cheng, W. E, A. Shi, and P. Zhang, "A numerical method for the study of nucleation of ordered phases," *J. Comput. Phys.* **229**, 1797–1809 (2010).

¹³N. Mousseau and G. Barkema, "Traveling through potential energy surfaces of disordered materials: The activation-relaxation technique," *Phys. Rev. E* **57**, 419 (1998).

¹⁴A. Onuki, "Phase transitions of fluids in shear flow," *J. Phys.: Condens. Matter* **9**, 6119–6157 (1997).

¹⁵W. Ren and E. Vanden-Eijnden, "A climbing string method for saddle point search," *J. Chem. Phys.* **138**, 34105 (2013).

¹⁶A. Roy, J. M. Rickman, J. D. Gunton, and K. R. Elder, "Simulation study of nucleation in a phase-field model with nonlocal interactions," *Phys. Rev. E* **57**, 2610–2617 (1998).

¹⁷A. Samanta and W. E, "Atomistic simulations of rare events using gentlest ascent dynamics," *J. Chem. Phys.* **136**, 24104 (2012).

¹⁸D. J. Wales, *Energy Landscapes with Application to Clusters, Biomolecules and Glasses* (Cambridge University Press, 2003).

¹⁹X. Wan, H. Yu, and W. E, "Model the nonlinear instability of wall-bounded shear flows as a rare event: A study on two-dimensional Poiseuille flow," *Nonlinearity* **28**, 409 (2015).

²⁰X. Wan, X. Zhou, and W. E, "Study of noise-induced transition and the exploration of the configuration space for the Kuromoto-Sivachinsky equation using the minimum action method," *Nonlinearity* **23**, 475–493 (2010).

²¹J. Zhang and Q. Du, "Shrinking dimer dynamics and its applications to saddle point search," *SIAM J. Numer. Anal.* **50**, 1899–1921 (2012).

²²L. Zhang, L. Q. Chen, and Q. Du, "Morphology of critical nuclei in solid state phase transformations," *Phys. Rev. Lett.* **98**, 65703 (2007).