Ann. Appl. Math.
doi: 10.4208/aam.OA-2023-0030

Vol. **40**, No. 1, pp. 1-20
February 2024

Probabilistic Error Estimate for Numerical Discretization of High-Index Saddle Dynamics with Inaccurate Models

Lei Zhang[1], Pingwen Zhang[2,3] and Xiangcheng Zheng[4,*]

[1] *Beijing International Center for Mathematical Research, Center for Machine Learning Research, Center for Quantitative Biology, Peking University, Beijing 100871, China*
[2] *School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, China*
[3] *School of Mathematical Sciences, Laboratory of Mathematics and Applied Mathematics, Peking University, Beijing 100871, China*
[4] *School of Mathematics, Shandong University, Jinan, Shandong 250100, China*

Received 11 September 2023; Accepted (in revised version) 30 November 2023

Abstract. We prove probabilistic error estimates for high-index saddle dynamics with or without constraints to account for the inaccurate values of the model, which could be encountered in various scenarios such as model uncertainties or surrogate model algorithms via machine learning methods. The main contribution lies in incorporating the probabilistic error bound of the model values with the conventional error estimate methods for high-index saddle dynamics. The derived results generalize the error analysis of deterministic saddle dynamics and characterize the affect of the inaccuracy of the model on the convergence rate.

AMS subject classifications: 37M05, 65L20, 60G15

Key words: Saddle point, saddle dynamics, solution landscape, Gaussian process, probabilistic error estimate.

---

*Corresponding author.
Emails: `zhangl@math.pku.edu.cn` (L. Zhang), `pzhang@pku.edu.cn` (P. Zhang), `xzheng@sdu.edu.cn` (X. Zheng)

1 Introduction

High-index saddle points of complex systems contain ample physical and chemical information and thus attract extensive attentions [4, 15, 27, 46]. Here the index of the saddle point refers to the Morse index characterized by the maximal dimension of a subspace on which its Hessian operator is negative definite [28]. There exist several successful algorithms for finding saddle points [5, 6, 8, 10, 11, 19, 47, 50]. For instance, the search extension method [2, 37] has been applied to find multiple solutions of nonlinear problems. The iterative minimization formulation [9] and the local minimax method [20, 21, 23] have been developed to search high-index saddle points. Recently, a high-index saddle dynamics is proposed in [43] to compute an index-k saddle point

$$\begin{cases} \dfrac{dx}{dt}=\beta\left(I-2\displaystyle\sum_{j=1}^{k}v_jv_j^\top\right)F(x), \\ \dfrac{dv_i}{dt}=\gamma\left(I-v_iv_i^\top-2\displaystyle\sum_{j=1}^{i-1}v_jv_j^\top\right)J(x)v_i, \quad 1\leq i\leq k. \end{cases} \tag{1.1}$$

Here $x\in\mathbb{R}^d$ represents the state variable, $\{v_i\}_{i=1}^k$ are directional variables, β, $\gamma>0$ are relaxation parameters, $F(x):\mathbb{R}^d\to\mathbb{R}^d$ represents the force generated from the energy $E(x)$ by $F(x)=-\nabla E(x)$ and $J(x)$ is the negative Hessian of $E(x)$, i.e., $J(x)=-\nabla^2E(x)$. This high-index saddle dynamics could be further combined with the downward and upward algorithms [42] to construct solution landscapes of complex systems, the pathway map consisting of all stationary points and their connections [36], that arises several successful applications [13, 14, 24, 30, 38, 40, 41, 44, 45, 48, 49, 51].

In most previous studies, exact model values such as F and J used in the high-index saddle dynamics (1.1) are assumed to be given a priori. However, this is not the case in many practical problems. For instance, a surrogate model based saddle dynamics is proposed in [12, 55] to reduce the number of queries of model values that may be expensive or time-consuming, where the model values are predicted via the Gaussian process learning. In this scenario, the model value may not be accurate but instead follows a probabilistic distribution. For such complicated cases, deterministic error estimates for numerical approximations to high-index saddle dynamics in, e.g., [52, 54] are not applicable and instead the probabilistic error estimates are natural to be considered, which motivates the current study.

In this work we prove probabilistic error estimates for high-index saddle dynamics with or without constraints. The main contribution lies in incorporating the probabilistic error bound of the model values with the conventional error estimate methods for high-index saddle dynamics. The derived error estimates could ensure the dynamical pathway convergence of the numerical scheme, that is, to ensure that the discrete high-index saddle dynamics evolves along the dynamical pathway of continuous high-index saddle dynamics, which in turn ensures that the numerical scheme converges to the same target saddle point of continuous high-index saddle dynamics. Furthermore, compared with traditional estimates that control the errors with respect to the time step size τ, the developed error estimate results such as (3.1) and (4.5) also characterize the impacts of the uncertainty and inaccuracy of the model values on the computational accuracy of the schemes.

The rest of the work is organized as follows: In Section 2 we present numerical discretization of high-index saddle dynamics and the assumptions and their explanations. In Section 3 we prove probabilistic error estimate for the numerical scheme of high-index saddle dynamics. In Section 4 we extend the results to study the probabilistic error estimate for the numerical scheme of high-index saddle dynamics constrained on the high-dimensional unit sphere. We finally address concluding remarks in the last section.

2 Numerical discretization with inaccurate models

2.1 Numerical scheme

For a fixed time step size $\tau>0$ and time steps $t_n=n\tau$ with the terminal time $T=t_N$ for some N, we apply the Euler discretization for first-order derivatives in (1.1) to obtain a reference equation

$$\begin{cases} x(t_n)=x(t_{n-1})+\tau\beta\Big(I-2\sum_{j=1}^{k}v_j(t_{n-1})v_j(t_{n-1})^\top\Big)F(x(t_{n-1}))+\mathcal{O}(\tau^2),\\ v_i(t_n)=v_i(t_{n-1})+\tau\gamma\Big(I-v_i(t_{n-1})v_i(t_{n-1})^\top\\ \qquad -2\sum_{j=1}^{i-1}v_j(t_{n-1})v_j(t_{n-1})^\top\Big)J(x(t_{n-1}))v_i(t_{n-1})+\mathcal{O}(\tau^2), \quad 1\le i\le k. \end{cases} \tag{2.1}$$

In general, we could drop truncation errors in (2.1) to obtain the explicit scheme of (1.1). However, exact values of F and J may not be available in practice due to various reasons such as modeling inaccuracies, experimental errors and uncertainties, and instead the inaccurate values are more common. For this concern, the numerical scheme with inaccurate $F(x)$ and $J(x)$, which are denoted by $\hat{F}(x)$ and

$\hat{J}(x)$, respectfully, is proposed as follows with the approximations $\{x_n,v_{i,n}\}_{n=1,i=1}^{N,k}$ to $\{x(t_n),v_i(t_n)\}_{n=1,i=1}^{N,k}$

$$\begin{cases} x_n = x_{n-1} + \tau\beta\left(I - 2\sum_{j=1}^{k} v_{j,n-1}v_{j,n-1}^\top\right)\hat{F}(x_{n-1}), \\ \tilde{v}_{i,n} = v_{i,n-1} + \tau\gamma\bigg(I - v_{i,n-1}v_{i,n-1}^\top \\ \qquad\qquad -2\sum_{j=1}^{i-1} v_{j,n-1}v_{j,n-1}^\top\bigg)\hat{J}(x_{n-1})v_{i,n-1}, \qquad 1 \le i \le k, \\ v_{i,n} = \dfrac{1}{Y_{i,n}}\left(\tilde{v}_{i,n} - \sum_{j=1}^{i-1}(\tilde{v}_{i,n}^\top v_{j,n})v_{j,n}\right), \qquad 1 \le i \le k, \end{cases} \tag{2.2}$$

for $1 \le n \le N$ with the normalization constants

$$\begin{aligned} Y_{i,n} &:= \left\|\tilde{v}_{i,n} - \sum_{j=1}^{i-1}(\tilde{v}_{i,n}^\top v_{j,n})v_{j,n}\right\| \\ &= \left(\|\tilde{v}_{i,n}\|^2 - 2\sum_{j=1}^{i-1}(\tilde{v}_{i,n}^\top v_{j,n})^2 + \sum_{j=1}^{i-1}(\tilde{v}_{i,n}^\top v_{j,n})v_{j,n}^\top \cdot \sum_{j=1}^{i-1}(\tilde{v}_{i,n}^\top v_{j,n})v_{j,n}\right)^{1/2} \\ &= \left(\|\tilde{v}_{i,n}\|^2 - \sum_{j=1}^{i-1}(\tilde{v}_{i,n}^\top v_{j,n})^2\right)^{1/2}, \end{aligned}$$

equipped with the initial conditions

$$x_0 = x(0), \quad v_{i,0} = v_i(0) \quad \text{for} \quad 1 \le i \le k \quad \text{satisfying} \quad v_{i,n}^\top v_{j,n} = \delta_{i,j} \quad \text{for} \quad 1 \le i,j \le k.$$

The third equation of (2.2) is indeed the Gram-Schmidt normalized orthogonalization procedure in order to preserve the orthonormal property of directional vectors as in the continuous problem (1.1) [43, 52].

2.2 Assumptions and motivations

We make the following assumptions for model properties and the degree of model inaccuracy throughout this work.

Assumption A. There exists a constant $L > 0$ such that the following linearly growth and Lipschitz conditions hold under the standard l^2 norm $\|\cdot\|$ of a vector or a matrix

$$\begin{aligned}&\|J(x_2)-J(x_1)\|+\|F(x_2)-F(x_1)\|\leq L\|x_2-x_1\|,\\&\|F(x)\|\leq L(1+\|x\|),\quad x,x_1,x_2\in\mathbb{R}^d.\end{aligned}$$

Assumption B. For a given compact set $\mathbb{X}\subset\mathbb{R}^d$, there exist constants $Q_0>0$, $0<\delta<1$ and $\varepsilon>0$ such that the following boundedness for inaccurate values of F and J

$$\max\{\|\hat{F}(x)\|,\|\hat{J}(x)\|\}\leq Q_0,\quad \forall x\in\mathbb{X},$$

and the probabilistic error bound

$$\begin{aligned}&P\bigg(\sum_{i=1}^m\|F(x_i)-\hat{F}(x_i)\|+\sum_{j=1}^n\|J(x_{m+j})-\hat{J}(x_{m+j})\|\leq(m+n)\varepsilon\bigg)\\&\geq 1-\delta,\quad \forall\{x_1,\cdots,x_{m+n}\}\subset\mathbb{X},\quad m,n\geq 0,\end{aligned}$$

hold.

In many applications, there exist various energy functions E such that the corresponding F and J satisfy the **Assumption A**, e.g., the Minyaev-Quapp surface [29] and the Eckhardt surface [7]. Furthermore, the **Assumption A** is also natural and commonly used for numerical analysis [52, 53] since if the dynamics is convergent to some saddle point, then the trajectory of the dynamics could certainly lie within a bounded domain such that a truncated F or J could be designed to satisfy the **Assumption A** without any impact on the analysis. However, the **Assumption B** is rarely encountered in the literature as the accurate values of models are usually supposed to be given a priori. In real applications the **Assumption B** is indeed much more practical as we will show in the following scenarios.

For problems with complicated underlying mechanisms the exact forms of F and J are not given a priori in some cases such that we need to either investigate the modeling of the underlying processes or perform experiments in order to obtain the inquired values of the model in high-index saddle dynamics. In practice, modeling complex problems may be difficult or inaccurate, while the experiments or simulations are often expensive that restrict the application of high-index saddle dynamics for computing saddle points.

A potential remedy is the data-driven approach replacing F and J in the original high-index saddle dynamics by surrogate models trained from, e.g., the Gassian process learning. In the past few decades, the Gaussian process has been widely employed in extensive applications for constructing the surrogate models from the training data [32]. In particular, there exist some recent works on combining the Gaussian process with searching algorithms of saddle points [3, 12, 17, 55].

Gaussian process regression is a Bayesian machine learning method based on the assumption that any finite collection of random variables $y_i \in \mathbb{R}$ follows a joint Gaussian distribution with prior mean 0 and covariance kernel $k(x,x')$ [33]. Therefore, the variables y_i are observations of a sample function $f: \mathbb{X} \subset \mathbb{R}^d \to \mathbb{R}$ of the Gaussian process distribution perturbed by the Gaussian noise with zero mean and variance σ^2. By concatenating M input data points x_i in a matrix X_M the elements of the Gaussian process kernel matrix $K(X_M,X_M)$ are defined as $K_{ij}=k(x_i,x_j)$, $i,j=1,\ldots,M$ and $k(X_M,x)$ denotes the kernel vector defined analogously. The probability distribution of the Gaussian process at a point x conditioned on the training data concatenated in X_M and Y_M is then given as a normal distribution [18,33] with mean

$$\nu(x)=k(x,X_M)\big(K(X_M,X_M)+\sigma^2 I_M\big)^{-1}Y_M \tag{2.3}$$

and variance

$$\sigma_*^2(x,x')=k(x,x')-k(x,X_M)\big(K(X_M,X_M)+\sigma^2 I_M\big)^{-1}k(X_M,x').$$

The mean ν serves as the input value of the original model in practical computations. The following theorem proved in [18, Theorem 3.1] provides a uniform probabilistic error bound for the Gaussian process regression.

Theorem 2.1. *Consider a zero mean Gaussian process defined through the continuous covariance kernel $k(\cdot,\cdot)$ with Lipschitz constant L_k on the compact set $\mathbb{X}$ defined as*

$$L_k := \max_{x,x'\in\mathbb{X}} \left\| \left[\begin{array}{ccc} \dfrac{\partial k(x,x')}{\partial x_1} & \cdots & \dfrac{\partial k(x,x')}{\partial x_d} \end{array} \right]^\top \right\|.$$

Furthermore, consider a continuous unknown function $f:\mathbb{X}\to\mathbb{R}$ with Lipschitz constant L_f and M observations y_i satisfying the assumption that $f(\cdot)$ is a sample from a Gaussian process $\mathcal{GP}(0,k(x,x'))$ and observations $y=f(x)+\epsilon$ are perturbed by zero mean i.i.d. Gaussian noise ϵ with variance σ^2. Then, the posterior mean function $\nu(\cdot)$ and standard deviation $\sigma_*(\cdot)$ of a Gaussian process conditioned on the training data $\{(x_i,y_i)\}_{i=1}^M$ are continuous with Lipschitz constant L_ν and modulus of continuity $\omega_{\sigma_*}(\cdot)$ on $\mathbb{X}$ such that

$$L_\nu \le L_k\sqrt{M}\left\|\big(K(X_M,X_M)+\sigma^2 I_M\big)^{-1}Y_M\right\|, \tag{2.4a}$$

$$\omega_{\sigma_*}(\tau)\le\sqrt{2\tau L_k\left(1+M\left\|(K(X_M,X_M)+\sigma^2 I_M)^{-1}\right\|\max_{x,x'\in\mathbb{X}}k(x,x')\right)}. \tag{2.4b}$$

Moreover, pick $\delta\in(0,1)$, $\tau>0$ *and set*

$$\beta(\tau)=2\log\left(\frac{\mathcal{M}(\tau,\mathbb{X})}{\delta}\right), \tag{2.5a}$$

$$\gamma(\tau)=(L_\nu+L_f)\tau+\sqrt{\beta(\tau)}\omega_{\sigma_*}(\tau), \tag{2.5b}$$

where the τ*-covering number* $\mathcal{M}(\tau,\mathbb{X})$ *of a set* $\mathbb{X}$ *(with respect to the Euclidean metric) is defined as the minimum number of spherical balls with radius* τ*, which is required to completely cover* $\mathbb{X}$*. Then, it holds that*

$$P\Big(|f(x)-\nu(x)|\leq\sqrt{\beta(\tau)}\sigma_*(x)+\gamma(\tau),\ \forall x\in\mathbb{X}\Big)\geq 1-\delta. \tag{2.6}$$

Based on this theorem, we could take $\bar{\sigma}:=\max_{x\in\mathbb{X}}\sigma_*(x)$ to derive from (2.6) that

$$P\Big(|f(x)-\nu(x)|\leq\sqrt{\beta(\tau)}\bar{\sigma}+\gamma(\tau),\ \forall x\in\mathbb{X}\Big)\geq 1-\delta. \tag{2.7}$$

Furthermore, we obtain from (2.3) that

$$|\nu(x)|\leq\max_{x\in\mathbb{X}}\|k(x,X_M)\|\|\big(K(X_M,X_M)+\sigma^2 I_M\big)^{-1}Y_M\|\leq Q, \tag{2.8}$$

where we use Q to denote a generic positive constant that may assume different values at different occurrences.

To clarify the relations between the above two relations and the **Assumption B**, we consider a simple case of the **Assumption B**, where $m=1$, $n=0$ and F or J is a scalar-valued function. In this case, the **Assumption B** is imposing the following conditions:

$$\begin{aligned}&\max\{\|\hat{F}(x)\|,\|\hat{J}(x)\|\}\leq Q_0, && \forall x\in\mathbb{X},\\ &P\Big(\|F(x)-\hat{F}(x)\|\leq\varepsilon\Big)\geq 1-\delta, && \forall x\subset\mathbb{X},\end{aligned}$$

which is valid if we select $\hat{F}(x)$ and $\hat{J}(x)$ as the posterior mean functions of a Gaussian process based on the training data of $F(x)$ and $J(x)$, respectively, and then apply (2.7) and (2.8). Similarly, another simple case of the **Assumption B** where $m=0$, $n=1$ and F or J is a scalar-valued function also holds, which indicates that the **Assumption B** may be appropriate.

As commented in [34, Section 6], for the vector-valued functions such as F and J, we could simply emulate each entry independently as the scalar-valued case such that the aforementioned simple cases of **Assumption B** could be applied for each entry. In many applications, however, it is natural to assume that the entries are correlated, and a better emulator could be constructed by including this correlation in the emulator [34]. For this more physically-relevant approach, one could follow [22, 33] to perform the vector-valued Gaussian process regression to predict F and J based on the training data or employ the derivative properties of the Gaussian process to predict F and J from the observations of E as [12]. As the corresponding error analysis such as the Theorem 2.1 for the scalar-valued case is not available in the literature, we impose the **Assumption B** for the sake of numerical analysis in this work.

3 Error estimate

In this section we prove error estimates for the scheme (2.2) by assuming that the trajectory of discrete high-index saddle dynamics lies within a compact set $\mathbb{X}$. This restriction is proposed as the approximation property of the Gaussian process regression is proved only for functions defined on a compact set $\mathbb{X}$ as shown in Theorem 2.1. Theoretically, this assumption is reasonable as if the high-index saddle dynamics algorithm is convergent then its trajectory would certainly lie within some $\mathbb{X}$. However, as the volume of $\mathbb{X}$ becomes larger, the bound of $f(x)-\nu(x)$ grows, cf. (2.6) such that the accuracy of the prediction of the Gaussian process regression decreases. Increasing training data could help to recover the prediction accuracy but may lead to an increment of cost. Nevertheless, this issue can be resolved in practice by applying the sequential learning algorithm [55], which divides the learning-based optimization into several trust region optimization subproblems such that each suboptimization is performed within a (relatively small) specified region that fulfills our presumption.

We first introduce an auxiliary estimate to support the error estimates. As The following lemma could be proved by exactly the same procedure as [52, Lemma 4.2] due to the fact that the trajectory of discrete high-index saddle dynamics lies within a compact set $\mathbb{X}$ (as assumed above) implies the boundedness of x_n. Thus we only present the result of the lemma without proof.

Lemma 3.1. *Under the* ***Assumptions*** A-B*, the following estimate holds for* $1\leq n\leq N$ *and* τ *sufficiently small*

$$\|v_{i,n}-\tilde{v}_{i,n}\|\leq Q\tau^2,\quad 1\leq i\leq k.$$

Here the positive constant Q *depends on* Q_0 *but is independent from* n, N *and* τ.

Based on these auxiliary results, we intend to prove error estimates for the numerical scheme (2.2) and characterize the affect of inaccurate model values on the convergence rate.

Theorem 3.1. *Suppose the* ***Assumptions*** A-B *hold. Then the following probabilistic error estimate holds for the scheme* (2.2) *for* τ *sufficiently small and for some* $Q>0$

$$P\big(\|x(t_n)-x_n\|\leq Q\varepsilon+Q\tau\big)\geq 1-\delta, \quad 1\leq n\leq N. \tag{3.1}$$

Here Q *depends on* k, L, T, β, γ *and* Q_0 *but is independent from* τ, n, N, ε *and* δ.

Proof. Define the errors

$$e_n^x := x(t_n)-x_n, \quad e_n^{v_i} := v_i(t_n)-v_{i,n}, \quad 1\leq n\leq N, \quad 1\leq i\leq k.$$

We subtract the second equation of (2.1) from that of (2.2) and apply $v_i(t_n)-\tilde{v}_{i,n}$ as $(v_i(t_n)-v_{i,n})+(v_{i,n}-\tilde{v}_{i,n})=e_n^{v_i}+(v_{i,n}-\tilde{v}_{i,n})$ to obtain

$$\begin{aligned} e_n^{v_i} =& e_{n-1}^{v_i}+\tau\gamma\Bigg(I-v_i(t_{n-1})v_i(t_{n-1})^\top-2\sum_{j=1}^{i-1}v_j(t_{n-1})v_j(t_{n-1})^\top\Bigg)J(x(t_{n-1}))v_i(t_{n-1}) \\ &-\tau\gamma\Bigg(I-v_{i,n-1}v_{i,n-1}^\top-2\sum_{j=1}^{i-1}v_{j,n-1}v_{j,n-1}^\top\Bigg)\hat{J}(x_{n-1})v_{i,n-1} \\ &-(v_{i,n}-\tilde{v}_{i,n})+\mathcal{O}(\tau^2), \end{aligned} \tag{3.2}$$

which, together with Lemma 3.1, implies

$$\begin{aligned} \|e_n^{v_i}\|\leq& \|e_{n-1}^{v_i}\|+\tau\gamma\Bigg\|I-v_i(t_{n-1})v_i(t_{n-1})^\top-2\sum_{j=1}^{i-1}v_j(t_{n-1})v_j(t_{n-1})^\top \\ &\quad -I+v_{i,n-1}v_{i,n-1}^\top+2\sum_{j=1}^{i-1}v_{j,n-1}v_{j,n-1}^\top\Bigg\|\big\|J(x(t_{n-1}))v_i(t_{n-1})\big\| \\ &\quad +\tau\gamma\Bigg\|I-v_{i,n-1}v_{i,n-1}^\top-2\sum_{j=1}^{i-1}v_{j,n-1}v_{j,n-1}^\top\Bigg\| \\ &\quad \times\big\|J(x(t_{n-1}))v_i(t_{n-1})-\hat{J}(x_{n-1})v_{i,n-1}\big\|+Q\tau^2. \end{aligned}$$

Direct calculations show that

$$\begin{aligned} &\big\|J(x(t_{n-1}))v_i(t_{n-1})-\hat{J}(x_{n-1})v_{i,n-1}\big\| \\ \leq& \big\|J(x(t_{n-1}))(v_i(t_{n-1})-v_{i,n-1})\big\| \\ &\quad +\big\|(J(x(t_{n-1}))-J(x_{n-1})+J(x_{n-1})-\hat{J}(x_{n-1}))v_{i,n-1}\big\| \\ \leq& Q\|e_{n-1}^{v_i}\|+Q\|e_{n-1}^x\|+\|J(x_{n-1})-\hat{J}(x_{n-1})\|, \end{aligned} \tag{3.3}$$

and we incorporate the above two equations to obtain

$$\|e_n^{v_i}\| \leq \|e_{n-1}^{v_i}\| + Q\tau\|e_{n-1}^x\| + Q\tau\sum_{j=1}^{i}\|e_{n-1}^{v_j}\| + Q\tau\|J(x_{n-1}) - \hat{J}(x_{n-1})\| + Q\tau^2.$$

Adding this equation from $i=1$ to k and denoting

$$E_n^v := \sum_{i=1}^{k}\|e_n^{v_i}\| \quad \text{for} \quad 1 \leq n \leq N,$$

yield a relation in terms of E_n^v

$$E_n^v \leq E_{n-1}^v + Q\tau\|e_{n-1}^x\| + Q\tau E_{n-1}^v + Q\tau\|J(x_{n-1}) - \hat{J}(x_{n-1})\| + Q\tau^2.$$

Adding this equation from $n=1$ to n_* leads to

$$E_{n_*}^v \leq Q\tau\sum_{n=1}^{n_*}E_{n-1}^v + Q\tau\sum_{n=1}^{n_*}\|e_{n-1}^x\| + Q\tau\sum_{n=1}^{n_*}\|J(x_{n-1}) - \hat{J}(x_{n-1})\| + Q\tau.$$

Then an application of the discrete Gronwall's inequality [31, Lemma 11.2] leads to

$$E_{n_*}^v \leq Q\tau\sum_{n=1}^{n_*}\|e_{n-1}^x\| + Q\tau\sum_{n=1}^{n_*}\|J(x_{n-1}) - \hat{J}(x_{n-1})\| + Q\tau \tag{3.4}$$

for $1 \leq n_* \leq N$.

We then subtract the equations of the state variable in (2.1) and (2.2) to obtain

$$\begin{aligned} e_n^x = e_{n-1}^x + \tau\beta\Bigg[&\left(I - 2\sum_{j=1}^{k}v_j(t_{n-1})v_j(t_{n-1})^\top\right)F(x(t_{n-1})) \\ &- \left(I - 2\sum_{j=1}^{k}v_{j,n-1}v_{j,n-1}^\top\right)\hat{F}(x_{n-1})\Bigg] + \mathcal{O}(\tau^2), \end{aligned} \tag{3.5}$$

and the key is to bound the difference in the above equation as follows

$$\begin{aligned} &\left\|\left(I - 2\sum_{j=1}^{k}v_j(t_{n-1})v_j(t_{n-1})^\top\right)F(x(t_{n-1})) - \left(I - 2\sum_{j=1}^{k}v_{j,n-1}v_{j,n-1}^\top\right)\hat{F}(x_{n-1})\right\| \\ \leq &\left\|\left(I - 2\sum_{j=1}^{k}v_j(t_{n-1})v_j(t_{n-1})^\top - I + 2\sum_{j=1}^{k}v_{j,n-1}v_{j,n-1}^\top\right)F(x(t_{n-1}))\right\| \\ &+ \left\|\left(I - 2\sum_{j=1}^{k}v_{j,n-1}v_{j,n-1}^\top\right)\big(F(x(t_{n-1})) - \hat{F}(x_{n-1})\big)\right\| \end{aligned}$$

$$
\begin{aligned}
&\leq Q\sum_{j=1}^{k}\|e_{n-1}^{v_j}\|+\|F(x(t_{n-1}))-F(x_{n-1})\|+\|F(x_{n-1})-\hat{F}(x_{n-1})\| \\
&\leq QE_{n-1}^{v}+Q\|e_{n-1}^{x}\|+\|F(x_{n-1})-\hat{F}(x_{n-1})\|.
\end{aligned}
$$

We incorporate the above two equations to obtain

$$
\|e_n^x\|\leq\|e_{n-1}^x\|+Q\tau E_{n-1}^v+Q\tau\|e_{n-1}^x\|+Q\tau\|F(x_{n-1})-\hat{F}(x_{n-1})\|+Q\tau^2.
$$

Summing this equation from $n=1$ to m we obtain

$$
\|e_m^x\|\leq Q\tau\sum_{n=1}^{m}E_{n-1}^v+Q\tau\sum_{n=1}^{m}\|e_{n-1}^x\|+Q\tau\sum_{n=1}^{m}\|F(x_{n-1})-\hat{F}(x_{n-1})\|+Q\tau.
$$

We invoke (3.4) and apply

$$
\begin{aligned}
\sum_{n=1}^{m}E_{n-1}^v&\leq\sum_{n=1}^{m}\left[Q\tau\sum_{q=1}^{n-1}\|e_{q-1}^x\|+Q\tau\sum_{q=1}^{n-1}\|J(x_{q-1})-\hat{J}(x_{q-1})\|+Q\tau\right] \\
&\leq Q\sum_{n=1}^{m-1}\|e_{n-1}^x\|+Q\sum_{n=1}^{m-1}\|J(x_{n-1})-\hat{J}(x_{n-1})\|+Q
\end{aligned}
$$

to reach

$$
\begin{aligned}
\|e_m^x\|\leq{}& Q\tau\sum_{n=1}^{m}\|e_{n-1}^x\|+Q\tau\sum_{n=1}^{m-1}\|J(x_{n-1})-\hat{J}(x_{n-1})\| \\
&+Q\tau\sum_{n=1}^{m}\|F(x_{n-1})-\hat{F}(x_{n-1})\|+Q\tau.
\end{aligned}
$$

An application of the discrete Gronwall's inequality leads to

$$
\|e_m^x\|\leq Q\tau\sum_{n=1}^{m-1}\|J(x_{n-1})-\hat{J}(x_{n-1})\|+Q\tau\sum_{n=1}^{m}\|F(x_{n-1})-\hat{F}(x_{n-1})\|+Q\tau,
$$

that is,

$$
\frac{\|e_m^x\|-Q\tau}{Q\tau}\leq\sum_{n=1}^{m-1}\|J(x_{n-1})-\hat{J}(x_{n-1})\|+\sum_{n=1}^{m}\|F(x_{n-1})-\hat{F}(x_{n-1})\|.
$$

Then we apply the **Assumptions B** to obtain

$$
\begin{aligned}
&P\left(\frac{\|e_m^x\|-Q\tau}{Q\tau}\leq(2m-1)\varepsilon\right)\\
\geq&P\left(\sum_{n=1}^{m-1}\|J(x_{n-1})-\hat{J}(x_{n-1})\|+\sum_{n=1}^{m}\|F(x_{n-1})-\hat{F}(x_{n-1})\|\leq(2m-1)\varepsilon\right)\geq 1-\delta,
\end{aligned}
$$

that is,

$$
P\big(\|e_m^x\|\leq Q\tau(2m-1)\varepsilon+Q\tau\big)\geq 1-\delta.
$$

As $\tau(2m-1)\leq 2T$, we reach the conclusion (3.1) that completes the proof of this theorem. □

4 Extension to constrained saddle dynamics

In many physical processes such as the Thomson problem [35] and the Bose–Einstein condensation [1], the energy functional is constrained on the high-dimensional unit sphere. To compute saddle points of such constrained problems, the constrained high-index saddle dynamics for an index-k saddle point on the unit sphere S^{d-1} is proposed in [39]

$$
\left\{
\begin{aligned}
&\frac{dx}{dt}=\Big(I-xx^\top-2\sum_{j=1}^{k}v_jv_j^\top\Big)F(x),\\
&\frac{dv_i}{dt}=\Big(I-xx^\top-v_iv_i^\top-2\sum_{j=1}^{i-1}v_jv_j^\top\Big)J(x)v_i+\beta xv_i^\top F(x),\quad 1\leq i\leq k.
\end{aligned}
\right.
\tag{4.1}
$$

We discretize the first-order derivative by the explicit Euler scheme to get the reference equations for constrained high-index saddle dynamics (4.1)

$$
\left\{
\begin{aligned}
x(t_n)=&x(t_{n-1})+\tau\Big(I-x(t_{n-1})x(t_{n-1})^\top\\
&-2\sum_{j=1}^{k}v_j(t_{n-1})v_j(t_{n-1})^\top\Big)F(x(t_{n-1}))+\mathcal{O}(\tau^2),\\
v_i(t_n)=&v_i(t_{n-1})+\tau\Big(I-x(t_{n-1})x(t_{n-1})^\top-v_i(t_{n-1})v_i(t_{n-1})^\top\\
&-2\sum_{j=1}^{i-1}v_j(t_{n-1})v_j(t_{n-1})^\top\Big)J(x(t_{n-1}))v_i(t_{n-1})\\
&+\tau x(t_{n-1})v_i(t_{n-1})^\top F(x(t_{n-1}))+\mathcal{O}(\tau^2),\quad 1\leq i\leq k.
\end{aligned}
\right.
\tag{4.2}
$$

Then we drop the truncation errors and take account of the inaccuracy of F and J to obtain a first-order scheme of (4.1)

$$\begin{cases} \tilde{x}_n = x_{n-1} + \tau\Big(I - x_{n-1}x_{n-1}^\top - 2\sum_{j=1}^{k} v_{j,n-1}v_{j,n-1}^\top\Big)\hat{F}(x_{n-1}), \\ x_n = \dfrac{\tilde{x}_n}{\|\tilde{x}_n\|}, \\ \tilde{v}_{i,n} = v_{i,n-1} + \tau\Big(I - x_{n-1}x_{n-1}^\top - v_{i,n-1}v_{i,n-1}^\top - 2\sum_{j=1}^{i-1} v_{j,n-1}v_{j,n-1}^\top\Big)\hat{J}(x_{n-1})v_{i,n-1} \\ \qquad + \tau x_{n-1}v_{i,n-1}^\top \hat{F}(x_{n-1}), \qquad 1\le i\le k, \\ \hat{v}_{i,n} = \tilde{v}_{i,n} - \tilde{v}_{i,n}^\top x_n x_n, \qquad 1\le i\le k, \\ v_{i,n} = \dfrac{1}{Y_{i,n}}\Big(\hat{v}_{i,n} - \sum_{j=1}^{i-1}(\hat{v}_{i,n}^\top v_{j,n})v_{j,n}\Big), \quad 1\le i\le k, \end{cases} \tag{4.3}$$

for $1\le n\le N$ and

$$x_0 = x(0), \quad v_{i,0} = v_i(0), \quad Y_{i,n} := \Big(\|\hat{v}_{i,n}\|^2 - \sum_{j=1}^{i-1}(\hat{v}_{i,n}^\top v_{j,n})^2\Big)^{1/2}, \quad 1\le i\le k,$$

such that

$$x_0^\top v_{i,0} = 0 \quad \text{and} \quad v_{i,0}^\top v_{j,0} = \delta_{i,j} \quad \text{for} \quad 1\le i,j\le k.$$

The second equation of (4.3) represents the retraction in order to ensure that $x_n \in S^{d-1}$. The last two equations of (4.3), which stand for the vector transport and the Gram-Schmidt orthonormalization procedure [39], respectively, aim to ensure the following properties as for the continuous problem

$$v_{i,n}^\top x_n = 0, \quad v_{i,n}^\top v_{j,n} = \delta_{ij}, \quad 1\le i,j\le k, \quad 0\le n\le N. \tag{4.4}$$

Compared with the numerical scheme (2.2) for the unconstrained high-index saddle dynamics, additional operations such as the retraction and vector transport in (4.3) caused from the sphere constraint make this scheme and the corresponding analysis more complicated.

We first introduce auxiliary estimates to support the error estimates. The following lemma could be proved by exactly the same procedure as [54, Lemmas 3.1-3.4] and thus we only present the results without proof.

Lemma 4.1. *Under the **Assumptions** A-B, the following estimates hold for τ sufficiently small*

$$\begin{aligned}
&\|x_n-\tilde{x}_n\|=\big|1-\|\tilde{x}_n\|\big|\le Q\tau^2, && 1\le n\le N,\\
&\|\hat{v}_{i,n}-\tilde{v}_{i,n}\|\le Q\tau^2, && 1\le i\le k, \quad 1\le n\le N,\\
&\|v_{i,n}-\hat{v}_{i,n}\|\le Q\tau^2, && 1\le i\le k, \quad 1\le n\le N.
\end{aligned}$$

Here Q depends on Q_0 but is independent from τ, n and N.

We then derive error estimates for the numerical scheme (4.3).

Theorem 4.1. *Suppose the **Assumptions** A-B hold. Then the following probabilistic error estimate holds for the scheme (4.3) for τ sufficiently small and for some $Q>0$*

$$P\big(\|x(t_n)-x_n\|\le Q\varepsilon+Q\tau\big)\ge 1-\delta, \quad 1\le n\le N. \tag{4.5}$$

Here Q depends on k, L, T, Q_0 but is independent from τ, n, N, ε and δ.

Proof. In general the proof could be performed following that of Theorem 3.1 and we thus only provide a sketch. We first subtract the equations of directional vectors in (4.2) and (4.3) and apply Lemma 4.1, which implies

$$v_i(t_n)-\tilde{v}_{i,n}=(v_i(t_n)-v_{i,n})+(v_{i,n}-\hat{v}_{i,n})+(\hat{v}_{i,n}-\tilde{v}_{i,n})=e_n^{v_i}+\mathcal{O}(\tau^2)$$

to obtain

$$\begin{aligned}
e_n^{v_i}=&e_{n-1}^{v_i}+\tau\Bigg(I-x(t_{n-1})x(t_{n-1})^\top-v_i(t_{n-1})v_i(t_{n-1})^\top\\
&-2\sum_{j=1}^{i-1}v_j(t_{n-1})v_j(t_{n-1})^\top\Bigg)J(x(t_{n-1}))v_i(t_{n-1})\\
&-\tau\Bigg(I-x_{n-1}x_{n-1}^\top-v_{i,n-1}v_{i,n-1}^\top-2\sum_{j=1}^{i-1}v_{j,n-1}v_{j,n-1}^\top\Bigg)\hat{J}(x_{n-1})v_{i,n-1}\\
&+\tau x(t_{n-1})v_i(t_{n-1})^\top F(x(t_{n-1}))-\tau x_{n-1}v_{i,n-1}^\top\hat{F}(x_{n-1})+\mathcal{O}(\tau^2).
\end{aligned}$$

Compared with (3.2), the newly encountered differences are

$$\tau x(t_{n-1})x(t_{n-1})^\top J(x(t_{n-1}))v_i(t_{n-1})-\tau x_{n-1}x_{n-1}^\top\hat{J}(x_{n-1})v_{i,n-1}$$

and

$$\tau x(t_{n-1})v_i(t_{n-1})^\top F(x(t_{n-1}))-\tau x_{n-1}v_{i,n-1}^\top\hat{F}(x_{n-1}),$$

which could be bounded by the splitting techniques as (3.3). Then we follow the derivations in (3.2)–(3.4) to reach a similar estimate as (3.4)

$$E^v_{n_*} \leq Q\tau \sum_{n=1}^{n_*} \|e^x_{n-1}\| + Q\tau \sum_{n=1}^{n_*} \|J(x_{n-1}) - \hat{J}(x_{n-1})\| \\ + Q\tau \sum_{n=1}^{n_*} \|F(x_{n-1}) - \hat{F}(x_{n-1})\| + Q\tau \tag{4.6}$$

for $1 \leq n_* \leq N$.

We could similarly subtract the equations of state variables in (4.2) and (4.3) and apply

$$x(t_n) - \tilde{x}_n = (x(t_n) - x_n) + (x_n - \tilde{x}_n) = e^x_n + \mathcal{O}(\tau^2)$$

to get the error equation

$$e^x_n = e^x_{n-1} + \tau \Bigg[\bigg(I - x(t_{n-1})x(t_{n-1})^\top - 2\sum_{j=1}^{k} v_j(t_{n-1})v_j(t_{n-1})^\top \bigg) F(x(t_{n-1})) \\ - \bigg(I - x_{n-1}x_{n-1}^\top - 2\sum_{j=1}^{k} v_{j,n-1}v_{j,n-1}^\top \bigg) \hat{F}(x_{n-1}) \Bigg] + \mathcal{O}(\tau^2).$$

Compared with (3.5), the newly encountered difference is

$$\tau x(t_{n-1})x(t_{n-1})^\top F(x(t_{n-1})) - \tau x_{n-1}x_{n-1}^\top \hat{F}(x_{n-1}),$$

which could be bounded by the splitting techniques as (3.3). Then we could estimate this error equation as that in the proof of Theorem 3.1 based on (4.6) and **Assumptions A-B** to complete the proof. □

5 Concluding remarks

We prove probabilistic error estimates for high-index saddle dynamics without constraints or with sphere constraint in order to account for the inaccurate values of the model, which could be encountered in various scenarios such as model uncertainties or surrogate model algorithms via machine learning methods. Therefore, the current study serves as a generalization of conventional numerical analysis results for deterministic high-index saddle dynamics.

There are potential extensions of the current work that deserve further exploration. For instance, the dimer method [16] with the dimer length l could be used in (2.2) and (4.3) to approximate the product of the Hessian matrix and the vector, i.e.,

$$J(x)v_i \approx \frac{F(x+lv_i)-F(x-lv_i)}{2l}$$

for efficient computation and storage, which leads to the shrinking-dimer high-index saddle dynamics [43,53]. Under this approximation, only the values of F are required in high-index saddle dynamics, which could simplify the implementation. However, this introduce additional errors that are not easy to be estimated due to, e.g., the low regularity of F generated from machine learning methods. Furthermore, it is meaningful but challenging to extend the ideas and techniques to analyze probabilistic error estimates for the numerical scheme for high-index saddle dynamics constrained by m equalities [39, Eq. 24]

$$\begin{cases} \dfrac{dx}{dt} = \left(I - 2\sum_{j=1}^{k} v_j v_j^\top\right) F(x), \\ \dfrac{dv_i}{dt} = \left(I - v_i v_i^\top - 2\sum_{j=1}^{i-1} v_j v_j^\top\right) \mathcal{H}(x)[v_i] \\ \qquad\qquad -A(x)\big(A(x)^\top A(x)\big)^{-1} \left(\nabla^2 c(x) \dfrac{dx}{dt}\right)^\top v_i, \quad 1 \le i \le k. \end{cases} \tag{5.1}$$

Here $c(x) = (c_1(x), \cdots, c_m(x)) = 0$ represents the m equality constraints and $A(x) = (\nabla c_1(x), \cdots, \nabla c_m(x))$. The sphere-constrained high-index saddle dynamics (4.1) is a special case of (5.1) with one equality constraint

$$c_1(x) = \|x\| - 1 = 0.$$

In the generalized constrained saddle dynamics (5.1), $\mathcal{H}(x)$ refers to the Riemannian Hessian [39], which is difficult to compute and approximate in practice.

Another interesting topic in optimization algorithms lies in performing the asymptotic stability analysis of the discretization. In [24], the asymptotic convergence rate is proved for the explicit scheme of the high-index saddle dynamics (1.1), following which the asymptotic convergence rate is proved for the explicit scheme of the accelerated high-index saddle dynamics that contains an additional momentum term in the dynamics of x [25]. Based on the methods in [24,25], the asymptotic stability analysis of the discretizations in this work will be considered in the future.

Apart from the explicit schemes, the semi-implicit schemes are developed and analyzed recently for high-index saddle dynamics (1.1) [26] and the constrained high-index saddle dynamics (4.1) [56], respectively. As shown in numerical experiments in [26], the semi-implicit scheme is more efficient than the explicit scheme in terms of the computational time and the number of queries of the model value. Thus we will combine the analysis techniques in [26, 56] and the current work to perform probabilistic error estimates of the semi-implicit schemes for the (constrained) high-index saddle dynamics with inaccurate models.

In the current setting, the x is assumed to be a d-dimensional vector with a finite dimension d. A more generalized version is the PDE case of the high-index saddle dynamics, i.e., the x becomes a vector of infinite dimension. For this situation, it seems that the developed analysis still works if we could impose suitable boundedness and Lipschitz assumptions on the model values like the **Assumptions A-B**. However, the validity of these assumptions for infinite dimensional problems remains to be considered and illustrated, which may lead to difficulties. A careful analysis is required to explore this interesting scenario.

Finally, there exists some numerical examples in [55] to substantiate the effectiveness of the surrogate-model based algorithm for high-index saddle dynamics without explicit expression of the model, a typical scenario that the model is inaccurate. However, to our best knowledge, numerical experiments to demonstrate the numerical analysis results in this work are not available in the literature. We will consider this interesting and meaningful topic in future works.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Nos. 12225102, T2321001, 12288101 and 12301555), the National Key R&D Program of China (Nos. 2021YFF1200500 and 2023YFA1008903), and the Taishan Scholars Program of Shandong Province (No. tsqn202306083).

References

[1] W. Bao and Y. Cai, Mathematical theory and numerical methods for Bose–Einstein condensation, Kinet. Relat. Models, 6 (2013), 1–135.

[2] C. Chen and Z. Xie, Search extension method for multiple solutions of a nonlinear problem, Comput. Math. Appl., 47 (2004), 327–343.

[3] A. Denzel and J. Kastner, Gaussian process regression for transition state search, J. Chem. Theory Comput., 14 (2018), 5777–5786.

[4] J. Doye and D. Wales, Saddle points and dynamics of Lennard-Jones clusters, solids, and supercooled liquids, J. Chem. Phys., 116 (2002), 3777–3788.

[5] W. E and X. Zhou, The gentlest ascent dynamics, Nonlinearity, 24 (2011), 1831–1842.

[6] W. E and E. Vanden-Eijnden, Transition-path theory and path-finding algorithms for the study of rare events, Annu. Rev. Phys. Chem., 61 (2010), 391–420.

[7] B. Eckhardt, Irregular scattering, Phys. D, 33 (1988), 89–98.

[8] P. E. Farrell, Á. Birkisson, and S. W. Funke, Deflation techniques for finding distinct solutions of nonlinear partial differential equations, SIAM J. Sci. Comput., 37 (2015), A2026–A2045.

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

[10] W. Grantham, Gradient transformation trajectory following algorithms for determining stationary min-max saddle points, in Advances in Dynamic Game Theory, Ann. Internat. Soc. Dynam. Games 9, Birkhauser Boston, Boston, MA, 2007, 639–657.

[11] N. Gould, C. Ortner and D. Packwood, A dimer-type saddle search algorithm with preconditioning and linesearch, Math. Comput., 85 (2016), 2939–2966.

[12] S. Gu, H. Wang and X. Zhou, Active learning for transition state calculation, J. Sci. Comput., 93 (2022), 78.

[13] Y. Han, J. Yin, P. Zhang, A. Majumdar, and L. Zhang, Solution landscape of a reduced Landau–de Gennes model on a hexagon, Nonlinearity, 34 (2021), 2048–2069.

[14] Y. Han, J. Yin, Y. Hu, A. Majumdar and L. Zhang, Solution landscapes of the simplified Ericksen-Leslie model and its comparison with the reduced Landau-de Gennes model, Proceedings of the Royal Society A, 477 (2021), 20210458.

[15] D. Heidrich and W. Quapp, Saddle points of index 2 on potential energy surfaces and their role in theoretical reactivity investigations, Theor. Chim. Acta, 70 (1986), 89–98.

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

[17] O. Koistinen, F. Dagbjartsdóttir, V. Ásgeirsson, A. Vehtari and H. Jónsson, Nudged elastic band calculations accelerated with Gaussian process regression, J. Chem. Phys., 147 (2017), 152720.

[18] A. Lederer, J. Umlauft and S. Hirche, Uniform error bounds for Gaussian process regression with application to safe control, in Proceedings of 33rd Conference on Neural Information Processing Systems, 657–667, Vancouver, Canada, 2019.

[19] A. Levitt and C. Ortner, Convergence and cycling in walker-type saddle search algorithms, SIAM J. Numer. Anal., 55 (2017), 2204–2227.

[20] Y. Li and J. Zhou, A minimax method for finding multiple critical points and its applications to semilinear PDEs, SIAM J. Sci. Comput., 23 (2001), 840–865.

[21] Z. Li and J. Zhou, A local minimax method using virtual geometric objects: Part II–For finding equality constrained saddles, J. Sci. Comput., 78 (2019), 226–245.

[22] H. Liu, J. Cai and Y. Ong, Remarks on multi-output Gaussian process regression, Knowledge-Based Syst., 144 (2018), 102–121.

[23] W. Liu, Z. Xie and W. Yi, Normalized Goldstein-type local minimax method for finding multiple unstable solutions of semilinear elliptic PDEs, Commun. Math. Sci.,

19 (2021), 147–174.

[24] Y. Luo, X. Zheng, X. Cheng and L. Zhang, Convergence analysis for discrete high-index saddle dynamics, SIAM J. Numer. Anal., 60 (2022), 2731–2750.

[25] Y. Luo, X. Zheng and L. Zhang, Accelerated high-index saddle dynamics method for searching high-index saddle points, arXiv:2310.04656.

[26] Y. Luo, L. Zhang, P. Zhang, Z. Zhang and X. Zheng, Semi-implicit method of high-index saddle dynamics and application to construct solution landscape, arXiv:2310.04643.

[27] D. Mehta, Finding all the stationary points of a potential-energy landscape via numerical polynomial-homotopy-continuation method, Phys. Rev. E, 84 (2011), 025702.

[28] J. W. Milnor, Morse Theory, Princeton University Press, 1963.

[29] R. M. Minyaev, W. Quapp, G. Subramanian, P. R. Schleyer and Y. Ho, Internal conrotation and disrotation in H2BCH2BH2 and diborylmethane 1, 3 H exchange, J. Comput. Chem., 18 (1997), 1792–1803.

[30] Q. Nie, L. Qiao, Y. Qiu, L. Zhang and W. Zhao, Noise control and utility: from regulatory network to spatial patterning, Sci. China Math., 63 (2020), 425–440.

[31] A. Quarteroni, R. Sacco and F. Saleri, Numerical Mathematics, Texts in Applied Mathematics 37, New York, Springer, 2007.

[32] M. Raissi, P. Perdikaris and G. Karniadakis, Machine learning of linear differential equations using Gaussian processes, J. Comput. Phys., 348 (2017), 683–693.

[33] C. Rasmussen and C. Williams, Gaussian Processes for Machine Learning, MIT Press, Cambridge, MA, 2006.

[34] A. Stuart and A. Teckentrup, Posterior consistency for Gaussian process approximations of Bayesian posterior distributions, Math. Comput., 87 (2018), 721–753.

[35] J. Thomson, XXIV. On the structure of the atom: an investigation of the stability and periods of oscillation of a number of corpuscles arranged at equal intervals around the circumference of a circle; with application of the results to the theory of atomic structure, London, Edinburgh, Dublin Phil. Mag. J. Sci., 7 (1904), 237–265.

[36] W. Wang, L. Zhang and P. Zhang, Modelling and computation of liquid crystals, Acta Numer., 30 (2021), 765–851.

[37] Z. Xie, C. Chen and Y. Xu, An improved search-extension method for computing multiple solutions of semilinear PDEs, IMA J. Numer. Anal., 25 (2005), 549–576.

[38] Z. Xu, Y. Han, J. Yin, B. Yu, Y. Nishiura and L. Zhang, Solution landscapes of the diblock copolymer-homopolymer model under two-dimensional confinement, Phys. Rev. E, 104 (2021), 014505.

[39] J. Yin, Z. Huang and L. Zhang, Constrained high-index saddle dynamics for the solution landscape with equality constraints, J. Sci. Comput., 91 (2022), 62.

[40] J. Yin, Y. Wang, J. Chen, P. Zhang and L. Zhang, Construction of a pathway map on a complicated energy landscape, Phys. Rev. Lett., 124 (2020), 090601.

[41] J. Yin, K. Jiang, A.-C. Shi, P. Zhang and L. Zhang, Transition pathways connecting crystals and quasicrystals, Proc. Natl. Acad. Sci., 118 (2021), e2106230118.

[42] J. Yin, B. Yu and L. Zhang, Searching the solution landscape by generalized high-

index saddle dynamics, Sci. China Math., 64 (2021), 1801.
[43] J. Yin, L. Zhang and P. Zhang, High-index optimization-based shrinking dimer method for finding high-index saddle points, SIAM J. Sci. Comput., 41 (2019), A3576–A3595.
[44] J. Yin, L. Zhang and P. Zhang, Solution landscape of the Onsager model identifies non-axisymmetric critical points, Phys. D, 430 (2022), 133081.
[45] B. Yu, X. Zheng, P. Zhang and L. Zhang, Computing solution landscape of nonlinear space-fractional problems via fast approximation algorithm, J. Comput. Phys., 468 (2022), 111513.
[46] H. Yu, G. Ji and P. Zhang, A nonhomogeneous kinetic model of liquid crystal polymers and its thermodynamic closure approximation, Commun. Comput. Phys., 7 (2010), 383–402.
[47] J. Zhang and Q. Du, Shrinking dimer dynamics and its applications to saddle point search, SIAM J. Numer. Anal., 50 (2012), 1899–1921.
[48] L. Zhang, L. Chen and Q. Du, Morphology of critical nuclei in solid-state phase transformations, Phys. Rev. Lett., 98 (2007), 265703.
[49] L. Zhang, L. Chen and Q. Du, Simultaneous prediction of morphologies of a critical nucleus and an equilibrium precipitate in solids, Commun. Comput. Phys., 7 (2010), 674–682.
[50] L. Zhang, Q. Du and Z. Zheng, Optimization-based shrinking dimer method for finding transition states, SIAM J. Sci. Comput., 38 (2016), A528–A544.
[51] L. Zhang, W. Ren, A. Samanta and Q. Du, Recent developments in computational modelling of nucleation in phase transformations, NPJ Comput. Mater., 2 (2016), 16003.
[52] L. Zhang, P. Zhang and X. Zheng, Error estimates of Euler discretization of high-index saddle dynamics, SIAM J. Numer. Anal., 60 (2022), 2925–2944.
[53] L. Zhang, P. Zhang and X. Zheng, Mathematical and numerical analysis to shrinking-dimer saddle dynamics with local Lipschitz conditions, CSIAM Trans. Appl. Math., 4 (2023), 157–176.
[54] L. Zhang, P. Zhang and X. Zheng, Discretization and index-robust error analysis for constrained high-index saddle dynamics on high-dimensional sphere, Sci. China Math., 66 (2023), 2347–2360.
[55] L. Zhang, P. Zhang and X. Zheng, A model-free shrinking-dimer saddle dynamics for finding saddle point and solution landscape, Japan J. Indust. Appl. Math., 40 (2023), 1677–1693.
[56] L. Zhang, P. Zhang and X. Zheng, Error estimate for semi-implicit method of sphere-constrained high-index saddle dynamics, Chin. Ann. Math. Ser. B, 44 (2023), 765–780.