SIAM J. SCI. COMPUT.
Vol. 38, No. 1, pp. A528–A544




# OPTIMIZATION-BASED SHRINKING DIMER METHOD FOR FINDING TRANSITION STATES*

LEI ZHANG†, QIANG DU‡, AND ZHENZHEN ZHENG§

**Abstract.** We present a new and efficient optimization-based shrinking dimer (OSD) algorithm for finding transition states associated with a given energy landscape. The transition states are given by index-one saddle points of the energy surface. By constructing the rotation and translation steps in the classical dimer method under an optimization framework, we are able to take advantage of more powerful optimization methods to substantially speed up the computation of transition states. Specifically, the Barzilai–Borwein gradient method is proposed as an effective implementation of OSD. We show that the OSD method is the generalized formulation of the original shrinking dimer dynamics (SDD) and enjoys superlinear convergence. We test various numerical examples, including some standard lower dimensional test examples, a cluster of seven particles, seven-atom island on the (111) surface of an FCC crystal, and nucleation in phase transformations. The results demonstrate that the OSD method with the proposed Barzilai–Borwein stepsize is very effective and provides a more efficient implementation than the SDD algorithm and some other dimer-type methods. Our new algorithms offer potential to significantly advance the state-of-the-art in dimer-type methods for transition state search.

**Key words.** saddle point, transition state, shrinking dimer dynamics, dimer method, Barzilai and Borwein gradient method

**AMS subject classifications.** 37M05, 49K35, 37N30, 34K28, 65P99

**DOI.** 10.1137/140972676

## 1. Introduction.

Finding transition states on a complex energy landscape that are given by index-one saddle points of the energy is a challenging problem that arises in various scientific fields like physics, chemistry, biology, and materials science. It has attracted much attention in many interesting applications such as chemical kinetics [1, 29], protein folding [3, 17, 32], nucleation during phase transformations [7, 36, 37], and so on. Extensive studies have focused on developing or refining methods for more efficient and more reliable saddle points search. Typically, most methods can be divided into two classes: chain-of-state methods and surface walking methods. The former is suitable when both the initial and final points on the energy surface are available, and it employs a discrete string connecting the initial and final points to find the minimum energy path and the saddle points. The most notable examples are the nudged elastic band method [21] and the string method [12]. The latter class utilizes, single point with the local information such as the gradient or Hessian to search saddle points. The representative methods include the gentlest ascent method [8, 13, 15], the eigenvector-following method [6], the activation-relaxation technique

*Submitted to the journal's Methods and Algorithms for Scientific Computing section June 12, 2014; accepted for publication (in revised form) November 24, 2015; published electronically February 18, 2016. This research was partially supported by U.S. NSF-DMS-1318586 and China NSFC 11421110001 and 91430217.

http://www.siam.org/journals/sisc/38-1/97267.html

†Beijing International Center for Mathematical Research, Center for Quantitative Biology, Peking University, Beijing 100871, China (zhangl@math.pku.edu.cn).

‡Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10027 (qd2125@columbia.edu).

§Beijing Computational Science Research Center, Beijing 100084, China. Current address: Science China Press, Beijing 100717, China (zhengzz@gmail.com).

[4, 26, 23], the step and slide method [25], the dimer method [20, 21, 27], the shrinking dimer dynamics (SDD) [34], and so on.

Zhang and Du proposed the SDD in [34] to compute index-one saddle points, which may be viewed as the continuous limit of the original dimer method popularized by [20]. By introducing a dynamic reduction of the dimer length, the SDD gradually reduces the length to zero when the search is close to the saddle point so that the converged steady state solution of the new dynamic system is consistent with the saddle point. Furthermore, the constrained SDD was developed in [35] to search for constrained index-one saddle points on the energy surface subject to general equality constraints, and then the generalized SDD was proposed in [38] as a new formulation of an SDD analogue of the Cahn–Hilliard dynamics associated with the free energy. The time discretization of the original SDD in [34] can be viewed as a modified gradient search. The first-order Euler-types schemes are implemented, and linear convergence and optimal stepsize are achieved in [34] when constant time steps are taken. It is obvious that adaptive time steps may improve the search efficiency, and the key is to establish a sensible adaptivity criteria. It is interesting to explore if a more sophisticated optimization-based strategy can be of any help.

Barzilai and Borwein [2] proposed a two-point stepsize gradient (BB) method for unconstrained optimization problems. For positive definite quadratic objective function in two dimensions, Barzilai and Borwein established the R-superlinear convergence of the BB method. For the general $n$-dimensional strictly convex quadratic functions, Raydan [30] proved that the BB method is globally convergent. In [9], Dai and Liao proved that the BB methods with two stepsize formulas are R-linearly convergent in any dimension for strictly convex quadratic functions and strongly convex quadratic functions, respectively. Due to its simplicity and numerical efficiency, the BB method has been widely used for large-scale unconstrained minimization problems. However, so far no attempts have been made to apply the BB method in the saddle point problems.

Inspired by the SDD, we propose a novel optimization-based shrinking dimer (OSD) method by utilizing the stepsize selection given by the BB gradient method. That is, we first formulate the dimer rotation and translation steps in the original dimer method to respective optimization problems and then utilize the BB gradient method to pick the stepsize in the OSD. The resulting algorithm substantially improves the local convergence of the saddle point search process over the original implementation of SDD with a constant stepsize and yields a method that converges to index-one saddle points (transition states on the energy landscape). The superlinear convergence is analyzed for a representative functional and is confirmed in numerous numerical tests, including some standard lower dimensional benchmark examples such as a problem for a cluster of seven particles with Lennard–Jones interactions, and the nonclassical nucleation in phase transformations. Comparisons with some existing dimer-type methods such as the classical dimer methods [21] and the recent improved dimer method [16] are also made, demonstrating the significant improvement of the OSD algorithm in computational efficiency. This work leads to new advances in the state of the art in using dimer-type approaches for transition state search and also opens up new possibilities of adopting other popular optimization-based strategies for further improvement.

The paper is organized as follows. In section 2, we present a full description of the OSD method and use the BB stepsize formula for the implementation of the OSD. Stability analysis and a comparison of OSD and SDD are also addressed in this

section. Various numerical examples are presented to illustrate the proposed method in section 3. Further discussion and final conclusion are given in section 4.

## 2. Optimization-based shrinking dimer method.

**2.1. Formulation of the OSD.** Given an energy functional $E$ on a Hilbert space $\mathcal{H}$ which is compactly embedded in a Hilbert space $\mathcal{L}$, we let $\nabla E(x)$ denote the gradient of $E$ at $x \in \mathcal{H}$ defined in the Fréchet sense with respect to an inner product in $\mathcal{L}$.

The dimer is a pair of points $x_1$ and $x_2$ in $\mathcal{H}$ with the 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 defined by

$$x_\alpha = (1-\alpha)x_1 + \alpha x_2, \tag{2.1}$$

where the parameter $\alpha \in [0,1]$ gives us the freedom to choose a point other than the geometric center (the midpoint of the dimer corresponding to $\alpha = 1/2$). It follows from (2.1) that $x_1 = x_\alpha + \alpha l v$ and $x_2 = x_\alpha + (\alpha - 1)lv$. For notation convenience, let

$$F_i = -\nabla E(x_i), \quad i = 1, 2,$$

be natural forces at the two endpoints of the dimer, and let

$$F = (1-\alpha)F_1 + \alpha F_2$$

be the approximated natural force at the dimer rotation center.

In the original dimer method [20], each time the dimer follows a rotation step, i.e., the dimer is rotated toward the minimum energy configuration that is equivalent to finding its lowest curvature mode, and a translation step, i.e., it allows the dimer to climb up to the energy ascent direction characterized by the orientation vector $v$. Motivated by the gentlest ascent dynamics [13] and the SDD [34], we formulate the dimer rotation and translation steps to the corresponding optimization problems so that we can take advantage of the optimization methods to compute the saddle points more efficiently.

For the dimer rotation, we need to compute the direction corresponding to the smallest eigenvalue of the Hessian and make this direction an ascent direction for the saddle point search. Therefore, finding the direction $v$ can be translated to solve the following constrained optimization problem:

$$\begin{aligned} &\min v^T H v \\ &\text{subject to } \ v^T v = 1. \end{aligned} \tag{2.2}$$

Then the Lagrangian function for (2.2) is defined as

$$\mathcal{L}(v, \lambda) := v^T H v + \lambda(1 - v^T v), \tag{2.3}$$

where $\lambda$ is Lagrange multiplier, which corresponds to the eigenvalue of Hessian matrix $H$, i.e., $\lambda = v^T H v$. To avoid direct computation of the Hessian, $\frac{F_2 - F_1}{l}$ is used in the dimer system to approximate action of the Hessian at the dimer center along the direction $v$.

One direct way for solving (2.3) is to use the gradient method:

$$v_{k+1} = v_k - \gamma_k \left[ \frac{(F_k^2 - F_k^1)}{l} - v_k v_k^T \frac{(F_k^2 - F_k^1)}{l} \right], \tag{2.4}$$

where $\gamma_k$ is a stepsize, and $\lambda_k$ is replaced by $v_k^T \frac{(F_k^2 - F_k^1)}{l}$. The notation $vv^T$ is interpreted as $vv^T y = v^T y v$ for any $v \in \mathcal{H}$ and $y \in \mathcal{H}^*$, where $\mathcal{H}^*$ is the dual space of $\mathcal{H}$ with respect to the inner product in $\mathcal{L}$. In the classical steepest descent method, the stepsize $\gamma_k$ is obtained by performing an exact linesearch.

For the dimer translation, we decompose every $x \in \mathcal{H}$ by

$$x = (x_V, x_{V^\perp}),$$

where $x_V$ is the projection of $x$ on subspace $V = span\{v\}$, and $x_{V^\perp}$is the projection of $x$ on $V^\perp = \mathcal{H}/V$. Since the saddle point is a maximum along the lowest curvature mode and a minimum along all other modes, it can be achieved by solving the following minmax optimization:

$$\min_{x_{V^\perp} \in V^\perp} \max_{x_V \in V} E(x_V, x_{V^\perp}). \tag{2.5}$$

We can solve (2.5) by using two searching directions:

$$x_{k+1} = x_k + \beta_k^u u_k + \beta_k^w w_k, \tag{2.6}$$

where $u_k = -v_k v_k^T F_k$ is an ascent direction and $w_k = F_k - v_k v_k^T F_k$ is a descent direction with stepsize $\beta_k^u > 0, \beta_k^w > 0$. One possible choice to obtain the stepsizes $(\beta_k^u, \beta_k^w)$ is to apply the alternating direction search method along the two directions $u_k$ and $w_k$. While classical alternating direction methods have many successful applications [22, 31, 33], it takes more computational effort to perform an exact alternating direction method. Alternatively, we may take a modified force $F_k^m = u_k + w_k = (I - 2v_k v_k^T)F_k$, which is equivalent to choosing $\beta_k^u = \beta_k^w = \beta_k$. The modified force allows the dimer to climb up to the ascent direction characterized by the orientation vector. Then, the gradient method for solving (2.5) gives

$$x_{k+1} = x_k + \beta_k (I - 2v_k v_k^T) F_k. \tag{2.7}$$

In the same spirit as the SDD, following the dimer rotation and translation steps, we shrink the dimer length $l(t)$ and force it to approach zero in order to guarantee the convergence [34]. Generally, dynamics of the dimer length may follow a simple gradient flow system with $l = 0$ being the unique stable solution. There are many possible choices leading to different decay rates in the dimer length. In this paper, we choose an explicit sequence of dimer lengths for simplicity to satisfy

$$l_k \to 0 \quad \text{as} \quad k \to \infty. \tag{2.8}$$

Combining the optimization problems (2.2) and (2.5) with the shrinking dimer length (2.8), we get a complete OSD formulation.

**2.2. BB gradient method.** One direct way to solve the OSD is to apply the steepest descent method. Despite its simplicity and the optimal property, the steepest descent method often performs poorly. It converges slowly and is badly affected by ill conditioning. Different methods have been applied to accelerate the simulation of the dimer dynamics, especially the dimer rotation. The most popular ones include conjugate gradient methods [20] and L-BFGS [5]. Here we propose the BB gradient method to solve the OSD.

The BB method chooses $H_k = \alpha_k I$ as an approximation of the inverse of the Hessian at $x_k$ and imposes some quasi-Newton property on $H_k$. Denote $\triangle x_k =$

$x_k - x_{k-1}$ and $\triangle g_k = g_k - g_{k-1}$ (where $g_k$ denotes the gradient at $x_k$). By solving the optimization problem $\min \|\triangle x_k - H_k \triangle g_k\|$, the BB stepsize can be obtained by

$$\alpha_k = \frac{\triangle x_k^T \triangle g_k}{\triangle g_k^T \triangle g_k}. \tag{2.9}$$

Barzilai and Borwein [2] also considered the following symmetric optimization problem $\min \|H_k^{-1} \triangle x_k - \triangle g_k\|$ and obtained another choice of BB stepsize:

$$\alpha_k = \frac{\triangle x_k^T \triangle x_k}{\triangle x_k^T \triangle g_k}. \tag{2.10}$$

For the dimer rotation, we implement the BB method to solve (2.2) and it yields

$$v_{k+1} = v_k - \gamma_k (I - v_k v_k^T) \frac{F_k^2 - F_k^1}{l_k}, \tag{2.11}$$

with the BB stepsize

$$\gamma_k = \frac{\triangle v_k^T \triangle d_k}{\triangle d_k^T \triangle d_k}, \tag{2.12}$$

with $\triangle v_k = v_k - v_{k-1}$ and $\triangle d_k = (I - v_k v_k^T) \frac{F_k^2 - F_k^1}{l_k} - (I - v_{k-1} v_{k-1}^T) \frac{F_{k-1}^2 - F_{k-1}^1}{l_{k-1}}$.

In a similar manner, for the dimer translation, we solve (2.5) by

$$x_{k+1} = x_k + \beta_k (I - 2 v_k v_k^T) F_k \tag{2.13}$$

with the BB stepsize

$$\begin{aligned} \beta_k &= \frac{\triangle x_k^T \triangle g_k}{\triangle g_k^T \triangle g_k} \\ &= \frac{(x_k - x_{k-1})^T [-(I - 2 v_k v_k^T) F_k + (I - 2 v_{k-1} v_{k-1}^T) F_{k-1}]}{\| -(I - 2 v_k v_k^T) F_k + (I - 2 v_{k-1} v_{k-1}^T) F_{k-1} \|^2}. \end{aligned} \tag{2.14}$$

Combining the BB methods for both (2.11) and (2.13) with the shrinking dimer (2.8), we summarize the OSD algorithm stated in Algorithm (2.15) as follows.

ALGORITHM 1 (the OSD algorithm based on the BB method). *Given the initial conditions $x_0$, $v_0$, $l_0$, $\triangle t$, and in the kth iteration, we apply the following equations to update $x_{k+1}$, $v_{k+1}$, $l_{k+1}$:*

$$\begin{cases} x_{k+1} = x_k + \beta_k (I - 2 v_k v_k^T) F_k \,, \\ v_{k+1} = v_k - \gamma_k (I - v_k v_k^T) \dfrac{F_k^2 - F_k^1}{l_k} \,, \\ l_{k+1} = l_k / (1 + \triangle t). \end{cases} \tag{2.15}$$

*Here, for $k = 0$, we choose $\beta_0 = \gamma_0 = \triangle t$; for $k > 0$, $\beta_k$ and $\gamma_k$ are computed by* (2.14) *and* (2.12)*, respectively.*

*The iteration stops when $\|F_{k+1}\| < Tol$.*

We notice that the numerical scheme for the second equation in (2.15) does not preserve the unit length of $v_k$ exactly. In practice, we use a simple normalization to ensure $||v_{k_1}|| = 1$ holds in the discrete scheme [34]. Furthermore, the Hessian of the energy functionals at the saddle point include negative eigenvalues, and the numerator of $\beta_k$ in (2.14) $\triangle x_k^T \triangle g_k = \triangle x_k^T H_k \triangle x_k$ could be negative. Thus, the absolute values are taken in (2.14) and (2.12) as a precaution to avoid the negative stepsize. In practice, we observe that the signs of $\beta_k$ and $\gamma_k$ are always positive once we are close to a saddle point. More choices of the BB stepsize are available in [10].

**2.3. Local stability and superlinear convergence.** The existing convergence analyses of the BB method are hard to find because of the general nonmonotone behavior for generic objective functions. It is thus also challenging to give a full convergence theory of the BB method for the saddle point problem. Here, we make some insightful illustrations by using a simple energy defined by

$$E(x) = -\frac{x_1^2}{2} + \sum_{i=2}^{d} \frac{x_i^2}{2} \tag{2.16}$$

for $x = (x_1, z_x, \ldots, x_d)$. The Hessian matrix at the iteration point $x_k$ is

$$H = \begin{pmatrix} -1 & 0 & 0 & ... & 0 \\ 0 & 1 & 0 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 1 \end{pmatrix}_{d\times d},$$

and the gradient at the iteration point $x_k$ is $F_k = Hx_k$. As shown in [34], the saddle point is at the origin, i.e., $x^* = (0, \ldots, 0)$, $v^* = (1, 0, \ldots, 0)$ or $(-1, 0, \ldots, 0)$ with the negative eigenvalue $\lambda^* = -1$ and the equilibrium dimer length $l^* = 0$.

Let $Q_{v_k} = (I - 2v_k v_k^T)$ and $P_{v_k} = (I - v_k v_k^T)$; from (2.14), we have

$$\beta_k = \frac{(x_k - x_{k-1})^T (Q_{v_k} H x_k - Q_{v_{k-1}} H x_{k-1})}{(Q_{v_k} H x_k - Q_{v_{k-1}} H x_{k-1})^T (Q_{v_k} H x_k - Q_{v_{k-1}} H x_{k-1})}.$$

Assuming the convergence of $(x_k, v_k, l_k)$, it is straightforward to get

$$\lim_{k\to\infty} Q_{v_k} H = \lim_{k\to\infty} (I - 2v_k v_k^T) H = H^2 = I,$$

and it leads to

$$\beta^* = \lim_{k\to\infty} \beta_k = 1.$$

Notice that $\lim_{l\to 0} \frac{F_k^2 - F_k^1}{l_k} = Hv_k$, and in (2.12),

$$\triangle d_k = P_{v_k} H v_k - P_{v_{k-1}} H v_{k-1} = 2(M_k v_k - M_{k-1} v_{k-1}),$$

with $M_k$ being a diagonal matrix with entries $(v_{k,1}^2 - 1, v_{k,1}^2, \ldots, v_{k,1}^2)$. In a similar manner, we can get

$$\gamma^* = \lim_{k\to\infty} \gamma_k = \frac{1}{2}.$$

Next, we show the analysis of local stability and superlinear convergence of the OSD in (2.15) for the above energy function (2.16).

THEOREM 2.1. *The numerical scheme of the OSD given in* (2.15) *is asymptotically stable and locally superlinear convergence at the steady state of the energy function in* (2.16).

*Proof.* The numerical scheme given by (2.15) is asymptotically stable at the steady state $(x^*, v^*, l^* = 0)$ if and only if all eigenvalues of the Jacobian matrix $J$ at

$(x^*, v^*, 0)$ stays inside the unit disc. Since $\beta$ and $\gamma$ are not constant, we extend the OSD in (2.15) to a new system $(x_k, v_k, x_{k-1}, v_{k-1}, \beta_k, \gamma_k)$ as follows:

$$
\begin{cases}
x_k^{new} = x_k + \beta_k Q_{v_k} F_k \,, \\
v_k^{new} = v_k - \gamma_k P_{v_k} \dfrac{F_k^2 - F_k^1}{l_k} \,, \\
x_{k-1}^{new} = x_k \,, \\
v_{k-1}^{new} = v_k \,, \\
\beta_k^{new} = \dfrac{(x_k - x_{k-1})^T (Q_{v_k} H x_k - Q_{v_{k-1}} H x_{k-1})}{(Q_{v_k} H x_k - Q_{v_{k-1}} H x_{k-1})^T (Q_{v_k} H x_k - Q_{v_{k-1}} H x_{k-1})} \,, \\
\gamma_k^{new} = \dfrac{(v_k - v_{k-1})^T (P_{v_k} H v_k - P_{v_{k-1}} H v_{k-1})}{(P_{v_k} H v_k - P_{v_{k-1}} H v_{k-1})^T (P_{v_k} H v_k - P_{v_{k-1}} H v_{k-1})} \,.
\end{cases} \tag{2.17}
$$

We note that since the dimer length superlinearly converges to zero and is independent of the other variables, it is not included in the above map. The Jacobian matrix $J$ associated with (2.17) has the following form:

$$
J = \begin{pmatrix}
I + \beta_k Q_{v_k} \nabla_{x_k} F_k & \beta_k \nabla_{v_k} (Q_{v_k} F_k) & 0 & 0 & Q_{v_k} F_k & 0 \\
-\gamma_k P_{v_k} \nabla_{x_k} H v_k & I - \gamma_k \nabla_{v_k} (P_{v_k} H v_k) & 0 & 0 & 0 & P_{v_k} H v_k \\
I & 0 & 0 & 0 & 0 & 0 \\
0 & I & 0 & 0 & 0 & 0 \\
J_{51} & J_{52} & J_{53} & J_{54} & 0 & 0 \\
0 & J_{62} & 0 & J_{64} & 0 & 0
\end{pmatrix}.
$$

At the steady state $(x^*, v^*, l^*, \beta^*, \gamma^*)$, $F_k = 0$, $P_{v_k} H v_k = 0$, and $\nabla_{x_k} F_k = -H$. Following the similar discussion in [34], it is easy to check

$$
\begin{aligned}
J_{11}^* &= I + \beta^* 2\lambda^* v^* (v^*)^T - H = 0 \,, \\
J_{12}^* &= \beta_k \nabla_{v_k} Q_{v_k} F_k + \beta_k Q_{v_k} \nabla_{v_k} F_k = 0 \,, \\
J_{22}^* &= I + \gamma^* (2\lambda^* v^* (v^*)^T - H + \lambda^* I) = 0 \,.
\end{aligned}
$$

Therefore, at the steady state (saddle point), the Jacobian can be simplified as

$$
J^* = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
* & 0 & 0 & 0 & 0 & 0 \\
I & 0 & 0 & 0 & 0 & 0 \\
0 & I & 0 & 0 & 0 & 0 \\
* & * & * & * & 0 & 0 \\
0 & * & 0 & * & 0 & 0
\end{pmatrix}. \tag{2.18}
$$

Since all diagonal entries of $J^*$ are zeroes, the local stability and superlinear convergence at the steady state are obtained. ∎

**2.4. Comparison with the other dimer-like approaches.** The proposed OSD is a hybrid of discrete SDD with the BB approach. The original SDD can be viewed as a dynamic system, similar in spirit to the gradient flow of the energy, and it utilizes the translation and rotation steps together with a shrinking dimer so that the convergent solution of the SDD leads to a saddle point. The SDD system has the

following form:

$$\begin{cases} \mu_1 \dot{x} = (I - 2vv^T)F, \\ \mu_2 \dot{v} = (I - vv^T)\dfrac{(F_1 - F_2)}{l}, \\ \mu_3 \dot{l} = -\nabla V(l), \end{cases} \tag{2.19}$$

where $\mu_1, \mu_2, \mu_3$ are nonnegative relaxation constants. The first equation in (2.19) represents the translation step of the dimer that can be viewed as a transformed gradient flow. The second equation in (2.19) is the rotation step which allows the dimer to align along the direction with the gentlest ascent. The third equation in (2.19) is the shrinking step of the dimer that uses a monotonically increasing function $V(l)$ to control the dimer length so that $l \equiv 0$ is the unique steady state.

The OSD formulation takes the advantage of the original SDD and translates it to an optimization framework, which gives much freedom to choose suitable optimization methods to speed up the computation of the saddle point. Furthermore, the OSD may be viewed as the generalized SDD. The dynamics of dimer translation in the OSD can be obtained by applying the gradient method with an ascent direction in $V$ and a descent direction in $V^{\perp}$, i.e.,

$$\frac{dx}{dt} = -\beta_V vv^T F + \beta_{V^{\perp}}(F - vv^T F). \tag{2.20}$$

The translation step of the original SDD in (2.19) is just the case that $\beta_V = \beta_{V^{\perp}}$ in (2.20). Moreover, if we take the gradient method in (2.3) to find the orientation vector $v$, the dynamic equation for $v$ in the OSD is the same as the rotation step in the original SDD (2.19).

For the time discretization of the SDD, the explicit Euler or modified Euler schemes were implemented, and linear convergence and optimal stepsize can be achieved [34]. Instead of using a fixed time stepsize in the original SDD, the OSD utilizes an automatically updated BB stepsize. Since there are no matrix computations in the BB stepsize formulas and only preexisting information at the current iteration point and previous one point is used, the BB method can greatly speed up the convergence of the gradient method. In comparison with other optimization-based dimer-like iterations, conjugate gradient–based approaches are also known for linear convergence though they may significantly reduce the possible number of iterations, but the BB approach is able to achieve superlinear convergence. In comparison with other quasi-Newton-based implementation, it is among the simplest strategies that avoid matrix computation.

**3. Numerical examples.** In this section, we show some numerical examples to demonstrate the efficiency and reliability of the OSD algorithm. Following similar examples in [34] and other existing works in the literature, we focus on the results obtained from both the original SDD with the modified Euler scheme and the OSD algorithm with the BB stepsize. To compare the performance of the SDD and OSD, the same initial guess is used in each example, and the error is defined by the L2 norm of the gradient at the saddle point because the exact saddle point $x$ and unstable direction $v$ are not available in most cases.

**3.1. Two-dimensional examples.** As the first illustration, we consider a simple two-dimensional energy function

$$E(x, y) = (x^2 - 1)^2 + y^2, \tag{3.1}$$

so that $(0,0)$ is the index-one saddle point and $(\pm 1,0)$ is the corresponding unstable direction. We implement both the original SDD and the OSD for comparison. As shown in [34], we choose the optimal time stepsize for the most significant error reduction, i.e., $\triangle t = 0.25$, in the modified Euler method of the SDD. Numerical simulations in Figure 1 show that the OSD leads to superlinear convergence, while the SDD is only linearly convergent. It turns out the OSD significantly improves the performance of SDD and OSD only needs 15 iteration steps compared to 67 iteration steps for SDD.

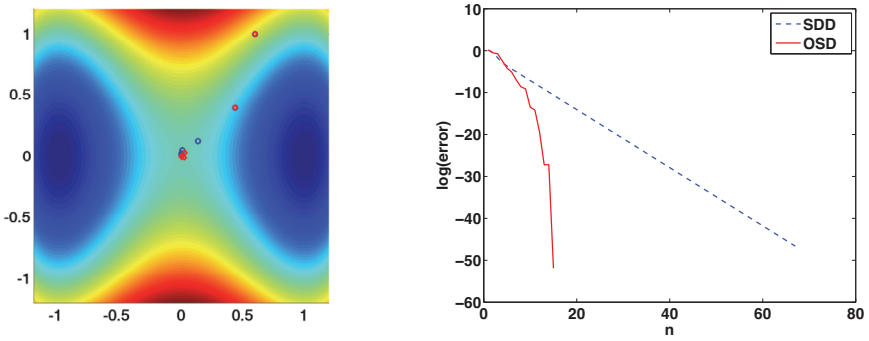


FIG. 1. *Comparison of SDD and OSD for the energy* $E(x,y) = (x^2-1)^2 + y^2$. *Left: dimer snapshots on the energy surface. SDD: blue circle; OSD: red diamond. Right: error reduction versus iteration steps.*

Next we test some other classical benchmark problems. One example is the Minyaev–Quapp surface [24] given by

$$E(x,y) = \cos(2x) + \cos(2y) + 0.57\cos(2x-2y). \tag{3.2}$$

In Figure 2, we show that both SDD and OSD are able to locate the saddle point successfully. Again, we take the optimal time stepsize $\triangle t = 0.25$ for SDD and it needs 108 iteration steps in order to reach the error tolerance $Tol = 10^{-12}$. OSD shows superlinear convergence and takes only 21 iteration steps for the convergence.

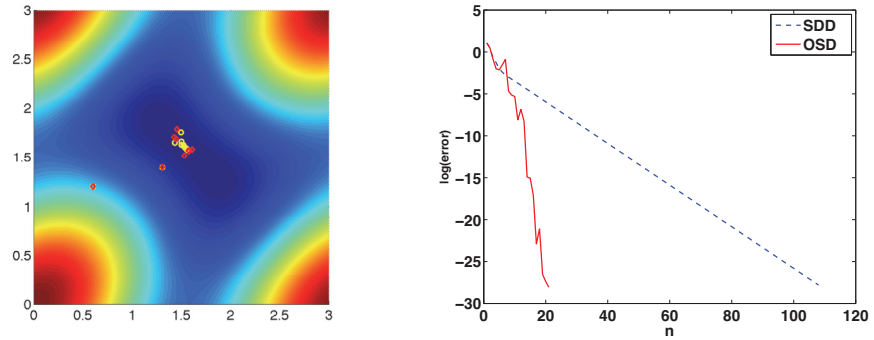


FIG. 2. *Comparison of SDD and OSD for the Minyaev–Quapp energy. Left: dimer snapshots on the energy surface. SDD: yellow circle; OSD: red diamond. Right: error reduction versus iteration steps.*

We also choose the stingray function [18] given by

$$E(x, y) = x^2 + c(x - 1)y^2, \tag{3.3}$$

where $c = 100$ is taken. The surface has a flat ridge which makes it difficult to search the saddle point. In Figure 3, we choose the initial position at $\mathbf{x}_0 = (0.5, 0.8)$ and four different initial orientation vectors $\mathbf{v}_0 = (-1, 2), (-1, 0.03), (-1, -1)$, and $(-1, -5)$. Numerical results show that SDD generally drives the dimer to climb up to the ridge and change direction to move along the ridge, then it slowly converges to the saddle point, which requires over 1500 iteration steps by using the largest time step we could choose. By contrast, OSD doesn't need to follow the ridge direction but converges to the saddle point very quickly. As a result, the minimum iteration step of the OSD is 27 and the maximum is 38 in these four cases, which indicates the OSD is very efficient and robust to different initial conditions.

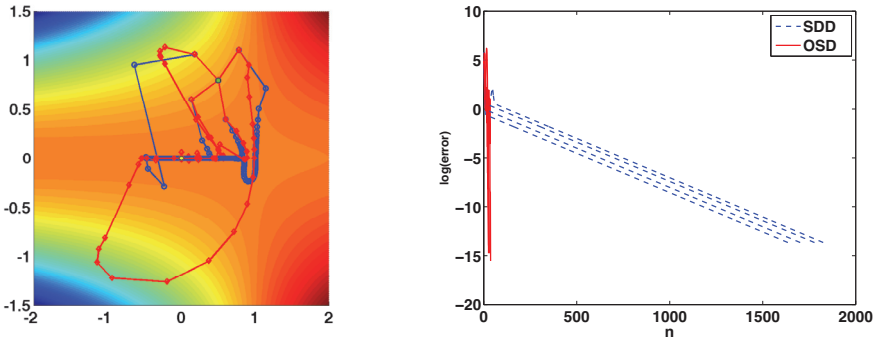


FIG. 3. *Comparison of SDD and OSD for the stingray energy. Left: dimer snapshots on the energy surface. SDD: blue circle lines; OSD: red diamond lines; initial position: green dot; saddle point: yellow dot. Right: error reduction versus iteration steps.*

Another example to be tested is the Eckhardt surface [14] given by

$$E(x, y) = \exp(-x^2-(y+1)^2)+\exp(-x^2-(y-1)^2)+4\exp(-3(x^2+y^2)/2)+y^2/2. \tag{3.4}$$

In Figure 4, we choose the starting point $\mathbf{x}_0 = (1.5, 1.2)$ and four different initial orientation vectors $\mathbf{v}_0 = (-1, 0.1), (-1, 2), (-1, 5)$, and $(-1, 7)$, respectively. Numerical results show that the SDD is more dependent on the initial orientation than the OSD. Bad initial orientation may drive the SDD to another saddle point which is far away from the closest one. Thus, the OSD with the BB method can not only accelerate the convergence but also keep the dimer from jumping over too far.

There are two BB stepsize formulas (2.9) and (2.10) proposed in [2]. In the above numerical simulations we adopted one of them, i.e., (2.14) and (2.12). We also test the second formula, i.e.,

$$\beta_k = \frac{\triangle x_k^T \triangle x_k}{\triangle x_k^T \triangle g_k}, \qquad \gamma_k = \frac{\triangle v_k^T \triangle v_k}{\triangle v_k^T \triangle d_k}, \tag{3.5}$$

and compare the two sets of formulas in the number of iterations.

From Table 1 we can see that the overall performance of the OSD algorithm with stepsizes (2.14) and (2.12) (i.e., OSD-1) is slightly better than the OSD algorithm with stepsizes (3.5) (i.e., OSD-2). They have an equal number of iterations, for example, one, but OSD-1 needs fewer iterations than OSD-2 for the other three examples.

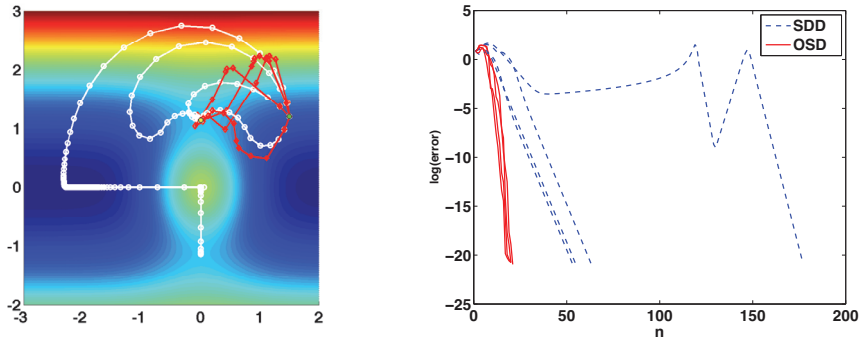


FIG. 4. *Comparison of SDD and OSD for the Eckhardt energy. Left: dimer snapshots on the energy surface. SDD: white circle lines; OSD: red diamond lines; initial position: green dot; saddle point: yellow dot. Right: error reduction versus iteration steps.*

TABLE 1

*Number of iterations of OSD-1 and OSD-2 applied to the above four examples. OSD-1 represents the OSD algorithm with the BB stepsizes* (2.14) *and* (2.12). *OSD-2 represents the OSD algorithm with the BB stepsizes* (3.5). *The four energy functions and the parameter settings are as follows: Case 1. The two-dimensional energy* (3.1)*: Tol* $= 10^{-20}, \triangle t = 0.25, x_0 = [0.6, 1]^T, v_0 = [1, 0.2]^T, l_0 = 0.0001$. *Case 2. The Minyaev–Quapp surface* (3.2)*: Tol* $= 10^{-12}, \triangle t = 0.25, x_0 = [0.6, 1.2]^T, v_0 = [1, -0.9]^T, l_0 = 0.0001$. *Case 3. The stingray function* (3.3)*: Tol* $= 10^{-6}, \triangle t = 0.004, x_0 = [0.5, 0.8]^T, v_0 = [-1, 2]^T, l_0 = 0.001$. *Case 4. The Eckhardt surface* (3.4)*: Tol* $= 10^{-9}, \triangle t = 0.1, x_0 = [1.5, 1.2]^T, v_0 = [-1, 0.1]^T, l_0 = 0.0001$.

| | Example 1 | Example 2 | Example 3 | Example 4 |
|---|---|---|---|---|
| OSD-1 | 15 | 21 | 26 | 21 |
| OSD-2 | 15 | 24 | 44 | 27 |

**3.2. Cluster of seven particles.** Another test is the cluster of seven particles in two dimensions [11]. The particles interact via the Lennard–Jones potential

$$V = \sum_{i<j} 4\epsilon \left[ \left( \frac{\sigma}{r_{ij}} \right)^{12} - \left( \frac{\sigma}{r_{ij}} \right)^{6} \right], \tag{3.6}$$

where $r_{ij}$ is the distance between particle $i$ and particle $j$. Distances are measured in units of $\sigma$ and energies in units of $\epsilon$. In fact, this is a 14-dimensional problem with respect to the locations of seven particles in two dimensions. As pointed in [11], during the migration of a particle from the center of the cluster to its surface, we list five possible saddle points with their energies (in units of $\epsilon$) as follows: A ($-11.040$), B ($-11.037$), C ($-10.841$), D ($-10.807$), and E ($-10.799$) (see Figure 5). Since the analytical solutions of the saddle points are not available, we numerically calculate the optimal stepsize in the original SDD for five saddle points in order to compare with the OSD algorithm. The chosen stepsizes in our simulations fall within the interval [0.002, 0.007] with the given initial position and direction of the system.

Figure 5 shows the computed particle configurations at the five saddle points and corresponding errors of SDD/OSD algorithms. It shows that although the OSD algorithm cannot guarantee the monotonic decrease of the error reduction, the overall performance of the OSD algorithm is about three to five times faster than the SDD algorithm with the optimal stepsize. In practice, without a priori information of

optimal stepsize in the SDD algorithm, the OSD algorithm is much more efficient than the SDD algorithm for finding possible saddle points of cluster of multiple particles.

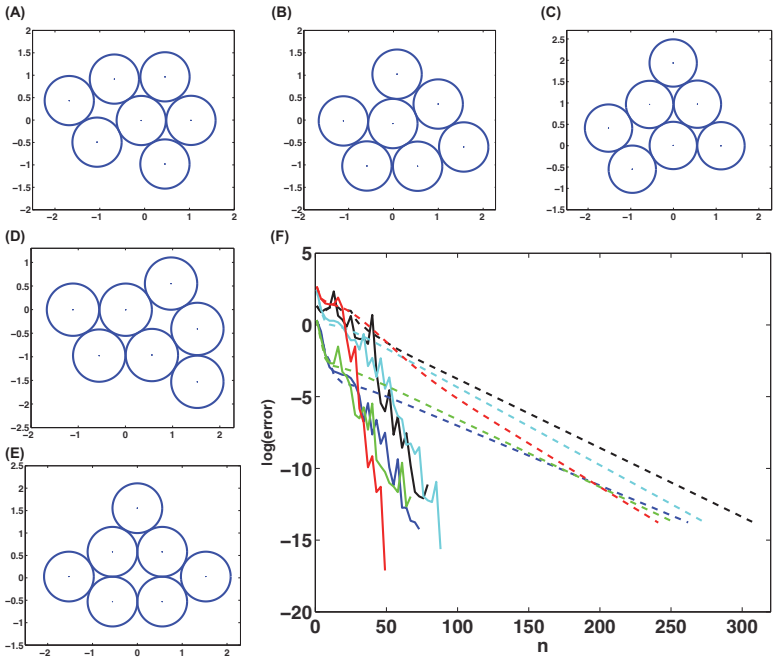


FIG. 5. *Particle configurations at saddle points and errors for the Lennard–Jones potential* $V = \sum_{i<j} 4\epsilon[(\frac{\sigma}{r_{ij}})^{12} - (\frac{\sigma}{r_{ij}})^{6}]$. (A)–(E) *Five particle configurations at the saddle points with energies (in units of* $\epsilon$*)* $-11.040, -11.037, -10.841, -10.807, -10.799$*, respectively.* (F) *Log(error) versus iteration numbers of the SDD algorithm (dashed curves) and OSD algorithm (solid curves). The black, cyan, blue, green, and red curves represent errors corresponding to the saddle points A, B, C, D, and E, respectively. Errors are only shown for one in four iterations to maintain visual clarity.*

**3.3. Seven-atom island on the (111) surface of an FCC crystal.** We next implement a high-dimensional example by using the seven-atom island on the (111) surface of an FCC crystal in three dimensions, and this system serves as a popular benchmark problem as it has been extensively studied in [15, 19, 23, 27]. The cluster consists of seven atoms with the substrate made by a six-layer slab, each of which contains 56 atoms. The bottom three layers in the slab are fixed. Periodic boundary conditions are used in the $x$ and $y$ directions, respectively. A total of $56 \times 3 + 7 = 175$ atoms are free to move during the saddle point searches. The atoms interact via a pairwise additive Morse potential

$$V(r) = D(e^{-2\alpha(r-r_0)} - 2e^{-\alpha(r-r_0)}). \tag{3.7}$$

Some parameters needed to reproduce diffusion barriers on Pt surfaces are $D = 0.7102$, $\alpha = 1.6047$, and $r_0 = 2.8970$. The potential was cut and shifted at $r_c = 9.5$, and the minimum energy lattice constant for the FCC solid is 2.74412.

We first start the minimum energy structure of this system and test the OSD method to find possible saddle points around the minimum. The initial condition of the atom position is determined by either randomly perturbing the minimum configuration or randomly moving one atom around the cluster, and the initial direction is

randomly chosen for the seven-atom cluster and 0 for the substrate. Following similar processes in [19], we are able to locate different types of saddle points around the minimum and list 13 saddle configurations with energy less than or equal to 1.513 eV in Figure 6, including two uniform translations of the intact island with the lowest energy (saddles 1 and 2), three saddle points corresponding to a pair of atoms sliding to adjacent FCC sites (saddles 3-5), four translations corresponding to that the pair of atoms slides to the adjacent HCP sites and the remaining five atoms slide in the opposite direction to HCP sites (saddles 6–9), two saddle points corresponding to that a row of three edge atoms slides into adjacent FCC sites (saddles 10–11), one saddle point for a pair of edge atoms moving in such a way that one of the atoms is displaced away from the island while the other atom takes its place (saddle 12), and one saddle point for the displacement of a single atom (saddle 13).

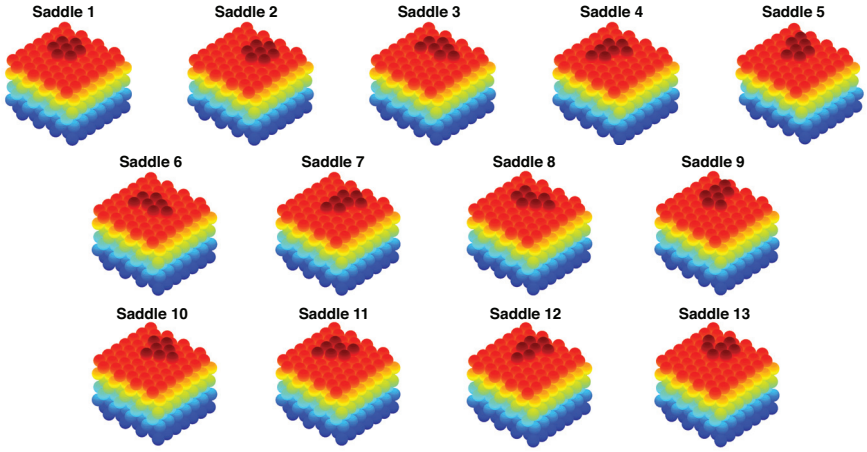


FIG. 6. *Thirteen saddle point configurations for the seven-atom island on the* (111) *surface of an FCC crystal.*

To demonstrate the efficiency of the proposed algorithm, we compare a series of the dimer methods, from the classical dimer method in [20] to the very recent development of an efficient dimer method proposed by Gould, Ortner, and Packwood in [16], as well as the SDD method and the OSD method. In [16], the dimer method is implemented by introducing a preconditioner technique in the dimer iteration and a linesearch technique for finding the stepsize to achieve better efficiency and convergence. For comparison, we take two versions proposed in [16]: (1) a dimer method with a preconditioner (denoted by Dimer P in Table 2) and (2) a linesearch dimer method with both preconditioner and static rotation (denoted by Dimer P&L in Table 2). The same initial conditions are used for different numerical algorithms. The results in Table 2 show that the classical dimer method and the SDD method have a similar performance. The dimer method coupled with a preconditioner requires less force calculation in most cases and in particular is very efficient for saddles 1 and 2. Yet it may also underperform the classical dimer method in some cases (saddles 4 and 11). For this particular example, the linesearch dimer method with both preconditioner and static rotation does not show any superior efficiency, resulting in more force evaluations except for Saddles 1 and 2. Nevertheless, one may clearly see that the OSD method proposed in this work can significantly reduce the number of force

calculations and improve the efficiency in all cases. It yields an average speedup of five times in comparison to the SDD method.

TABLE 2

*Comparison of dimer method, SDD method, dimer method with preconditioner (Dimer P), dimer method with preconditioner and linsearch (Dimer P&L), and OSD method. For comparison, the numbers of force calculations for reaching a saddle point are collected with a* 0.001 *tolerance in the force evaluation, and the same initial conditions are used for different numerical algorithms.*

| Saddle | Dimer | SDD | Dimer P | Dimer P&L | OSD |
|---|---|---|---|---|---|
| #1 | 786 | 928 | 234 | 608 | 172 |
| #2 | 1018 | 1040 | 366 | 552 | 284 |
| #3 | 1400 | 976 | 964 | 2432 | 232 |
| #4 | 1170 | 1216 | 1338 | 2594 | 270 |
| #5 | 804 | 748 | 632 | 1846 | 126 |
| #6 | 924 | 868 | 836 | 1902 | 136 |
| #7 | 906 | 860 | 822 | 1338 | 164 |
| #8 | 900 | 852 | 856 | 2888 | 182 |
| #9 | 960 | 864 | 836 | 2440 | 176 |
| #10 | 1030 | 948 | 902 | 2480 | 186 |
| #11 | 1166 | 1076 | 1190 | 4226 | 184 |
| #12 | 856 | 800 | 718 | 2170 | 178 |
| #13 | 976 | 924 | 760 | 1438 | 182 |

**3.4. Nucleation in phase transformation.** Nucleation is one of the most common physical phenomena in nature, and finding the critical nucleus has attracted much attention in recent years. As an illustration in infinite-dimensional space, we consider the diffuse-interface model proposed in [36]. Only a structural transition is assumed with no compositional changes and the structural difference between the parent phase and the nucleating phase is also assumed to be sufficiently described by a single-order parameter $\eta$. At a given temperature, the free energy dependence on $\eta$ is described by a double-well potential

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

with local energy wells at $\eta = 1$ and $\eta = -1$, respectively, so that $\lambda$ determines the well depth difference.

The total free energy of an inhomogeneous system described by a spatial distribution of $\eta$ is given by

$$E = \int_\Omega \left( f(\eta) + \frac{1}{2}|M\nabla\eta|^2 \right) d\mathbf{x} \,. \tag{3.8}$$

Here, the domain $\Omega = (-1, 1)^d$ is used with $d$ being the space dimension and a periodic boundary condition is used for the order parameter $\eta$. $M$ is the interfacial energy coefficient which is a constant diagonal tensor in $\Omega$ for isotropic interfacial energy.

By the energy variational principle, the natural force is given by

$$F(\eta) = M\triangle\eta + (1 - \eta^2)\left(\eta + \frac{3\lambda}{4}\right). \tag{3.9}$$

In numerical implementation, we apply the Fourier spectral approximations for the spatial discretization and a semi-implicit scheme for the temporal discretization.

TABLE 3
*Iteration steps of original SDD and OSD for finding the critical nucleus in 2D with* $Tol = 10^{-5}$.

| $\triangle t$ | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1.0 |
|---|---|---|---|---|---|---|---|---|---|---|
| SDD | 5717 | 2968 | 2053 | 1596 | 1322 | 1140 | 1010 | 912 | 837 | unstable |
| OSD | 550 | 546 | 566 | 552 | 552 | 576 | 549 | 556 | 539 | 551 |

The computational domain is $\Omega = [-1, 1]^2$ and the number of Fourier modes is $512 \times 512$. Some other parameters are $M = 4 \times 10^{-4}$, $\lambda = 0.1$, $\mu_1 = \mu_2 = \mu_3 = 1$ in the original SDD. The initial condition of the critical nucleus is taken as a tanh profile with a circular shape, and the initial direction is $v = 1 - \tanh^2(x)$.

Numerical simulations show that the critical nucleus has a circular shape with $+1$ inside and $-1$ outside of the nucleus [36]. Table 3 shows that the performance of the original SDD highly depends on the time step. Since the optimal time step generally cannot be found a priori, we often have to take a small time step in practice to avoid the instability of the numerical scheme. On the other hand, although the BB method for the OSD uses the fixed stepsize $\triangle t$ at the first step to get the BB stepsize, different time steps $\triangle t$ barely affect the overall performance of the OSD. Moreover, the OSD requires fewer time steps than the SDD with the largest time step, which indicates the effectiveness of the OSD by using the BB method.

**4. Conclusions.** In this paper, we develop an optimization-based shrinking dimer method for finding saddle points. The dimer rotation and translation steps in the classical dimer method are converted to the corresponding optimization problems so that we can apply the sophisticated optimization methods to speed up the calculation of the saddle point. To keep the coding simple and robust, we begin with a study of the dynamic selection of stepsize, guided by optimization approaches, that requires little additional computation. Specifically, we apply the BB gradient method to implement the OSD system. The BB method is originally designed for solving unconstrained minimization problems. By taking advantage of the original SDD, we are able to extend the BB method to solve the saddle point problem. Although the BB method does not guarantee the energy functional decreases at each iteration, this nonmonotonicity does not harm the fast convergence of the OSD in practice. Various numerical examples demonstrate that the OSD method with the BB stepsize shows superlinear convergence and is significantly faster than the original SDD algorithm. For the analysis, we verify the local stability and superlinear convergence for a specific but representative energy function. The analysis of convergence for more general energy landscapes will be interesting to carry out in the future. Furthermore, the excellent performance of the OSD framework encourages us to incorporate other optimization techniques to effectively capture saddle points and solve more application problems.

**Acknowledgments.** The authors would like to thank Dr. Zaiwen Wen for providing much assistance on the optimization algorithm. We greatly appreciate Dr. Christoph Ortner for the assistance of the dimer algorithm. We also thank the referees for their valuable suggestions.

REFERENCES

[1] J. BAKER, *An algorithm for the location of transition states*, J. Comput. Chem., 7 (1986), pp. 385–395.

[2] J. BARZILAI AND J. M. BORWEIN, *Two point step size gradient methods*, IMA J. Numer. Anal., 8 (1988), pp. 141–148.

[3] R. Burton, G. Huang, M. Daugherty, T. Calderone, and T. Oas, *The energy landscape of a fast-folding protein mapped by Ala → Gly substitutions*, Nature Structural Molecular Biol., 4 (1997), pp. 305–310.
[4] E. Cances, F. Legoll, M.-C. Marinica, K. Minoukadeh, and F. Willaime, *Some improvements of the activation-relaxation technique method for finding transition pathways on potential energy surfaces*, J. Chem. Phys., 130 (2009), 114711.
[5] J. M. Carr and D. J. Wales *Global optimization and folding pathways of selected α-helical proteins*, J. Chem. Phys., 123 (2005), 234901.
[6] C. J. Cerjan and W. H. Miller, *On finding transition states*, J. Chem. Phys., 75 (1981), pp. 2800–2806.
[7] X. Y. Cheng, L. Lin, W. E, P. W. Zhang, and A.-C. Shi, *Nucleation of ordered phases in block copolymers*, Phys. Rev. Lett., 104 (2010), 148301.
[8] G. Crippen and H. Scheraga, *Minimization of polypeptide energy* XI. *The method of gentlest ascent*, Arch. Biochem. Biophys., 144 (1971), pp. 462–466.
[9] Y. H. Dai and L. Z. Liao, *R-linear convergence of the Barzilai and Borwein gradient method*, IMA J. Numer. Anal., 22 (2002), pp. 1–10.
[10] Y. H. Dai, J. Yuan, and Y. X. Yuan, *Modified two-point stepsize gradient methods for unconstrained optimization*, Comput. Optim. Appl., 22 (2002), pp. 103–109.
[11] C. Dellago, P. G. Bolhuis, and D. Chandler, *Efficient transition path sampling: Application to Lennard-Jones cluster rearrangements*, J. Chem. Phys., 108 (1998), pp. 9236–9245.
[12] W. E, W. Ren, and E. Vanden-Eijnden, *String method for the study of rare events*, Phys. Rev. B, 66 (2002), 052301.
[13] W. E and X. Zhou, *The gentlest ascent dynamics*, Nonlinearity, 24 (2011), pp. 1831–1842.
[14] B. Eckhardt, *Irregular scattering*, Phys. D, 33 (1988), pp. 89–98.
[15] W. Gao, J. Leng, and X. Zhou, *An Iterative Minimization Formulation for Saddle-Point Search*, preprint, arXiv:1406.1991, 2014.
[16] N. Gould, C. Ortner, and D. Packwood, *An Efficient Dimer Method with Preconditioning and Linesearch*, preprint, arXiv:1407.2817, 2014.
[17] N. Govind, M. Petersen, G. Fitzgerald, D. King-Smith, and J. Andzelm, *Generalized synchronous transit method for transition state location*, Comput. Materials Sci., 28 (2003), pp. 250–258.
[18] 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, pp. 639–657.
[19] G. Henkelman, G. Jóhannesson, and H. Jónsson, *Methods for finding saddle points and minimum energy paths*, in Progress on Theoretical Chemistry and Physics, S. D. Schwartz, ed., Kluwer, Dordrecht, the Netherlands, 2000, pp. 269–300.
[20] 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.
[21] G. Henkelman and H. Jónsson, *Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points*, J. Chem. Phys., 113 (2000), 9978.
[22] R. Hooke and T. A. Jeeves, *Direct search solution of numerical and statistical problems*, J. ACM, 8 (1961), pp. 212–229.
[23] E. Machado-Charry, L. Béland, D. Caliste, L. Genovese, T. Deutsch, N. Mousseau, and P. Pochet, *Optimized energy landscape exploration using the ab initio based activation-relaxation technique*, J. Chem. Phys., 135 (2011), 034102.
[24] 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), pp. 1792–1803.
[25] R. Miron and K. Fichthorn, *The step and slide method for finding saddle points on multidimensional potential surfaces*, J. Chem. Phys., 115 (2001), pp. 8742–8750.
[26] N. Mousseau and G. T. Barkema, *Traveling through potential energy landscapes of disordered materials: The activation-relaxation technique*, Phys. Rev. E, 57 (1998), pp. 2419–2424.
[27] R. A. Olsen, G. J. Kroes, G. Henkelman, A. Arnaldsson, and H. Jónsson, *Comparison of methods for finding saddle points without knowledge of the final states*, J. Chem. Phys., 121 (2004), pp. 9776–9792.
[28] A. Pedersen, S. Hafstein, and H. Jónsson, *Efficient sampling of saddle points with the minimum-mode following method*, SIAM J. Sci. Comput., 33 (2011), pp. 633–652.
[29] C. Peng, P. Ayala, H. Schlegel, and M. Frisch *Using redundant internal coordinates to optimize equilibrium geometries and transition states*, J. Comput. Chem., 17 (1996), pp. 49–56.

[30] M. Raydan, *On the Barzilai and Borwein choice of steplength for the gradient method*, IMA J. Numer. Anal., 13 (1993), pp. 321–326.
[31] H. H. Rosenbrock, *An automatic method for finding the greatest or least value of a function*, Computer J., 3 (1960), pp. 175–184.
[32] D. Wales, *Energy Landscapes: Applications to Clusters, Biomolecules and Glasses*, Cambridge University Press, Cambridge, UK, 2003.
[33] Y. X. Yuan and Z. Z. Zheng, *Disturbed sparse linear equations over the* 0-1 *finite field*, J. Comput. Math., 24 (2006), pp. 357–364.
[34] J. Y. Zhang and Q. Du, *Shrinking dimer dynamics and its applications to saddle point search*, SIAM J. Numer. Anal., 50 (2012), pp. 1899–1921.
[35] J. Y. Zhang and Q. Du, *Constrained shrinking dimer dynamics for saddle point search with constraints*, J. Comput. Phys., 231 (2012), pp. 4745–4758.
[36] L. Zhang, L.-Q. Chen, and Q. Du, *Morphology of critical nuclei in solid state phase transformations*, Phys. Rev. Lett., 98 (2007), 265703.
[37] L. Zhang, L.-Q. Chen, and Q. Du, *Simultaneous prediction of morphologies of a critical nucleus and an equilibrium precipitate in solids*, Commun. Comput. Phys., 7 (2010), pp. 674–682.
[38] L. Zhang, J. Y. Zhang, and Q. Du, *Finding critical nuclei in phase transformations by shrinking dimer dynamics and its variants*, Commun. Comput. Phys., 16 (2014), pp. 781–798.