Constrained Shrinking Dimer Dynamics for Saddle Point Search with Constraints

Jingyan Zhang and Qiang Du

Department of Mathematics, Pennsylvania State University, PA 16802, USA
(jzhang@math.psu.edu, qdu@math.psu.edu)

Abstract
In this paper, we study the constrained shrinking dimer dynamics (CSDD) which leads to numerical procedures for locating saddle points (transition states) associated with an energy functional defined on a constrained manifold. We focus on the most generic case corresponding to a constrained stationary point where the projected Hessian of the energy onto the tangent hyperplane of the constrained manifold has only one unstable direction and demonstrate, in this case, the local stability of the CSDD. We examine various numerical implementation issues and consider some interesting applications of CSDD including the computation of periodic centroidal Voronoi tessellations, generalized Thomson problems about particle/charge distribution on the unit sphere, and the critical nuclei morphology in binary phase transformations.

Keywords: Saddle point, transition state, potential energy surface, shrinking dimer dynamics, constrained saddle point, periodic centroidal Voronoi tessellations, generalized Thomson problem, critical nuclei

1. Introduction
Exploring the complex landscape of an energy surface and searching for local equilibria and transition states/saddle points are of much interests in many scientific applications. For example, it is often stated that many aspects of computational chemistry are related to the understanding of potential energy surfaces [33]. In comparison with local equilibria computation, the search for saddle points is considerably more difficult, given their unstable nature. Yet, saddle points that represent transition states and their energy values are very important in determining various physical and chemical properties, such as nu-

1This research is supported in part by NSF-DMS 1016073 and NSF-IIP 1034965 Center for Computational Materials Design (CCMD).
cleation and reaction rates. Such calculations are often used in studies of structure stabilities and phase transitions in materials as well as reaction pathways in biological and chemical systems [17, 22, 24, 27, 29, 31, 33, 36]. Locating saddle points (transition states) of an energy surface is particularly helpful to the study of rare thermally-activated transitions between different equilibria and metastable states in a stochastically perturbed dissipative dynamic process associated with the energy surface. This has led to extensive studies of effective saddle point search algorithms that can be roughly divided into two categories [23]: one being chain-of-state methods such as nudged elastic band [21] and string method [14, 15], while the other being surface-walking methods such as gentlest ascent method [8] and dimer method [20]. We refer to relevant reviews [6, 33, 36] for more detailed discussions on existing works.

In many applications, the challenging task of searching for saddle points may be further complicated by additional constraints on the state variables, for instance, a fixed bond length between two atoms in chemical reactions, or a conserved order parameter field in nucleation studies [13]. Though an obvious approach is to reparametrize via a set of unconstrained variables, this is not always feasible nor computationally efficient in practice. The main objective of the current study is to develop an effective algorithm that can deal with constrained transition state/saddle point search by overcoming difficulties due to the unstable nature of saddle points, the lack of a priori information on their stable and unstable directions, and the complications due to constraints.

The method we present here is based on our recent development of the shrinking dimer dynamics (SDD) [12]. The SDD is a dynamic system reformulation of the popular dimer method originally formulated by Henkelman and Jonsson [20] for unconstrained saddle point search. The basic idea behind the dimer method and thus the SDD is to consider an extended system involving not only the saddle location but also the associated unstable directions. For an index-1 saddle with only one unstable direction, this can be realized using a dimer system though both its center and its orientation, with the former determining the saddle position and the latter representing the unique unstable direction. It is worthy noting that such index-1 saddle points are often most interesting in practice as they correspond to the most generic transition states along minimum energy paths connecting near-by local energy wells. Here, we
present a dynamic system similar to the SDD, named as the constrained shrinking dimer dynamics (CSDD), to search for constrained index-1 saddle points on a potential energy surface subject to equality constraints. More precisely, we search for a constrained critical point at which the projected Hessian of the energy onto the tangent hyperplane of the constrained manifold has a single unstable direction, or equivalently, only one negative eigenvalue. To incorporate constraints, the method of Lagrange multipliers is seamlessly integrated into the SDD (or CSDD to be precise) framework. Through stability analysis of the new dynamic system, we demonstrate rigorously that locally stable steady states of CSDD corresponds to index-1 saddle points of the energy. Hence, CSDD leads to a theoretically reliable constrained saddle search algorithm which requires only the evaluation of forces (gradients) associated with the energy while avoiding its Hessian. These features make CSDD a very attractive approach for constrained saddle point search in many practical applications. Furthermore, we also discuss how one may avoid computing Hessians of the constraints. Problems involving linear and quadratic constraints are given special attention. The relevant dynamic system and algorithms are discussed in details in later sections, together with some numerical experiments related to several interesting applications of CSDD. One application is the computation of a special type of centroidal Voronoi tessellations which have been used in many areas of science and engineering [9]. Another application is on generalized Thomson problems where we compute energy diagrams for interacting particles on a unit sphere which can help us better understand the geometric and topological frustrations in such systems. In addition, we consider the phase field critical nuclei computation which has become an effective approach for studying the long-standing question of critical nucleation in solid state phase transformations [38].

The rest of this paper is organized as follows: we present the formulation of CSDD and briefly discuss its convergence to the constrained index-1 saddle points in section 2. Discussions on the stability analysis and several computational issues are presented. The latter includes some discussion on the efficient treatment of constraints and the associated Lagrange multipliers. We then present some numerical experiments and applications of CSDD to illustrate the performance of the proposed numerical methods and to reveal some interesting observations and new findings based on numerical simulations.
2. Constrained Shrinking Dimer Dynamics

We begin by briefly recalling the basic theory for constrained saddle point and the original SDD studied in [12]. We then discuss the modification of SDD to incorporate constraints, i.e., the constrained shrinking dimer dynamics.

2.1. Constrained saddle point

We first introduce some notation. Given an energy functional $E$ on a Hilbert space $H$ with $\nabla E(x)$ denoting the gradient of $E$ at $x \in H$ defined in the Frechet sense with respect to an inner product (duality pairing) in $L$, a Hilbert space containing $H$. Similarly, $H_E(x)$ denotes the Hessian operator $E$ at $x \in H$.

We consider the index-1 saddle point of $E$ subject to the constraint:

$$G(x) = 0 \quad (1)$$

where $G : H \to B$ is an operator from $H$ to a Hilbert space $B$.

To avoid technical complications, we assume that $G$ is smooth and denote $\nabla G$ the gradient of $G$ defined in the Frechet sense, again with respect to duality pairing in $L$. For any constant vector $b \in B$, $b^T \nabla G = \nabla G^T b$ can also be viewed as the gradient operator for the real-valued functional $b^T G = G^T b$ (interpreted as the inner product of $G$ and $b$ in $B$), that is,

$$y^T (b^T \nabla G) = y^T \nabla (b^T G) = y^T \nabla (G^T b) = y^T \nabla (G^T) b = (\nabla (G^T) b)^T y.$$

We note that $G^T b$ is taken as the inner product in space $B$ while $y^T x$ is taken as the inner product or the duality pairing between $y \in H$ and $x$ is the dual space of $H$. Thus, in general we need to distinguish $\nabla G^T$ from $(\nabla G)^T$. Similar as for the energy, we let $H_G(x)$ denote the Hessian operator $G$ at $x \in H$.

Computationally, a special case of interests is given by $B = \mathbb{R}^m$ so that

$$G(x) = (G^1(x), G^2(x), ..., G^m(x))^T = 0 \quad (2)$$

represents a finite number of constraints, with $G$ being a vector-valued function of $m$ dimensions. For $H = \mathbb{R}^n$ and $B = \mathbb{R}^m$, we let $L = \mathbb{R}^n$ and take all inner products as the standard inner product in the Euclidean spaces. Then, $\nabla G$ is simply the standard gradient of the vector $G$, consisting of $m$ row vectors and with size $m \times n$ and $H_G$ becomes a third order tensor. The general Hilbert space setting allows us to also work with infinite dimensional energy landscapes.
To incorporate constraints into SDD, we need to define the projected natural forces, i.e., the negative gradient forces projected on the tangential hyperplane (space) of constraints. We consider the projection:

$$P(G) := I - \nabla G^T [(\nabla G)^T \nabla G]^T - 1 (\nabla G)^T.$$ (3)

Here, we assume that the constraints are independent so that $[(\nabla G)^T \nabla G]^{-1}$ is well-defined in a suitable space. Note that $P$ is viewed as an operator dependent on $G$, so that if $G$ is changed in (3) then $P(G)$ is changed accordingly.

An index-1 saddle point of $E = E(x)$ subject to $G(x) = 0$, denoted by $x^* \in \mathcal{H}$, satisfies the following conditions [25]:

- First-order KKT condition: there is some $\eta(x^*) \in \mathcal{B}$ such that
  $$F(x^*) - \nabla G^T(x^*) \eta(x^*) = F(x^*) - \eta^T(x^*) \nabla G(x^*) = 0,$$ (4)
  with
  $$\eta(x^*) = [(\nabla G(x^*))^T \nabla G^T(x^*)]^{-1} (\nabla G(x^*))^T F(x^*)$$ (5)
  being the Lagrangian multiplier. Using the $P(G(x^*))$ defined in (3), the condition (4) is equivalent to $P(G(x^*)) F(x^*) = 0$ which avoids the explicit reference to the Lagrange multiplier $\eta$.

- Second-order KKT condition: for the Hessian of the Lagrangian $E - G^T \eta = E - \eta^T G$ given by
  $$\tilde{H}(x) = H_E(x) - \eta(x)^T H_G(x),$$ (6)
  we let the projected Hessian be
  $$\tilde{H}_G(x) = P(G(x)) \tilde{H}(x) P(G(x)),$$ (7)
  then $\tilde{H}_G(x^*)$ has one negative eigenvalue with an eigenvector normal to constraint at $x^*$, some zero eigenvalues with eigenvectors in the tangent space of $G$ at $x^*$, and rest eigenvalues all being positive.

To construct the dimer system, we let $x_1$ and $x_2$ be two end points of a dimer (a vector connecting $x_1$ and $x_2$) in $\mathcal{H}$ with a 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

$$x_\alpha = (2 - \alpha)x_1 + (\alpha - 1)x_2,$$

(8)

where the parameter $\alpha \in [1, 2]$ gives us the freedom to choose a point other than the geometric center (the midpoint of the dimer corresponding to $\alpha = 3/2$). For notation convenience, let

$$F_i = -\nabla E(x_i), \quad \text{and} \quad F_\alpha = -\nabla E(x_\alpha)$$

(9)

be natural forces at the two endpoints and the rotating center of the dimer. Similarly, for $i = 1, 2$, we let

$$G_i = G(x_i), \quad \text{and} \quad \nabla G_i = \nabla G(x_i),$$

(10)

$$G_\alpha = G(x_\alpha) = G((2 - \alpha)x_1 + (1 - \alpha)x_2)$$

(11)

and

$$\bar{G}_\alpha = (2 - \alpha)G_1 + (1 - \alpha)G_2,$$

(12)

which is an interpolation of $G_\alpha$. Moreover, we let $H_{G_\alpha} = H_G(x_\alpha)$ and $H_{G_i} = H_G(x_i)$ be the Hessian operators of $G$ at $x_\alpha$ and $x_i$ ($i = 1, 2$) respectively.

To control the dimer length, we introduce an auxiliary function $E_l = E_l(l)$ defined on $\mathbb{R}_+ = [0, +\infty)$ with $l = 0$ being its global minimum. Following [12], we take $E_l(l) = l^2/2$ which, as seen later for the SDD and CSDD, gives an exponential decay of the dimer length. Other choices, such as $E_l(l) = l^3$ for a more gradual polynomial decay, are also possible.

### 2.2. Formulation of the original shrinking dimer dynamics

The SDD developed in [12] for computing an index-1 saddle point is a dynamic system for the variables $(x_\alpha, v, l) = (x_\alpha(t), v(t), l(t)) : \mathbb{R}_+ \rightarrow \mathcal{H} \times \mathcal{H} \times \mathbb{R}_+$ that describe the dimer evolution. The SDD, in the absence of any constraint, is given by:

$$\begin{cases}
\mu_1 \dot{x}_\alpha = (I - 2vv^T)((2 - \alpha)F_1 + (1 - \alpha)F_2), \\
\mu_2 \dot{v} = (I - vv^T)(F_1 - F_2)/l, \\
\mu_3 \dot{l} = -E'_l(l),
\end{cases}$$

(13)

where $\mu_1, \mu_2, \mu_3$ are nonnegative relaxation constants and $\alpha \in [1, 2]$. The system (13) is coupled with the initial condition

$$x_\alpha(0) = x_0, \quad v(0) = v_0, \quad l(0) = l_0,$$

(14)
where \( l_0 > 0, x_0, v_0 \in \mathcal{H} \) with \( \|v_0\| = 1 \) in \( \mathcal{L} \).

As explained in [12], the SDD is a dynamic system formulation of the dimer method originally developed in [20], but subject to a diminishing dimer length in time. The first two equations in (13) represent the translation step and the rotation step, respectively. The operator \((I - 2vv^T)\) is the Householder mirror reflection which reverses the component of the natural force along \( v \), the dimer orientation direction. The operator \((I - vv^T)\) is a projection that makes \( v \) of unit length. Rather than using a dimer with a small but constant length as in the original dimer method in [20], the third equation in (13) follows the gradient flow of \( E_l = l^2 \) to allow the dynamic shrinking of the dimer length by forcing it to approach zero at the steady state. This ensures the converged steady state to be an exact saddle point. Moreover, some rigorous stability results have been presented in [12] for the SDD, showing that the only bounded stable equilibria of the SDD are those corresponding to index-1 saddle points of the original energy. By examining different time marching schemes for discretizing SDD, one also gets a systematic framework for analyzing the convergence of iterative procedures like the dimer method with a shrinking dimer length.

The SDD approach is similar to the recent studies of a dynamic system formulation of the dimer methods given by Poddey and Blochl [29] and the gentlest ascent dynamics (GAD) formulated by E and Zhou [16] for the gentlest ascent iteration given by Crippen and Scheraga [8]. The dynamic system view presented in [16, 29] provided a better angle to understand mathematically the saddle point search algorithms. Other dynamic system formulations of the saddle point search can be found in [4, 19]. In terms of the dynamic interpretations, SDD is very close to the GAD studied in [16, 32]. By formally taking the limit of zero dimer length, that is \( l \to 0 \) uniformly in time, SDD reduces to GAD in the limit but the latter formulation as well as the original gentlest ascent iteration given in [8] explicitly requires the Hessian of the energy. The original dimer method [20] and SDD, on the other hand, completely avoid computing the Hessian which is attractive in practical implementations especially for high dimensional complex energy surfaces. SDD provides an efficient way to approximate actions of the Hessian along the dimer orientation direction which is the only unstable direction around an index-1 saddle point. Indeed, utilizing forces (gradient vectors of the energy) computed at the dimer ends to estimate...
the Hessian action along the dimer orientation direction is a highlight of the
dimer approach which is also naturally inherited by SDD and the dynamic sys-
tem given in [29]. While the latter two approaches both avoid explicit Hessian
evaluation, SDD differs from the formulation in [29] in several aspects: first,
it employs a dynamically shrinking dimer length which has been shown to be
crucial for the convergence; secondly, it is a first order (in time) system in the
extended configuration space as opposed to the second order system derived in
[29] using a Lagrangian mechanics formulation coupled with the Rayleigh dis-
sipation so the numerical approximations of SDD are also simpler; in addition,
same as the original dimer method, the formulation in [29] corresponds to a
constant dimer length.

Note that for high dimensional problems where \( x \) has a large number of
components, the computational cost of the right hand side of (13) is essentially
associated with the two force evaluations at both ends of the dimer. It is possible
to replace the interpolated force \( \bar{F}_\alpha \) by the natural force \( F_\alpha = -\nabla E(x_\alpha) \) at the
dimer rotation center but this would require one additional force evaluation.

2.3. Formulation of the constrained shrinking dimer dynamics

With a given constraint \( G(x) = 0 \), let us pick

\[
\hat{F}_i = P(G)F_i, \quad i = 1, 2
\]

so that the projection is with respect to tangential space of the constraints at
the rotating center \( x \) of the dimer. The interpolated projected force, which is
an approximation of \( \hat{F}_\alpha = P(G)F_\alpha \) at the rotating center \( x_\alpha \) is given by

\[
\hat{F}_\alpha = (2 - \alpha)\hat{F}_1 + (\alpha - 1)\hat{F}_2.
\]

The constrained shrinking dimer dynamics (CSDD) then takes on the fol-
lowing form:

\[
\begin{cases}
\mu_1 \dot{x}_\alpha = (I - 2vv^T)\hat{F}_\alpha, \\
\mu_2 \dot{v} = (I - vv^T)(\hat{F}_1 - \hat{F}_2)/l + \nabla G^T \beta, \\
\mu_3 \dot{l} = -E'_l(l),
\end{cases}
\]

with the relaxation constants \( \mu_1, \mu_2, \mu_3 \), the functional \( E_l(l) \) all defined as in
Section 2, and an initial condition given by (14) except that we make the fol-
lowing compatibility assumption:

\[
G(x_0) = 0, \quad \|v_0\| = 1 \quad \text{and} \quad (\nabla G(x_0))^Tv_0 = 0.
\]
As for the dynamically modified Lagrange multiplier $\beta$ appearing in (17), we have several possible formulations. In the most general case, we let

$$\beta = \mu_1^{-1}\mu_2[(\nabla G_\alpha)^T \nabla G_\alpha |^{-1} v^T H_{G_\alpha}[\hat{F}_\alpha - 2v^T \hat{F}_\alpha v]]$$ (19)

In practice, there are other efficient approximations that avoid computing the Hessian of constraints. This is to be re-examined later in the discussion on the numerical implementation.

2.4. Analysis of stable equilibria of CSDD

Let $G, E, \nabla G_1, \nabla G_2, \nabla G, \hat{F}_1, \hat{F}_2, \hat{F}, P, \eta$ and $E_l$ be defined as in Section 2.3. For initial conditions given by (18), we first have

**Lemma 2.1.** For any $t > 0$, the dynamics in (17) satisfies

$$\|v\| = 1,$$ (20)

$$(\nabla G_\alpha)^T v = 0, \quad i.e. \quad P(G_\alpha)v = v,$$ (21)

$$G(x_\alpha) \equiv 0.$$ (22)

**Proof.** Using the formulation of CSDD and the expression of the multiplier $\beta$ given by (19), it can be checked that

$$(\nabla G_\alpha)^T \dot{v} + \dot{x}_\alpha^T H_{G_\alpha} v = l^{-1} v^T(\hat{F}_2 - \hat{F}_1)(\nabla G_\alpha)^T v / \mu_2.$$

That is

$$(\nabla G_\alpha)^T v \dot{t} = c(t)(\nabla G(x_\alpha))^T v,$$

where $c(t) = l^{-1} v^T(\hat{F}_2 - \hat{F}_1)/\mu_2$, which indicates that $(\nabla G_\alpha)^T v = 0$, and thus $P(G_\alpha)v = v$, if the initial data satisfy (18) and $\beta$ is given by (19).

Now using $(\nabla G_\alpha)^T v = 0$ and $(\nabla G_\alpha)^T \hat{F}_\alpha = 0$, we can also get

$$(\nabla G_\alpha)^T \dot{x} = 0$$

by multiplying $(\nabla G_\alpha)^T$ on both sides of the first equation of (17). This then implies that $G(x_\alpha(t)) = 0$ with initial data specified in the above. Finally, by multiplying $v^T$ to both sides of (17) and using $(\nabla G(x_\alpha))^T v = 0$, we can get $v^T \dot{v} = 0$. Thus, by the condition that $\|v_0\| = 1$, we get $\|v\| = 1$ for $t > 0$. □

Next, we demonstrate that steady states of (17) must correspond to critical points of $E = E(x)$ subject to $G(x) = 0$. Similar results for the unconstrained case can be found in [12].
Lemma 2.2. \((x^*, v^*, l^*)\) is a steady state of the dynamic system (17) if and only if the following conditions hold:

1. \(x^*\) is a critical point of the energy \(E\) subject to the constraint \(G(x) = 0\);
2. \(v^*\) is an eigenvector of \(\hat{H} = P\hat{H}P\) where \(\hat{H}\) is defined in (6);
3. \(l^* = 0\).

**Proof.** First, we show the necessary part. If \((x^*, v^*, l^*)\) is a steady state, then obviously \(l^* = 0\) by the construction of \(E\). We note that, since \((I - 2vv^T)\) is orthogonal (Householder transform), the right-hand-side of the first equation of (17) vanishes with \(x_1 = x_2 = x^*\) if and only if \(\tilde{F} = 0\) at the steady state. Since \(v\) is initially chosen to stay in the tangential space of the constraint, by Lemma 2.1, \(G(x^*) = 0\) and \(Pv = v\). Together, this indicates that \(x^*\) must be a critical point of the energy \(E\) with the constraint \(G(x) = 0\) due to the first-order KKT condition (4).

Note that the first term of the right-hand-side of the second equation of (17) are in the tangential space of \(G\) and third term is in the subspace spanned by \(\nabla G\). Setting the second equation equal to zero, we get

\[-P\hat{H}Pv + v^TP\hat{H}Pv = 0\quad\text{and}\quad\nabla G^T\beta = 0\]

at the steady state, since \(Pv = v\) and \(P = P^T\). This implies that \(v^*\) is an eigenvector of \(P\hat{H}P\). Similarly, we can also show the sufficiency. \(\square\)

We now show that linearly stable equilibria of (17) must correspond to index-1 saddle points of \(E = E(x)\) subject to \(G(x) = 0\), see [12, 16, 29] for similar computations.

**Theorem 2.3.** \((x^*, v^*, l^*)\) is a linearly stable steady state of the dynamic system (17) with \(\mu_1, \mu_2, \mu_3 > 0\) if and only if the following conditions hold:

1. \(x^*\) is an index-1 saddle point of the energy \(E\) with the constraint \(G(x) = 0\);
2. \(v^*\) is an eigenvector of \(\hat{H} = P\hat{H}P\) that corresponds to the smallest and only negative eigenvalue \(\lambda^*\). Here \(\hat{H}\) is defined in (6);
3. \(l^* = 0\).

**Proof.** Based on Lemma 2.2, we only need to verify conditions on the linear stability.
We start by examining the Jacobian (operator) of the right-hand-side of (17), denoted by $J$. At the steady state, $J$ takes on the following form:

$$
J = (J_{ij}) = \begin{pmatrix}
2\lambda^* v^* v^* T - \hat{H}(x^*) & 0 & 0 \\
0 & 2\lambda^* v^* v^* T - \hat{H}(x^*) + \lambda^* I & 0 \\
0 & 0 & -H_l
\end{pmatrix},
$$

where $\{J_{ij}\}$ denote the corresponding blocks of $J$ and $H_l$ is the Hessian (second derivative) of $E_l = E_l(l)$ at $l = 0$ which is positive.

We can show that the eigenvalues of $J_{11}$ are: $\lambda^*$ and the additive inverse of all other eigenvalues of $\hat{H}(x^*)$ and the corresponding eigenvectors of $J_{11}$ are the same as those of $\hat{H}(x^*)$. Also, for $J_{22}$, the eigenvalues are $2\lambda^*$ and $\lambda^* - \lambda$ where $\lambda$ is any of the remaining eigenvalues of $\hat{H}(x^*)$.

Thus, $(x^*, v^*, l^*)$ is a linearly stable steady state of (17) if and only if all eigenvalues of $J$ are negative. Equivalently, $\lambda^*$, as the eigenvalue of $\hat{H}(x^*)$ corresponding to $v^*$, must be the smallest and the only negative one for the energy $E$ with the constraint $G(x) = 0$. This completes the proof. □

3. Numerical Implementation of the CSDD

The theorem given in the previous section implies that the constrained saddle point search can be transformed into a dynamic path following of the CSDD to its stable steady state. Thus the CSDD turns an often unstable saddle computation in the original configuration manifold to a stable one in the extended configuration space that couples the manifold with its tangent hyperplane. We now discuss how CSDD can be used to develop discrete numerical iterations.

3.1. Time discretization

Similar to the original SDD, the CSDD can be discretized in time in various ways. Since the equation governing the reduction in the dimer length is independent from other equations, we may either implement the analytic solution or use a stable marching scheme. As for equations for the dimer center and orientation, the most straightforward discretization approach is an explicit marching in time scheme like the standard forward Euler method.

In [12], detailed stability analysis and optimal time step selection have been discussed for discretized SDD. Similar discussions for CSDD can be made but
they would be too technical for our objective here. Instead, we discuss some other important features associated with the treatment of constraints.

3.2. The case of linear constraints
Linear constraints are among the most popular conditions imposed in many practical applications. Both the PCVT and nucleation computation discussed later fall into this category.

If constraints are linear, then $\nabla G$ and thus $P(G)$ are independent of $x$. Moreover, $H_{G_\alpha} = 0$ so that $\beta = 0$ and the CSDD (17) is reduced to:

$$
\begin{cases}
\mu_1 \dot{\alpha} = (I - 2vv^T)\tilde{F}_\alpha, \\
\mu_2 \dot{\nu} = (I - vv^T)(\hat{F}_1 - \hat{F}_2)/l, \\
\mu_3 \dot{l} = -E'_l.
\end{cases}
$$

(23)

This is particularly elegant as the formulation of CSDD in this case remains essentially the same as that of SDD where natural forces $F_i$’s are simply replaced by $\tilde{F}_i$’s which are projections onto the tangent space of constraints.

3.3. Adopting an approximate Hessian of constraints
To compute the dynamic Lagrange multiplier $\beta$ given in (19) without the Hessian of constraints, we may mimic the idea of dimer system that uses differences of the forces to approximate actions of Hessian.

To this end, we may replace $v^T H_{G_\alpha}$ by $(\nabla G_2 - \nabla G_1)^T/l$ in (19) to get

$$
\beta = (l\mu_1)^{-1}\mu_2[(\nabla G_\alpha)^T \nabla G_\alpha]^{-1}(\nabla G_2 - \nabla G_1)^T[\tilde{F}_\alpha - 2v^T\tilde{F}_\alpha v].
$$

(24)

As both $\nabla G_1$ and $\nabla G_2$ are used in (24), it is advantageous to further avoid the computation of $\nabla G(x_\alpha)$, then, the new formula for $\beta$ becomes:

$$
\beta = (l\mu_1)^{-1}\mu_2[(\nabla G_\alpha)^T \nabla G_\alpha]^{-1}(\nabla G_2 - \nabla G_1)^T[\tilde{F}_\alpha - 2v^T\tilde{F}_\alpha v].
$$

(25)

Note that if the constraint $G$ is quadratic, then

$$
v^T H_{G_\alpha} = (\nabla G_2 - \nabla G_1)^T/l
$$

is exact. Moreover,

$$
\nabla \tilde{G}_\alpha = (2 - \alpha)\nabla G_1 + (\alpha - 1)\nabla G_2 = \nabla G(x_\alpha)
$$

also holds, therefore, (25) is equivalent to (19). For general constraints, (25) can be viewed as an approximation which, while not assuring the properties given in the lemma 2.1 at all time, becomes more accurate as $l$ goes down to 0.
3.4. An splitting strategy

For general constraints, we have also tested the following splitting strategy which worked very well in test problems. Let

\[ \psi_1(x, v, l) = (I - 2vv^T)\tilde{F}_\alpha, \quad \psi_2(x, v, l) = (I - vv^T)(\tilde{F}_1 - \tilde{F}_2)/l. \]

Given \((x^k, v^k, l^k)\), an approximation of \((x(t), v(t), l(t))\) at \(t = t_k\), the approximation at \(t_{k+1} = t_k + \Delta t\) is computed by the following:

1. find \(\tilde{x}^{k+1}\) by the forward Euler approximation to the first equation in (17):
   \[ \tilde{x}^{k+1} = x^k + \Delta t\psi_1(x^k, v^k, l^k); \]
2. update \(x^{k+1}\) by finding the closest point to \(\tilde{x}^{k+1}\) on \(G(x) = 0\), that is,
   \[ x^{k+1} = \text{argmin}\{ (x - \tilde{x}^{k+1})^T (x - \tilde{x}^{k+1}) \mid G(x) = 0 \}; \]
3. compute \(\tilde{v}^{k+1}\) by
   \[ \tilde{v}^{k+1} = v^k + \Delta t\psi_2(x^{k+1}, v^k, l^k); \]
4. project \(\tilde{v}^{k+1}\) onto the tangential space by \(\hat{v}^{k+1} = P(G(x^{k+1}))\tilde{v}^{k+1};\)
5. normalize \(\hat{v}^{k+1}\) to get \(v^{k+1} = \hat{v}^{k+1}/\|\hat{v}^{k+1}\|;\)
6. update \(l^{k+1}\), and go back to 1.

3.5. Parameter tuning

The parameters involved in the numerical implementation of the CSDD include the weight parameter \(\alpha\) which picks the center of the dimer, the time step and relaxation constants.

We take the choice of \(\alpha = 1.5\) here to get the best precision in approximating the projected Hessian along the dimer. For discussions on other choices of \(\alpha\) in the context of SDD and the discrete dimer methods, we refer to [12, 23].

In [12], linear stability analysis has been conducted on the maximum time-step allowed for the Euler marching scheme. Optimal step size has been investigated there for the most effective error reduction in the time iteration. These results provide some theoretical insight and guidance. However, as the analysis depends on the spectral properties of the Hessian at the saddle point which, in the most general case, is unavailable, so often the actual time steps taken will be problem dependent and trial based.
As for relaxation constants $\mu_i$, following discussions in [12], we notice that there are generally two possibilities, one with $\mu_2$ being close to zero, which means to first rotate the dimer until convergence then perform the translation step [23]. This tends to offer more effective error reduction at the cost of extra work at each iteration. In cases where estimations of the extreme eigenvalues of the Hessian at the constrained saddle point are available, another possibility is to achieve a similar error reduction rate for a nonzero $\mu_2$ whose value is chosen based on $\mu_1$ and $\Delta t$. A detailed analysis of the latter situation in the unconstrained case can be found in [12].

Picking a right initial guess for $x_0$ and $v_0$ is certainly crucial for the efficient time integration of CSDD. Naturally, an initial guess that is closer to the actual solution tends to require less time marching steps for the convergence. For a given $x_0$, if it is close a local equilibrium, it is a good strategy to take $v_0$ in the negative gradient force direction to maximize the chance of escaping the local well. On the other hand, if it is close to the saddle, then it is safer to use the gentlest ascend direction to avoid overshooting. In general, when no prior information is available, we simply pick them at random. By earlier analysis, we may expect eventual convergence so long the solutions stay bounded.

4. Numerical Experiments and Applications

In this section, we show how CSDD can be implemented to solve a couple of interesting application problems. Ideally, it would be desirable to test CSDD on a set of well-defined bench-mark problems, like the case for SDD presented in [12] and in comparison with other existing approaches such as the constrained string method [13]. However, a widely accepted set of constrained saddle point search problem has not been developed in the literature. We thus focus on testing on a few problems that can lead to not only performance check on the numerical procedures but also interesting conclusions based on the numerical simulations. The first application is a finite dimensional example with a linear constraint related to the popular concept of centroidal Voronoi tessellations, while the second example is a case of quadratic constraints related to the famous Thomson’s problem and its generalizations. The third example is given by an infinite dimensional case related to the critical nucleation in phase transition problems.
4.1. Application to periodic centroidal Voronoi tessellations

Centroidal Voronoi tessellations (CVTs) are Voronoi tessellations whose generators coincide with the mass centroids of the respective Voronoi regions [9]. In [37], periodic centroidal Voronoi tessellations (PCVTs) have been studied. They are defined with respect to a unit cell $U$, which, in $\mathbb{R}^2$, may be defined as

$$U = \{ \mathbf{x} \in \mathbb{R}^2 \mid \mathbf{x} = x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2, \quad 0 \leq x_1, x_2 < 1 \}.$$

for two linearly independent basis vectors $\{ \mathbf{a}_i \in \mathbb{R}^2 \}_{i=1}^2$. A PCVT, simply put, is the restriction in $U$ of a translational invariant CVT, with respect to $U$, defined in the whole space. Similar as the case of CVTs, we may associate the PCVTs as those associated with the critical points of an energy for a set of generators $\hat{\mathbf{Y}} = \{ \hat{\mathbf{y}}_i \}$ in $U$:

$$\mathcal{E}(\hat{\mathbf{Y}}) = \sum_{i=1}^n \int_{\hat{V}_i} \rho(\mathbf{x}) \| \mathbf{x} - \hat{\mathbf{y}}_i \|^2_U \, d\mathbf{x}, \quad (26)$$

where $\hat{\mathbf{V}} = \{ \hat{V}_i \}_{i=1}^n$ are the restrictions in the unit cell $U$ of the Voronoi tessellation associated with points in $\hat{\mathbf{Y}}$ and their respective periodic replica in the whole space ($\hat{V}_i$ is effectively the union of all restricted Voronoi regions associated with $\hat{\mathbf{y}}_i$ and all of its periodic replicas.) For any two points $\mathbf{x}, \mathbf{y} \in U$, this definition is equivalent to define their $U$-distance, with respect to the periodicity of unit cell as:

$$\| \mathbf{x} - \mathbf{y} \|^2_U = \min \{ \| \mathbf{x} - \mathbf{y} + k_1 \mathbf{a}_1 + k_2 \mathbf{a}_2 \|, \quad \forall \mathbf{k} = (k_1, k_2)^T \in \mathbb{Z}^2 \}.$$

For PCVTs, given the global translation invariance, one may impose a constraint on the set of generators in the unit cell,

$$\sum_{i=1}^n (\mathbf{y}_i - 0.5 \mathbf{a}_i) = 0 \quad (27)$$

which leads to the application of CSDD.

In [37], the CSDD was primarily used to explore the landscape of the PCVT energy and finding saddle points, along with some interesting applications to the optimal design of photonic crystals. The performance of the algorithm, as well as its dependence on the parameters was not carefully examined. Here, to better assess the convergence behavior of the CSDD algorithm, we take a special case to conduct some bench-mark tests. The case under consideration
corresponds to the unit square being the unit cell and with only two distinct
generators per unit cell, that is, \( n = 2 \), in the above discussions. Given the
constraint (27), the total degree of freedom is effectively two so that we may visualize the energy surface easily [37].

First of all, as discussed earlier, the parameter \( \mu_2 \) can affect the performance of the CSDD method. In Fig. 1, we plot, in the left picture, a PCVT corresponding to a saddle point of the energy. Notice that the generators are the centroids of the corresponding Voronoi regions in the whole space, which differ from the conventional CVTs defined in the domain. In the center picture of Fig. 1, different trajectories are presented for one of the generators \((x_1, x_2) = (0.25, 0.25)\), starting at the same initial condition, using \( \mu_1 = 0.2, \mu_3 = 1, l_0 = 0.01, \Delta t = 0.5 \) and \( E_l(s) = (10s)^3 \), but with \( \mu_2 \) varying from 0.2 to 0.25. The corresponding time histories of the energy are also shown for all trajectories in the picture on the right picture of the figure. The results show that the trajectories wander around more wildly for larger \( \mu_2 \) and take more direct paths to the saddle point for smaller \( \mu_2 \) which is consistent with the theoretical discussion given in [12] for the original SDD.

![Figure 1: On the left, PCVT with \( n = 2 \) and \( \rho = 1 \). In the center, some CSDD trajectories for the generator \((x_1, x_2) = (0.25, 0.25)\) with the horizontal axis for \( x_1 \) and the vertical axis for \( x_2 \). On the right, the evolution of the associated energy (vertical axis) changes in time (horizontal axis).](image)

As the goal is to find the steady state effectively rather than accurately computation of the time history, to help quantify the convergence behavior, in figure 2, we plot the ratios of the distances between the time-dependent solutions and the limiting saddle point at consecutive time iterations. Starting at the same initial condition, using the same parameters \( \mu_1, \mu_2, \mu_3, l_0 \) as before but with a few different values of the time-step \( \Delta t \) with the largest time step

16
being $\Delta t = 2$. The left picture corresponds to the choice of $E_l(s) = (10s)^3$ while the right picture corresponds to $E_l(s) = 5s^2$. The quantitative behavior of the two pictures remains similar, but it is obvious that the larger step size leads to a more effective error reduction per step and thus shows faster convergence to the steady state (saddle point). The rate of reduction in error matches well with the form $1 - c\Delta t$ for some positive constant $c$, which is consistent to a similar analysis given in \[12\] for some discretized SDD. One can see that for very small $\Delta t$, the rate of error reduction saturates near the value 1.

Due to the use of explicit time marching for both dimer location and orientation, a stability analysis similar to that given in \[12\] would imply that one cannot take $\Delta t$ too large. As a numerical illustration, we plot in Fig.3 some trajectories, starting at the same position but with two different initial orientation vectors (shown in two separate pictures) with $\Delta t = 1$ and $\Delta t = 8$ for each initial choice while using the same parameters as before for the trajectories shown in the figure 1. It is evident that the larger $\Delta t$ fails to produce a saddle point but rather develops into a periodic orbit. We refer to \[37\] for other interesting observations made with PCVT computations using CSDD.

4.2. Application to the generalized Thomson problem

The original Thomson problem \[35\] asks for the minimal-energy configuration of $N$ classical particles/charges confined to the unit sphere $S^2$ which interact via a Coulomb potential. The generalized Thomson problem \[7\] is an
extension where the potential energy for \( Y = \{ x_i \} \subset S^2 \) is defined as

\[
E_f(Y) = \frac{1}{2} \sum_{x_i, x_j \in S^2, x_i \neq x_j} f(|x_i - x_j|).
\]  

(28)

where \( f \) is a decreasing and convex function. There have also been experimental studies on the energy minimum configurations in the literature [1] as well as the mathematical analysis and computational studies. [2, 7]. The observed physical phenomena are likely related not only to the minimum energy configurations but also the transition states. To offer further insight to the energy landscape, we apply the CSDD to the energy function (28). Here, we only present some results for \( N = 5 \) with \( f(r) = 1/r^s \) which is considered as a very interesting but also peculiar case, as described in [7] and studied in [34] with computer-assisted mathematical proofs. It has been claimed that, based on available computational evidence, the configuration associated with vertices of the triangular bipolar pyramid (TBP) gives a local energy minimizer for \( s < s_1 \approx 21.147123 \) which is also a global energy minimizer for \( s < s_2 \approx 15.040808 \). The TBP has two vertices at two antipodal points and the remaining 3 points forming an equilateral triangle on the equator midway between the two antipodal points. The associated energy value can then be determined by \( g(s) = f(1/2) + 3f(\sqrt{3}) + 6f(\sqrt{2}) \). It is also believed that for \( s > s_2 \) the global minimizer is a square pyramid (SP), that is, a pyramid having a square base with the precise configuration depending on \( s \) in this range [34]. In the literature, there has not been detailed descriptions of the phase or energy diagrams for different stationary configura-
tions of the energy even in this simple $N = 5$ case. Applying the CSDD, we find a new saddle point configuration for some values of $s$ within $(s_2, s_1)$ which corresponds to a double-tetrahedron (DT) that can be obtained from the SP by folding the base along one of the diagonals. Thus, we find three branches of critical points for the constrained energy.

To describe in details, we first discuss the implementation. The problem is obviously subject to quadratic constraints at each points on the sphere. Due to the rotation symmetry, we specify one of the points to be $(0, 0, 1)$ and also specify an additional point to lie on the $y-z$ plane. Effectively, this reduces the problem to a energy functional defined for 11 variables subject to 4 quadratic constraints. We take either random initial guesses in the CSDD when no good initial guess is available or a choice based on parameter continuations, in particular, with respect to $s$. We use unit relaxation constants, together with a time step that ranges from $\Delta t = 1e - 2$ for $s$ around 0.5 to larger step sizes on the order of $\Delta t = 10$ for $s$ close to $-35.5$. We give three different solutions corresponding to $s = 15.5$ in Figs.4-6 respectively. Different viewing angles are provided. In all of these solutions, one may see that there exist three points forming an isosceles triangle on a great circle. We refer to the plane formed by these three points the center plane as it also acts as a mirror of symmetry so that the other two points, which are called the top and bottom vertices, are located symmetrically on both sides of the center plane. Moreover, the top and bottom vertices also form a second plane of symmetry, together with the apex of the isosceles triangle, that is the vertex sharing the two equal edges. The line connecting the top and bottom vertices goes through the center of the equilateral triangle on the center plane (which is also the center of the sphere), for the TBP configuration, while it intersects with the edge opposed to the apex of the isosceles triangle for the SP configuration so that the four co-planar vertices form the square base of the pyramid. For the DT configuration corresponding to the saddle point, it is fitting to see the line connecting the top and bottom vertices sits in between the same lines corresponding to the TBP and SP configurations. Similar observation can be made for the edge of the isosceles opposing the apex. Moreover, for the DT configuration, the negative eigenvalue of the Hessian (and the projected Hessian) corresponding to the unstable direction is numerically estimated to be $-8.215e - 2$ (and $-5.616e - 4$ respectively).
Figure 4: The TBP configuration, as a local minimum of the energy for $s = -15.5$, is shown from different angles. The rightmost plot shows the $x-y$ projection.

Figure 5: The SP configuration, as the global minimum of the energy for $s = -15.5$, is shown from different angles.

In Fig. 7, for different values of $s$, the graph of $g = g(s)$ (green curve) is given together in the left picture with the numerically computed energy values associated with the solutions on the three different branches. It can be seen that the energy values are very close to each other when the different solution branches coexist. Thus, in the right picture, we plot the differences of the energy between the SP branch and the TBP branch and also the differences of the energy between the DT branch and the TBP branch. This clearly indicates the change of the global minimizer for $s$ close to $s \approx 15$.

Similar to the case of $s = 15.5$, all saddle point DT configurations computed for different values of $s$ share the same symmetry properties, see Fig. 8 for DT configurations corresponding to $s = -14.5$, $-16.5$ and $-18.5$ respectively. Moreover, one can see that as $s$ gets larger, the distance between the line joining the top and bottom vertices and the edge (line segment) of the isosceles triangle opposing to the apex gets smaller. In two separate limiting cases, such a distance has a minimum value of 0 for the SP configuration in one limit, while its maximum value is equal to 1 corresponding to the TBP configuration for the other limit. More studies on systems involving larger number of particles and
more complex interaction potentials are currently underway.

4.3. Application to nucleation in solid state phase transformation

As the second example, the CSDD is applied to the computation of critical nuclei in phase transformations under the diffuse-interface framework developed in [38] which is based on the nonclassical nucleation theory developed in [3]. For illustration, we consider the case where a single phase field variable $\phi$ is used to model the compositional difference between the parent phase and the nucleating phase in a binary system. Let $\Omega = (-1, 1)^2$, the total free energy is given by

$$E(\phi) = \int_{\Omega} \frac{\alpha}{2} |\nabla \phi|^2 + f(\phi)|dx,$$

where $\alpha$ is a constant to describe the surface energy and $f(c) = \frac{1}{4}(c^2 - 1)^2 - \frac{1}{4}(c_0^2 - 1)^2 - (c - c_0)(c_0^3 - c_0)$ is a double well potential with $c_0$ being the average composition. The natural force of $E$ is given by $F(\phi) = \alpha \Delta \phi - (\phi^3 - \phi - c_0^3 + c_0)$. 
For the critical nuclei in a conserved field, the composition profile is subject to the constraint

$$\int_{\Omega} (\phi - c_0) dx = 0.$$  \hspace{1cm} (30)

With regards to the Hilbert space setting used earlier, we let $H$ be the space of functions in $H^1(\Omega)$ with the periodic boundary condition and $L = L^2(\Omega)$. We then compute the constrained saddle point of $E$ subject to (30) by first discretizing the functional in space with Fourier spectral methods [38]. Due to the high dimensionality associated with the spatial discretization, to reduce the stiffness of the time integration of the CSDD, we adopt a semi-implicit integration scheme given by

$$\begin{align*}
\phi^{n+1} &= \phi^n + \frac{\Delta t}{\mu_1} \left[ \alpha \Delta \phi^{n+1} - 2\phi^{n+1} - (\phi^n)^3 + 3\phi^n - 3(\ell^n)^2 \phi^n (v^n)^2 \right] \\
&\quad + \frac{\Delta t}{4\mu_1} \int_{\Omega} \left[ (\phi^n)^3 - \phi^n + 3(\ell^n)^2 \phi^n (v^n)^2 \right] dx \\
&\quad + \frac{2\Delta t}{\mu_1} v^n \int_{\Omega} \left[ \alpha \nabla \phi^n \nabla v^n + ((\phi^n)^3 - \phi^n) v^n \right] dx,
\end{align*}$$

and we allow the dimer length to change by $\ell^{n+1} = \ell^n / (1 + \Delta t / \mu_3)$ which corresponds to a backward Euler discretization with $E_l(s) = s^2/2$. The semi-implicit discretization may also be viewed as an operator splitting strategy aimed at improving the efficiency of the CSDD. More detailed studies on optimal splitting method are left for future works. Here, as in [38], we use periodic boundary conditions and Fourier spectral methods. Fig.9 and Fig.10 shows the critical

Figure 8: DT configurations corresponding to $s = -14.5, -16.5$ and $-18.5$ respectively, with only the $x - y$ projections shown in all cases.
nuclei and their respective unstable directions for $c_0 = -0.85$ with $\alpha = 0.0025$ and $\alpha = 0.0004$ respectively. Note that in each of the plots, only the portion of the solution near the center of the domain is shown. As $c_0$ is close to the spinodal point, the interface of critical nucleus diffuse. The composition value at the center of a critical nucleus decreases and is smaller than the nucleating composition of equilibrium solution, which is consistent with the non-classical nucleation theory as well as the computation given in [39] using a constrained string method [13]. The CSDD also allows us to provide information on the unstable directions which have not been reported in the literature before.

5. Conclusion

In this paper, a constrained shrinking dimer dynamics is developed to locate index-1 saddle points associated with an energy functional defined on a constrained manifold. Both theoretical analysis and computational experiments are carried out. In comparison with the original SDD with no constraints, the CSDD system works by forcing the dimer center to stay on the constrained manifold and the dimer orientation to stay in the tangent space of constraints. This is achieved via tangent projections of the natural forces and suitable choices of the Lagrange multipliers. Similar to the constraint-free SDD, CSDD requires
only the evaluation of the projected natural forces at two closely lying points while avoiding the calculation of the Hessian of the energy functional. Moreover, we also show how one may approximate the projected Hessian of the constraints which may be exact for linear and quadratic constraints. CSDD is also discretized and applied to problems including the computation of periodic centroidal Voronoi tessellations, generalized Thomson problems on particle distribution on the unit sphere and critical nuclei morphologies in binary phase transformations. Some numerical implementation issues are also discussed, such as parameter tuning and choices of initial guesses, to demonstrate how they affect the performance of the algorithm. The examples are preliminary studies of respective problems as they are only presented to demonstrate the effectiveness of the CSDD as a computational tool for constrained saddle point search.

There are naturally various issues worthy further investigation. On the algorithmic side, how to best approximate the Hessian of constraints will be of significant practical interests for more general constraints. The remedy provided in this paper works effectively for linear and quadratic cases where the approximations remain exact. It would be desirable to find out effective numerical approximations that, for more general cases, even if they may cause the unstable direction to get off the tangential space of the constraints during the time evolution, the constraints would remain satisfied at the steady state. While the Lagrange multiplier formulation is adopted here, one may explore if some augmented Lagrangian formulation can also be developed for the CSDD [13, 18]. The extension to constrained saddle of higher indices is surely possible and may be investigated further. Other time-stepping strategies as well as adaptive time integrations may also help improving the efficiency. Theoretical analysis of stability properties and global convergence properties of various time discretization schemes, such as the splitting strategy documented in section 3.4, can also be complementary. On the application side, there are surely more interesting questions to be answered via the CSDD to explore more complex energy landscapes including those similar to the applications considered here. For instance, one may consider other practically used potentials for the generalized Thomson problems that may model colloidal particles [1] and quantized vortices [11] which would help us to understand the nature and properties of the geometric and topologically frustrated systems. Another important task is to
develop an extensive list of bench-mark problems related to the constrained saddle point search so that more detailed comparison studies can be carried out by implementing both the CSDD type algorithms and other existing algorithms developed in the literature. This will further demonstrate the effectiveness and/or limitations of the CSDD in practical applications. Finally, the current work only treat equality constraints, how to deal with inequality constraints will also be an interesting problem to be studied further.

Acknowledgment

The authors would like to thank the referees for their careful reading and suggestions that have helped us improving the manuscript.

References


[6] C. Cerjan and W. Miller, On finding transition states,


[34] R. Schwartz, The 5 electron case of Thomson’s problem, preprint, 2010


