

**Original citation:**

Gould, Nicholas, Ortner, Christoph and Packwood, David. (2016) A dimer-type saddle search algorithm with preconditioning and linesearch. *Mathematics of Computation*, 85. pp. 2939-2966.

**Permanent WRAP URL:**

<http://wrap.warwick.ac.uk/86334>

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

**Publisher's statement:**

First published in *Mathematics of Computation* published by the American Mathematical Society. © 2016 American Mathematical Society.

**A note on versions:**

The version presented here may differ from the published version or, version of record, if you wish to cite this item you are advised to consult the publisher's version. Please see the 'permanent WRAP URL' above for details on accessing the published version and note that access may require a subscription.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk)

# A DIMER-TYPE SADDLE SEARCH ALGORITHM WITH PRECONDITIONING AND LINESEARCH

N. GOULD, C. ORTNER, AND D. PACKWOOD

**ABSTRACT.** The dimer method is a Hessian-free algorithm for computing saddle points. We augment the method with a linesearch mechanism for automatic step size selection as well as preconditioning capabilities. We prove local linear convergence. A series of numerical tests demonstrate significant performance gains.

## 1. INTRODUCTION

The problem of determining saddle points on high dimensional surfaces has received a great deal of attention from the chemical physics community over the past few decades. These surfaces arise, in particular, as potential energies of molecules or materials. The local minima of such functions describe stable atomistic configurations, while saddle points provide information about the transition rates between minima in the harmonic approximation of transition state theory. Independently, they are useful for mapping the energy landscape and are used to inform accelerated MD type schemes such as hyperdynamics [25, 23] or kinetic Monte Carlo (KMC) [26].

While the problem of determining the minima of such an energy function is well known in the numerical analysis community, the problem of locating saddles point has received little attention. Saddle search algorithms can be broadly categorised into two groups.

The first group has been called ‘chain of states’ methods. A chain of ‘images’ are placed on the energy surface, often the two end points of the chain are placed at two different local minima, for which the connecting saddle is being sought. The chain is then ‘relaxed’ by some dynamics for which the minimum energy path (MEP) is (thought to be) an attractor. Two archetypical methods of this class are the nudged elastic band (NEB) method [12] and the string method [27, 28].

The second group of methods for finding the saddle have been called ‘walker’ methods. Here a single ‘image’ moves from its initial point (sometimes, but not obligatorily, a local minimum) until it becomes sufficiently close to a saddle point. The first method to work in this framework was Rational Function Optimization (RFO) and later its derivative, the Partitioned RFO (PRFO)[7, 22, 3]. Here, the full eigenstructure of the Hessian is explicitly calculated and then one or more eigenvalues are manually shifted. In particular, if the minimum eigenvalue is shifted in the correct manner, and a Newton step is applied using the resultant modified Hessian, then the walker moves uphill in the direction corresponding to the lowest eigenvector and downhill in all other directions. If the Hessian is expensive to calculate,

---

*Date:* July 15, 2015.

*2000 Mathematics Subject Classification.* 65K99, 90C06, 65Z05.

*Key words and phrases.* saddle search, preconditioning, convergence, dimer method.

This work was supported in part by EPSRC grants EP/J021377/1 and EP/J022055/1.

or even unavailable, it, or at least its action, can be approximated as the computation proceeds by any variety of techniques, for example the symmetric rank-one approximation [19]. Of course any useful Hessian approximation should necessarily have the flexibility to be indefinite. Other walker type techniques are satisfied with computing the lowest eigenpair only. One such technique is the Activation Relaxation Technique (ART) nouveau [17, 16, 18, 6]. The original ART method used an ascent step not along the minimum eigenvector, but along a line drawn between the image and a known local minimum [4, 5]. In ART nouveau this is replaced by the minimum eigenpair which is calculated by means of the Lanczos [14] method.

The technique which forms the basis of the present paper, is the *dimer method* [10, 11]. In this method a pair of ‘walkers’ is placed on the energy surface and aligned with the minimum eigenvector (irrespective of the sign of the corresponding eigenvalue) by minimizing the sum of the energies at the two end points. This can be thought of as the computation of the minimal eigenvalue using a finite difference approximation to the action of the Hessian matrix. In practice this ‘rotation step’ is not converged to great precision. More advanced modifications can be used to improve walker search directions, e.g., an L-BFGS [15] scaling, rather than a default steepest descent type scheme [13].

In the only rigorous analysis of the dimer method that we are aware of Zhang and Du [29] prove local convergence of a variant where the ‘dimer length’ (the separation distance between the two walkers) shrinks to zero. In that work the dimer evolution is treated as a dynamical system, and the stability of different types of equilibria is investigated.

In the present paper we present three new results:

- (1) We augment the dimer method with preconditioning capabilities to improve its efficiency for ill-conditioned problems, in particular with an eye to high-dimensional molecular energy landscapes. This modification is based on the elementary observation, common in numerical optimization and linear algebra, that the dimer method can be formulated with respect to an arbitrary inner product. Previously the  $\ell^2$ -inner product was used almost exclusively; the only exception we are aware of being the use of the  $H^{-1}$  inner product in order to mimic the conserved dynamics of the Cahn-Hilliard equation [30].
- (2) We introduce a linesearch procedure. To that end, the main difficulty is the absence of a merit function for saddles. Instead, we propose a *local merit function*, which we minimise at each dimer iteration using traditional linesearch strategies from optimisation, and which is updated between steps. We remark that [6] introduces linesearch to the relaxation step of the ART nouveau method. By contrast, our linesearch procedure is applied to combined ascent/descent directions.
- (3) In the analysis of Zhang and Du [29] the dimer length,  $h$ , is shrunk to zero to ensure that the dimer converges to a saddle. As already noted in [29], due to round-off error this shrinking cannot be done to an arbitrary level and may need to be adaptively controlled in practice. We present a variation of the analysis in [29] showing that, if it is kept fixed, then the dimer walkers converge to a point that lies within  $O(h^2)$  of a saddle. We also extend this analysis to incorporate preconditioning and linesearch.

Concerning (2), it would of course be preferable to construct a global merit function as this would provide a path towards constructing a globally convergent scheme.

Indeed, our (non-trivial) generalisation of the convergence analysis to the linesearch variant of the dimer method only yields local results, and we even present a (somewhat academic) counterexample to global convergence.

The paper is organised as follows: In §2, after introducing preliminary concepts, we describe the basic dimer method, and establish its local convergence. In §3, a linesearch enhancement is proposed, and its local convergence behaviour is analysed. Numerical experiments illustrating the advantages of the linesearch are given in §4. We conclude in §5. Full details of our analysis are given in Appendix A.

## 2. LOCAL CONVERGENCE OF THE DIMER METHOD

**2.1. Preliminaries.** Let  $X$  be a Hilbert space with norm  $\|x\|$  and inner product  $x \cdot y$ . We write  $x \perp y$  if  $x \cdot y = 0$ .  $I : X \rightarrow X$  denotes the identity. For  $x, y \in X$ ,  $x \otimes y : X \rightarrow X$  denotes the operator defined by  $(x \otimes y)z = (y \cdot z)x$ . The unit sphere is denoted by  $S_X := \{v \in X \mid \|v\| = 1\}$ .

Given two real functions  $f$  and  $g$  defined in some neighbourhood  $\mathcal{N}$  of the origin, we say that  $f(x) = O(g(x))$  as  $x \rightarrow 0$  if  $|f(x)| \leq C|g(x)|$  for some constant  $C > 0$  and all  $x \in \mathcal{N}$ .

For a bounded linear operator  $A \in L(X)$  we denote its spectrum by  $\sigma(A)$ . We say that  $(\lambda, v) \in \mathbb{R} \times X$  is an eigenpair if  $Av = \lambda v$ . If  $(\lambda, v)$  is an eigenpair and  $\lambda = \inf \sigma(A)$ , then we call it a minimal eigenpair. We say that  $A$  has *index-1 saddle structure* if there exists a unique minimal eigenpair  $(\lambda, v)$  with  $\lambda < 0$  and  $A$  is positive definite in  $\{v\}^\perp$ .

If  $F : X \rightarrow \mathbb{R}$  is Fréchet differentiable at a point  $x$  then we denote its *gradient* by  $\nabla F(x)$ , i.e.,

$$\nabla F(x) \cdot y = \lim_{t \rightarrow 0} t^{-1}(F(x + ty) - F(x)).$$

(Note that  $\nabla F(x)$  is the Riesz representation of the first variation  $\delta F(x) \in X^*$ .) Similarly, if  $F : X \rightarrow X$  is Fréchet differentiable at  $x$ , then  $\nabla F(x) \in L(X)$  is a bounded linear operator satisfying  $\nabla F(x)u = \lim_{t \rightarrow 0} t^{-1}(F(x + tu) - F(x))$ . In particular, if  $F : X \rightarrow \mathbb{R}$ , then the Hessian  $\nabla^2 F(x) \in L(X)$  (rather than  $\nabla^2 F(x) : X \rightarrow X^*$ ). Higher derivatives are defined analogously, but we shall avoid their explicit use as much as possible.

Let  $E \in C^4(X)$ . We say that  $x_*$  is an *index-1 saddle* of  $E$  if

$$\nabla E(x_*) = 0 \quad \text{and} \quad \nabla^2 E(x_*) \text{ has index-1 saddle structure.} \quad (1)$$

With slight abuse of notation, we shall also call  $(x_*, v_*, \lambda_*)$  an index-1 saddle if  $x_*$  is an index-1 saddle and  $(v_*, \lambda_*)$  the associated minimal eigenpair.

Given a dimer length  $h$  and a vector  $v \in S_X$ , we define

$$\mathcal{E}_h(x, v) := \frac{1}{2}(E(x + hv) + E(x - hv)).$$

Finally, we observe that

$$\nabla_x \mathcal{E}_h(x, v) = \frac{1}{2}(\nabla E(x + hv) + \nabla E(x - hv)) = \nabla E(x) + O(h^2), \quad (2)$$

$$\nabla_x^2 \mathcal{E}_h(x, v) = \frac{1}{2}(\nabla^2 E(x + hv) + \nabla^2 E(x - hv)) = \nabla^2 E(x) + O(h^2), \quad (3)$$

$$\nabla_v \mathcal{E}_h(x, v) = \frac{h}{2}(\nabla E(x + hv) - \nabla E(x - hv)) = h^2 \nabla^2 E(x)v + O(h^4) \quad \text{and} \quad (4)$$

$$\nabla_v^2 \mathcal{E}_h(x, v) = \frac{h^2}{2}(\nabla^2 E(x + hv) + \nabla^2 E(x - hv)) = h^2 \nabla^2 E(x) + O(h^4), \quad (5)$$

$$\nabla_x \nabla_v \mathcal{E}_h(x, v) = h^2 \nabla^3 E(x) \cdot v + O(h^4), \quad (6)$$

where we note that these errors are uniform whenever  $x$  remains in a bounded set. For future reference, we define the discrete Hessian action operator

$$H_h(x; v) := h^{-2} \nabla_v \mathcal{E}_h(x, v). \quad (7)$$

**2.2. A simple dimer variant.** We now formulate a simple variant of the dimer method. This is a variation of the original dimer method [10, 21], alternating steps in the position ( $x_k$ ) and direction ( $s_k$ ) variables, but employing a modification proposed by [29]. Indeed, the following algorithm can be thought of as [29] with  $\lambda$  ( $h$  in our case) taken to be constant instead of  $h \rightarrow 0$  as  $k \rightarrow \infty$ .

### Simple Dimer Algorithm

- (0) **Input:**  $x_0 \in X, v_0 \in S_X, h > 0, (\alpha_k)_{k \in \mathbb{N}}, (\beta_k)_{k \in \mathbb{N}}$ .
- (1) For  $n = 0, 1, 2, \dots$  do
- (2)  $s_k := -(I - v_k \otimes v_k) h^{-2} \nabla_v \mathcal{E}_h(x_k, v_k)$
- (3)  $v_{k+1} := \cos(\|s_k\| \beta_k) v_k + \sin(\|s_k\| \beta_k) \frac{s_k}{\|s_k\|}$
- (4)  $x_{k+1} := x_k - \alpha_k (I - 2v_k \otimes v_k) \nabla_x \mathcal{E}_h(x_k, v_k)$ .

**Remark 1.** Another natural variation of the Simple Dimer Algorithm is to replace step (4) with

$$x_{k+1} := x_k - \alpha_k (I - 2v_k \otimes v_k) \nabla E(x_k),$$

i.e., to replace the averaged gradient with the centered gradient. This has the advantage that the method would converge to an exact saddle rather than an approximate saddle within an  $O(h^2)$  neighbourhood (cf. § 2).

For the sake of simplicity, we do not consider these variants, but we note that (i) all our results can be extended to these variants, and (ii) it seems to us that this has minor effects on the accuracy and efficiency of the algorithm, with the exception that it requires an additional gradient evaluation at each iteration. By employing a one-sided finite difference instead of a centered finite difference, this additional evaluation could again be removed, but at the cost of an  $O(h)$  accurate rotation instead of  $O(h^2)$ . This trade-off is well known [29].

However, it might be useful to “post-process” the dimer algorithms (including the Linesearch Dimer Algorithm in § 3.2).  $\square$

**2.3. The dimer saddle.** Our first observation is that the Simple Dimer Method approximates the action of the Hessian by a finite difference and the gradient by an average. Therefore, the iterates  $(x_k, v_k)$  with fixed dimer length  $h$  *cannot* in general converge to a saddle but only to a critical point  $(x_h, v_h)$  near a saddle, satisfying

$$\nabla_x \mathcal{E}_h(x_h, v_h) = 0 \quad \text{and} \quad (I - v_h \otimes v_h) \nabla_v \mathcal{E}_h(x_h, v_h) = 0. \quad (8)$$

The existence (and local uniqueness) of such critical points is established in the following result.

**Proposition 2.** *Let  $(x_*, v_*, \lambda_*)$  be an index-1 saddle, then there exists  $h_0 > 0$  such that, for all  $h \leq h_0$ , there exist  $x_h, v_h \in X$ ,  $\lambda_h \in \mathbb{R}$  and a constant  $C$ , such that*

$$\begin{aligned} \nabla_x \mathcal{E}_h(x_h, v_h) &\equiv \frac{1}{2} (\nabla E(x_h + hv_h) + \nabla E(x_h - hv_h)) = 0, \\ \frac{1}{h^2} \nabla_v \mathcal{E}_h(x_h, v_h) &\equiv \frac{1}{2h} (\nabla E(x_h + hv_h) - \nabla E(x_h - hv_h)) = \lambda_h v_h, \\ \|v_h\|^2 &= 1, \end{aligned} \quad (9)$$

and moreover

$$\|x_h - x_*\| + \|v_h - v_*\| + |\lambda_h - \lambda_*| \leq Ch^2. \quad (10)$$

*Idea of proof.* The result is a consequence of the inverse function theorem. Comparing (9) with the exact saddle  $(x_*, v_*, \lambda_*)$  the Taylor expansions (2)–(5) show that the residual is of order  $O(h^2)$  and that the linearisation is  $O(h^2)$  close (in operator norm) to the linearisation of the exact saddle system  $\nabla E(x_*) = 0, \nabla^2 E(x_*)v_* = \lambda_* v_*, \|v_*\| = 1$ . The linearisation of the latter is an isomorphism by the assumption that  $x_*$  is an index-1 saddle. The complete proof is given in A.1.  $\square$

We shall refer to a triple  $(x_h, v_h, \lambda_h) \in X \times X \times \mathbb{R}$  that satisfies (9) as a *dimer saddle*.

**2.4. Local convergence.** We now state a local convergence result for the Simple Dimer Algorithm.

**Theorem 3.** *Let  $(x_*, v_*, \lambda_*)$  be an index-1 saddle. Then there exists a radius  $r$ , a maximal dimer length  $h_0$  and maximal step sizes  $\bar{\alpha}$  and  $\bar{\beta}$  (independent of one another) as well as a dimer saddle  $(x_h, v_h, \lambda_h)$  satisfying (9) such that the following hold for all  $h \leq h_0$ :*

*Let  $x_0 \in B_r(x_*)$ ,  $v_0 \in B_r(v_*)$ ,  $\sup_k \alpha_k \leq \bar{\alpha}$ ,  $\sup_k \beta_k \leq \bar{\alpha}$ ,  $\inf_k \alpha_k > 0$ ,  $\inf_k \beta_k > 0$ , and let  $(x_k, v_k)$  be the iterates generated by the Simple Dimer Algorithm, then there exist  $C > 0, \eta \in (0, 1)$  such that*

$$\|x_k - x_h\| + \|v_k - v_h\| \leq C\eta^k (\|x_0 - x_h\| + \|v_0 - v_h\|). \quad (11)$$

*Idea of proof.* The proof is a modification of the proofs of [29, Thm. 2.1 and Thm. 3.1]. Upon linearisation of the updates about the exact saddle  $(x_*, v_*)$ , the updates can be re-written as (see § A.2 for the proof)

$$\begin{pmatrix} x_{k+1} - x_h \\ v_{k+1} - v_h \end{pmatrix} = \left[ I - \begin{pmatrix} \alpha_k A & 0 \\ \beta_k B & \beta_k C \end{pmatrix} \right] \begin{pmatrix} x_k - x_h \\ v_k - v_h \end{pmatrix} + O((\alpha_k + \beta_k)(h^2 + r_k)r_k), \quad (12)$$

$$\text{where } r_k^2 := \|x_k - x_h\|^2 + \|v_k - v_h\|^2,$$

$$A = (I - 2v_* \otimes v_*)\nabla^2 E(x_*), \quad C = (I - v_* \otimes v_*)\nabla^2 E(x_*) - \lambda_* I,$$

and  $B$  is a bounded linear operator whose precise form is unimportant.

Clearly,  $A, C$  are both symmetric and positive definite, hence the spectrum of  $\mathbf{A} = (\alpha A, 0; \beta B, \beta C)$  is strictly positive. If we chose  $\alpha_k \equiv \alpha, \beta_k \equiv \beta$  constant, then (11) follows from standard stability results for dynamical systems. The (straightforward) generalisation to non-uniform step sizes is given in [9].  $\square$

### 3. A DIMER ALGORITHM WITH LINESEARCH

In continuous optimisation, linesearch ensures both robustness and efficiency of algorithms. Our aim is to incorporate such a mechanism into the dimer method.

**3.1. Motivation: a local merit function.** A rotation step (Steps (2, 3) of the Simple Dimer Algorithm) is a descent step on the unit sphere, for which it is straightforward to implement a linesearch. By contrast, it is not obvious how to do this for the translation (Step (4) in the Simple Dimer Algorithm), which is an ascent step in the  $v_k$  direction but a descent step in the  $\{v_k\}^\perp$  space. A natural idea is to employ a merit function.

Let  $x_* \in X$  be an index-1 saddle with minimal eigenpair  $(v_*, \lambda_*)$ , and consider the modified energy functional

$$F(x) := E(x) + \frac{\kappa}{2}(v_* \cdot (x - x_*))^2.$$

Then,  $\nabla F(x_*) = 0$  and  $\nabla^2 F(x_*) = (I + \kappa v_* \otimes v_*) \nabla^2 E(x_*)$ , which is positive definite if and only if  $\kappa > -\lambda_*$ . It follows that  $x_*$  is a strict local minimizer of  $F$ .

Analogously, if  $(x_h, v_h, \lambda_h)$  is a dimer saddle point (cf. Proposition 2) and we define a modified energy functional

$$F_h(x) := \mathcal{E}_h(x, v_h) + \frac{\kappa}{2}(v_h \cdot (x - x_h))^2,$$

then choosing  $\kappa > -\lambda_*$  and  $h$  sufficiently small guarantees that  $x_h$  is a local minimizer of  $F_h$ . We can think of this procedure as ‘stabilising’ the saddle.

Since the dimer saddle  $(x_h, v_h)$  is unknown,  $F_h$  cannot be employed as a merit function. Instead, we construct a merit function that is updated at each dimer iteration to employ the best possible information available about the saddle. Given an iterate  $(x_k, v_k)$  we make the ansatz

$$F_k(x) := \mathcal{E}_h(x, v_k) + g_k \cdot (x - x_k) + \frac{\kappa_k}{2}(v_k \cdot (x - x_k))^2.$$

This merit function should have the property that the steepest descent direction at  $x = x_k$  is the dimer search direction, i.e.,

$$\nabla F_k(x_k) = (I - 2v_k \otimes v_k) \nabla_x \mathcal{E}_h(x_k, v_k),$$

which is achieved for the choice

$$g_k := -2(v_k \otimes v_k) \nabla_x \mathcal{E}_h(x_k, v_k).$$

Secondly, minimising  $F_k$  should yield an update  $y_k$  that is a substantial improvement over  $x_k$ . For  $(x_k, v_k)$  sufficiently close to  $(x_h, v_h)$  the inverse function theorem readily yields existence of a point  $\tilde{y}_k = x_h + O(h^2)$  such that  $\nabla_x \mathcal{E}_h(\tilde{y}_k, v_k) = 0$ . We now estimate the residual

$$\begin{aligned} \nabla F_k(\tilde{y}_k) &= \nabla_x \mathcal{E}_h(\tilde{y}_k, v_k) + g_k + \kappa_k(v_k \otimes v_k)(\tilde{y}_k - x_k) \\ &= g_k + \kappa_k(v_k \otimes v_k) \nabla_x^2 \mathcal{E}_h(x_k, v_k)^{-1} (\nabla_x \mathcal{E}_h(\tilde{y}_k, v_k) - \nabla_x \mathcal{E}_h(x_k, v_k)) + O(\|\tilde{y}_k - x_k\|^2) \\ &= g_k - \frac{\kappa_k}{\lambda_k}(v_k \otimes v_k) \nabla_x \mathcal{E}_h(x_k, v_k) + O(\|\tilde{y}_k - x_k\|^2 + h^2 \|\tilde{y}_k - x_k\|) \end{aligned}$$

where  $\lambda_k = H_h(x_k; v_k) \cdot v_k$  and we assumed, for simplicity, that  $\nabla_v \mathcal{E}_h(x_k, v_k) = 0$  which implies that  $\nabla_x^2 \mathcal{E}_h(x_k, v_k) v_k = \lambda_k v_k + O(h^2)$ . Recalling that  $g_k = -2(v_k \otimes v_k) \nabla_x \mathcal{E}_h(x_k, v_k)$ , the choice  $\kappa_k = -2\lambda_k$  guarantees that minimising  $F_k$  is comparable to performing a Newton step towards  $\tilde{y}_k$  and hence towards  $x_h$ .

**3.2. Dimer algorithm with linesearch.** Given an iterate  $x_k$ ,  $v_k$  and  $\lambda_k := v_k \cdot H_h(x_k; v_k)$ , we define the auxiliary functional  $F_k \in C^4(X)$ ,

$$\begin{aligned} F_k(x) &:= \mathcal{E}_h(x, v_k) - 2[(v_k \otimes v_k) \nabla_x \mathcal{E}_h(x_k, v_k)] \cdot (x - x_k) - \lambda_k \|(v_k \otimes v_k)(x - x_k)\|^2 \\ &= \mathcal{E}_h(x, v_k) - 2((v_k \cdot \nabla_x \mathcal{E}_h(x_k, v_k)) (v_k \cdot (x - x_k)) - \lambda_k (v_k \cdot (x - x_k))^2), \end{aligned} \quad (13)$$

motivated by the discussion in §3.1. Instead of locally minimising  $F_k$  we only perform a minimisation step in the steepest descent direction, using a standard linesearch procedure augmented with the following sanity check: For a trial  $x^t = x_k - \alpha \nabla F_k(x_k)$  we require that  $v_k$  is still a reasonable dimer orientation for  $x^t$  by checking the residual  $\|(I - v_k \otimes v_k)H_h(x^t; v_k)\|$ . If this residual falls above a certain tolerance then we reject the step and reduce the step size.

**Linesearch Dimer Algorithm:**

- (1) **Input:**  $x_0, v_{-1}, h$   
**Parameters:**  $\beta_{-1}, \alpha_0, \alpha_{\max} > 0, \Theta \in (0, 1), \Psi > 1$
- (2) **For**  $k = 0, 1, 2, \dots$  **do**
  - %% Rotation %%
  - (3)  $[v_k, \beta_k] := \mathbf{Rotation}[x_k, v_{k-1}, \beta_{k-1}]$
  - %% Translation %%
  - (4)  $p := -\nabla F_k(x_k)$
  - (5)  $\alpha := \min(\alpha_{\max}, 2\alpha_{k-1})$
  - (6) **While**  $(F_k(x_k + \alpha p) > F_k(x_k) - \Theta \alpha \|p\|^2)$   
**or**  $(\|(I - v_k \otimes v_k)H_h(x_k + \alpha p; v_k)\| > \Psi \|\nabla_x \mathcal{E}_h(x_k, v_{k-1})\|)$  **do**
  - (7)  $\alpha := \alpha/2$
  - (8)  $x_{k+1} := x_k + \alpha p; \alpha_k := \alpha$

It remains to specify step (3) of the Linesearch Dimer Algorithm. Any method computing an update  $v_k$  satisfying  $\|(I - v_k \otimes v_k)H_h(x_k; v_k)\| \leq \text{TOL}$ , for given TOL, is suitable. A basic choice is the following projected steepest descent algorithm.

**Rotation:**

- (1) **Input:**  $x, v, \beta$   
**Parameters:**  $\text{TOL} = \|\nabla_x \mathcal{E}_h(x, v)\|, \beta_{\max} > 0, \Theta \in (0, 1)$
- (2) **While**  $\|(I - v \otimes v)H_h(x; v)\| > \text{TOL}$  **do**
  - (3)  $s := -(I - v \otimes v)H_h(x; v)$
  - (4)  $r := \|s\|; \beta := \max(\beta_{\max}, 2\beta)$
  - (5)  $v_\beta := \cos(\beta r)v + \sin(\beta r)\frac{s}{r}$
  - (6) **While**  $\mathcal{E}_h(x, v_\beta) > \mathcal{E}_h(x, v) - \Theta \beta \|s\|^2$  **do**
  - (7)  $\beta := \beta/2$
  - (8)  $v := v_\beta$
- (9) **Output:**  $v, \beta$

**Proposition 4.** *The Linesearch Dimer Algorithm is well-defined in that the rotation step (3) as well as the linesearch loop (6, 7) both terminate after a finite number of iterations, the latter provided that  $\nabla_x \mathcal{E}_h(x_k, v_k) \neq 0$ .*



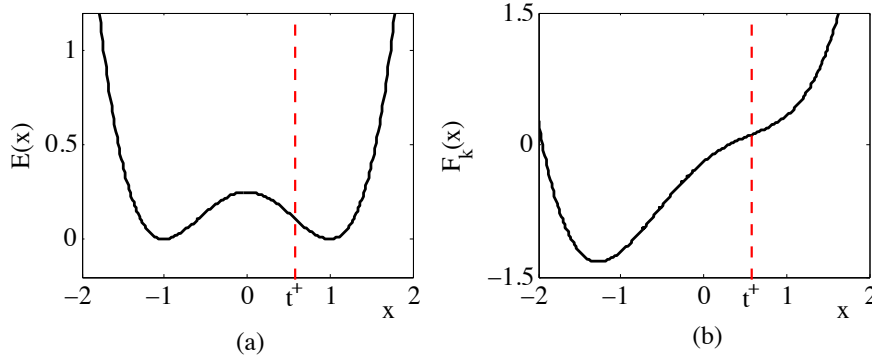


FIGURE 1. (a) Double-well energy defined in (14). (b) The auxiliary functional  $F_k(x)$  with  $x_k = t_h^+$ ; cf. §3.3. The second turning point  $t_h^- = -t_h^+$  is an admissible descent step for  $F_k$ , hence the dimer method can potentially cycle.

*Proof.* The Rotation Algorithm employed in step (3) of Linesearch Dimer Algorithm terminates for any starting guess due to the fact that it is a steepest descent algorithm on a Stiefel manifold (the unit sphere) with a backtracking linesearch employing the Armijo condition [24]. Convergence of this iteration to a critical point is well known [1, Chap.4]. The loop (6,7) terminates after a finite number of iterations [20] since  $p$  is a descent direction for  $F_k \in C^4$ , that is,  $F_k(x_k + \alpha p) = F_k(x_k) - \alpha \|p\|^2 + O(\alpha^2)$ .  $\square$

**Remark 5.** The two basic backtracking linesearch loops (5)–(8) and (11)–(12) can (and should) be replaced with more effective linesearch routines in practise, in particular choosing more effective initial guesses for the steps and using polynomial interpolation to compute linesearch steps. However, the discussion in §3.3 indicates that a Wolfe-type termination criterion might be inappropriate.  $\square$

**3.3. Failure of global convergence.** The modifications of the original dimer algorithms that we have in the Linesearch Dimer Algorithm would, in the case of optimisation, yield a globally convergent scheme. Unfortunately, this is not the case in saddle search. To see this, consider a one-dimensional double-well example,

$$E(x) = \frac{1}{4}(1 - x^2)^2 = \frac{1}{4}x^4 - \frac{1}{2}x^2 + \frac{1}{4}; \quad (14)$$

cf. Figure 1(a). There are only two possible (equivalent) dimer orientation  $v = \pm 1$ , and therefore the rotation steps in Linesearch Dimer Algorithm are ignored. We always take  $v = 1$  without loss of generality. The translation search direction at step  $k$  is always given by  $p = -(1 - 2)\nabla_x \mathcal{E}_h(x_k, 1) = \nabla_x \mathcal{E}_h(x_k, 1)$ , i.e., an ascent direction.

It is easy to see that  $x_* = 0$  is an index-1 saddle (i.e., a maximum), and that there are two turning points  $t_h^\pm = \pm 3^{-1/2}$ . Thus, there exist “discrete turning points”  $t_h^\pm = \pm 3^{-1/2} + O(h^2)$  such that  $\lambda(t_h^\pm) = 0$ , where  $\lambda(x) = H_h(x; 1) \cdot 1 = \frac{1}{2h^2}(E'(x+h) - E'(x-h))$ .

Suppose that we have an iterate  $x_k = t_h^+$ , then the translation search direction is  $p^+ = \nabla_x \mathcal{E}_h(t_h^+, 1) < 0$ . Since  $\mathcal{E}_h(t_h^-, 1) = \mathcal{E}_h(t_h^+, 1)$  it follows that

$$F_k(t_h^-) = \mathcal{E}_h(t_h^-, 1) - 2p^+(t_h^- - t_h^+) < \mathcal{E}_h(t_h^-, 1) = F_k(t_h^+).$$

Thus, for  $\Theta$  sufficiently small, the update  $x_{k+1} = t_h^-$  satisfies all the conditions for termination of the loop (11)–(12) in Linesearch Dimer Algorithm. See also Figure 1 (b), where  $F_k$  is visualised.

We therefore conclude that our newly proposed variant of the dimer algorithm does not exclude cycling behaviour. We also remark that the example is not exclusively one-dimensional, but that analogous constructions can be readily made in any dimension.

**3.4. Local convergence.** We now establish a local convergence result.

**Theorem 6.** *Let  $(x_*, v_*, \lambda_*)$  be an index-1 saddle, let  $(x_h, v_h, \lambda_h)$  denote the dimer saddle associated with  $(x_*, v_*, \lambda_*)$  (cf. Theorem 2) and let  $x_k, v_k$  be the iterates generated by the Linesearch Dimer Algorithm. Then there exist  $r, h_0, C > 0$  and  $\gamma \in (0, 1)$  such that, for  $x_0 \in B_r(x_*)$ ,  $v_{-1} \in B_r(v_*) \cap S_X$  and  $h \leq h_0$ , one of the following alternatives are true:*

- (i) *If  $\nabla_x \mathcal{E}_h(x_k, v_{k-1}) = 0$  for some  $k \in \mathbb{N}$ , then  $\|x_k - x_h\| \leq Ch^2$ .*
- (ii) *If  $\nabla_x \mathcal{E}_h(x_k, v_{k-1}) \neq 0$  for all  $k \in \mathbb{N}$ , then*

$$\|x_k - x_h\| + \|v_k - v_h\| \leq C\gamma^k (\|x_0 - x_h\| + h^2 \|v_{-1} - v_h\|). \quad (15)$$

*Sketch of proof.* Case (i) merely serves to exclude an unlikely situation, in which the Rotation algorithm is ill-defined. We do not discuss this case here, but treat it in §A.4.4. In the following assume Case (ii).

0. Let  $r_k = \|x_k - x_h\|$  and  $s_k := \|v_k - v_h\|$ . We recall basic contraction results for Armijo-based linesearch methods both in a general Hilbert space and for iterates constrained to lie on the unit sphere in §A.3.

1. As a first proper step we establish that, under the termination criterion  $\|(1 - v_k \otimes v_k)H_h(x_k; v_k)\| \leq \|\nabla_x \mathcal{E}_h(x_k, v_{k-1})\|$  for the rotation step, it follows that  $s_k \lesssim r_k + h^2 s_{k-1}$ . This is proven in Lemma 15 and Lemma 16.

2. Next, we use this result to establish that there exists a local minimizer  $y_k$  of  $F_k$  satisfying  $\|y_k - x_h\| \lesssim r_k^2 + h^2 r_k + h^4 s_{k-1}$ . This is established in Lemma 17.

3. The linesearch procedure and the upper bound on the step length ensure that the step of  $x_k$  to  $x_{k+1}$  contracts towards  $y_k$ , that is,  $\|x_{k+1} - y_k\|_* \leq \gamma_* \|x_k - y_k\|_*$  for some  $\gamma_* \in (0, 1)$  and  $\|\cdot\|_*$  the *energy* norm induced by  $(I - 2v_* \otimes v_*)\nabla^2 E(x_*) \approx \nabla^2 F_k(y_k)$ . This is obtained in Lemma 18.

4. The three preceding steps can then be combined to establish that, for  $r_0, s_{-1}, h$  sufficiently small, there exists a constant  $\gamma_3 \in (\gamma_*, 1)$  such that

$$r_{k+1}^* + h^2 s_k \leq \gamma_3 (r_k^* + h^2 s_{k-1}), \quad (16)$$

where  $r_k^* := \|x_k - x_h\|_*$ . This contraction result readily implies the result of the theorem.

The complete proof is given in §A.4. □

## 4. NUMERICAL TESTS

**4.1. Remarks on the implementation.** Throughout our analysis and in the descriptions of the Simple Dimer Algorithm (§2.2) and the Linesearch Dimer Algorithm (§3.2) the notion of gradient  $\nabla$  and tensor product  $\otimes$  take into account the choice of inner product (preconditioning). Therefore, we next describe how preconditioning is implemented in practise and give details how our actual implementation (slightly)

deviates from the theoretical formulations of the Simple Dimer Algorithm and the Linesearch Dimer Algorithm.

In all cases the underlying space is  $X = \mathbb{R}^N$  for some  $N \in \mathbb{N}$ . The main deviation from the algorithms stated in §2.2 and §3.2 is that we admit inner products that may change from one step to another,

$$\|u\| = \sqrt{u^T M_k u}, \quad \text{and} \quad u \cdot v = u^T M_k v,$$

where  $M_k$  is symmetric and positive definite. That is, our implementation is a *variable metric variant*; see also Remark 8 below.

Let  $E \in C^4(X) = C^4(\mathbb{R}^N)$ , and let  $\nabla'$  denote the standard gradient and  $\otimes'$  the standard tensor product (i.e., the gradient and tensor products with respect to the  $\ell^2$ -norm), then the gradient and tensor products in step  $k$  become

$$\nabla E(x) = M_k^{-1} \nabla' E(x), \quad \text{and} \quad (v \otimes v) \nabla E(x) = (v \otimes' v) \nabla' E(x).$$

The variable metric variant of the Simple Dimer Algorithm, augmented with a termination criterion, is given below. Note that here the rotation step is performed by a tangential descent step followed by a projection, rather than a step on the manifold.

**Simple Dimer Algorithm (VM):**

- (1) **Input:**  $x_0, v_0 \in X$ ,  $h > 0, \alpha, \beta > 0, \text{TOL}^x, \text{TOL}^v > 0$ ;  $k := 0$ ;
- (2) **While**  $\|M_k^{-1/2} \nabla'_x \mathcal{E}_h(x_k, v_k)\|_{\ell^2} > \text{TOL}^x$   
**or**  $\|(M_k^{-1/2} - M_k^{1/2} v_k \otimes' v_k) h^{-2} \nabla'_v \mathcal{E}_h(x_k, v_k)\|_{\ell^2} > \text{TOL}^v$  **do**  
%% Metric %%
- (3) Compute a spd matrix  $M_k \in \mathbb{R}^{N \times N}$ ;
- (4)  $v_k := v_k / \|M_k^{1/2} v_k\|_{\ell^2}$ ;
- (5)  $v_{k+1} := v_k - \beta(M_k^{-1} - v_k \otimes v_k) h^{-2} \nabla'_v \mathcal{E}_h(x_k, v_k)$
- (6)  $x_{k+1} := x_k - \alpha(M_k^{-1} - 2v_k \otimes v_k) \nabla'_x \mathcal{E}_h(x_k, v_k)$ .
- (7)  $k := k + 1$

**Remark 7.** In our experiments we observe that the rotation residual decreases more quickly than the translation residual, hence the convergence criteria could be based on the translation residual only, without affecting the results.  $\square$

**Remark 8.** Our analysis of both the Simple Dimer Algorithm and of the Linesearch Dimer Algorithm is readily extended to their variable metric variants, provided that the metric  $M_k$  at iterate  $k$  is a smooth function of the state, i.e.,  $M_k = \mathbf{M}(x_k, v_k)$ , where  $\mathbf{M} \in C^2(B_r(x_*) \times S_X; L(X))$ , for some  $r > 0$ . This is the case in all examples that we consider below. A more general convergence theory, e.g., employing quasi-Newton type Hessian updates requires additional work.  $\square$

Analogous modifications are made to the Linesearch Dimer Algorithm. The auxiliary functional  $F_k$  now reads

$$F_k(x) = \mathcal{E}_h(x; v_k) - 2(v_k^T \nabla'_x \mathcal{E}_h(x_k, v_k))(v_k^T M_k(x - x_k)) + \lambda_k (v_k^T M_k(x - x_k))^2,$$

$$\lambda_k = h^{-2} v_k^T \nabla'_v \mathcal{E}_h(x_k, v_k),$$

$$\nabla'_x \mathcal{E}_h(x, v) = \frac{1}{2}(\nabla' E(x + hv) + \nabla' E(x - hv)),$$

$$\nabla'_v \mathcal{E}_h(x, v) = \frac{h}{2}(\nabla' E(x + hv) - \nabla' E(x - hv)),$$

where we recall that  $\nabla'$  denotes the standard gradient (i.e., the gradient with respect to the  $\ell^2$ -norm). See Remark 9 on additional improvements.

**Linesearch Dimer Algorithm (VM):**

- (1) **Input:**  $x_0, v_0 \in X$ ,  $h > 0, \text{TOL}^x, \text{TOL}^v > 0$ ;  $k := 0$
- (2) **While**  $\|M_k^{-1/2} \nabla'_x \mathcal{E}_h(x_k, v_k)\|_{\ell^2} > \text{TOL}^x$ 
  - %% Metric %%
  - (3) Compute a spd matrix  $M_k \in \mathbb{R}^{N \times N}$
  - (4)  $v'_k := v_k / \|M_k^{1/2} v_{k-1}\|$
  - %% Rotation %%
  - (5)  $[v_{k+1}, \beta] := \mathbf{Rotation (VM)}[x_k, v'_k, \beta, M_k]$
  - %% Translation %%
  - (6)  $p_M := -(M_k^{-1} - 2v_{k+1} \otimes v_{k+1}) \nabla'_x \mathcal{E}_h(x_k; v_{k+1})$
  - (7)  $\alpha := \min(\alpha_{\max}, 2\alpha)$
  - (8) **While**  $(F_k(x_k + \alpha p_M) > F_k(x_k) - \Theta \alpha p_M^T M_k p_M)$ 
    - or**  $(\|M_k^{1/2}(M_k^{-1} - v_{k+1} \otimes v_{k+1})h^{-2} \nabla'_v \mathcal{E}_h(x_k + \alpha p_M; v_{k+1})\|_{\ell^2} >$
    - $\Psi \|M_k^{1/2}(M_k^{-1} - v_{k+1} \otimes v_{k+1})h^{-2} \nabla'_v \mathcal{E}_h(x_k; v_{k+1})\|_{\ell^2})$  **do**
    - (9)  $\alpha := \alpha/2$
    - (10)  $x_{k+1} := x_k + \alpha p_M$
    - (11)  $k := k + 1$

**Rotation (VM):**

- (1) **Input:**  $x, v, \beta, M_k$
- Parameters:**  $\text{TOL} = \max(\|M_k^{-1/2} \nabla'_x \mathcal{E}_h(x, v)\|_{\ell^2}, \text{TOL}^v)$ ,  $\Theta \in (0, 1)$ ,  $\beta_{\max}$ ;
- (2) **While**  $\|M_k^{1/2}(M_k^{-1} - v \otimes v)h^{-2} \nabla'_v \mathcal{E}_h(x; v)\| > \text{TOL}$  **do**
- (3)  $s := -(M_k^{-1} - v \otimes v)h^{-2} \nabla'_v \mathcal{E}_h(x; v)$
- (4)  $t := \|M_k^{1/2} s\|_{\ell^2}$ ;  $\beta := \min(\beta_{\max}, 2\beta)$
- (5)  $v_\beta := \cos(t\beta)v + \sin(t\beta)t^{-1}s$
- (6) **While**  $\mathcal{E}_h(x, v_\beta) > \mathcal{E}_h(x, v) - \Theta \beta t^2$  **do**
- (7)  $\beta := \beta/2$
- (8)  $v := v_\beta$
- (9) **Output:**  $v, \beta$

**Remark 9.** A modification that can give significant performance gains is to employ a different heuristic for the initial guess of  $\alpha$  in Step (7) of Linesearch Dimer Algorithm (VM): With  $p_{M,k} := -(M_k^{-1} - 2v_k \otimes v_k) \nabla'_x \mathcal{E}_h(x_k; v_k)$  and  $p_{I,k} := -(I - 2v_k \otimes v_k) \nabla'_x \mathcal{E}_h(x_k, v_k)$  let, for  $k \geq 2$ ,  $\gamma_k := (p_{M,k-1} \cdot' p_{I,k-1}) / (p_{M,k} \cdot' p_{I,k})$ , then for  $k \geq 2$  we replace Step (7) with

$$\alpha := \min(\text{avg}(\gamma_{\max(2,k-4)}, \dots, \gamma_k), 2\alpha, \alpha_{\max}).$$

An analogous modification can be made for the rotation algorithm.  $\square$

In all numerical tests we use the following parameters:  $h = 10^{-3}$ ,  $\Theta = \sqrt{0.1}$ ,  $\text{TOL}^x = 10^{-5}$ ,  $\text{TOL}^v = 10^{-1}$ ,  $\alpha_{\max} = 1$  and  $\Psi = 100$ . We briefly discuss these choices:

- $h$  should be small enough such that the dimer saddle is sufficiently close to the true saddle (with respect to the length scales of the given problem), while large enough that numerical robustness does not become a problem for the rotation. In all our tests,  $h = 10^{-3}$  was a good compromise.
- $\Theta$  should be sufficiently large (though,  $\leq 1/2$ ) to ensure that the linesearch method finds steps which give a large decrease in dimer energy. It is often chosen much smaller than our choice of  $\Theta = \sqrt{0.1}$  to immediately accept steps that make some progress. Our experience is that, with preconditioned search direction, our more stringent choice gives better performance as it leads to an improved initial steplength guess for the next iteration.
- The choice of  $\text{TOL}^x$  simply controls the desired level of convergence to the dimer saddle.
- The parameter  $\text{TOL}^v$  should be chosen as weakly as possible such that either algorithm converges to the saddle. In the Linesearch Dimer Algorithm (VM) rotations are performed such that the rotation residual is at least as good as the translation residual until it moves below this value. Subsequent translations may increase the rotation residual such that further applications of the rotation algorithm are needed. In practise this means that the rotation algorithm is performed at every iteration of the Linesearch Dimer Algorithm (VM) for the first few steps, then only sporadically or not at all once the rotation residual reaches  $\text{TOL}^v$ . The use of this parameter then decreases the overall number of gradient evaluations needed to find the dimer saddle, by only performing the rotation as necessary.
- The maximum step  $\alpha_{\max}$  should principally be chosen such that the dimer cannot translate into non-physical regimes for the given problem.
- The parameter  $\Psi$  should be chosen  $> 1$  and restricts the translation step from moving the dimer to a point where it becomes too badly orientated. In our numerical tests this parameter is set sufficiently large that this termination criteria for the translation never occurs (the translation always terminates by finding a sufficient decrease in the auxiliary functional  $F_k$ ).

**Remark 10.** We observe during numerical testing that the rotation component of the linesearch dimer is somewhat vulnerable to rounding error in the objective function  $E$ . As the dimer becomes increasingly well orientated,  $\nabla E$  becomes almost orthogonal to the dimer orientation and any small rotation may result in a zero change (to numerical precision) in the dimer energy. In the numerical examples presented in this section, this never occurs since we use a relatively high value for  $\text{TOL}^v$ , that is the rotation is only ever weakly converged. In our examples this is sufficient for the the dimer to converge to the saddle. If a stronger level of convergence were required, another technique should be used to improve the rotation residual further, such as changing to a gradient based method or simply making fixed steps. This could, for example, be performed in a post-processing step (cf. Remark 1 on post-processing  $x$ ).  $\square$

**4.2. Test 1: Vacancy Diffusion.** Our first test case is a standard example from molecular physics. A single atom is removed from a 2D lattice and a neighbouring atom is moved partway into the gap. Atoms within a certain radius of the vacancy are allowed to move, while those beyond that radius are fixed. This configuration is illustrated in Figure 2(A). The final computed configuration is given in Figure 3.

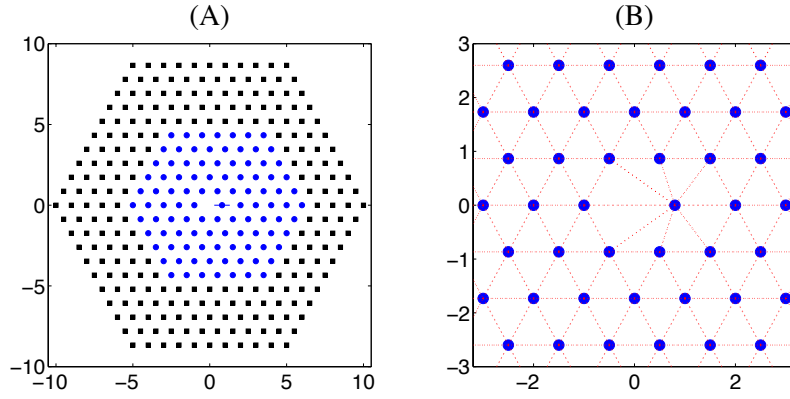


FIGURE 2. Initial configuration of the atoms in the vacancy diffusion problem (Test 1). Black squares are fixed atoms while blue circles are atoms which move freely. (A) The initial dimer orientation is selected so that the translated atom has an orientation along the  $y = 0$  direction, and is zero for all other atoms. (B) The Delaunay  $\mathcal{T}_k$  triangulation used for the connectivity norm.

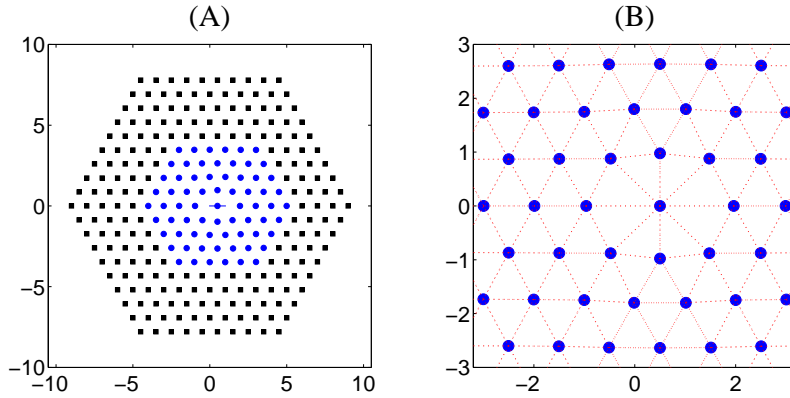


FIGURE 3. Final configuration of the atoms in the vacancy diffusion problem (Test 1). Black squares are fixed atoms while blue circles are atoms which move freely. (A) In the final configuration an atom moves to the midpoint between two ‘basins’. (B) The Delaunay triangulation  $\mathcal{T}_k$  used for the connectivity norm.

The energy function is given by the simple Morse potential,

$$E(\{x_i\}) = \sum_{i,j} V(\|x_i - x_j\|_2), \quad V(r) = e^{-2a(r-1)} - 2e^{-a(r-1)}, \quad (17)$$

with stiffness parameter  $a = 4$ .

The experiment is run both using the generic  $\ell_2$  norm (no preconditioner), as well as a ‘connectivity’ norm. Such a norm can be defined based on the Delaunay

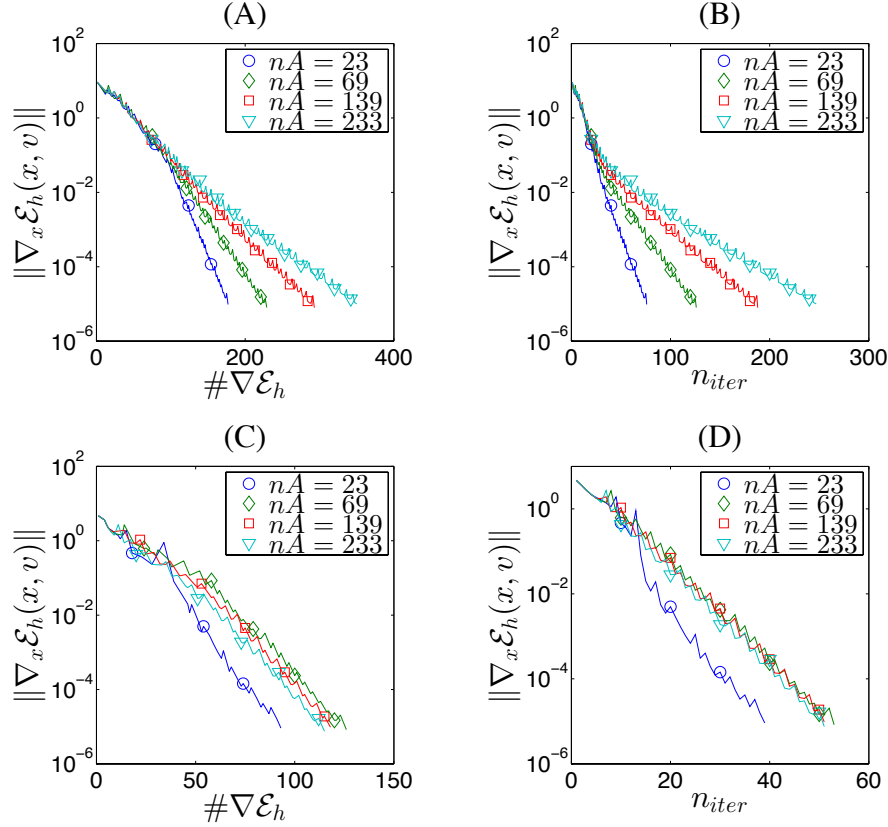


FIGURE 4. Convergence of the linesearch dimer to the saddle in the vacancy diffusion problem (Test 1) with (A),(B) the  $\ell_2$  norm and (C),(D) connectivity norm versus the number of force evaluations and dimer iterations for increasing numbers of free atoms.

triangulation of the atomistic positions (Figure 2(B))

$$\langle M_k u, u \rangle = \int |\nabla I_{\mathcal{T}_k} u|^2,$$

where  $\mathcal{T}_k$  is the triangulation depicted in the figure and  $I_{\mathcal{T}_k}$  the associated nodal interpolant.

Figure 4 demonstrates the convergence to the saddle with different numbers of free atoms  $nA$  (giving different dimensionality of the system) in the two norms for the linesearch dimer. We can also observe the benefit of the linesearch vs a simple dimer scheme when using the connectivity norm (Figure 5). The linesearch dimer selects very efficient stepsizes with no a-priori information, while the simple dimer method might exhibit either slow convergence, or no convergence, if the fixed steps are poorly chosen.

**4.3. Test 2: A Phase Field Example.** Our second example is based on a simple phase field model where the global energy is given by,

$$E(u) = \int_{\Omega} \frac{\epsilon}{2} |\nabla u|^2 + \frac{1}{2\epsilon} (u^2 - 1)^2, \quad (18)$$

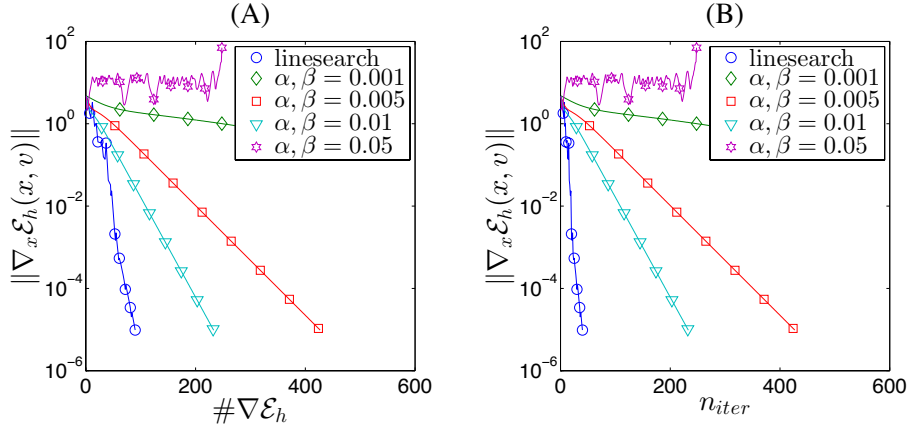


FIGURE 5. Convergence of the linesearch dimer vs the simple dimer method for Test 1 (vacancy), for several choices of the simple dimer step sizes, with  $nA = 69$ , using the connectivity norm as a preconditioner.

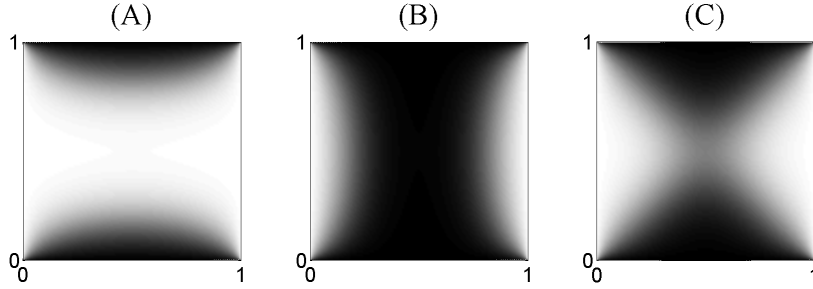


FIGURE 6. Minima (A,B) and saddle point (C) of the phase field problem (Test 2) with  $\epsilon = 1/10$ . The shading is linearly interpolated between white(-1) and black(1).

where  $\Omega = (0, 1)^2$ , and the boundary conditions are,

$$u(x) = \begin{cases} -1, & x_1 \in \{0, 1\} \\ 1, & x_2 \in \{0, 1\}. \end{cases} \quad (19)$$

There are 2 minima of such an energy, these are given in Figure 6(A),(B). The saddle between these two minima is given in Figure 6(C).

A possible choice for a preconditioner for this system is a stabilized Laplacian,

$$P = \epsilon \Delta + \frac{1}{\epsilon} I. \quad (20)$$

In order to compute either a minimum or a saddle point for such a system we triangulate the domain into a variable number of elements, thereby creating a discrete system of variable dimensionality. In our tests we take the initial dimer point as a small random perturbation of one of the local minima, and the initial dimer orientation is the metric inverted against a vector of ones.



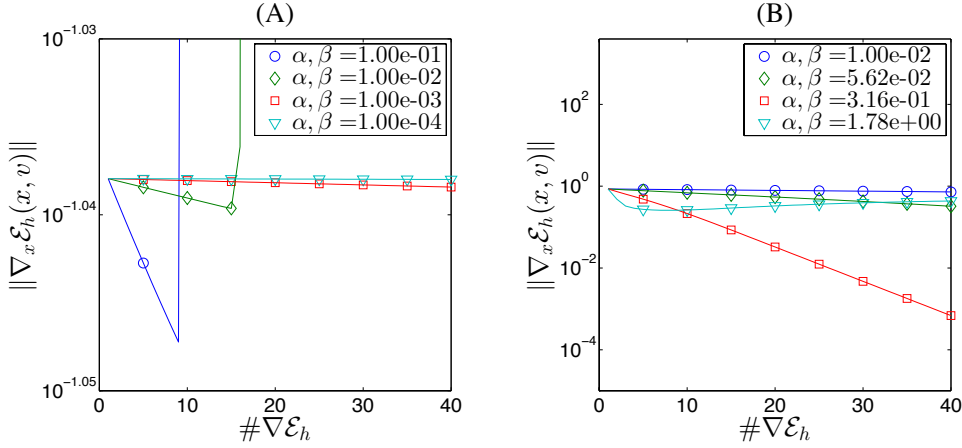


FIGURE 7. Convergence of the simple dimer to the saddle in the phase field problem (Test 2) with (A) the  $\ell_2$  metric and (B) the stabilized Laplacian metric where  $\epsilon = 1/10$  for a triangulation with 3485 degrees of freedom.

In Figure 7 we demonstrate the necessity of using a preconditioner to solve this problem using the simple dimer method. When using the preconditioner (20), the algorithm performs well when the step size is chosen appropriately. We observe the expected behaviour, that there exists an optimal step size where convergence is fastest, and beyond that step size the dimer diverges. In fact we observe that the stabilized Laplacian metric is so effective, that the optimal step size seems very close to the unit step. If the  $\ell_2$  norm (identity preconditioner) is used then for all step sizes tested the dimer diverges, indicating that at best a very small step would need to be chosen for convergence.

In Figure 8 we demonstrate that the use of the scaled Laplacian metric for different system sizes. We observe that the use of this metric gives almost perfect scale invariance.

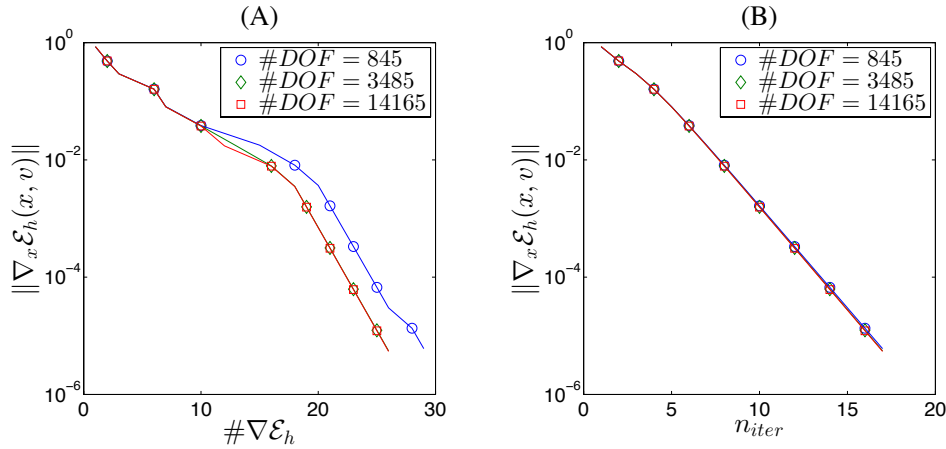


FIGURE 8. Convergence of the linesearch dimer to the saddle in the phase field problem (Test 2) with the stabilized Laplacian metric and triangulations of varying coarseness.

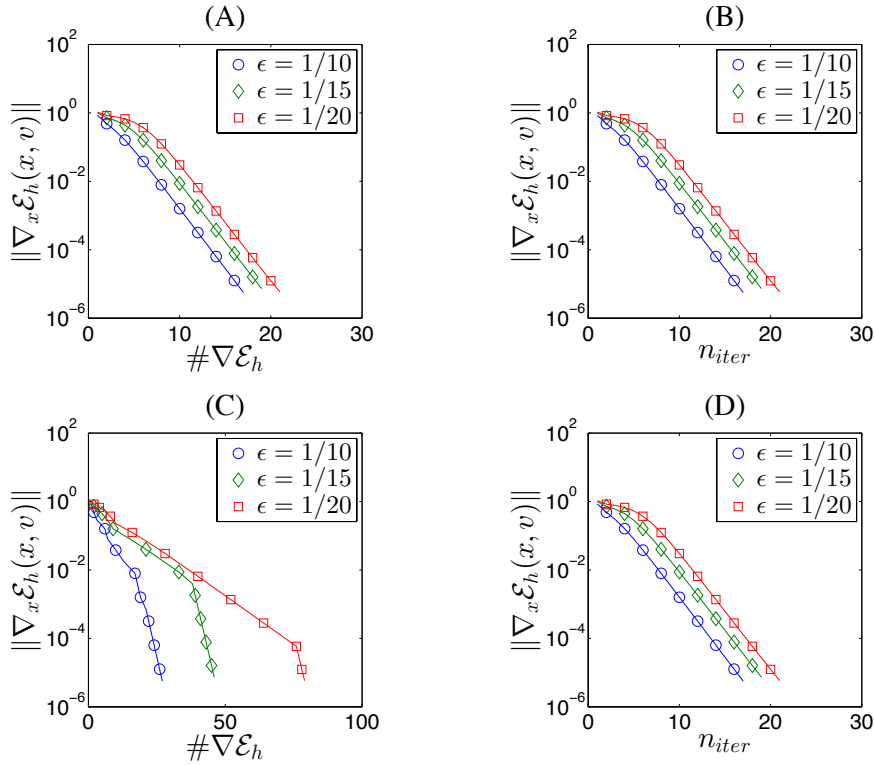


FIGURE 9. Convergence to the saddle in the phase field problem (Test 2) using the stabilized Laplacian metric with (A),(B) the simple dimer with unit step length and (C),(D) the linesearch dimer for a triangulation with 2405,9805,22205 degrees of freedom for the respective choices of  $\epsilon$ .

In Figure 9 we give the results of applying the simple and linesearch dimers with varying  $\epsilon$ ; the coarseness of the discretization in each experiment is chosen such that  $\Delta x \approx \epsilon/5$ . In some of these cases the linesearch dimer fails due to rounding error. Specifically, due to rounding error in the naive implementation of the energy function (simple summation over the elements), the translation step fails to find a sufficient decrease in the dimer energy, the step size selected shrinks to zero (to rounding error) and the method stagnates. In order to correct this a more robust method of evaluating the energy or a more advanced optimization algorithm should be implemented which can either choose better linesearch directions or more robustly deal with numerically zero energy changes.

We also observe, in the case  $\epsilon = 1/30$  that the rate of convergence of even the simple dimer changes once the residual moves below a certain value. We are unable to give a satisfactory explanation for this effect, but speculate that the singularity in the boundary condition (which excludes admissible  $H^1$ -states) might be the case. (In particular, we observed that this behaviour is independent of the mesh coarseness and of the dimer length.)

## 5. CONCLUSIONS

We have described a dimer method for finding a saddle point in which the dimer length  $h$  is not required to shrink to zero, but which converges to a point that lies within  $O(h^2)$  of a saddle. We have enhanced this algorithm with a linesearch to improve its robustness and efficiency, and use the observation that the dimer method may be formulated and applied in a general Hilbert space to allow preconditioning that improves the method's efficiency. The linesearch uses a local merit function. Unfortunately our particular merit function may not lead to global convergence of the iterates, and it is an open question as to whether there is another merit function that ensures global convergence. We have illustrated the positive effects of our algorithms on two realistic examples.

## APPENDIX A. PROOFS

We remark that all proofs in this section are independent of the choice of norm  $\|\cdot\|$  (not necessarily the  $\ell^2$ -norm) and are hence valid for the preconditioned version of the algorithm.

**A.1. Proof of Proposition 2.** We prove the result using the inverse function theorem. We write (9) as  $F(x_h, v_h, \lambda_h) = 0$  and show that  $\|F(x_*, v_*, \lambda_*)\| \leq Ch^2$  and that  $\nabla F(x_*, v_*, \lambda_*)$  is an isomorphism with bounds independent of  $h$ . The inverse function theorem then yields the stated result.

*Residual estimate.* Let the residual components be

$$\begin{aligned} r_x &:= F_x(x_*, v_*, \lambda_*) = \nabla_x \mathcal{E}_h(x_*, v_*), \\ r_v &:= F_v(x_*, v_*, \lambda_*) = h^{-2} \nabla_v \mathcal{E}_h(x_*, v_*) - \lambda_* v_*, \\ r_\lambda &:= F_\lambda(x_*, v_*, \lambda_*) = \frac{1}{2}(\|v_*\|^2 - 1). \end{aligned}$$

Then (2) and (4) imply that  $r_x, r_v, r_\lambda = O(h^2)$  and hence  $\|F(x_*, v_*, \lambda_*)\| \leq Ch^2$ .

*Stability.*  $\nabla F(x_*, v_*, \lambda_*)$  can be written in the form, using

$$\begin{aligned} \nabla F(x_*, v_*, \lambda_*) &= \begin{bmatrix} \nabla_x^2 \mathcal{E}_h(x_*, v_*) & \nabla_v \nabla_x \mathcal{E}_h(x_*, v_*) & 0 \\ h^{-2} \nabla_v \nabla_x \mathcal{E}_h(x_*, v_*) & h^{-2} \nabla_v^2 \mathcal{E}_h(x_*, v_*) - \lambda_* I & -v_* \\ 0 & -v_*^T & 0 \end{bmatrix} \\ &= \begin{bmatrix} \nabla^2 E(x_*) & 0 & 0 \\ \nabla^3 E(x_*) \cdot v_* & \nabla^2 E(x_*) - \lambda_* I & -v_* \\ 0 & v_*^T & 0 \end{bmatrix} + O(h^2) =: \mathbf{A} + O(h^2), \end{aligned}$$

where we used (3), (4) and (6). By assumption,  $\nabla^2 E(x_*)$  is an isomorphism on  $X$ . Since, also by assumption,  $\lambda_*$  is a simple eigenvalue, the block

$$\begin{bmatrix} \nabla^2 E(x_*) - \lambda_* I & -v_* \\ v_*^T & 0 \end{bmatrix} \quad (21)$$

is an isomorphism on  $X \times \mathbb{R}$  as well. Thus,  $\mathbf{A}$  is an isomorphism on  $X \times X \times \mathbb{R}$  and consequently, for all  $h$  sufficiently small,  $\nabla F(x_*, v_*, \lambda_*) = \mathbf{A} + O(h^2)$  is also an isomorphism, with a uniform bound on its inverse.

Thus, the inverse function theorem shows that there exist a radius  $r_0 > 0$  and a dimer length  $h_0 > 0$ , such that, for  $h \leq h_0$ , there exists a unique solution  $(x_h, v_h, \lambda_h)$  to (9) in a ball of radius  $r_0$  about  $(x_*, v_*, \lambda_*)$ , satisfying the estimate (10).

**A.2. Proof of Theorem 3.** Fix  $r$  and  $h_0$  sufficiently small so that Theorem 2 applies. Let  $e_k := x_k - x_h$ ,  $f_k := v_k - v_h$  and  $r_k := \sqrt{\|e_k\|^2 + \|f_k\|^2}$ , so that trivially  $\|e_k\| \leq r_k$  and  $\|f_k\| \leq r_k$ .

**Lemma 11.** *Let  $p := -(I - 2v_k \otimes v_k) \nabla_x \mathcal{E}_h(x_k, v_k)$  and  $s := -(I - v_k \otimes v_k) H_h(x_k; v_k)$ , then, under the assumptions of Theorem 3,*

$$p = -Ae_k + O(r_k^2 + h^2 r_k), \quad \text{and} \quad (22)$$

$$s = -Be_k - Cf_k + O(r_k^2 + h^2 r_k), \quad (23)$$

where the operators  $A$  and  $C$  are defined in (12) and  $B$  is a bounded linear operator.

*Proof.* We begin by noting the elementary identities which are easy to establish:

$$\begin{aligned} \nabla_x \mathcal{E}_h(x_k, v_k) - \nabla_x \mathcal{E}_h(x_h, v_h) &= O(r_k), \\ v_k \otimes v_k - v_h \otimes v_h &= O(r_k), \\ \nabla_x^2 \mathcal{E}_h(x_h, v_h) &= \nabla^2 E(x_h) + O(h^2) = \nabla^2 E(x_*) + O(h^2) \end{aligned} \quad (24)$$

Using the identities (24), as well as (2), (6), we can expand

$$\begin{aligned} p &= -(I - 2v_k \otimes v_k) (\nabla_x \mathcal{E}_h(x_k, v_k) - \nabla_x \mathcal{E}_h(x_h, v_h)), \\ &= -(I - 2v_h \otimes v_h) (\nabla_x^2 \mathcal{E}_h(x_h, v_h) e_k + \nabla_x \nabla_v \mathcal{E}_h(x_h, v_h) f_k) + O(r_k^2) \\ &= -(I - 2v_h \otimes v_h) \nabla^2 E(x_h) e_k + O(r_k^2 + h^2 r_k) \\ &= -(I - 2v_* \otimes v_*) \nabla^2 E(x_*) e_k + O(r_k^2 + h^2 r_k) \\ &= -Ae_k + O(r_k^2 + h^2 r_k). \end{aligned}$$

To prove (23), we first note that, with  $\|v\| = 1$ ,

$$\begin{aligned} H_h(x; v) &= \int_{-1}^1 \nabla^2 E(x + thv) dt v = \nabla^2 E(x) v + O(h^2), \\ H_h(x_h; v_h) &= \nabla^2 E(x_h) v_h + O(h^2) = \nabla^2 E(x_*) v_* + O(h^2), \\ H_h(x_k; v_k) - H_h(x_h; v_h) &= \int_{-1}^1 \left( \nabla^2 E(x_k + thv_k) - \nabla^2 E(x_h + thv_h) \right) dt v_k \\ &\quad + \int_{-1}^1 \nabla^2 E(x_h + thv_h) dt (v_k - v_h) \\ &= \int_{-1}^1 \left( \nabla^3 E(x_h + thv_h) [(x_k - x_h) + th(v_k - v_h)] dt v_h \right. \\ &\quad \left. + \nabla^2 E(x_*) (v_k - v_h) + O(r_k^2 + h^2 r_k) \right) \\ &= (\nabla^3 E(x_*) v_*) e_k + \nabla^2 E(x_*) f_k + O(h^2 r_k + r_k^2), \end{aligned}$$

where we interpret  $\nabla^3 E(x) \cdot v \in L(X)$  via the action  $w \cdot ((\nabla^3 E(x) \cdot v)z) = \lim_{t \rightarrow 0} t^{-1} w \cdot ((\nabla^2 E(x + tv) - \nabla^2 E(x))z)$ . Finally, we also have

$$\begin{aligned}
(v_k \otimes v_k - v_h \otimes v_h)H_h(x_h; v_h) &= (v_k \otimes v_k - v_h \otimes v_h)\nabla^2 E(x_*)v_* + O(h^2 r_k) \\
&= \lambda_*(v_k \otimes v_k - v_h \otimes v_h)v_* + O(h^2 r_k) \\
&= \lambda_*(v_k \otimes v_k - v_h \otimes v_h)v_h + O(h^2 r_k) \\
&= \lambda_*(v_k - v_h) + \lambda_* v_k((v_k - v_h) \cdot v_h) + O(h^2 r_k) \\
&= \lambda_* f_k + O(r_k^2 + h^2 r_k).
\end{aligned}$$

In the very last line we also used the fact that  $v_k \cdot v_h - 1 = -\frac{1}{2}\|v_k - v_h\|^2$ .

Using these identities, we can compute

$$\begin{aligned}
s &= -(I - v_k \otimes v_k)H_h(x_k; v_k) \\
&= (I - v_h \otimes v_h)H_h(x_h; v_h) - (I - v_k \otimes v_k)H_h(x_k; v_k) \\
&= -(I - v_k \otimes v_k)(H_h(x_k; v_k) - H_h(x_h; v_h)) + (v_k \otimes v_k - v_h \otimes v_h)H_h(x_h; v_h) \\
&= -(I - v_k \otimes v_k)((\nabla^3 E(x_*)v_*)e_k + \nabla^2 E(x_*)f_k) + O(h^2 r_k + r_k^2) \\
&\quad + \lambda_* f_k + O(r_k^2 + h^2 r_k) \\
&=: -Be_k + [\lambda_* I - (I - v_* \otimes v_*)\nabla^2 E(x_*)]f_k + O(r_k^2 + h^2 r_k) \\
&= -Be_k - Cf_k + O(r_k^2 + h^2 r_k).
\end{aligned}$$

□

From Lemma 11 it follows in particular that  $s = O(r_k)$ . Hence, Taylor expansions of sine and cosine in the identity

$$v_{k+1} = \cos(\|s\|\beta_k)v_k + \sin(\|s\|\beta_k)\frac{s}{\|s\|},$$

yield

$$f_{k+1} = f_k + \beta_k s + O(\beta_k^2 s_k^2)$$

Using Lemma 11, the identity  $e_{k+1} = e_k + \alpha_k p$ , and the fact that  $\beta_k$  is bounded, we therefore obtain identity (12) in the proof outline in §2.4.

Since  $A, C$  are positive definite, it follows that, for  $\alpha_k, \beta_k$  sufficiently small, the spectrum of the operator

$$\left[ I - \begin{pmatrix} \alpha_k A & 0 \\ \beta_k B & \beta_k C \end{pmatrix} \right]$$

lies in some interval  $[0, \mu]$  for  $\mu < 1$ . Thus, for fixed steps  $\alpha_k \equiv \alpha$  and  $\beta_k \equiv \beta$ , (11) follows from standard linearised stability arguments for discrete dynamical systems. The straightforward generalisation to non-uniform steps is presented in [9].

**A.3. Contraction of steepest descent with linesearch.** In the section following this one, we will use statements about the steepest descent method with backtracking that we suspect are well known. Since we have been unable to find precisely the statement that we require, we state both below, and give complete proofs in [9].

**Lemma 12.** *Let  $X$  be a Hilbert space,  $F \in C^3(X)$ , and  $x_* \in X$  with  $\nabla F(x_*) = 0$  and  $\nabla^2 F(x_*)$  positive definite, i.e.,  $u \cdot (\nabla^2 F(x_*)u) \geq \mu \|u\|^2$  for  $\mu > 0$ . Let  $\|u\|_*^2 := u \cdot (\nabla^2 F(x_*)u)$ . Further, let  $\bar{\alpha} > \underline{\alpha} > 0$ ,  $\Theta \in (0, 1)$ .*

Then, there exists  $r > 0$  and  $\gamma \in (0, 1)$ , depending only on  $\underline{\alpha}, \bar{\alpha}, \mu, \|\nabla^j F(x)\|$  for  $x \in B_1(x_*)$ , such that, for all  $\alpha \in [\underline{\alpha}, \bar{\alpha}]$  and for all  $x \in B_r(x_*)$  satisfying the Armijo condition

$$F(x - \alpha \nabla F(x)) \leq F(x) - \Theta \alpha \|\nabla F(x)\|^2,$$

we have

$$\| [x - \alpha \nabla F(x)] - x_* \|_* \leq \gamma \|x - x_*\|_*.$$

We now generalize the foregoing result to steepest descent on the unit sphere. Convergence results for many methods on manifolds are given by [1, Chap.4]. See specifically [1, Thm.4.5.6] and [2].

**Lemma 13.** *Let  $X$  be a Hilbert space,  $P_v := v \otimes v$  and  $P'_v := I - P_v$  for  $v \in S_X$ . Let  $F \in C^3(X)$ ,*

$$g(v) := P'_v \nabla F(v) \quad \text{and} \quad H(v) := P'_v \nabla^2 F(v) P'_v - (\nabla F(v) \cdot v) I.$$

We assume that there exists  $v_* \in S_X$  and  $\mu > 0$  such that

$$g(v_*) = 0 \quad \text{and} \quad u \cdot (H(v_*)u) \geq \mu \|u\|^2 \quad \forall u \in X. \quad (25)$$

Let  $\|u\|_* := \sqrt{u \cdot (H(v_*)u)}$ .

Let  $\bar{\alpha} > 0$ ,  $\Theta \in (0, 1)$ , and for  $v \in S_X$  and  $\alpha \in \mathbb{R}$ , denote

$$v_\alpha := \cos(\alpha \|g(v)\|) v - \sin(\alpha \|g(v)\|) \frac{g(v)}{\|g(v)\|}.$$

Then, there exists  $r > 0$  such that, for all  $v \in B_r(v_*) \cap S_X$  and  $\alpha \in (0, \bar{\alpha}]$  satisfying the Armijo condition

$$F(v_\alpha) \leq F(v) - \Theta \alpha \|g(v)\|^2,$$

there exists a constant  $\gamma(\alpha) \in [0, 1)$  such that

$$\|v_\alpha - v_*\|_* \leq \gamma(\alpha) \|v - v_*\|_*.$$

The contraction factor  $\gamma(\alpha)$  depends on  $\alpha, \mu$  and on  $\|\nabla^j F(x)\|, x \in B_1(v_*)$ . Moreover, for any  $\underline{\alpha} \in (0, \bar{\alpha}]$ ,  $\sup_{\alpha \in [\underline{\alpha}, \bar{\alpha}]} \gamma(\alpha) < 1$ .

**A.4. Proof of Theorem 6.** Throughout this proof, we fix an index-1 saddle  $(x_*, v_*, \lambda_*)$ , and assume that  $h_0$  is small enough so that Proposition 2 ensures the existence of a dimer saddle  $(x_h, v_h, \lambda_h)$  in an  $O(h^2)$  neighbourhood of  $(x_*, v_*, \lambda_*)$ .

Until we state otherwise (namely, in §A.4.4) we assume that  $\nabla_x \mathcal{E}_h(x_k, v_{k-1}) \neq 0$  for all  $k$ . In particular, the Linesearch Dimer Algorithm is then well-defined and produces a sequence of iterates  $(x_k, v_k)_{k \in \mathbb{N}}$ . That is, we are in Case (ii) of Theorem 6. The alternative, Case (i), is treated in §A.4.4.

**A.4.1. Analysis of the rotation.** We begin by establishing an auxiliary result concerning existence of minimisers of  $v \mapsto \mathcal{E}_h(x, v)$ . Let

$$V(x) := \arg \min_{v \in S_X} \mathcal{E}_h(x, v),$$

whenever this minimiser exists and is unique. While, in general this minimum need not exist (or be unique), we show in the following lemma that it is unique and a smooth function of  $x$  in the neighbourhood of an index-1 saddle.

**Lemma 14.** *Let  $(x_*, v_*, \lambda_*)$  be an index-1 saddle, then there exist  $r > 0, h_0 > 0$  (chosen independently of one another) such that, for all  $x \in B_r(x_*)$  and  $h \in (0, h_0]$ ,  $V(x)$  is well-defined and moreover  $x \mapsto V(x) \in C^1(B_r(x_*))$ .*

*Proof.* For  $r$  sufficiently small, if  $x \in B_r(x_*)$  then  $\nabla^2 E(x)$  also has index-1 saddle structure and, if  $(\lambda, v)$  is the smallest eigenpair of  $\nabla^2 E(x)$ , then  $\lambda \leq \lambda_*/2$  and  $(\nabla^2 E(x)w) \cdot w \geq \mu_*/2 \|w\|^2$  for  $w \perp v$ , where  $\mu_* := \inf_{\|w\|=1, w \perp v_*} (\nabla^2 E(x_*)w) \cdot w > 0$ . (This statement is a straightforward consequence of the local Lipschitz continuity of  $\nabla^2 E$ , which follows since  $E \in C^4(X)$ .)

The statement of the Lemma is then proven similarly as Proposition 2, provided  $h_0$  is chosen sufficiently small (depending on  $\lambda_*, \mu_*$  and on derivatives of  $E$  in  $B_{2r}(x_*)$ ). The  $C^1$ -dependence of  $V(x)$  on  $x$  is a consequence of the implicit function theorem.  $\square$

Next, we obtain a bound on  $v_k - v_h$  in terms of  $x_k - x_h$  and the residual of  $v_k$ .

**Lemma 15.** *There exist  $r, h_0, C_1 > 0$  such that, for  $h \in (0, h_0]$ ,  $x \in B_r(x_*)$  and  $v \in B_r(v_*)$  with  $\|v\| = 1$ , we have*

$$\|v - v_h\| \leq \frac{1}{2} C_1 (\|x - x_h\| + \|(I - v \otimes v)H_h(x; v)\|).$$

*Proof.* Let  $\lambda := H_h(x; v) \cdot v$ , then

$$\begin{aligned} H_h(x_h; v) &= \lambda v + s, \\ \frac{1}{2} \|v\|^2 &= \frac{1}{2}, \end{aligned} \tag{26}$$

where

$$s = (H_h(x_h; v) - H_h(x; v)) + (I - v \otimes v)H_h(x; v).$$

Since  $v_h$  solves (26) with  $s = 0$ , and since

$$\|s\| \leq C_2 (\|x - x_h\| + \|(I - v \otimes v)H_h(x; v)\|),$$

the stated result follows from the Lipschitz continuity of  $H_h(\cdot; v)$  and an application of the inverse function theorem, in a similar spirit as the proof in §A.1.  $\square$

Next, we present a result ensuring that the rotation step of Algorithm 3 not only terminates but also produces a new dimer orientation  $v_k$  which remains in a small neighbourhood of the “exact” orientation  $v_h$ .

**Lemma 16.** *There exist  $r, h_0, C_2 > 0, C_3 \geq 1$  such that, if  $h \in (0, h_0]$ ,  $x_k \in B_r(x_*)$ ,  $v_{k-1} \in B_{C_3 r}(v_*)$ ,  $\|v_{k-1}\| = 1$ , then Step (3) of the Linesearch Dimer Algorithm terminates with outputs  $v_k \in B_{C_3 r}(v_*)$ ,  $\|v_k\| = 1$ ,  $\beta_k > 0$ , satisfying*

$$\|v_k - v_h\| \leq C_2 (\|x_k - x_h\| + h^2 \|v_{k-1} - v_h\|). \tag{27}$$

*Proof.* Let  $G(v) := h^{-2}(\mathcal{E}_h(x_k; v) - \mathcal{E}_h(x_k; V(x_k)))$ , then each step of the Rotation Algorithm is a steepest descent step of  $G$  on the manifold  $S_X$ . We need to ensure that these iterations do not “escape” from the minimiser  $V(x_k)$  (cf. Lemma 14).

Lemma 13 (with  $F(v) = G(v)$  and  $v_* = V(x_k)$ ) implies that each such step is a contraction towards  $V(x_k)$  with respect to the norm  $\|\cdot\|_H$  induced by the operator

$$H := (I - V \otimes V) \nabla^2 G(V) (I - V \otimes V) - (\nabla G(V) \cdot V) I,$$

where  $V \equiv V(x_k)$ ; provided that  $r$  is sufficiently small and  $H$  is positive definite.

To see that the latter is indeed true, we recall from (4) and (5) that

$$\nabla G(V) = \nabla^2 E(x_k)V + O(h^2) \quad \text{and} \quad \nabla^2 G(V) = \nabla^2 E(x_k) + O(h^2)$$

and from Proposition 2 and Lemma 14 that

$$V(x_k) = v_* + O(h^2 + r), \quad (28)$$

from which we can deduce that

$$\begin{aligned} H &= (I - v_* \otimes v_*) \nabla^2 E(x_*) (I - v_* \otimes v_*) - ((\nabla^2 E(x_*) v_*) \cdot v_*) I + O(h^2 + r) \\ &= (I - v_* \otimes v_*) \nabla^2 E(x_*) - \lambda_* I + O(h^2 + r). \end{aligned}$$

Since  $(x_*, v_*, \lambda_*)$  is an index-1 saddle,  $(I - v_* \otimes v_*) \nabla^2 E(x_*)$  is positive definite in  $\{v_*\}^\perp$ , and  $\lambda_* < 0$ . Thus, for  $h, r$  sufficiently small,  $H$  is positive definite as required.

From Lemma 13, it follows that all iterates  $v_k^{(j)}$  of the Rotation Algorithm satisfy  $\|v_k^{(j)} - V(x_k)\|_H \leq \|v_{k-1} - V(x_k)\|_H$ . Since the eigenvalues of  $H$  are uniformly bounded below and above, the norms  $\|\cdot\|_H, \|\cdot\|$  are equivalent, and hence in particular

$$\|v_k - V(x_k)\| \leq C_7 \|v_{k-1} - V(x_k)\| \leq C_7 (\|v_{k-1} - v_*\| + \|V(x_k) - v_*\|) = O(h^2 + r)$$

for some constant  $C_7 > 0$ , since  $v_{k-1} \in B_{C_{3r}}(v_*)$  and using (28). Combining this with (28) and choosing  $h_0^2 \leq r$ , we deduce that the Rotation Algorithm terminates with an iterate  $v_k$  such that

$$\|v_* - v_k\| \leq \|v_* - V(x_k)\| + \|v_k - V(x_k)\| \leq C_4 r$$

for some constant that depends only on  $r$  but is independent of  $v_{k-1}$  and remains bounded as  $r \rightarrow 0$ .

At termination the Rotation Algorithm guarantees the estimate

$$\|(I - v_k \otimes v_k) H_h(x_k; v_k)\| \leq \|\nabla_x \mathcal{E}_h(x_k, v_{k-1})\|.$$

We set  $x^t = (1 - t)x_h + tx_k$ ,  $v^t = v_h + tv_{k-1}$  and expand

$$\begin{aligned} \|\nabla_x \mathcal{E}_h(x_k, v_{k-1})\| &= \left\| \int_0^1 \left( \nabla_x^2 \mathcal{E}_h(x^t, v^t)(x_k - x_h) + \nabla_v \nabla_x \mathcal{E}_h(x^t, v^t)(v_{k-1} - v_h) \right) dt \right\| \\ &\leq C_2' (\|x_k - x_h\| + h^2 \|v_{k-1} - v_h\|). \end{aligned}$$

Combined with Lemma 15 this yields the estimate (27).

The statement that  $v_k \in B_{C_{3r}}(v_*)$  (instead of only  $B_{C_{4r}}(v_*)$ ) is an immediate consequence of (27) by ensuring that  $C_3 \geq C_2 + C_3 h^2 + C' h^4$ , where  $\|v_h - v_*\| \leq C' h^2$  for all  $h \leq h_0$  from Proposition 2. While there is an interdependence between  $C_3$  and  $C_2$ , for  $r$  and  $h_0$  sufficiently small, this is clearly achievable.  $\square$

**A.4.2. Analysis of the translation.** We first establish the existence of a minimiser of the auxiliary functional  $F_k$  under the conditions ensured by the rotation step of the Linesearch Dimer Algorithm.

**Lemma 17.** *Under the conditions of Lemma 16, possibly after choosing a smaller  $r, h_0$ , there exists a constant  $C_4 > 0$ , such that the functional  $F_k$  defined in (13) has a unique minimiser  $y_k \in B_r(x_*)$  satisfying*

$$\|y_k - x_h\| \leq C_4 (r_k^2 + h^2 r_k + h^4 s_{k-1}). \quad (29)$$



*Proof.* We begin by estimating the residual

$$\nabla F_k(x_h) = \nabla_x \mathcal{E}_h(x_h, v_k) - 2(\nabla_x \mathcal{E}_h(x_k, v_k) \cdot v_k)v_k + 2\lambda_k((x_k - x_h) \cdot v_k)v_k,$$

where  $\lambda_k = H_h(x_k; v_k) \cdot v_k$ . We consider each constituent term in this expression in turn; we expand about  $(x_h, v_h)$ , and use the identities (7), (9) and (24). This gives

$$\begin{aligned} v_k &= v_h + O(s_k) \\ \nabla_x \mathcal{E}_h(x_h, v_k) &= \nabla_x \nabla_v \mathcal{E}_h(x_h, v_h)(v_k - v_h) + O(s_k^2) \\ \nabla_x \mathcal{E}_h(x_k, v_k) &= \nabla_x^2 \mathcal{E}_h(x_h, v_h)(x_k - x_h) + \nabla_x \nabla_v \mathcal{E}_h(x_h, v_h)(v_k - v_h) \\ &\quad + O(r_k^2) + O(s_k^2) \\ &= \nabla_x^2 \mathcal{E}_h(x_h, v_h)(x_k - x_h) + O(h^2 s_k) + O(r_k^2) + O(s_k^2), \\ \nabla_x \mathcal{E}_h(x_k, v_k) \cdot v_k v_k &= (\nabla_x^2 \mathcal{E}_h(x_h, v_h)(x_k - x_h) + \nabla_x \nabla_v \mathcal{E}_h(x_h, v_h)(v_k - v_h)) \cdot v_h v_h \\ &\quad + O(r_k^2) + O(s_k^2) + O(r_k s_k) \\ &= \nabla_x^2 \mathcal{E}_h(x_h, v_h)(x_k - x_h) \cdot v_h v_h \\ &\quad + O(h^2 s_k) + O(r_k^2) + O(s_k^2) + O(r_k s_k) \\ H_h(x_k; v_k) &= H_h(x_h; v_h) + O(r_k) + O(s_k) \\ \lambda_k &= \lambda_h + v_k \cdot H_h(x_k; v_k) - v_h \cdot H_h(x_h; v_h) = \lambda_h + O(r_k) + O(s_k) \\ \lambda_k((x_k - x_h) \cdot v_k)v_k &= (\lambda_h + O(r_k) + O(s_k))((x_k - x_h) \cdot v_k)v_k \\ &= \lambda_h((x_k - x_h) \cdot v_h)v_h + O(r_k^2) + O(r_k s_k). \end{aligned}$$

Thus since (10) and our assumption that  $v_{k-1} \in B_{C_3 r}(v_*)$  ensure that  $s_{k-1} = O(1 + h_0^2)$ , while (27) implies that  $s_k = O(r_k) + O(h^2 s_{k-1})$ , we combine the above to obtain

$$\begin{aligned} \nabla F_k(x_h) &= -2[(\nabla_x^2 \mathcal{E}_h(x_h, v_h)(x_k - x_h)) \cdot v_h]v_h + 2\lambda_h((x_k - x_h) \cdot v_h)v_h \\ &\quad + O(r_k^2 + h^2 r_k + h^4 s_{k-1}), \end{aligned}$$

Next, we note that, by definition of  $\mathcal{E}_h$ ,  $\nabla_x^2 \mathcal{E}_h(x_h, v_h)v_h = \nabla^2 E(x_h)v_h + O(h^2)$ , and thus from (4) that  $\nabla_x^2 \mathcal{E}_h(x_h, v_h)v_h = H_h(x_h; v_h) + O(h^2)$ . Hence applying (9),

$$\nabla F_k(x_h) = [-2H_h(x_h; v_h) \cdot (x_k - x_h) + 2\lambda_h(x_k - x_h) \cdot v_h]v_h \quad (30)$$

$$\begin{aligned} &\quad + O(r_k^2 + h^2 r_k + h^4 s_{k-1}) \\ &= O(r_k^2 + h^2 r_k + h^4 s_{k-1}). \end{aligned} \quad (31)$$

Finally, we observe that  $\nabla^2 F_k(x_h)$  is positive definite, since

$$\begin{aligned} \nabla^2 F_k(x_h) &= \nabla_x^2 \mathcal{E}_h(x_h, v_k) - 2\lambda_k v_k \otimes v_k \\ &= \nabla_x^2 \mathcal{E}_h(x_h, v_h) - 2\lambda_h v_h \otimes v_h + O(r_k) \\ &= \nabla^2 E(x_*) - 2\lambda_* v_* \otimes v_* + O(h^2 + r_k), \end{aligned} \quad (32)$$

which immediately implies that, for  $r, h_0$  sufficiently small,  $\nabla^2 F_k(x_h)$  is an isomorphism with uniformly bounded inverse.

Thus an application of the inverse function theorem to  $\nabla F_k$  at  $y_k$  using (31) yields the stated result.  $\square$

We now turn towards analysing the linesearch for  $x$ . Recall the definition of the energy norm  $\|u\|_* := \sqrt{u \cdot ((I - 2v_* \otimes v_*)\nabla^2 E(x_*)u)}$ , which is equivalent to  $\|\cdot\|$ . In

particular,

$$\mu^{1/2}\|u\| \leq \|u\|_* \leq \|\nabla^2 E(x_*)\| \|u\| \quad \text{where} \quad \mu := \min(-\lambda_*, \mu_*) > 0. \quad (33)$$

**Lemma 18.** *There exists  $r, h_0, \underline{\alpha} \in (0, \alpha_0]$  and  $\gamma_* \in (0, 1)$ , such that, if  $h \in (0, h_0]$ ,  $x_k \in B_r(x_*)$ ,  $v_k \in B_{C_3 r}(v_*)$  and  $\alpha_{k-1} \geq \underline{\alpha}$ , then*

$$\alpha_k \geq \underline{\alpha} \quad \text{and} \quad \|x_{k+1} - y_k\|_* \leq \gamma_* \|x_k - y_k\|_*,$$

where  $y_k$  is the minimiser of  $F_k$  established in Lemma 17.

*Proof.* We begin by noting that, for any  $r > 0$ , the norms  $\|\nabla^2 F_k(x)\|$  are uniformly bounded among all choices of  $x_k \in B_r(x_*)$ ,  $x \in B_{r+1}(x_*)$ . This is straightforward to establish.

Therefore, there exists  $\underline{\alpha} > 0$  such that, for  $x_k \in B_r(x_*)$  and for any  $\alpha \in (0, 2\underline{\alpha}]$ , the conditions in Step (6) of the Linesearch Dimer Algorithm are met (this includes an Armijo condition for  $F_k$ ) since  $\nabla F_k$  is Lipschitz in a neighbourhood of  $x_k$  [8, Thm.2.1]. It is no restriction of generality to require  $\underline{\alpha} \leq \alpha_0$ . In particular,  $\alpha_k \geq \underline{\alpha}$ .

For  $r, h_0$  sufficiently small, we have  $y_k \in B_r(x_*)$  as well. Upon choosing  $r$  sufficiently small,  $u \cdot (\nabla^2 F_k(y)u) \geq \mu/2\|u\|^2$  for all  $u \in X$ ,  $y \in B_r(x_*)$ . Thus, we can apply Lemma 12 (with  $x_* \equiv y_k$ ) to deduce that, for  $r$  sufficiently small, the step  $x_{k+1} = x_k - \alpha_k \nabla F(x_k)$  is a contraction with a constant  $\gamma_1$  that is independent of  $x_k, v_k$ . That is,

$$(x_{k+1} - y_k) \cdot [\nabla^2 F_k(y_k)(x_{k+1} - y_k)] \leq \gamma_1^2 (x_k - y_k) \cdot [\nabla^2 F_k(y_k)(x_k - y_k)],$$

Recalling from (29) and (32) that  $\nabla^2 F_k(y_k) = (I - 2v_* \otimes v_*)\nabla^2 E(x_*) + O(r + h^2)$  we find that, for  $r, h_0$  sufficiently small,

$$\|x_{k+1} - y_k\|_* \leq \gamma_* \|x_k - y_k\|_*, \quad (34)$$

where  $\gamma_* \in [\gamma_1, 1)$ , again independent of  $x_k, v_k$ , but depending on  $r, h_0$ .  $\square$

**A.4.3. Proof of Case (ii).** We have assembled all prerequisites to complete the proof of Theorem 6, Case (ii).

Inspired by Lemma 18, our aim is to prove that, for  $r$  sufficiently small, there exists  $\gamma \in (0, 1)$  such that, for all  $j \geq 0$ ,

$$r_j^* + h^2 s_{j-1} \leq \gamma^j (r_0^* + h^2 s_{-1}) =: \gamma^j t_0, \quad (35)$$

where  $\gamma := \frac{1}{2}(\gamma_* + 1)$ ,  $r_k^* := \|x_k - x_h\|_*$  and  $s_k := \|v_k - v_h\|$ .

A consequence of (35) would be that there exists a constant  $c$  such that  $\|x_j - x_*\| \leq cr =: \hat{r}$ . Thus, under the assumptions of the Theorem, let  $r, h_0$  be chosen sufficiently small so that Proposition 2, and Lemmas 15, 16, 17 and 18 apply with  $r$  replaced by  $\hat{r}$ .

We now begin the induction argument adding to (35) the conditions that

$$v_{j-1} \in B_{C_3 r}(v_*) \quad \text{and} \quad \alpha_j \geq \underline{\alpha}, \quad (36)$$

where  $C_3 \geq 1$  is the constant from Lemma 16 and  $\underline{\alpha}$  the constant from Lemma 18. Clearly (35) and (36) hold for  $j = 0$ . Suppose that they hold for  $j = 0, \dots, k$ , where  $k \geq 0$ .

The choice of  $r$  implies that  $x_k \in B_r(x_*)$  again, and Lemma 16 implies that  $v_k \in B_{C_3 r}(v_*)$ . Thus, the first condition in (36) is established for  $j = k + 1$ .

Applying Lemma 18 we obtain the second condition in (36) for  $j = k + 1$ , and in addition that

$$\|x_{k+1} - y_k\|_* \leq \gamma_* \|x_k - y_k\|_*,$$

where  $y_k$  is the minimiser of  $F_k$  established in Lemma 17. Using (34), the fact that  $\gamma_* < 1$  and Lemma 17 we therefore deduce that there exists a constant  $C_5$  which depends on  $C_4$  and on the norm-equivalence between  $\|\cdot\|$  and  $\|\cdot\|_*$ , such that

$$\begin{aligned} \|x_{k+1} - x_h\|_* &\leq \|x_{k+1} - y_k\|_* + \|y_k - x_h\|_* \\ &\leq \gamma_* \|x_k - y_k\|_* + \|y_k - x_h\|_* \\ &\leq \gamma_* \|x_k - x_h\|_* + 2\|y_k - x_h\|_* \\ &\leq (\gamma_* + C_5 h^2 + C_5 r_k) \|x_k - x_h\|_* + C_5 h^4 \|v_{k-1} - v_h\|. \end{aligned}$$

Adding  $h^2 \|v_k - v_h\|$  to both sides of the inequality and applying (27) and (33) we thus obtain

$$\begin{aligned} r_{k+1}^* + h^2 s_k &\leq (\gamma_* + C_5 h^2 + C_5 r_k) r_k^* + h^2 s_k + C_5 h^4 s_{k-1} \\ &\leq (\gamma_* + C_5 h^2 + \mu^{-1/2} C_2 h^2 + C_5 (c+1) r) r_k^* + (C_5 + C_2) h^4 s_{k-1}. \end{aligned}$$

Recalling that  $\gamma = \frac{1}{2}(\gamma_* + 1)$ , choosing  $h_0, r$  sufficiently small, we obtain that

$$r_{k+1}^* + h^2 s_k \leq \gamma(r_k^* + h^2 s_{k-1}).$$

This establishes (35) for  $j = k + 1$  and thus completes the induction argument.

In summary, we have proven that (35) and (36) hold for all  $j \geq 0$ . As a first consequence, we obtain that  $r_k := \|x_k - x_h\| \leq \mu^{-1/2} \|\nabla^2 E(x_*)\| \gamma^k (r_0 + h^2 s_{-1})$  using (33), which in particular establishes the first part of (15).

To obtain a convergence rate for  $v_k$  we combine (27) and (35), to obtain

$$\|v_k - v_h\| \leq C_6 (r_k^* + h^2 s_{k-1}) \leq C_6 \gamma^k t_0 \leq C_6 \|\nabla^2 E(x_*)\| \gamma^k (r_0 + h^2 s_{-1}),$$

for a constant  $C_6$ . Choosing  $C = 2 \max(C_6, \mu^{-1/2} \|\nabla^2 E(x_*)\|)$  completes the proof of Theorem 6, Case (ii).

**A.4.4. Proof of Case (i).** The proof of Case (ii) establishes that, for as long as  $\nabla_x \mathcal{E}_h(x_k, v_{k-1}) \neq 0$ , the iterates are well-defined and  $\|x_k - x_h\| + \|v_k - v_h\| \leq Cr$  for some suitable constant  $C$ . We now drop this assumption and instead suppose that, at the  $\ell$ th iterate,  $\nabla_x \mathcal{E}_h(x_\ell, v_{\ell-1}) = 0$ . In this case, we can apply the following lemma.

**Lemma 19.** *Let  $(x_*, v_*, \lambda_*)$  be an index-1 saddle, then there exist  $r, h_0, C > 0$  such that, for all  $h \in (0, h_0]$  and for all  $v \in S_X$ , there exists a unique  $x_{h,v} \in B_r(x_*)$  such that  $\nabla_x \mathcal{E}_h(x_{h,v}, v) = 0$ . Moreover,  $\|x_{h,v} - x_h\| \leq Ch^2$ .*

*Proof.* This is an immediate corollary of (2) and the inverse function theorem.  $\square$

Since  $\|x_\ell - x_h\| \leq Cr$ , Lemma 19 implies that, in fact  $\|x_\ell - x_h\| \leq C'h^2$  for some other constants  $C'$ , provided that  $r, h$  are chosen sufficiently small.

This concludes the proof of Theorem 6, Case (i).

## REFERENCES

- [1] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, Princeton, USA, 2008.
- [2] P.-A. Absil, R. Mahony, and J. Trumpf. An extrinsic look at the Riemannian Hessian. In F. Nielsen and F. Barbaresco, editors, *Geometric Science of Information*, number 8005 in Lecture Notes in Computer Science, pages 361–368, Heidelberg, Berlin, New York, 2013. Springer Verlag.
- [3] A. Banerjee, N. Adams, J. Simons, and R. Shepard. Search for stationary points on surfaces. *The Journal of Physical Chemistry*, 89:52–57, 1985.
- [4] G.T. Barkema and N. Mousseau. Event-based relaxation of continuous disordered systems. *Physical Review Letters*, 77:4358, 1996.
- [5] G.T. Barkema and N. Mousseau. The activation-relation technique: an efficient algorithm for sampling energy landscapes. *Computational Materials Science*, 20(3–4):285–292, 2001.
- [6] E. Cancès, 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. *The Journal of Chemical Physics*, 130(114711), 2009.
- [7] C.J. Cerjan and W.H. Miller. On finding transition states. *Journal of Chemical Physics*, 75:2800, 1981.
- [8] N. I. M. Gould and S. Leyffer. An introduction to algorithms for nonlinear optimization. In A. W. Craig J. F. Blowey and T. Shardlow, editors, *Frontiers in Numerical Analysis (Durham 2002)*, pages 109–197, Heidelberg, Berlin, New York, 2003. Springer Verlag.
- [9] N. I. M. Gould, C. Ortner, and D. Packwood. An efficient dimer method with preconditioning and linesearch. ArXiv:1407.2817v2.
- [10] G. Henkelman and H. Jónsson. A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives. *Journal of Chemical Physics*, 111(5):7010–7022, 1999.
- [11] A. Heyden, A. T. Bell, and F. J. Keil. Efficient methods for finding transition states in chemical reactions: Comparison of improved dimer method and partitioned rational function optimization method. *Journal of Chemical Physics*, 123(224101), 2005.
- [12] H. Jónsson, G. Mills, and K. W. Jacobsen. Nudged elastic band for finding minimum energy paths of transitions. In G. Ciccotti B. J. Berne and D. F. Coker, editors, *Classical and quantum dynamics in condensed phase simulations*, volume 385. World Scientific, 1998.
- [13] J. Kästner and P. Sherwood. Superlinearly converging dimer method for transition state search. *The Journal of Chemical Physics*, 128(014106), 2008.
- [14] C. Lanczos. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *Journal of research of the National Bureau of Standards B*, 45:225–280, 1950.
- [15] D. Liu and J. Nocedal. On the limited memory BFGS method for large scale optimization. *Mathematical Programming, Series B*, 45(3):503–528, 1989.
- [16] E. Machado-Charry, L.K. Beland, D. Caliste, Luigi Genovese, T. Deutsch, N. Mousseau, and P. Pochet. Optimized energy landscape exploration using the ab initio based activation-relaxation technique. *Journal of Chemical Physics*, 135(034102), 2011.
- [17] M.C. Marinica, F. Willaime, and N. Mousseau. Energy landscape of small clusters of self-interstitial dumbbells in iron. *Physical Review B*, 83(094119), 2011.
- [18] N. Mousseau, L.K. Beland, P. Brommer, J.F. Joly, F. El-Mellouhi, E. Machado-Charry, M.C. Marinica, and P. Pochet. The activation-relaxation technique: Art nouveau and kinetic art. *Journal of Atomic, Molecular and Optical Physics*, 2012(952278), 2012.
- [19] B. A. Murtagh and R. W. H. Sargent. Computational experience with quadratically convergent minimisation methods. *The Computer Journal*, 13:185–194, 1970.
- [20] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, 1999.
- [21] R. A. Olsen, G. J. Kroes, G. Henkelman, A. Arnaldsson, and H. Jnsson. Comparison of methods for finding saddle points without knowledge of the final states. *Journal of Chemical Physics*, 121:9776, 2004.
- [22] J. Simons, P. Joergensen, H. Taylor, and J. Ozment. Walking on potential energy surfaces. *The Journal of Physical Chemistry*, 87:2745–2753, 1983.

- [23] B. P. Uberuaga, F. Montalenti, T. C. Germann, and A. F. Voter. Accelerated molecular dynamics methods. In S. Yip, editor, *Handbook of Materials Modelling, Part A- Methods*, page 629. Springer, 2005.
- [24] A. E. Perekatov V. S. Mikhalevich, N. N. Redkovskii. Methods of minimization of functions on a sphere and their applications. *Cybernetics and Systems Analysis*, 23(6):721–730, 1987.
- [25] A. F. Voter. Accelerated molecular dynamics of infrequent events. *Physical Review Letters*, 78(3908), 1997.
- [26] A. F. Voter. Introduction to the kinetic Monte Carlo method. In K. E. Sickafus, E. A. Kotomin, and B. P. Uberuaga, editors, *Radiation Effects in Solids*, volume 235 of *NATO Science Series*, pages 1–23. Springer Netherlands, 2007.
- [27] E. Weinan, W. Ren, and E. Vanden-Eijnden. String method for the study of rare events. *Physical Review B*, 66(052301), 2002.
- [28] E. Weinan, W. Ren, and E. Vanden-Eijnden. Simplified and improved string method for computing the minimum energy path in barrier-crossing events. *Journal of Chemical Physics*, 126(164103), 2007.
- [29] J. Zhang and Q. Du. Shrinking dimer dynamics and its applications to saddle point search. *SIAM Journal of Numerical Analysis*, 50(4):1899–1921, 2012.
- [30] L. Zhang, J. Zhang, and Q. Du. Finding critical nuclei in phase transformations by shrinking dimer dynamics and its variants. *Communications in Computational Physics*, 16(3):781–798, 2014.

N.I.M. GOULD, SCIENTIFIC COMPUTING DEPARTMENT, STFC-RUTHERFORD APPLETON LABORATORY, CHILTON, OX11 0QX, UK

*E-mail address:* `nick.gould@stfc.ac.uk`

C. ORTNER, MATHEMATICS INSTITUTE, ZEEMAN BUILDING, UNIVERSITY OF WARWICK, COVENTRY CV4 7AL, UK

*E-mail address:* `christoph.ortner@warwick.ac.uk`

D. PACKWOOD, MATHEMATICS INSTITUTE, ZEEMAN BUILDING, UNIVERSITY OF WARWICK, COVENTRY CV4 7AL, UK

*E-mail address:* `d.packwood@warwick.ac.uk`