RESEARCH ARTICLE | NOVEMBER 22 2004

Comparison of methods for finding saddle points without knowledge of the final states

R. A. Olsen; G. J. Kroes; G. Henkelman; A. Arnaldsson; H. Jónsson

J. Chem. Phys. 121, 9776–9792 (2004)
https://doi.org/10.1063/1.1809574

Comparison of methods for finding saddle points without knowledge of the final states

R. A. Olsen and G. J. Kroes
Leiden Institute of Chemistry, Gorlaeus Laboratories, Leiden University, P.O. Box 9502, 2300 RA Leiden, The Netherlands

G. Henkelman
Department of Chemistry and Biochemistry, The University of Texas at Austin, 1 University Station A5300, Austin, Texas 78712-0165

A. Arnaldsson
Department of Chemistry, University of Washington, P.O. Box 351700, Seattle, Washington 98195-1700

H. Jónsson
Department of Chemistry, University of Washington, P.O. Box 351700, Seattle, Washington 98195-1700 and Faculty of Science, VR-II, University of Iceland, 107 Reykjavik, Iceland

(Received 14 July 2004; accepted 3 September 2004)

Within the harmonic approximation to transition state theory, the biggest challenge involved in finding the mechanism or rate of transitions is the location of the relevant saddle points on the multidimensional potential energy surface. The saddle point search is particularly challenging when the final state of the transition is not specified. In this article we report on a comparison of several methods for locating saddle points under these conditions and compare, in particular, the well-established rational function optimization (RFO) methods using either exact or approximate Hessians with the more recently proposed minimum mode following methods where only the minimum eigenvalue mode is found, either by the dimer or the Lanczos method. A test problem involving transitions in a seven-atom Pt island on a Pt(111) surface using a simple Morse pairwise potential function is used and the number of degrees of freedom varied by varying the number of movable atoms. In the full system, 175 atoms can move so 525 degrees of freedom need to be optimized to find the saddle points. For testing purposes, we have also restricted the number of movable atoms to 7 and 1. Our results indicate that if attempting to make a map of all relevant saddle points for a large system (as would be necessary when simulating the long time scale evolution of a thermal system) the minimum mode following methods are preferred. The minimum mode following methods are also more efficient when searching for the lowest saddle points in a large system, and if the force can be obtained cheaply. However, if only the lowest saddle points are sought and the calculation of the force is expensive but a good approximation for the Hessian at the starting position of the search can be obtained at low cost, then the RFO approaches employing an approximate Hessian represent the preferred choice. For small and medium sized systems where the force is expensive to calculate, the RFO approaches employing an approximate Hessian is also the more efficient, but when the force and Hessian can be obtained cheaply and only the lowest saddle points are sought the RFO approach using an exact Hessian is the better choice. These conclusions have been reached based on a comparison of the total computational effort needed to find the saddle points and the number of saddle points found for each of the methods. The RFO methods do not perform very well with respect to the latter aspect, but starting the searches further away from the initial minimum or using the hybrid RFO version presented here improves this behavior considerably in most cases. [DOI: 10.1063/1.1809574]

I. INTRODUCTION

Within the harmonic approximation to transition state theory (whether used explicitly or implicitly), finding the mechanism or rate of transition for a chemical reaction or for the diffusion of one or more atoms on a surface or in a bulk system, requires the location of saddle points on the potential energy surface (PES) governing the transition. In these cases the transition rates are largely determined by the region around the saddle points and their energy difference with the starting minimum. It is therefore not surprising that a large number of studies have focused on developing or refining methods for locating saddle points on PESs (see, e.g., Refs. 1–27 or references found in the reviews Refs. 28 and 29). However, in many of these methods it is assumed that a reasonable guess for the saddle point can be made and/or that the reaction proceeds to a known final state. If one is looking for saddle points describing yet to be discovered reaction mechanisms or one would like to map out as many saddle points as possible irrespectively of the final states, these methods will not suffice.

An illustration of the importance of having methods that

can be used to systematically walk from a given minimum towards saddle points, without assuming any knowledge of the final states, is the discovery by Feibelman in 1990 that an Al atom does not diffuse on the Al(100) surface by repeated hops from one site to another, as had been previously assumed, but rather by a concerted displacement of two atoms.[30] In chemistry, there are many reactions where the mechanism and reaction intermediates are unknown, which could in principle be tackled using transition state theory. Examples include the photo-oxidation of water on TiO_2,[31,32] which is relevant in the clean production of hydrogen, the conversion of $NaAlH_4$ to Na_3AlH_6,[33] which is relevant to hydrogen storage, and several reactions important to biochemistry.[34–39] The number of methods currently available allowing one to walk from a minimum to a saddle point without any knowledge of the final state are limited, and, to the best of our knowledge, a systematic comparison of such methods has not been reported. The goal of the present study is to provide such a comparison.

Many of the traditional methods used within the chemistry community that are able to start from a minimum and converge on a saddle point without using any knowledge of the final state are variations on the pioneering efforts of Cerjan and Miller[1] and Simons *et al.*[3,5] We have investigated the performance of some of these modified Newton-Raphson approaches (we will refer to them as all-mode following methods), which are the rational function optimization methods using either exact or approximate Hessians (the second derivative matrix of the potential energy with respect to the nuclear coordinates). More recently, a different approach has been proposed where only the minimum eigenvalue mode is required. Rather than constructing the full Hessian matrix and diagonalizing it, only the minimum mode is found. In this study we will focus on two representatives of what we will refer to as minimum mode following methods, the dimer[17] and Lanczos methods.[19] Other methods could also be investigated, such as various preconditioners[24,40] or modified eigenvector following methods,[16] but our intention is to provide a comparison between two main classes of saddle point search methods and this goal can be reached by considering representative members of each class. An important factor in determining which method is the more successful is the number of force evaluations and geometry steps needed to reach the saddle points. Also the ability of the methods to discover as many saddle points as possible might play an important role. The computational effort per relevant saddle point found will eventually decide which method will be the preferred one. In this study all these aspects will be discussed and the results might serve as a basis for further methodological improvements.

All our tests have been performed on a model system involving transitions in a seven-atom Pt island on a Pt(111) surface using a simple Morse pairwise potential function. This test system was previously used to study the efficiency of various methods which rely on knowledge of both the initial and final state.[28] A relatively simple model system is chosen in these studies to make it easy for others to reproduce the results, apply other methods to the same system, and to allow for a rather extensive study including three different system sizes. In the full system, 175 atoms can move so 525 degrees of freedom need to be optimized to find the saddle points. In the medium sized system, only the seven atoms of the island are free to move (21 degrees of freedom). In the small system, only one of these seven atoms can move (three degrees of freedom). Even though the model interactions are of a simple Morse form, the resulting PESs describe a wide variety of realistic transition mechanisms and provide challenging tasks for the PES walkers.

The outline of the paper is the following: In Sec. II the different methods considered are described. This is followed by a short description of the model PESs in Sec. III and the results and a detailed discussion of them are given in Sec. IV. Conclusions are presented in Sec. V.

II. THE POTENTIAL ENERGY SURFACE WALKERS

All methods described below rely on a local approximation to the PES using information about the gradient $\mathbf{g}$ and the exact or approximated Hessian matrix $\mathbf{H}$ at the current configuration of the system, $\mathbf{x}_l$. Using all or parts of the available information a step vector, $\Delta\mathbf{x}_l$, is calculated and the system moved to a new configuration $\mathbf{x}_{l+1}=\mathbf{x}_l+\Delta\mathbf{x}_l$. The series of geometry iterations is in our case started close to a local minimum of the PES, $\mathbf{x}_1$, and continued until all components of the gradient vector fall below a given threshold, $\delta g_{\max}$. From a set of initial starting configurations the goal is to locate as many of the saddle points directly connected to the minimum as possible.

The methods we consider can be separated in two main classes.

(i) In the minimum mode following methods only the lowest eigenvalue and the corresponding eigenvector of the Hessian are sought and subsequently used together with the gradient to determine the step vector. Either the Lanczos iterative method (Sec. II A) or the dimer method (Sec. II B) can be used to find the minimum mode. In the form they are presented here they use exactly the same way of calculating the step vector (Sec. II C), the only difference lies in how the lowest Hessian eigenvalue and the corresponding eigenvector are obtained.

(ii) In the all-mode following methods the full Hessian matrix is calculated or approximated, and all eigenvalues and eigenvectors are used in conjunction with the gradient to determine the step vector. Here we have used a rational function optimization approach as an example of an all-mode following method. Two versions, one using the exact Hessian (Sec. II D) and the other using approximate Hessians (Sec. II E), are outlined below. We also investigate the behavior of a hybrid method (Sec. II F).

A. A Lanczos iterative method for finding the minimum mode

The activation-relaxation technique nouveau (ARTn) employed in Ref. 19 was the first saddle point search method that used a Lanczos iterative approach for calculating the lowest eigenvalue of the Hessian matrix. In this method the system is moved stepwise from a local minimum along a random direction until a negative Hessian eigenvalue is encountered. For each step the total energy of the configuration is minimized in the hyperplane perpendicular to the step di-

rection. Once a negative eigenvalue has been found, the system is pushed against the force along the eigenvector of the Hessian corresponding to the negative eigenvalue, while minimizing the force in all other directions. Unless the lowest eigenvalue turns positive, this procedure ensures convergence to a first-order saddle point, and as such it is an improvement of the earlier version of the activation-relaxation technique presented in Refs. 11 and 15.

Two methods similar in spirit to ARTn were previously introduced by Munro and Wales.[16]

(i) In one approach a shifted power iteration scheme[16,41] is used to find the eigenvector corresponding to the lowest eigenvalue of the Hessian matrix. A step uphill along the eigenvector corresponding to the lowest eigenvalue is then taken, while a conjugate gradient method is used to minimize the total energy in all other directions. Although the method only uses information about the eigenvector corresponding to the lowest eigenvalue, the full Hessian matrix needs to be constructed in each step. But what has been gained is that there is no longer a need for a full diagonalization of the Hessian matrix, an operation that can become prohibitive for large systems.

(ii) The second approach relies on a variational technique where the Rayleigh-Ritz ratio[16,41] is minimized through a conjugated gradient method. In this way the lowest eigenvalue and the corresponding eigenvector can be obtained without having to build the full Hessian matrix. The obtained information is then used in the same way as in the first approach to step towards a saddle point.

In Ref. 24 the Davidson method[40] was used to efficiently characterize the stationary points on a PES by computing the lowest eigenvalue of the Hessian matrix. Although the method was only used to determine whether a stationary point was a local minimum or a saddle point, it was suggested that the method could be used in conjunction with eigenvector-following methods as an efficient way of optimizing transition states. As such it could be another member of the class we refer to as minimum mode following methods.

In the following we will describe in some detail how one of the schemes used here, the Lanczos scheme, can be used to calculate the minimum mode. The Hessian matrix is real and symmetric and can therefore be reduced to a tridiagonal form by an orthogonal similarity transformation, $\mathbf{T}=\mathbf{Q}^t\mathbf{H}\mathbf{Q}$ (t denotes the transpose).[42] After obtaining the eigenvalue and eigenvector pairs, $\{\lambda_i,\mathbf{v}_i^{\mathbf{T}}\}_{i=1,\ldots,n}$ of $\mathbf{T}$, the similarity ensures that the eigenpairs of $\mathbf{H}$ are given by $\{\lambda_i,\mathbf{v}_i=\mathbf{Q}\mathbf{v}_i^{\mathbf{T}}\}_{i=1,\ldots,n}$ (the superscript $\mathbf{T}$ is used to indicate that the eigenvector is obtained from the tridiagonal matrix $\mathbf{T}$). If, as in our case, one is interested in the lowest eigenpair only, $\{\lambda_1,\mathbf{v}_1\}$, a Lanczos scheme can be used to construct a set of tridiagonal matrices, $\{\mathbf{T}_2,\ldots,\mathbf{T}_j\}$, whose lowest eigenvalue will converge to λ_1 as j increases.[41,43,44] The advantages of the Lanczos scheme are that it (i) replaces the diagonalization of the $n\times n$ Hessian matrix by the diagonalization of a tridiagonal matrix $\mathbf{T}_j$, where $j\ll n$ when n is large, and (ii) requires $\mathbf{H}$ to be known only in a small, j-dimensional basis of the Lanczos vectors in which $\mathbf{H}$ is tridiagonal, rather than in an n-dimensional basis of a given set of primitive vectors. The tridiagonal matrices are constructed according to[41,43,44]

$$\mathbf{T}_j=\begin{bmatrix}\alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \beta_2 & \alpha_3 & \ddots & \\ & & \ddots & \ddots & \beta_{j-1} \\ & & & \beta_{j-1} & \alpha_j\end{bmatrix}, \tag{1}$$

where $\alpha_{k=1,\ldots,j}$ and $\beta_{k=1,\ldots,j-1}$ are obtained through an iterative procedure. Beginning with a vector $\mathbf{r}_0$ (if it is the first geometry iteration a random nonzero vector is chosen, otherwise the eigenvector found in the previous geometry cycle is used) and setting $\beta_0=\|\mathbf{r}_0\|$ and $\mathbf{q}_0=\mathbf{0}$ the following steps are repeated ($k=1,\ldots,j$):

$$\mathbf{q}_k=\frac{\mathbf{r}_{k-1}}{\beta_{k-1}}, \tag{2}$$

$$\mathbf{u}_k=\mathbf{H}\mathbf{q}_k, \tag{3}$$

$$\mathbf{r}_k=\mathbf{u}_k-\beta_{k-1}\mathbf{q}_{k-1}, \tag{4}$$

$$\alpha_k=\mathbf{q}_k^t\mathbf{r}_k, \tag{5}$$

$$\mathbf{r}_k=\mathbf{r}_k-\alpha_k\mathbf{q}_k, \tag{6}$$

$$\beta_k=\|\mathbf{r}_k\|. \tag{7}$$

For each new set $\{\alpha_k,\beta_{k-1}\}$ the lowest eigenvalue $\lambda_1^{\mathbf{T}_k}$ of $\mathbf{T}_k$ is found. Once

$$\left|\frac{\lambda_1^{\mathbf{T}_k}-\lambda_1^{\mathbf{T}_{k-1}}}{\lambda_1^{\mathbf{T}_{k-1}}}\right|<\delta\lambda_{\mathrm{L}} \tag{8}$$

the eigenvalue is considered converged to $\lambda_1=\lambda_1^{\mathbf{T}_{k=j}}$. The choice of $\delta\lambda_{\mathrm{L}}$ will be discussed later. We also consider the possibility of terminating the iteration cycles after $n_{\mathrm{L}}^{\max}$ iterations. If the full Hessian is not known in a given basis of primitive vectors, the second step in the iteration cycle above [Eq. (3)] can be replaced by the finite difference approximation

$$\mathbf{u}_k=\frac{\mathbf{g}(\mathbf{x}_k)-\mathbf{g}(\mathbf{x}_l)}{\delta x_{\mathrm{L}}}, \tag{9}$$

with

$$\mathbf{x}_k=\mathbf{x}_l+\delta x_{\mathrm{L}}\mathbf{q}_k, \tag{10}$$

where $\mathbf{x}_l$ is the current configuration of the system and $\mathbf{g}$ the gradient at the configuration indicated. The choice of δx_{L} will be discussed later. As can be seen from Eqs. (9) and (10), the second derivative of the potential energy only needs to be calculated along the Lanczos vectors. Note that the iteration formulas, Eqs. (2) through (7), do not contain a reorthogonalization step.[41,43,44] In the cases we have considered here the convergence is fast enough to ensure that the orthogonality of the Lanczos vectors, $\{\mathbf{q}_1,\ldots,\mathbf{q}_j\}$, is not lost.

The eigenvalues of the matrices $\mathbf{T}_{k=1,\ldots,j}$ can be obtained efficiently using a standard QL algorithm with implicit shifts.[42] Once the lowest eigenvalue is considered converged to $\lambda_1=\lambda_1^{\mathbf{T}_j}$, the corresponding eigenvector $\mathbf{v}_1^{\mathbf{T}_j}$, can be found

by inverse iteration. To be able to use the very efficient Cholesky factorization scheme[42] the eigenvalue spectrum of $\mathbf{T}_j$ is shifted so that all eigenvalues are positive. Furthermore, to ensure fast convergence of the inverse iteration, the lowest eigenvalue is set to a small positive number $\lambda_{\text{small}}^{\mathbf{T}_j}$ (we have used $\lambda_{\text{small}}^{\mathbf{T}_j}=10^{-4}$ eV/Å^2). Finally, since the set of orthonormal Lanczos vectors $\{\mathbf{q}_1,\ldots,\mathbf{q}_j\}$ form the column vectors of the matrix $\mathbf{Q}$, the eigenvector corresponding to the lowest eigenvalue λ_1 of $\mathbf{H}$ can be obtained through $\mathbf{v}_1=\mathbf{Q}\mathbf{v}_1^{\mathbf{T}_j}$ (note that $\mathbf{Q}$ is of dimension $n\times j$, and $\mathbf{v}_1^{\mathbf{T}_j}$ of dimension $j\times 1$).

From the outline of the Lanczos method given above and the method used to calculate the geometry step (Sec. II C), it is clear that there are many parameters available for tuning to reach optimal performance. However, very good performance can be reached by adopting ''standard'' settings. This will be discussed in more detail in Sec. IV.

B. The dimer method for finding the minimum mode

Since a full presentation of the dimer method was given in Ref. 17 we will only review the important aspects of the method here together with some modifications. The dimer, consisting of two auxiliary configurations (images) of the system, is defined by

$$\mathbf{x}_l^1=\mathbf{x}_l+\delta x_{\mathrm{D}}\hat{\mathbf{N}}, \tag{11}$$

$$\mathbf{x}_l^2=\mathbf{x}_l-\delta x_{\mathrm{D}}\hat{\mathbf{N}}, \tag{12}$$

where $\mathbf{x}_l$ is the current configuration of the system and δx_{D} determines the displacement of the two images along a unit vector $\hat{\mathbf{N}}$ (in the first geometry iteration a random nonzero unit vector is chosen for $\hat{\mathbf{N}}$, while for the following geometry iterations the lowest Hessian eigenvector found in the previous cycle is used. The choice of δx_{D} will be discussed later). Next, the dimer energy is defined as the sum of the energies for the two images, $V_{\mathrm{D}}=V_1+V_2$. The essential feature of the method is that when V_{D} is minimized under the constraint of fixed $\mathbf{x}_l$ and δx_{D}, i.e., a rotation of the dimer around the midpoint, the dimer will align itself along the eigenvector of the Hessian matrix corresponding to the lowest eigenvalue.

The following set of iterative operations will accomplish this: First, the forces $\mathbf{F}_l=-\mathbf{g}(\mathbf{x}_l)$ and $\mathbf{F}_1=-\mathbf{g}(\mathbf{x}_l^1)$ are calculated, and then used to approximate the force at the second image through $\mathbf{F}_2=2\mathbf{F}_l-\mathbf{F}_1$. Next, the scaled rotational force acting on the dimer is obtained by $\mathbf{F}^{\perp}=(\mathbf{F}_1^{\perp}-\mathbf{F}_2^{\perp})/\delta x_{\mathrm{D}}$, where $\mathbf{F}_i^{\perp}\equiv\mathbf{F}_i-(\mathbf{F}_i\cdot\hat{\mathbf{N}})\,\hat{\mathbf{N}}$ for $i=1,2$. A second unit vector (normal to the first by construction) is defined as $\hat{\mathbf{\Theta}}=\mathbf{F}^{\perp}/\|\mathbf{F}^{\perp}\|$, and subsequently $\hat{\mathbf{N}}$ and $\hat{\mathbf{\Theta}}$ are rotated through an angle $\delta\theta_{\mathrm{D}}$ within the plane spanned by the two unit vectors, giving two new orthonormal vectors $\hat{\mathbf{N}}^*$ and $\hat{\mathbf{\Theta}}^*$. Then a second dimer is formed

$$\mathbf{x}_l^{1*}=\mathbf{x}_l+\delta x_{\mathrm{D}}\hat{\mathbf{N}}^*, \tag{13}$$

$$\mathbf{x}_l^{2*}=\mathbf{x}_l-\delta x_{\mathrm{D}}\hat{\mathbf{N}}^*, \tag{14}$$

and the forces $\mathbf{F}_{1*}=-\mathbf{g}(\mathbf{x}_l^{1*})$, $\mathbf{F}_{2*}=2\mathbf{F}_l-\mathbf{F}_{1*}$, and $\mathbf{F}^{*\perp}=(\mathbf{F}_{1*}^{\perp}-\mathbf{F}_{2*}^{\perp})/\delta x_{\mathrm{D}}$ calculated. In the following step the magnitude of the rotational force at $\delta\theta_{\mathrm{D}}/2$ is found through $F=(\mathbf{F}^{*\perp}\cdot\hat{\mathbf{\Theta}}^*+\mathbf{F}^{\perp}\cdot\hat{\mathbf{\Theta}})/2$ and a finite difference approximation to the change in the rotational force (at $\delta\theta_{\mathrm{D}}/2$) is obtained by

$$F'=\frac{\mathbf{F}^{*\perp}\cdot\hat{\mathbf{\Theta}}^*-\mathbf{F}^{\perp}\cdot\hat{\mathbf{\Theta}}}{\delta\theta_{\mathrm{D}}}. \tag{15}$$

Subsequently, the angle with which the *second* dimer now has to be rotated to minimize the dimer energy is given by[17]

$$\Delta\theta=-\frac{1}{2}\arctan\left(\frac{2F}{F'}\right)-\delta\theta_{\mathrm{D}}/2. \tag{16}$$

Finally, if the rotational curvature F' is negative the rotation by $\Delta\theta$ will move the dimer towards a (local) maximum for the dimer energy. In this case adding $\pi/2$ to $\Delta\theta$ ensures that the rotation is towards the minimum. Note that the term $\delta\theta_{\mathrm{D}}/2$ in Eq. (16) is appropriate since the rotational curvature is estimated most accurately by Eq. (15) for the midpoint of the two dimers, using central differencing.

If the dimer method is used to fully converge the lowest eigenmode the procedure can be summarized as follows: (i) The rotational force on the first dimer is calculated. If it is below our chosen convergence criterion, $F<\delta F_{\mathrm{D}}$, the eigenvector corresponding to the lowest eigenvalue is considered converged to $\mathbf{v}_1=\hat{\mathbf{N}}$ and the eigenvalue calculated from $\lambda_1=[(\mathbf{F}_2-\mathbf{F}_1)\cdot\hat{\mathbf{N}}]/2\delta x_{\mathrm{D}}$. (ii) If the rotational force is not below the chosen criterion, the rotational force is obtained for the second dimer, the angle with which the second dimer needs to be rotated is calculated using Eq. (16), and the second dimer is rotated accordingly. The procedure is repeated until the convergence criterion in (i) is met. Note that for each rotation of the dimer, two force evaluations are needed.

The given procedure can be used to fully converge the lowest eigenmode, but as we will see in the following the dimer method needs considerably more force calls to reach convergence than the Lanczos method using this approach. However, by limiting the number of dimer rotations in combination with a less strict convergence criterion for the rotational force, the number of force evaluations can be drastically reduced. We do this by adding to the scheme outline above: (iii) If we have reached the maximum number of allowed dimer rotations $n_{\mathrm{D}}^{\max}$ the eigenvector is then approximated by the normalized vector along the rotated dimer. Otherwise we repeat the procedure starting from (i). The eigenvalue corresponding to the obtained eigenvector is calculated differently depending on in which step the iteration cycle is stopped: If the criterion in (i) is met the eigenvalue can be obtained as in the paragraph above, $\lambda_1=[(\mathbf{F}_2-\mathbf{F}_1)\cdot\hat{\mathbf{N}}]/2\delta x_{\mathrm{D}}$. However, if the iteration cycle is terminated at step (iii), a bit more effort is needed. This is because we would like to avoid performing an extra force evaluation to get a reasonably accurate estimate of the eigenvalue [note that after the rotation of the dimer in step (ii) we have no direct information about the forces on the rotated dimer]. First, a curvature estimate along $\hat{\mathbf{N}}^*$ is obtained through $C=[(\mathbf{F}_2^*-\mathbf{F}_1^*)\cdot\hat{\mathbf{N}}^*]/2\delta x_{\mathrm{D}}$. Then, using the local quadratic approximation as indicated in Ref. 17, the obtained curvature

estimate can be corrected to give a reasonable estimate for the curvature (eigenvalue) along the rotated dimer through

$$\lambda_1 = C - \tfrac{1}{2}\|\mathbf{F}^{*\perp}\|\tan(\Delta\theta - \delta\theta_D/2). \tag{17}$$

It is important to note that using a modified conjugated gradient approach to determine the rotational plane (as described in Ref. 17) improves the performance of the method, especially for large systems. The improvement the conjugated gradient method offers compared to the steepest decent method is due to the former using the force at previous iterations in addition to the force at the current iteration to determine the optimal direction of minimization.[42]

The above outline indicates that there are many parameters which can be adjusted when using the dimer method. However, in practice there are simple strategies which can be used to restrict the parameter space. This will be discussed in more detail in Sec. IV.

C. Determining the geometry step using the minimum mode only

Compared to the effort needed to obtain the lowest eigenpair of the Hessian, calculating the step vector for the minimum mode following methods is relatively straightforward. As noted in Ref. 17, the force at the current configuration will tend to pull the system towards the minimum, but simply inverting the component of the force along the lowest eigenmode will tend to move the system towards a saddle point on the PES. Another useful trick is to force the system to move only along the lowest (approximate) eigenmode in convex regions of the PES, resulting in the system leaving the convex region faster. Both strategies can be combined by using a modified force to determine the direction of the step vector as follows:

$$\mathbf{F}^{\dagger} = \begin{cases} -(\mathbf{F}_l \cdot \mathbf{v}_1)\mathbf{v}_1 & \text{if } \lambda_1 > 0 \\ \mathbf{F}_l - 2(\mathbf{F}_l \cdot \mathbf{v}_1)\mathbf{v}_1 & \text{if } \lambda_1 < 0 \end{cases}, \tag{18}$$

where λ_1 is the lowest eigenvalue and $\mathbf{v}_1$ the corresponding normalized eigenvector of the Hessian at the current position $\mathbf{x}_l$. Next, the length of the step should be determined. Our approach is to do a line search along the direction of the modified force by evaluating the force at $\mathbf{x}_l^* = \mathbf{x}_l + \delta x_{\mathrm{lm}}\mathbf{N}^{\dagger}$, with $\mathbf{N}^{\dagger} = \mathbf{F}^{\dagger}/\|\mathbf{F}^{\dagger}\|$ (the choice of δx_{lm} will be discussed later). Equation (18) is then used to calculate the modified force at $\mathbf{x}_l^*$ and the magnitude of the force and the curvature along the direction of displacement at $(\mathbf{x}_l^* + \mathbf{x}_l)/2$ are given by $F = (\mathbf{F}^{*\dagger} + \mathbf{F}^{\dagger}) \cdot \mathbf{N}^{\dagger}/2$ and $C_{\mathrm{lm}} = (\mathbf{F}^{*\dagger} - \mathbf{F}^{\dagger}) \cdot \mathbf{N}^{\dagger}/\delta x_{\mathrm{lm}}$, respectively. Finally, the step vector can be calculated from $\Delta\mathbf{x}_l = (-F/C_{\mathrm{lm}} + \delta x_{\mathrm{lm}}/2)\mathbf{N}^{\dagger}$. To avoid stepping too far the step length is not allowed to exceed $\Delta x^{\max}$ (to be discussed later), and to ensure that we are leaving the convex region as fast as possible the step length is always $\Delta x^{\max}$ in this region. A conjugated gradient approach to determine the direction of the step vector is used to reduce the number of force calls needed to reach a saddle point.

D. Rational function optimization: Exact Hessian

Based on a Taylor expansion of the PES around a current position and a constraint on the step length, Cerjan and Miller showed (through the use of a Lagrangian multiplier technique)[1] that a modified Newton-Raphson approach could be turned into an efficient PES walker. With steps determined by

$$\Delta x_i = -\frac{g_i}{\lambda_i - \gamma}, \tag{19}$$

where Δx_i and g_i are the components of the step and gradient vectors, respectively, in the Hessian eigenvector basis ($\Delta\mathbf{x} = \Sigma_{i=1}^{n}\Delta x_i\mathbf{v}_i$, $\mathbf{g} = \Sigma_{i=1}^{n} g_i\mathbf{v}_i$, $\mathbf{v}_i$ being the Hessian eigenvectors, and λ_i the Hessian eigenvalues), they outlined how γ could be chosen in order to walk efficiently from close to a minimum to a saddle point. In Ref. 3 a more detailed account on how to choose the optimal γ was given (leading to somewhat different recommendations than in Ref. 1). An important element of the strategy is that, if $\lambda_1 > 0$, γ can be chosen in such a way as to ensure an uphill walk along $\mathbf{v}_1$ and a downhill walk in all other directions. Then, in Ref. 5 it was shown how Eq. (19) can be obtained when a rational function optimization (RFO) approach is used to approximate the PES locally by

$$V(\mathbf{x}_{l+1}) - V(\mathbf{x}_l) = \frac{\mathbf{g}^t \cdot \Delta\mathbf{x}_l + \frac{1}{2}\Delta\mathbf{x}_l^t \cdot \mathbf{H} \cdot \Delta\mathbf{x}_l}{1 + \Delta\mathbf{x}_l^t \cdot \mathbf{S} \cdot \Delta\mathbf{x}_l} \tag{20}$$

and choosing $\mathbf{S} = \gamma\mathbf{1}$. Here we have chosen to work with the slightly more general form of $\mathbf{S}$ (also introduced in Ref. 5) where the components of the step vector are calculated through

$$\Delta x_i = -\frac{g_i}{\lambda_i - \gamma_i}, \tag{21}$$

with $\lambda_i - \gamma_i$ given by

$$\lambda_i - \gamma_i = \tfrac{1}{2} d_i(|\lambda_i| + \sqrt{\lambda_i^2 + 4g_i^2}),$$
$$d_1 = -1,\ d_i = 1 \ \text{ for } \ i = \{2,\ldots,n\}. \tag{22}$$

This defines a walker that will move uphill along the lowest eigenmode of the Hessian and downhill along all other modes. The particular choice for the rescaling of the Hessian eigenvalues has been motivated by the considerations in Refs. 5 and 6. We also ensure that the total length of the step vector $\|\Delta\mathbf{x}_l\| = \|\Sigma_{i=1}^{n}\Delta x_i\mathbf{v}_i\|$ does not exceed $\Delta x^{\max}$. Standard LAPACK (Ref. 45) routines for real symmetric matrices have been used to calculate all eigenvalues and corresponding eigenvectors.

E. Rational function optimization: Approximate Hessian

As already noted in Ref. 3 there is nothing in the procedure described above (Sec. II D) that requires the Hessian matrix to be exact—it can equally well be applied using an approximate Hessian. We have used two different updating schemes

$$\mathbf{H}_{l+1} = \mathbf{H}_l + \Delta\mathbf{H}_l, \tag{23}$$

one due to Powell[46] and the other to Bofill.[10] The Powell update is given by

$$\Delta \mathbf{H}_l^{\text{Powell}} = \frac{(\Delta \mathbf{g}_{l+1} - \mathbf{H}_l \cdot \Delta \mathbf{x}_l) \cdot \Delta \mathbf{x}_l^t + \Delta \mathbf{x}_l \cdot (\Delta \mathbf{g}_{l+1} - \mathbf{H}_l \cdot \Delta \mathbf{x}_l)^t}{\Delta \mathbf{x}_l^t \cdot \Delta \mathbf{x}_l} - (\Delta \mathbf{g}_{l+1} - \mathbf{H}_l \cdot \Delta \mathbf{x}_l)^t \cdot \Delta \mathbf{x}_l \frac{\Delta \mathbf{x}_l \cdot \Delta \mathbf{x}_l^t}{(\Delta \mathbf{x}_l^t \cdot \Delta \mathbf{x}_l)^2} \quad (24)$$

with $\Delta \mathbf{g}_{l+1} = \mathbf{g}(\mathbf{x}_{l+1}) - \mathbf{g}(\mathbf{x}_l)$, and the Bofill update by

$$\Delta \mathbf{H}_l^{\text{Bofill}} = \phi^{\text{Bofill}} \Delta \mathbf{H}_l^{\text{SR1}} + (1 - \phi^{\text{Bofill}}) \Delta \mathbf{H}_l^{\text{Powell}}, \quad (25)$$

where $\Delta \mathbf{H}_l^{\text{SR1}}$ is a symmetric rank one update[47]

$$\Delta \mathbf{H}_l^{\text{SR1}} = \frac{(\Delta \mathbf{g}_{l+1} - \mathbf{H}_l \cdot \Delta \mathbf{x}_l) \cdot (\Delta \mathbf{g}_{l+1} - \mathbf{H}_l \cdot \Delta \mathbf{x}_l)^t}{(\Delta \mathbf{g}_{l+1} - \mathbf{H}_l \cdot \Delta \mathbf{x}_l)^t \cdot \Delta \mathbf{x}_l}, \quad (26)$$

and the Bofill factor is given by

$$\phi^{\text{Bofill}} = \frac{[(\Delta \mathbf{g}_{l+1} - \mathbf{H}_l \cdot \Delta \mathbf{x}_l)^t \cdot \Delta \mathbf{x}_l]^2}{[(\Delta \mathbf{g}_{l+1} - \mathbf{H}_l \cdot \Delta \mathbf{x}_l)^t \cdot (\Delta \mathbf{g}_{l+1} - \mathbf{H}_l \cdot \Delta \mathbf{x}_l)](\Delta \mathbf{x}_l^t \cdot \Delta \mathbf{x}_l)}. \quad (27)$$

F. Rational function optimization combined with a minimum mode following method: A new hybrid approach

The RFO approaches (with an exact or approximate Hessian) can in some cases encounter problems with leaving the convex region. Also, by construction, they will tend to climb out of the convex region by following the lowest streambed. When trying to discover all saddle points around a minimum this tendency can become a weakness compared to the minimum mode following methods. Based on our experience with the minimum mode following methods we propose a new hybrid RFO approach which partly remedies this problem: In the convex region the step vector can be determined as in Sec. II C using only the lowest eigenvalue and the corresponding eigenvector of the Hessian. Thus the hybrid PES walker is a minimum mode following method when the system is located in a convex region of the PES (although the Hessian matrix is updated for each step when using approximate Hessians). Once the lowest eigenvalue of the Hessian becomes negative the traditional RFO approach as described above is used. This approach is similar in spirit to the two methods introduced by Munro and Wales[16] and the ARTn.[19] The main difference is that no optimization in the hyperplane perpendicular to the step direction is performed in the convex region, and the full Hessian matrix is constructed and diagonalized in each step. The method will therefore suffer from the same computational bottleneck as the traditional RFO methods with respect to the diagonalization of the Hessian matrix for large systems, but it will in most of the cases we have considered lead to appreciably more saddle points found.

III. MODEL SYSTEMS AND POTENTIAL ENERGY SURFACES

As in Ref. 28, all the PES walkers have been tested on a model system involving transitions in a seven-atom Pt island on a Pt(111) surface. The pairwise interaction between the atoms is given by the Morse potential

$$V(r) = D_e(e^{-2\alpha(r-r_0)} - 2e^{-\alpha(r-r_0)}) \quad (28)$$

with parameters chosen to reproduce diffusion barriers on Pt surfaces,[48] $D_e = 0.7102$ eV, $\alpha = 1.6047$ Å^{-1}, $r_0 = 2.8970$ Å. The potential is cut and shifted at 9.5 Å. The surface is represented by a six layer slab where each layer contains 56 independent atoms and periodic boundary conditions have been applied.

In order to study how the walkers perform on PESs of different dimensionality, the number of atoms allowed to move have been varied.

(i) In the full system the seven-atom island and the top three layers of the slab are free to move, the bottom three layers being kept frozen. There are 175 atoms free to move, thus the PES is 525-dimensional (525D).

(ii) In a reduced size system only the seven-atom island atoms are allowed to move, whereas all atoms in the underlying slab are kept frozen. This represents a 21D PES.

(iii) In the second reduced size system only one edge atom of the seven-atom island is free to move whereas all other atoms are kept frozen, giving rise to a 3D PES.

In all cases 500 initial configurations have been determined by displacing randomly the atoms of the seven-atom island that are free to move a distance Δx^{ran}. The searches are stopped when all components of the gradient vector falls below the threshold $\delta g_{\max} = 0.001$ eV/Å. In the following section the results are presented in the order of increasing dimensionality of the system, before some overall trends are outlined.

IV. RESULTS AND DISCUSSION

A. One atom free to move

A contour plot of the 2D PES obtained by minimizing the potential energy along the coordinate for motion normal to the surface for each lateral position of the edge atom that is free to move is shown in Fig. 1. There are five saddle points within 4 eV of the starting minimum that are directly connected to this minimum (given by all atoms in the seven-atom island being located in neighboring fcc sites, see Fig. 2). Two of the saddle points (at 1.693 and 1.978 eV above the minimum) are found when the atom is moving around the island keeping as close to the neighboring edge atoms as possible. A third saddle point (2.134 eV) is found when the atom is moving away from the six other island atoms. The fourth and fifth saddle points (3.665 and 3.667 eV) are found when moving the atom on top of the island.

The Lanczos results were obtained with settings shown in Tables I and II, together with the parameter ranges tested and found to give similar results. As seen from the tables and Secs. II A and II C, there are a number of parameters available for tuning. In this study we have made an effort to reach optimal search performance, and this has lead to some recommendations towards standard parameter settings that could be adopted by others. This will be discussed in more detail in Sec. IV D. The search results for the Lanczos method are given in Table II and we see that on average about five force calls per geometry step are needed to reach the saddle points: one is used to calculate the gradient at the

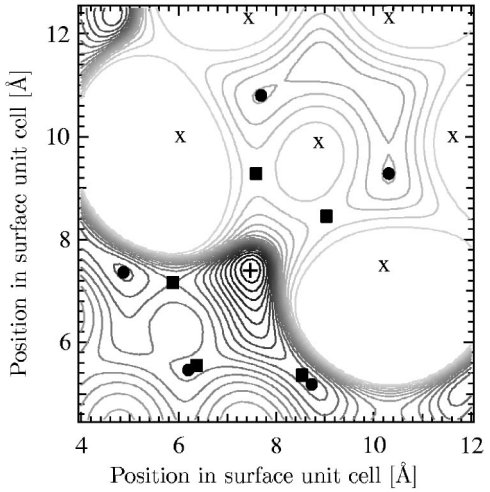

FIG. 1. A contour plot of the PES obtained by choosing the optimal height above the surface for the one atom that is free to move for each position in the surface plane is displayed. The positions of the five saddle points directly connected to the starting minimum are marked by filled squares with the corresponding final states indicated by filled circles. The starting minimum is indicated by +. The six other atoms in the seven-atom island that are kept frozen are indicated by the crosses (see also Fig. 2). The contour spacing is 0.2 eV.

current position, three are used to converge to the lowest Hessian eigenvalue, and the last one is used in the line minimization. We note that when using a maximum step length of 0.1 Å the third saddle point listed in Fig. 2 is not found within the 500 searches performed. This can be understood from looking at Fig. 1: When using a short maximum step length and a fully converged lowest eigenmode (this can be obtained with $\delta\lambda_L=0.01$ or lower), getting into the region which would allow for a convergence on the third saddle point is rather difficult. Using a longer maximum step length solves this problem.

The results for the dimer method were obtained with the settings given in Tables I and II. Parameter ranges tested and found to give similar results are also indicated. As for the Lanczos method, there are a number of parameters that can

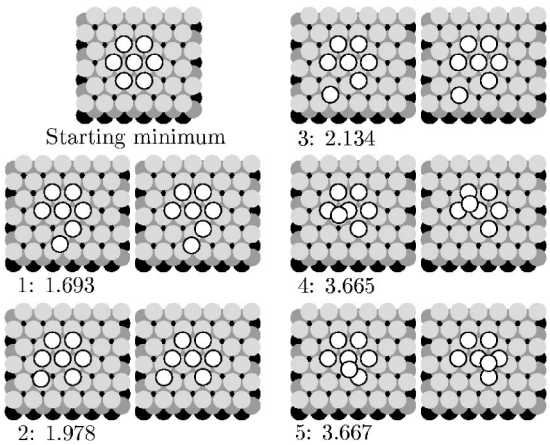

FIG. 2. The initial configuration without the random displacements, and the five saddle points (left) together with the corresponding final configurations (right), are shown for the 3D PES. The saddle point energies with respect to the initial configuration are also indicated (in eV).

TABLE I. The finite difference parameters used for the Lanczos and dimer methods. Also indicated are the parameter ranges tested and found to give similar results to the ones presented in this study.

Lanczos	δx_L (Å)	δx_{lm} (Å)	
Used	10^{-4}	10^{-3}	
Tested	$10^{-6}-10^{-2}$	$10^{-4}-10^{-1}$	
Dimer	δx_D (Å)	$\delta\theta_D$ (rad)	δx_{lm} (Å)
Used	10^{-4}	10^{-4}	10^{-3}
Tested	$10^{-5}-10^{-1}$	$10^{-5}-10^{-1}$	$10^{-4}-10^{-1}$

be tuned to reach optimal performance, and that has been done here. This has lead to some recommendations towards standard settings for others to use (see Sec. IV D). The dimer results are given in Table II and it is seen that the short maximum step length of 0.1 Å results in the third saddle point not being found, if the dimer method is used in the mode that the dimer is rotated until the lowest eigenmode is fully converged (the maximum rotational force allowed needs to be set to $\delta F_D=0.01$ eV/Å or lower to reach convergence). That this behavior is the same as for the Lanczos method is to be expected on the basis that the Lanczos and dimer methods use exactly the same algorithm for calculating the geometry step. The only difference between the two methods is how many force calls are needed to reach a converged lowest eigenmode. From Table II we see that when the dimer is allowed to rotate until the lowest Hessian eigenvalue is fully converged on average about 6.8 force evaluations per geometry step are needed. However, by allowing a maximum of only one dimer rotation per geometry step in combination with a less strict convergence criterion for the rotational force ($\delta F_D=1.0$ eV/Å), the average number of force calls needed is considerably reduced. With these search parameters the dimer method is somewhat more efficient than the Lanczos method with respect to the average number of force calls needed, and all the five saddle points are found. This is the way the dimer method should be employed—using the dimer method to fully converge the lowest Hessian eigenvalue is not a good strategy, as is demonstrated by the results in Table II. For the longer maximum step length $\Delta x^{max}=0.5$ Å good performance was reached with a maximum of one or two dimer rotation and $\delta F_D=1.0$ eV/Å. As for the shorter maximum step length the dimer method is somewhat more efficient than the Lanczos method with respect to the average number of force calls needed.

The results of the RFO approach using the Bofill and Powell updaters are very similar, and therefore only the Bofill results have been included in Table II. They show that there is a significant decrease in the average number of force evaluations needed to find the saddle points when increasing the maximum step length from 0.1 to 0.2 Å. It is also seen that increasing the step length beyond 0.2 Å leads to additional reduction in the number of force calls, but at the same time a rather large number of unwanted search results is produced. Somewhat surprisingly, it does not matter very much whether the RFO search using an approximate (updated) Hessian is started with an exact initial Hessian or a unit matrix.

TABLE II. The overall performance of the different PES walkers is shown for the case where one edge atom of the seven-atom island is free to move. $\Delta x^{\max}$ is the maximum step length allowed. n^{TS} is the number of saddle points found within 4 eV of the starting minimum and directly connected to it; there are five of them. ν indicates how many of the searches found one of the n^{TS} saddle points (out of 500 searches performed). $\langle f \rangle$ and $\langle s \rangle$ are the average number of force calls and geometry steps needed, respectively. $\tilde{\nu}$ indicates how many searches found saddle points within 4 eV of the starting minimum and not directly connected to it. $\delta\lambda_{\mathrm{L}}$ is the Lanczos eigenvalue convergence criterion, $n_{\mathrm{D}}^{\max}$ is the maximum number of dimer rotations, and δF_{D} is the convergence criterion used for the rotational force. For the approximate RFO methods searches have been started both with a unit matrix and an exact Hessian. All results are for a random initial displacement of $\Delta x^{\mathrm{ran}}=0.1$ Å.

Method	$\Delta x^{\max}$ (Å)	n^{TS}	ν	$\langle f \rangle$	$\langle s \rangle$	$\tilde{\nu}$
Lanczos (Secs. II A, II C), $\delta\lambda_{\mathrm{L}}=0.01$	0.1	4	458	144.9	29.8	0
$\delta\lambda_{\mathrm{L}}=0.01$	0.5	5	479	75.7	15.9	0
Dimer (Secs. II B, II C), $\delta F_{\mathrm{D}}=0.01$ eV/Å	0.1	4	458	203.6	29.8	0
$n_{\mathrm{D}}^{\max}=1$ and $\delta F_{\mathrm{D}}=1.0$ eV/Å	0.1	5	441	112.1	32.8	3
$n_{\mathrm{D}}^{\max}=1$ and $\delta F_{\mathrm{D}}=1.0$ eV/Å	0.5	5	323	60.5	17.6	119
$n_{\mathrm{D}}^{\max}=2$ and $\delta F_{\mathrm{D}}=1.0$ eV/Å	0.5	5	440	70.4	17.2	22
RFO (exact, Sec. II D)	0.1	5	500	25.9	25.9	0
	0.5	5	500	10.2	10.2	0
Hybrid RFO (exact, Secs. II D, II F)	0.1	5	458	25.0	25.0	0
	0.5	5	482	8.5	8.5	0
RFO (Bofill, Sec. II E)	0.1	5	498	30.2	30.2	2
Exact initial **H**	0.1	5	497	30.0	30.0	3
	0.2	5	487	21.2	21.2	11
Exact initial **H**	0.2	5	491	19.9	19.9	9
	0.5	5	292	18.5	18.5	146
Exact initial **H**	0.5	5	340	19.2	19.2	140
Hybrid RFO (Bofill, Secs. II E, II F)	0.1	5	427	35.2	35.2	7
Exact initial **H**	0.1	5	487	30.1	30.1	0
	0.2	5	298	23.3	23.3	77
Exact initial **H**	0.2	5	450	20.4	20.4	40
	0.5	5	203	19.1	19.1	101
Exact initial **H**	0.5	5	178	15.4	15.4	212

When limiting the maximum step length to 0.1 Å it is seen from Table II that the number of geometry steps needed to reach the saddle points is quite similar for all the different methods. But in terms of the average number of force calls needed the minimum mode following methods perform substantially worse than the RFO approaches. The reason is that for the minimum mode following methods a number of force calls are needed per geometry step to (partially) converge the lowest eigenmode of the Hessian and perform a line minimization step, whereas the RFO approaches only need one force call per geometry step. We also note that the approximate RFO scheme is almost as good as the exact one for a short maximum step length, indicating that the updating formulas result in a rather good approximation to the full Hessian matrix.

Increasing the maximum allowed step length reduces the average number of geometry steps and force calls needed to reach the saddle points for all methods tested, as seen from Table II. However, the reduction is the strongest for the RFO method with an exact Hessian. The higher reduction factor for the RFO method with an exact Hessian (about 2.5 when increasing the maximum step length from 0.1 to 0.5 Å) compared to the minimum mode following methods (about 1.9) indicates that the knowledge of all eigenmodes versus only the lowest eigenmode helps considerably in reducing the number of force calls needed to reach the saddle points. Furthermore, the results in the table suggest that the local, second order approximation to the 3D PES when making a step is rather good even for a maximum allowed step length of 0.5 Å—that the approximate RFO schemes lead to a large number of unwanted search results is due to the not sufficiently accurate approximation of the Hessian eigenmodes in these schemes.

The performance of the hybrid methods, both with exact and approximate Hessians, has been checked and found to be quite similar to the original ones with respect to the average number of force calls (Table II). For the hybrid RFO approach with an exact Hessian the number of force calls needed is reduced slightly compared to the RFO approach with an exact Hessian. For the hybrid RFO approach with approximate Hessians starting the searches with a unit matrix results in a slight increase in the number of force calls as compare to the RFO approach with approximate Hessians starting with a unit matrix. The number of force calls is about the same when comparing the traditional and hybrid approximate RFO schemes with an exact initial Hessian, except for $\Delta x^{\mathrm{ran}}=0.5$ Å where there is a decrease with the hybrid scheme. The hybrid approximate RFO schemes with a maximum step length of 0.5 Å both result in more than a half of the searches leading to unwanted results.

All results discussed above have been obtained with $\Delta x^{\mathrm{ran}}=0.1$ Å (the distance with which the initial configuration is randomly displaced away from the starting minimum). Tests varying Δx^{ran} in the interval 0.1–0.3 Å show that the average number of force calls needed to reach the different saddle points hardly changes, except for a slight increase for

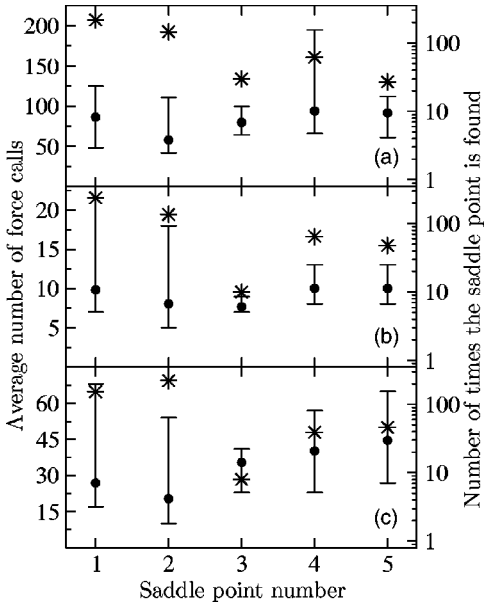

FIG. 3. In the figure the average number of force calls (filled circles) needed to find the saddle point (the saddle point corresponding to the saddle point number is indicated in Fig. 2) is shown together with the maximum and minimum number of force calls ("error" bars) for the 3D PES (left axis is used). The asterixes indicate how often the saddle points were found (out of the 500 searches performed, right axis is used). Note the logarithmic scale on the right axes. In (a) the results for the Lanczos method are displayed with $\Delta x^{\max}=0.5$ Å and $\Delta x^{\mathrm{ran}}=0.1$ Å, in (b) for the RFO method employing an exact Hessian with $\Delta x^{\max}=0.5$ Å and $\Delta x^{\mathrm{ran}}=0.3$ Å, and in (c) for the RFO method with a Bofill update starting with a unit matrix, $\Delta x^{\max}=0.15$ Å, and $\Delta x^{\mathrm{ran}}=0.3$ Å.

the approximate RFO schemes (results not shown). However, the choice of Δx^{ran} do to some extent influence the number of times the different saddle points are found. For all methods except Lanczos the number of times the third, fourth, and fifth saddle points are found increases with increasing Δx^{ran}, together with a decrease in the number of times the first and second saddle points are found. For the Lanczos method the third saddle point is found most often with $\Delta x^{\mathrm{ran}}=0.1$ Å. From the results presented in Fig. 3, we see that the Lanczos method performs somewhat better than the other methods with respect to the number of times the most difficult to locate saddle point (the third one) is found. Another way of increasing the number of times the third, fourth, and fifth saddle points are found for the RFO approaches is to use the hybrid version, but the Lanczos method still remains the more efficient at finding the most difficult to locate saddle point.

B. Seven atoms free to move

We have found more than 250 different saddle points within 4 eV of the starting minimum and directly connected to it. Even though they can be grouped together in classes describing similar transition mechanisms, there are more than 25 of these classes. Since our goal is to compare how the different saddle point search methods perform, and not to provide a full analysis of possible transitions mechanisms for our model system, we will not enter into further detail on this point.

The Lanczos and dimer results (Table III) were obtained after a considerable effort searching the large parameter space for optimal performance. The resulting settings are indicated in Tables I and III. As a results of the effort to reach optimal search performance, a set of standard parameter settings was developed that could be adopted by others. This will be discussed in more detail in Sec. IV D. When fully converging the lowest Hessian eigenvalue the Lanczos method needs on average about 8.7 force calls per geometry step for $\Delta x^{\max}=0.1$ Å ($\delta\lambda_{\mathrm{L}}$ must be 0.0001 or lower to reach convergence), whereas the dimer method requires about 16.4 force calls (the maximum rotational force allowed needs to be set to $\delta F_{\mathrm{D}}=0.01$ eV/Å or lower to obtain converged results). This clearly indicates that the Lanczos method is better in converging the lowest eigenmode than the dimer method for a given geometry. However, fully converging the lowest Hessian eigenvalue at each geometry step is not a good strategy when employing the dimer or Lanczos method. Optimal performance for the dimer method for $\Delta x^{\max}=0.1$ Å is obtained when limiting the number of rotations allowed for each geometry step ($n_{\mathrm{D}}^{\max}=1$) in combination with a less strict convergence criterion for the rotational force ($\delta F_{\mathrm{D}}=0.1$ eV/Å). As seen from Table III this leads to a larger number of saddle points being found, and a marked decrease in the average number of force calls needed per saddle point found that is directly connected to the starting minimum (if the goal is to locate as many saddle points directly connect to the starting minimum as possible, this is an important measure of the success of a series of searches). The performance of the Lanczos method for $\Delta x^{\max}=0.1$ Å can also be improved by limiting the number of Lanczos iterations for each geometry step ($n_{\mathrm{L}}^{\max}=3$) and using a less strict criterion for the Lanczos eigenvalue convergence ($\delta\lambda_{\mathrm{L}}=1.0$), even though the effect is smaller than the effect of limiting the number of rotations for the dimer method in conjunction with a less strict convergence criterion for the rotational force. From Table III we see that increasing the maximum step length $\Delta x^{\max}$ is an efficient way of increasing the number of saddle points found and decreasing the average number of force calls needed per saddle point found that is directly connected to the starting minimum. Starting the searches further away from the minimum (by increasing Δx^{ran}) also improves the efficiency of the minimum mode following methods for finding as many saddle points as possible. In this way, more saddle points are found with the minimum mode following methods than with the RFO methods. The results of Table III indicate that the two minimum mode following methods are rather similar both with respect to the average number of force calls needed to reach the saddle points and the number of saddle points found directly connected to the initial minimum.

The RFO approach with an exact Hessian is from Table III seen to be very efficient compared to the minimum mode following methods. For the same maximum step length considerably less geometry steps and force calls are needed to reach the saddle points. As in Sec. IV A, this indicates that the knowledge of all eigenmodes instead of the lowest eigenmode only helps considerably in reducing the number of geometry steps needed to reach the saddle points. However,

TABLE III. The overall performance of the different walkers for the 21D PES. n_L^{max} is the maximum on the number of Lanczos iterations and $\delta\lambda_L$ is the Lanczos eigenvalue convergence criterion. n_D^{max} is the maximum number of dimer rotations and δF_D is the convergence criterion used for the rotational force. $\langle f^{TS}\rangle$ is the average number of force calls needed per saddle point found that is directly connected to the starting minimum. The rest of the nomenclature is the same as in Table II. Each set of results is based on 500 searches. There are more than 250 saddle points within 4 eV of the starting minimum and directly connected to it.

Method	Δx^{max} (Å)	Δx^{ran} (Å)	n^{TS}	ν	$\langle f\rangle$	$\langle f^{TS}\rangle$	$\langle s\rangle$	$\tilde{\nu}$
Lanczos (Secs. II A, II C), $\delta\lambda_L=0.0001$	0.1	0.1	20	497	435.8	10375.9	50.0	1
$n_L^{max}=3$ and $\delta\lambda_L=1.0$	0.1	0.1	42	486	266.8	3175.9	54.2	14
$n_L^{max}=3$ and $\delta\lambda_L=1.0$	0.4	0.1	84	344	190.4	1133.5	38.9	150
$n_L^{max}=3$ and $\delta\lambda_L=1.0$	0.4	0.3	94	342	190.7	1014.2	38.9	122
Dimer (Secs. II B, II C), $\delta F_D=0.01$ eV/Å	0.1	0.1	20	497	820.5	20512.4	50.0	1
$n_D^{max}=1$ and $\delta F_D=0.1$ eV/Å	0.1	0.1	48	484	218.8	2279.4	57.6	14
$n_D^{max}=1$ and $\delta F_D=0.1$ eV/Å	0.3	0.1	84	249	164.8	960.9	44.0	221
$n_D^{max}=2$ and $\delta F_D=0.1$ eV/Å	0.45	0.1	82	335	197.3	1203.3	39.4	157
$n_D^{max}=1$ and $\delta F_D=0.1$ eV/Å	0.3	0.3	95	259	176.8	930.7	46.9	164
RFO (exact, Sec. II D)	0.1	0.1	17	499	32.8	964.8	32.8	0
	0.5	0.1	17	499	11.4	335.2	11.4	0
	0.5	0.3	59	470	11.9	100.7	11.9	28
Hybrid RFO (exact, Secs. II D, II F)	0.1	0.1	32	495	33.6	525.3	33.6	4
	0.5	0.1	40	493	11.1	139.1	11.1	7
	0.5	0.3	68	458	12.1	88.6	12.1	50
RFO (Bofill, Sec. II E)	0.1	0.1	6	499	60.6	5047.6	60.6	0
Exact initial **H**	0.1	0.1	16	497	49.6	1548.8	49.6	1
	0.2	0.1	10	500	47.8	2389.5	47.8	0
Exact initial **H**	0.2	0.1	13	497	36.8	1416.4	36.8	3
Exact initial **H**	0.2	0.3	60	402	73.9	616.1	73.9	94
Hybrid RFO (Bofill, Secs. II E, II F)	0.1	0.1	10	500	61.0	3049.9	61.0	0
Exact initial **H**	0.1	0.1	33	484	50.0	757.4	50.0	16
	0.2	0.1	10	492	50.1	2503.2	50.1	8
Exact initial **H**	0.2	0.1	35	470	39.1	558.1	39.1	28
Exact initial **H**	0.2	0.3	65	289	71.9	552.9	71.9	183

the number of saddle points found is smaller than for the minimum mode following methods. This can to some extent be improved by employing the hybrid RFO with an exact Hessian (note that the hybrid version is just as efficient as the traditional RFO approach employing an exact Hessian with respect to the average number of force calls). An even more efficient way of increasing the number of saddle points found is to increase Δx^{ran} (the distance with which the initial configuration is randomly displaced away from the starting minimum), but the number of saddle points found remains lower for the RFO approach with an exact Hessian than for the minimum mode following methods. However, the average number of force calls needed per saddle point found that is directly connected to the starting minimum is considerably lower than for the minimum mode following methods.

The RFO approach employing a Bofill update of the Hessian performs very well compared to the minimum mode following methods with respect to the average number of force calls needed, as is seen from Table III. In the cases where an exact initial Hessian is used the average number of geometry steps needed to reach the saddle points is slightly smaller than for the minimum mode following methods. If the search is started with a unit Hessian matrix the number of geometry steps needed does increase, but due to the use of only one force call per step, the RFO approach with a Bofill update requires on average considerably less force calls then the minimum mode following methods. However, as for the RFO approach with an exact Hessian, relative few saddle points are found when employing the RFO approach with a Bofill update and $\Delta x^{ran}=0.1$ Å. But from considering the average number of force calls needed per saddle point found that is directly connected to the starting minimum, we see that in all but one case a considerable improvement can be obtained when employing the hybrid version of the approximate RFO approach instead of the traditional one. Furthermore, increasing Δx^{ran} to 0.3 Å leads to a marked improvement for the traditional RFO approach employing a Bofill update, whereas the further improvement is marginal in this case of the hybrid version. From these results we also see that the approximate RFO approaches are better than the minimum mode following methods with respect to average number of force calls needed per saddle point found that is directly connected to the starting minimum only in the cases where the searches are started with a good initial Hessian.

The results for the RFO approach employing a Powell update are quite similar to those for the Bofill update and therefore not included in Table III. When the searches are started with a unit matrix the Powell update needs on average about ten force calls more than the Bofill update (for all Δx^{max}). Starting with an exact Hessian matrix the difference is smaller, the Powell update needing on average about two force calls more than the Bofill update. This indicates that

TABLE IV. The overall performance of the different walkers for the 525D PES. The nomenclature is the same as in Tables II and III. Each set of results is based on 500 searches. There are at least 170 saddle points within 4 eV of the starting minimum and directly connected to it.

Method	$\Delta x^{\max}$ (Å)	Δx^{ran} (Å)	n^{TS}	ν	$\langle f \rangle$	$\langle f^{\mathrm{TS}} \rangle$	$\langle s \rangle$	$\tilde{\nu}$
Lanczos (Secs. II A, II C) $n_{\mathrm{L}}^{\max}=3$ and $\delta\lambda_{\mathrm{L}}=1.0$	0.1	0.1	40	365	514.0	6424.8	103.6	132
$n_{\mathrm{L}}^{\max}=4$ and $\delta\lambda_{\mathrm{L}}=0.1$	0.5	0.1	69	214	378.3	2741.3	73.8	281
$n_{\mathrm{L}}^{\max}=4$ and $\delta\lambda_{\mathrm{L}}=0.1$	0.5	0.3	86	210	372.1	2163.4	72.7	265
Dimer (Secs. II B, II C) $n_{\mathrm{D}}^{\max}=1$ and $\delta F_{\mathrm{D}}=0.1$ eV/Å	0.1	0.1	52	375	405.4	3898.4	105.9	122
$n_{\mathrm{D}}^{\max}=1$ and $\delta F_{\mathrm{D}}=0.1$ eV/Å	0.3	0.1	60	145	314.4	2620.1	83.0	335
$n_{\mathrm{D}}^{\max}=2$ and $\delta F_{\mathrm{D}}=0.1$ eV/Å	0.45	0.1	66	196	360.7	2732.8	71.2	291
$n_{\mathrm{D}}^{\max}=1$ and $\delta F_{\mathrm{D}}=0.1$ eV/Å	0.15	0.3	78	276	335.1	2148.3	87.8	222
RFO (exact, Sec. II D)	0.1	0.1	11	492	40.9	1859.1	40.9	0
	0.5	0.1	10	488	15.0	750.0	15.0	0
	0.5	0.3	45	413	19.1	212.2	19.1	57
Hybrid RFO (exact, Secs. II D, II F)	0.1	0.1	18	500	36.3	1008.3	36.3	0
	0.5	0.1	51	476	14.9	146.1	14.9	24
	0.5	0.3	58	411	17.1	147.4	17.1	87
RFO (Bofill, Sec. II E)	0.1	0.1	6	449	1420.9	118408.3	1420.9	17
Exact initial **H**	0.1	0.1	5	494	311.3	31130.0	311.3	4
	0.2	0.1	16	353	1286.6	40206.3	1286.6	124
Exact initial **H**	0.2	0.1	10	432	342.6	17130.0	342.6	63
Exact initial **H**	0.1	0.3	27	474	539.2	9985.2	539.2	20
Hybrid RFO (Bofill, Secs. II E, II F)	0.1	0.1	9	441	1351.1	75005.6	1351.1	44
Exact initial **H**	0.1	0.1	11	481	282.6	12845.5	282.6	17
	0.2	0.1	17	294	1224.3	36008.8	1224.3	154
Exact initial **H**	0.2	0.1	15	380	291.4	9713.3	291.4	110
Exact initial **H**	0.1	0.3	30	406	465.2	7753.3	465.2	84

the Bofill updater is better in building an approximate Hessian from a poor starting situation than the Powell updater, whereas they are almost equally good in preserving a good approximation to the Hessian for this particular system.

C. 175 atoms free to move

We have found more than 170 different saddle points within 4 eV of the starting minimum and directly connected to it, belonging to more than 20 different classes describing similar transition mechanism. For the same reason as in the preceding section, we will not enter into further detail on this point.

The results of the searches with the Lanczos and dimer methods for the 525D PES are given in Table IV. As for the 3D and 21D PESs considerable effort went into finding parameters giving optimal performance (the parameters used are given in Tables I and IV). But, also as in the previous cases, close to optimal performance can be reached with standard settings (see Sec. IV D). Using the dimer method to fully converge the lowest Hessian eigenvalue at each geometry step is very inefficient (results not shown), and even though the Lanczos method performs considerably better in this respect (results also not shown), this is a strategy that should not be followed. Optimal performance for the dimer method is reached by limiting the number of rotations at each geometry step ($n_{\mathrm{D}}^{\max}=1$ or 2) in combination with a less strict convergence criterion for the rotational force ($\delta F_{\mathrm{D}}=0.1$ eV/Å). This also results in more saddle points being found and a large decrease in the number of force calls needed per saddle point found that is directly connected to the starting minimum (even though the number of unwanted results of the searches do increase). For the Lanczos method the lowest average number of force calls per saddle point found that is directly connected to the starting minimum is obtained when restricting the maximum number of Lanczos iterations per geometry step ($n_{\mathrm{L}}^{\max}=4$) in conjunction with a less strict criterion for the Lanczos eigenvalue convergence ($\delta\lambda_{\mathrm{L}}=0.1$). As also seen in the two preceding sections the two minimum mode following methods are rather similar both with respect to the average number of force calls needed to reach the saddle points and the number of saddle points found directly connected to the initial minimum.

Table IV shows that the average number of geometry steps taken to reach a saddle point by the RFO approach employing an exact Hessian is considerably smaller than for the minimum mode following methods. It is worth noting that the increase in the average number of steps taken is rather small when increasing the dimension of the PES from 21 to 525. This indicates that accurate Hessian information for all degrees of freedom can be used very efficiently by a PES walker. Employing the hybrid version of the RFO approach with an exact Hessian gives a marked improvement above the traditional RFO approach with an exact Hessian with respect to the number of saddle points found, resulting in less force calls needed per saddle point found that is directly connected to the starting minimum. Another way of increasing the efficiency for the traditional RFO approach with an exact Hessian in finding as many saddle points as possible is to increase Δx^{ran}, whereas in the case of the hybrid version no further gain in efficiency is achieved.

TABLE V. Parameter choices that will result in an efficient use of the minimum mode following methods. The parameters which are affecting the efficiency the most are listed first. For the finite difference parameters the smaller value should be used in conjunction with empirical (analytical) potentials and the larger value when the force evaluations contain numerical noise (see text). For small or medium sized systems a maximum of three Lanczos iterations should be employed, whereas a maximum of four Lanczos iterations is more appropriate for large systems.

Lanczos	$\Delta x^{\max}$ (Å)	Δx^{ran} (Å)	$n_{\mathrm{L}}^{\max}$	$\delta\lambda_{\mathrm{L}}$	δx_{L} (Å)	δx_{lm} (Å)	
	0.1–0.5	0.1–0.3	3, 4	0.1	0.001, 0.01	0.001, 0.01	
Dimer	$\Delta x^{\max}$ (Å)	Δx^{ran} (Å)	$n_{\mathrm{D}}^{\max}$	δF_{D} (eV/Å)	δx_{D} (Å)	$\delta\theta_{\mathrm{D}}$ (rad)	δx_{lm} (Å)
	0.1–0.5	0.1–0.3	1	0.1	0.001, 0.01	0.001, 0.01	0.001, 0.01

The RFO approach with a Bofill update started with an exact Hessian requires on average less force calls to reach a saddle point than the minimum mode following methods for $\Delta x^{\mathrm{ran}}=0.1$ Å, but at the same time we see from Table IV that the average number of geometry steps needed is considerably higher. Also, the number of saddle points found is low, and the number of force calls needed per saddle point found that is directly connected to the starting minimum is much higher than for the minimum mode following methods. If the searches are started with a unit matrix the RFO approach with an approximate Hessian performs considerably worse than the minimum mode following methods: Not only does it take a lot more force calls to home in on a saddle point, the number of force calls needed per saddle point found that is directly connected to the starting minimum is very high. Employing the hybrid version of the RFO approach with a Bofill update improves all aspects of the search performance compared to the traditional RFO approach with a Bofill update, but the minimum mode methods remain superior with respect to the number of force calls needed per saddle point found that is directly connected to the starting minimum. Furthermore, increasing Δx^{ran} does lower the number of force calls needed per saddle point found that is directly connected to the starting minimum, but not enough to compete with the minimum mode methods, and the advantage above the minimum mode methods with respect to the average number of force calls needed is lost. Thus when the dimension of the system increases the Bofill updater is struggling harder to buildup a reasonable approximation to the Hessian. Even when provided with an excellent (in this case exact) starting Hessian the update formulas are struggling to keep up with the changes in the Hessian as the system moves towards the saddle points.

For this 525D PES the Powell updater performs considerably worse than the Bofill updater—starting with an exact Hessian the Powell updater requires 200–250 more force calls than the Bofill updater and starting with a unit matrix it requires 350–500 force calls more (results not shown here).

D. Emerging guidelines for optimal searches

From Secs. II, IV A, IV B, and IV C it is clear that the number of tunable parameters is larger for the two minimum mode following methods than for the (hybrid) RFO methods. However, in most cases it will not be necessary to perform a full optimization with respect to all parameters. Through our experience with the current systems (but also other systems) we have developed a set of parameter settings that for most purposes will result in efficient saddle point searching. These parameters are displayed in Table V, where they have been listed in the order of importance with respect to the efficiency of the searches.

The easiest parameters to set are the finite difference parameters (δx_{L}, δx_{D}, $\delta\theta_{\mathrm{D}}$, and δx_{lm}). Optimization of these parameters beyond the values given in Table V will not result in appreciable changes in the outcome of the searches. In our experience search performance varies in the range $\pm 10\%$ for the parameter ranges displayed in Table I, and only those that have a deeper interest in the methodology of the Lanczos or dimer method might want to explore other settings than given in Table V. When using empirical (analytical) potentials, the finite difference distances (angles) can be set to the small value 0.001 Å (radian). If the calculated forces contain numerical noise, for example, when employing electronic structure calculation based on density functional theory, the distances (angles) should be chosen larger [0.01 Å (radian)]. Note that in the latter case the calculated forces may have to be converged more accurately than normally done for minimizations, because the minimum mode following methods calculate curvatures through finite force differences.

The minimum mode following methods are most efficient if the lowest Hessian eigenvector is not fully converged at the beginning of a search—there is little to be gained by finding the lowest curvature mode very accurately at a point in space far from a saddle point. In the convex region around minima, where all Hessian eigenvalues are positive, the minimum mode finding iterations should certainly not be highly optimized. Finding the lowest mode accurately in this region will typically result in a low energy, delocalized breathing mode (when considering high-dimensional systems). Once such a delocalized mode is found, the minimum mode following methods will rarely be able to find a saddle point. It is a far better strategy to ensure that the lowest Hessian eigenvector is gradually converged towards the lowest curvature mode as the system is moved from the starting configuration towards the saddle point. Actually, this gradual convergence is exactly what is accomplished by allowing only a few Hessian eigenvector optimization iterations per geometry step.

Varying the parameters that control the minimum mode optimization ($n_{\mathrm{L}}^{\max}$, $\delta\lambda_{\mathrm{L}}$, $n_{\mathrm{D}}^{\max}$, and δF_{D}) relative to the parameter values indicated in Table V may further enhance the

efficiency of the searches. However, the increased efficiency beyond that which can be achieved with those parameter values, may not warrant the extra effort spent in detailed optimization of these parameters. If such an optimization is undertaken, one should note that when imposing a maximum on the number Lanczos iterations $n_{\mathrm{L}}^{\max}$, a simultaneous adjustment of $\delta\lambda_{\mathrm{L}}$ can serve to reduce the number of actual iterations performed (when the lowest Hessian eigenvalue is converging fast enough to meet this criterion). The relation between the maximum number of dimer rotations $n_{\mathrm{D}}^{\max}$ and the maximum allowed rotational force δF_{D} is similar. The actual number of dimer rotations can be reduced if the maximum rotational force criterion is met before the maximum number of rotations have been performed. Increasing $n_{\mathrm{L}}^{\max}$ ($n_{\mathrm{D}}^{\max}$) and/or decreasing $\delta\lambda_{\mathrm{L}}$ (δF_{D}), while keeping all other parameters fixed, will tend to lead to better streambed following and therefore more conservative searches, i.e., often the lowest lying saddle points will be found.

A proper choice of $\Delta x^{\max}$ is important for the Lanczos and dimer methods, as it also is for the RFO approaches. It can be used to tune the efficiency of the searches in terms of the number of force calls needed and to optimize the variety of saddle points found. The optimal strategy will be different for different applications. If the lowest saddle point is desired, or only a few searches can be run because of computer limitations, a conservative setting is preferred. If many saddles are desired and many searches can be run, an aggressive, large value of the maximum step length will be better. This is demonstrated in Fig. 4 by displaying the performance of 500 Lanczos and dimer searches for varying maximum step lengths for the 525D system, employing the parameter settings given in Table V. For a small step length, few saddle points are found [Fig. 4(a)], but most of them are directly connected to the starting minimum. As the maximum length is increased the searches become more aggressive. Each search requires fewer force calls [Fig. 4(b)], reaching a minimum of 376.5 (303.1) at $\Delta x^{\max}=0.45$ (0.45) Å for the Lanczos (dimer) method. At the same time, more saddle points are found so that the number of force calls per saddle point which is directly connected to the starting minimum also drops, reaching a minimum of 2751.3 (2723.2) at $\Delta x^{\max}=0.35$ (0.30) Å [Fig. 4(c)]. For very aggressive searches with the maximum step length greater than 0.45 (0.45) Å the cost per search increases [Fig. 4(b)] due to a tendency for some searches to take many iterations at high energy before converging. The number of saddle points found, not necessarily directly connected to the starting minimum, increases with increasing maximum step length, reaching 286 (368) out of 500 searches at $\Delta x^{\max}=0.5$ (0.35) Å [Fig. 4(d)]. This large range of different saddle points can be valuable for some applications, but if one is interested in as many saddles points which are directly connected to the starting minimum as possible, a maximum step length of 0.35 (0.30) Å is optimal [Fig. 4(c)]. Note that from a comparison of results from Fig. 4 and Table IV, it is seen that little is lost by employing the standard settings from Table V as compared to the more fully optimized parameters used in Sec. IV C.

The parameter Δx^{ran} (the distance with which the system is randomly displaced away from the starting minimum) also

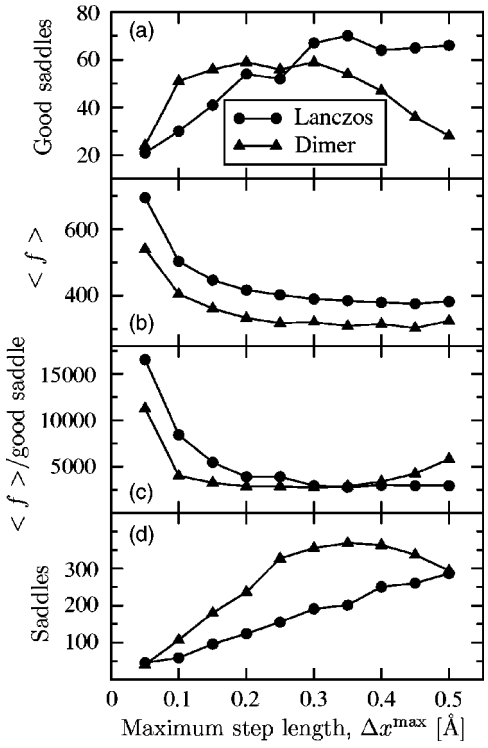

FIG. 4. Results of 500 Lanczos and dimer searches for varying maximum step lengths $\Delta x^{\max}$ for the 525D system, employing the parameter settings given in Table V. (a) The number of saddle points found within 4 eV of the starting minimum and directly connected to it. (b) The average number of force calls needed to reach them. (c) The average number of force calls per saddle point found within 4 eV of the starting minimum and directly connected to it. (d) The total number of saddle points found, including those that are not directly connected to the starting minimum.

influences the outcome of the searches. It is seen from Tables III and IV that similar considerations apply as for $\Delta x^{\max}$. If only the lowest saddle point is desired, or the number of searches need to be limited due to computational limitations, a small value (0.1 Å) is preferred. However, if one attempts to make a map of all relevant saddle points and many searches can be run, a larger value (0.2–0.3 Å) will in most cases be better.

It is important that a saddle point search method can be run with minimal adjustments of parameters. From the above considerations it is clear that for most applications only two parameters ($\Delta x^{\max}$,Δx^{ran}) need to be considered when employing the Lanczos or dimer method. These are the same two parameters that need to be optimized when employing the (hybrid) RFO methods, and one should therefore consider the minimum mode following methods as easy in use as the (hybrid) RFO methods. Thus, which method to choose can be decided purely on how well the different methods perform in a given situation.

The results presented above and in the three Secs. IV A, IV B, and IV C show that the Lanczos and dimer methods are rather similar both with respect to the average number of force calls needed to reach the saddle points and the number of saddle points found directly connected to the initial minimum. Also when considering the total computational effort they perform similarly, as seen from Fig. 5.

For the low-dimensional system there is little difference

TABLE VI. In the table the preferred method for optimal saddle point searches is indicated for different system sizes. It depends on whether the force is cheap to calculate (e.g., if the force is obtained analytically from a simple model potential as in the present case) or expensive to calculate (e.g., if it is obtained from an electronic structure calculation at the density functional theory or at higher levels). The preferred method also depends on whether the goal of the study is to find only the few lowest saddle points or a more complete mapping of all saddle points is intended. The guidelines are given based on a combination of the considerations presented in Sec. IV D. The word hybrid within parentheses indicates that both the standard and hybrid versions of the RFO method could be used. For the 21D and 525D systems, it is essential to used a good estimate of the initial Hessian for the (hybrid) RFO approach with a Bofill update. Otherwise the minimum mode following methods will be the preferred choice.

	Cheap force		Expensive force	
System size	Lowest saddle points	All saddle points	Lowest saddle points	All saddle points
3D	(Hybrid) RFO exact	Minimum mode or (Hybrid) RFO exact	(Hybrid) RFO Bofill	Minimum mode or (Hybrid) RFO Bofill
21D	(Hybrid) RFO exact	(Hybrid) RFO exact	(Hybrid) RFO Bofill	Minimum mode or (Hybrid) RFO Bofill
525D	Minimum mode	Minimum mode	Minimum mode or (Hybrid) RFO Bofill	Minimum mode

between the RFO methods using the two approximate Hessians, the Bofill and Powell updaters. But when increasing the dimension of the system the Bofill updater clearly performs better than the Powell updater. This holds for searches started with an exact initial Hessian as well as with a unit matrix. Therefore the RFO approach with a Bofill updater will be considered as the representative RFO method with an approximate Hessian in the following comparisons.

In determining the most efficient saddle point search method for the 525D system, a number of considerations must be made, and they lead to guidelines for optimal search methods as summarized in Table VI. When working with a rather simple model potential where the force can be calculated cheaply (through analytical derivatives, as in the present study), the cost of the full diagonalization of the Hessian matrix renders the (hybrid) RFO approaches far too computationally demanding compared to the minimum mode following methods (Fig. 5). However, if calculating the force becomes much more expensive, e.g., when it is obtained from an electronic structure calculation at the density functional theory or higher levels, the diagonalization may no longer be the dominant part. But since an increasing cost of calculating the force also implies that the calculation of the exact Hessian would become very demanding, the (hybrid) RFO approach employing an exact Hessian would still not be an alternative to the minimum mode following methods. This argument does not apply to the approximate (hybrid) RFO scheme, where the computational cost will be dominated by the force evaluations. From Table IV we see that the (hybrid) RFO approach with a Bofill update of the approximate Hessian starting from an exact initial Hessian locates the saddle points using on average less force calls than the minimum mode following methods, and it will therefore be favored with respect to the computational effort. However, the table also shows that the minimum mode following methods locate appreciably more saddle points and requires considerable less force calls per saddle point found that is directly connected to the starting minimum than the (hybrid) RFO approach with a Bofill update. Taken together this indicates that if only the lowest saddle points are sought, and one can start the search with a good initial estimate of the Hessian, and the cost of the diagonalization will not be dominating, the (hybrid) RFO approach with a Bofill updater is to be preferred above the minimum mode following methods. But in all other cases the minimum mode following methods clearly represent the better choice. Note that the requirement of starting with a good initial estimate of the Hessian for the (hybrid) RFO approach might further limit the use of this method in the case of high-dimensional systems.

Similar considerations can be made for the 21D system.

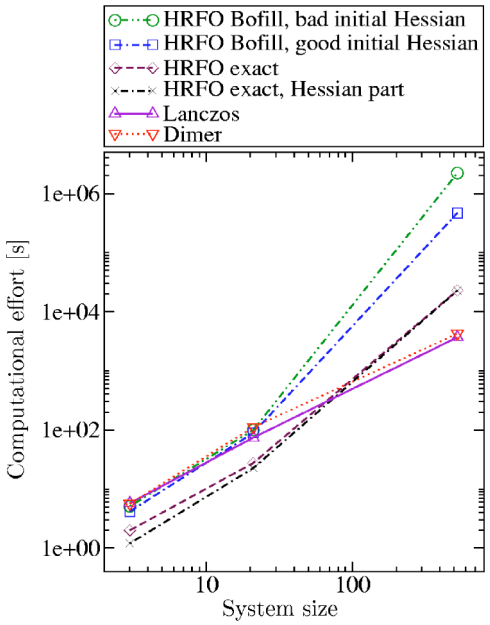

FIG. 5. (Color online) The total computational effort is given in CPU-seconds for the different system sizes. Representative results for each method are presented for $\Delta x^{\mathrm{ran}}=0.1$ Å. For the minimum mode following methods the effort is completely dominated by the force calls. For the hybrid RFO (HRFO) method employing an exact Hessian the computational effort of calculating the Hessian together with the diagonalization of it has been indicated separately. Since similar results are found for the RFO methods employing an approximate Hessian they have not been included in this figure. The lines are meant as a guide to the eye. Note the logarithmic scale on both axes.

If the force can be calculated cheaply, we see from Fig. 5 that the (hybrid) RFO approach using an exact Hessian requires the least total computational effort, and the method efficiently locates the lowest saddle points. From Table III we also see that this method is the most efficient with respect to the number of force calls needed per saddle point found that is directly connected to the starting minimum. Considering the case where the force is expensive to calculate, the (hybrid) RFO approach with an exact Hessian is no longer an alternative (see preceding paragraph). However, the (hybrid) RFO approach with a Bofill updater requires considerably less force calls than the minimum mode following methods and locates the lowest saddle points efficiently. But if the goal is to locate as many saddle points as possible, we see from Table III that only in the case of starting the searches with a good initial Hessian does the (hybrid) RFO approach with a Bofill updater require less force calls per saddle point found than the minimum mode methods. The guidelines for optimal search methods for the 21D system are summarized in Table VI.

From Fig. 5 we see that the (hybrid) RFO approach using an exact Hessian requires the least total computational effort for the 3D system when the force is cheap to calculate, and this will therefore be the preferred method if the lowest saddle points are sought. But even if all methods are able to locate all five saddle points, Fig. 3 indicates that the Lanczos method finds the most difficult to locate saddle point more often than the (hybrid) RFO approach with an exact Hessian, suggesting that less searches employing the Lanczos method would be needed as compared to the (hybrid) RFO approach with an exact Hessian when attempting to locate as many saddle points as possible. Taking this into account, the Lanczos method and (hybrid) RFO approach with an exact Hessian would perform almost equally well. For cases where the force is expensive to calculate the (hybrid) RFO approach using an exact Hessian would become too computationally demanding compared to the other methods (see above). But from Table II we see that the (hybrid) RFO approach with a Bofill update requires considerably less force calls to reach the saddle points than the minimum mode methods, making it the preferred method when seeking to locate the lowest saddle points. As in the case where the force is cheap to calculate, the different efficiencies in finding the most difficult to locate saddle point render the performance of the (hybrid) RFO approach with a Bofill update and the Lanczos method rather similar. In Table VI the guidelines for optimal search methods for the 3D system are summarized.

In closing this section, we would like to stress that a different maximum step length is required for optimal performance of the different methods.

V. CONCLUSIONS

Being able to locate saddle points on PESs is a prerequisite to the study of transitions when working within the harmonic approximation to transition state theory (whether used explicitly or implicitly). It is therefore not surprising that there has been a large number of studies in which different saddle point search methods have been developed or refined. However, in many cases it is assumed that a reasonable guess for the saddle point can be made and/or that the transition proceeds to a known final state. The number of studies describing methods allowing one to walk from a minimum to a saddle point without any knowledge of the final state are more limited, and, to the best of our knowledge, a systematic comparison of how such methods perform has not yet been made. In an effort to improve this situation we have set out to compare some traditional all-mode following methods with more recently developed minimum mode following methods.

As representative examples of the all-mode following methods we have investigated the performance of some modified Newton-Raphson approaches, the RFO methods pioneered by Cerjan and Miller[1] and Simons *et al.*[3,5] using either exact or approximate Hessians. In these methods all Hessian eigenvalues and eigenvectors are used together with the gradient vector to calculate the steps leading towards a saddle point. A comparison is made between the different versions of the all-mode following methods and also with two minimum mode following approaches, the Lanczos and dimer methods, where only the lowest eigenvalue and eigenvector are used in conjunction with the gradient vector to determine the steps. We have also tested the performance of hybrid versions of the RFO approaches in which the geometry steps are determined as in the minimum mode following methods in the convex regions of the PES (where all Hessian eigenvalues are positive) and as in the traditional RFO methods outside these regions. The PES walkers have been tested on a model system involving transitions in a seven-atom Pt island on a Pt(111) surface using a simple Morse pairwise potential function. Three different model system sizes have been considered—the full system with 175 moving atoms and two reduced size systems. The PES governing the transitions for the largest system is a 525D one, the reduced size systems are of 21D and 3D, respectively.

In terms of the average number of force calls needed to reach the different saddle points the RFO approach employing an exact Hessian is clearly better than the RFO approaches employing an approximate Hessian or the minimum mode following methods. However, due to the cost of calculating the exact Hessian and the cost of performing a full diagonalization of it, the total computational effort is in most cases unfavorable compared to the other methods. Only when the Hessian matrix can be calculated cheaply and the diagonalization of it does not dominate the overall effort, the RFO approach employing an exact Hessian is the most efficient of the methods we have tested. This is the case for small (3D) and medium (21D) sized systems where, e.g., a simple model potential like the one used here is employed and the Hessian matrix can be obtained analytically without much effort. Our results also show that the traditional RFO approach employing an exact Hessian mainly finds the lowest saddle points, but this behavior can be improved considerably by starting the search further away from the starting minimum or by employing the hybrid version introduced here.

With respect to the average number of force calls the RFO searches based on a Bofill update of the approximate Hessian also perform better than the minimum mode follow-

ing methods for the three systems sizes tested here. But due to the cost of performing a full diagonalization of Hessian, the overall computational effort will only be favorable in situations were the cost of calculating the force clearly dominates over the cost of the diagonalization. Furthermore, as for the RFO approach employing an exact Hessian, the RFO approach employing an approximate Hessian mainly locates the lower-lying saddle points. For the small (3D) and medium (21D) sized system starting the search further away from the initial minimum or using the hybrid version improves this behavior considerably, but in the case of the large (525D) system the improvement is less pronounced. Taken together this means that the (hybrid) RFO approach employing an approximate Hessian through a Bofill update is the most efficient of the methods tested here for the small and medium sized system, and when the cost of calculating the force represents the dominating part of the overall computational effort. For large systems the (hybrid) RFO approach is the most efficient only when the lowest saddle points are sought and the calculation of the force is dominating the overall computational effort. The latter is true if, e.g., the force is obtained through an electronic structure calculation at the density functional theory or higher levels. Starting the searches with a good initial approximation for the Hessian is not very important for the small (3D) systems, but a prerequisite for the medium (21D) and large (525D) systems.

The minimum mode following methods are to be preferred if a more complete mapping of all existing saddle points is desired for a large system. Also in the cases where only the lowest saddle points are sought for large systems and the force can be evaluated cheaply (e.g., if a simple model potential is used) the minimum mode following methods will be the preferred choice. Moreover, the minimum mode following methods will be preferred above the (hybrid) RFO approach employing an approximate Hessian through a Bofill update for large systems where the force is expensive to calculate and a good initial approximation for the Hessian is not available.

When working with the dimer method it is essential to limit the number of dimer rotations and to use a convergence criterion for the rotational force that is not too strict, i.e., not to work with a fully converged lowest eigenmode. Although somewhat less important for the Lanczos method, working with a not fully converged lowest eigenmode by limiting the number Lanczos iterations and using a less strict criterion for the Lanczos eigenvalue convergence, also improves the performance of the method considerably. To facilitate the use of the minimum mode following methods we have presented a set of ''safe'' parameter settings that in most cases will result in efficient saddle point searching. This renders the minimum mode following methods just as easy to use as the traditional RFO approach.

Finally, for all methods it is seen that the maximum length allowed for each geometry step is an important parameter, and a good choice can reduce the computational effort considerably and produce the wanted saddle point search behavior. A proper choice of the initial displacement away from the starting minimum is also important for obtaining the desired search results.

ACKNOWLEDGMENTS

We gratefully acknowledge help from Stan van Gisbergen, Drew McCormack, and Marcel Swart on different aspects of the rational function optimization approach, discussions with Andreas Heyden, Frerich J. Keil, and Michael Seth regarding the dimer method, and discussions with Fransesco Buda regarding the use of saddle point search methods in biochemistry. The research of RAO and GJK was financed by the Dutch National Research Council-Chemical Sciences (NWO-CW) through a PIONIER grant. G.H. is grateful for funding from the director's postdoctoral fellowship program at Los Alamos National Laboratory. The work has also been supported in part by the European Community's Human Potential Program under contract No. HPRN-CT-2002-00170. We also acknowledge support for computer time granted by the Dutch National Computing Facilities Foundation (NCF).

[1] C. J. Cerjan and W. H. Miller, J. Chem. Phys. **75**, 2800 (1981).
[2] H. B. Schlegel, J. Comput. Chem. **3**, 214 (1982).
[3] J. Simons, P. Jørgensen, H. Taylor, and J. Ozment, J. Phys. Chem. **87**, 2745 (1983).
[4] H. Taylor and J. Simons, J. Phys. Chem. **89**, 684 (1985).
[5] A. Banerjee, N. Adams, J. Simons, and R. Shepard, J. Phys. Chem. **89**, 52 (1985).
[6] J. Baker, J. Comput. Chem. **7**, 385 (1986).
[7] J. Baker, J. Comput. Chem. **8**, 563 (1987).
[8] J. Nichols, H. Taylor, P. Schmidt, and J. Simons, J. Chem. Phys. **92**, 340 (1989).
[9] J. Simons and J. Nichols, Int. J. Quantum Chem., Symp. **24**, 263 (1990).
[10] J. M. Bofill, J. Comput. Chem. **15**, 1 (1994).
[11] G. T. Barkema and N. Mousseau, Phys. Rev. Lett. **77**, 4358 (1996).
[12] J. M. Bofill, Chem. Phys. Lett. **260**, 359 (1996).
[13] J. P. K. Doye and D. J. Wales, Z. Phys. D: At., Mol. Clusters **40**, 194 (1997).
[14] W. Quapp, M. Hirsch, O. Imig, and D. Heidrich, J. Comput. Chem. **19**, 1087 (1998).
[15] N. Mousseau and G. T. Barkema, Phys. Rev. E **57**, 2419 (1998).
[16] L. J. Munro and D. J. Wales, Phys. Rev. B **59**, 3969 (1999).
[17] G. Henkelman and H. Jónsson, J. Chem. Phys. **111**, 7010 (1999); G. Henkelman, B. Uberuaga, and H. Jónsson, *ibid.* **113**, 9901 (2000).
[18] J. M. Anglada, E. Besalú, J. M. Bofill, and J. Rubio, J. Math. Chem. **25**, 85 (1999).
[19] R. Malek and N. Mousseau, Phys. Rev. E **62**, 7723 (2000).
[20] S. R. Billeter, A. J. Turner, and W. Thiel, Phys. Chem. Chem. Phys. **2**, 2177 (2000).
[21] J. M. Bofill and J. M. Anglada, Theor. Chem. Acc. **105**, 463 (2001).
[22] G. H. Jóhannesson and H. Jónsson, J. Chem. Phys. **115**, 9644 (2001).
[23] Ö. Farkas and H. B. Schlegel, Phys. Chem. Chem. Phys. **4**, 11 (2002).
[24] P. Deglmann and F. Furche, J. Chem. Phys. **117**, 9535 (2002).
[25] B. Peters, W. Liang, A. T. Bell, and A. Chakraborty, J. Chem. Phys. **118**, 9533 (2003).
[26] M. V. Fernández-Serra, E. Artacho, and J. M. Soler, Phys. Rev. B **67**, 100101(R) (2003).
[27] R. Crehuet and M. J. Field, J. Chem. Phys. **118**, 9563 (2003).
[28] G. Henkelman, G. Jóhannesson, and H. Jónsson, in *Progress in Theoretical Chemistry and Physics*, edited by S. D. Schwartz (Kluwer Academic, Dordrecht, 2000), Vol. 5.
[29] H. B. Schlegel, J. Comput. Chem. **24**, 1514 (2003).
[30] P. J. Feibelman, Phys. Rev. Lett. **65**, 729 (1990).
[31] P. Salvador, J. Phys. Chem. **89**, 3863 (1985).
[32] J. Augustynski, Struct. Bonding (Berlin) **69**, 1 (1988).
[33] K. J. Gross, S. Guthrie, S. Takara, and G. Thomas, J. Alloys Compd. **297**, 270 (2000).
[34] W. Nowak, R. Czerminski, and R. Elber, J. Am. Chem. Soc. **113**, 5627 (1991).
[35] M. Pavlov, P. E. M. Siegbahn, M. R. A. Blomberg, and R. H. Crabtree, J. Am. Chem. Soc. **120**, 548 (1998).

[36] C. Alhambra, J. Gao, J. C. Corchado, J. Villa, and D. G. Truhlar, J. Am. Chem. Soc. **121**, 2253 (1999).
[37] P. E. M. Siegbahn and M. R. A. Blomberg, Chem. Rev. (Washington, D.C.) **100**, 421 (2000).
[38] M. Stein, E. van Lenthe, E. J. Baerends, and W. Lubitz, J. Am. Chem. Soc. **123**, 5839 (2001).
[39] F. Himo and P. E. M. Siegbahn, Chem. Rev. (Washington, D.C.) **103**, 2421 (2003).
[40] E. R. Davidson, J. Comput. Phys. **17**, 87 (1975).
[41] M. T. Heath, *Scientific Computing: An Introductory Survey*, McGraw-Hill Series in Computer Science (McGraw-Hill, New York, 1997).
[42] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, Cambridge, 1992).
[43] B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall series in computational mathematics (Prentice-Hall, Englewood Cliffs, NJ, 1980).
[44] G. H. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed. (John Hopkins University Press, Baltimore, 1996).
[45] E. Anderson, Z. Bai, C. Bischof *et al.*, *LAPACK Users' Guide*, 3rd ed. (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1999), ISBN 0-89871-447-8 (paperback).
[46] M. J. D. Powell, *Nonlinear programing* (Academic, New York, 1970).
[47] B. A. Murtagh and R. W. H. Sargent, Comput. J. **13**, 185 (1970).
[48] D. W. Bassett and P. R. Webber, Surf. Sci. **70**, 520 (1978).