Efficient softest mode finding in transition states calculations

Jing Leng, Weiguo Gao, Cheng Shang, and Zhi-Pan Liu

Citation: J. Chem. Phys. **138**, 094110 (2013); doi: 10.1063/1.4792644
View online: http://dx.doi.org/10.1063/1.4792644
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v138/i9
Published by the American Institute of Physics.

Efficient softest mode finding in transition states calculations

Jing Leng,[1] Weiguo Gao,[1,a)] Cheng Shang,[2] and Zhi-Pan Liu[2]
[1]*School of Mathematical Sciences, MOE Laboratory of Mathematics for Nonlinear Science, Fudan University, Shanghai 200433, China*
[2]*Department of Chemistry, MOE Key Laboratory for Computational Physical Sciences, Fudan University, Shanghai 200433, China*

(Received 30 September 2012; accepted 5 February 2013; published online 6 March 2013)

Transition states are fundamental to understanding the reaction dynamics qualitatively in chemical physics. To date various methods of first principle location of the transition states have been developed. In the absence of the knowledge of the final structure, the softest-mode following method climbs up to a transition state without calculating the Hessian matrix. One weakness of this kind of approaches is that the number of rotations to determine the softest mode is usually unpredictable. In this paper, we propose a locally optimal search direction finding algorithm, namely LOR, which is an extension of the traditional conjugate gradient method without additional calculations of the forces. We also show that the translation of forces improves the numerical stability. Experiments for the Baker test system show that the proposed algorithm is much faster than the original dimer conjugate gradient method. © *2013 American Institute of Physics*. [http://dx.doi.org/10.1063/1.4792644]

I. INTRODUCTION

Transition state theory (TST) plays a central role in chemical kinetics. To determine TS and thus predict chemical activity based on TST is a major theme in modern theoretical simulation of chemical reactions. The algorithms for locating TS can be generally divided into two classes, namely, (i) chain-of-states methods and (ii) surface-walking methods. The former class locates the TS by simultaneously optimizing a few connected images on the potential energy surface (PES) to identify the minimum energy path (MEP). The representative methods include the nudged elastic band method (NEB),[1–5] doubly-nudged elastic band (DNEB),[6–9] and the string method.[10,11] The later class is much less demanding in computational power as only one structural image on the PES is required and the local information such as the gradient (Force) or the second derivative (Hessian) of PES is utilized to manipulate the image towards TS. Belonging to this category are the methods, such as the partitioned rational function optimizer (P-RFO),[12–15] the activation-relaxation technique (ART),[16,17] the hybrid eigenvector following method,[18–22] the dimer method,[23–27] the bond-length constrained minimization method,[28] the constrained-Broyden dimer method,[29] and the gentlest ascent dynamics method (GAD).[30]

Suppose N is the number of atoms in the system. The transition states are the first order saddle points of the energy surface in $3N$ potential space. The softest-mode following methods[16,20,23] in the later class need to approximate the softest mode of the Hessian at current image.

Let $\mathbf{R}_0$ be current image and $\mathbf{H}$ be the Hessian at $\mathbf{R}_0$. The softest mode is determined by the eigenvector corresponding to the smallest eigenvalue of $\mathbf{H}$, i.e., the unit vector $\mathbf{N}$ minimizing the curvature $C_{\mathbf{N}}$ as follows

$$\min_{\mathbf{N}^T\mathbf{N}=1} C_{\mathbf{N}} = \mathbf{N}^T\mathbf{H}\mathbf{N}. \tag{1}$$

Assume that the energy surface is almost quadratic within a fixed distance δR. We use $\mathbf{R}_1$ and $\mathbf{R}_2$ to denote the end points on the sphere centered at $\mathbf{R}_0$ along the unit vector $\mathbf{N}$, that is,

$$\mathbf{R}_1 = \mathbf{R}_0 + \delta R\mathbf{N}, \quad \mathbf{R}_2 = \mathbf{R}_0 - \delta R\mathbf{N}.$$

As the calculation of $\mathbf{H}$ is very expensive, we approximate the curvature by

$$C_{\mathbf{N}} \approx \frac{(\mathbf{F}_2 - \mathbf{F}_1)^T\mathbf{N}}{2\delta R},$$

where $\mathbf{F}_1$ and $\mathbf{F}_2$ are the forces acting on configurations $\mathbf{R}_1$ and $\mathbf{R}_2$, see Figure 1. This shows that minimizing $C_{\mathbf{N}}$ is "equivalent" to minimizing $(\mathbf{F}_2 - \mathbf{F}_1)^T\mathbf{N}$. An alternative way is to use $2\mathbf{F}_0 \approx \mathbf{F}_1 + \mathbf{F}_2$ to reduce the calculation of the forces.

Recalling that in the gradient-type method, we need to calculate the gradient of $C_{\mathbf{N}}$ with respect to $\mathbf{N}$ which gives

$$(\mathbf{I} - \mathbf{N}\mathbf{N}^T)\mathbf{H}\mathbf{N} \approx \frac{(\mathbf{I} - \mathbf{N}\mathbf{N}^T)(\mathbf{F}_2 - \mathbf{F}_1)}{2\delta R} = \frac{\mathbf{F}_2^{\perp} - \mathbf{F}_1^{\perp}}{2\delta R},$$

where $\mathbf{F}_i^{\perp} = (\mathbf{I} - \mathbf{N}\mathbf{N}^T)\mathbf{F}_i$ is orthogonal to $\mathbf{N}$. We remark that the gradient is also known as the residual corresponding to the direction vector $\mathbf{N}$.

Notably, the dimer method[23] is a softest mode following algorithm starting with an initial structure and mode, which uses only first derivatives of potential energy. Finding the softest mode is also known as the rotation step in the dimer method. Then the geometry step is to translate the midpoint of the dimer along a modified force

$$\mathbf{F}_{\text{tran}} = \begin{cases} -\mathbf{N}\mathbf{N}^T\mathbf{F}_0 & \text{if} \quad C_{\mathbf{N}} > 0, \\ (\mathbf{I} - 2\mathbf{N}\mathbf{N}^T)\mathbf{F}_0 & \text{if} \quad C_{\mathbf{N}} < 0. \end{cases}$$

And the step length is suggested to be

$$\delta x = \frac{1}{2}\frac{\mathbf{F}_{\text{tran}}^T\mathbf{T}}{|C_{\mathbf{T}}|},$$

a)Electronic mail: wggao@fudan.edu.cn.

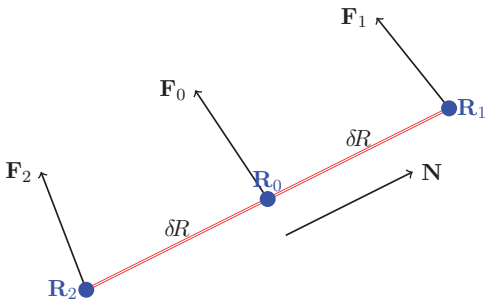

FIG. 1. Dimer model.

where **T** is the unit vector of current translation direction in Ref. 25.

Unlike the minimization procedure, the behavior of saddle point locating method is much more subtle. Reformulating the rotation and geometry steps into a dynamic system framework gains mathematical insight. For examples, the gentlest ascent dynamics[30,31] and furthermore the constrained shrinking dimer dynamics (CSDD)[32] present the nature of the dynamical behavior. However, efficient numerical schemes to accelerate the simulation of the dynamical system are still unclear.

An important factor in evaluating the success of the method is the number of total force calls. To reduce the force calls in the rotation step, maximum number of allowed rotations and less strict convergence criterion are considered.[24] However, limiting the number of rotations may lead to inaccurate translation direction and therefore miss the relevant saddle point on PES.

In this work, we developed a better approach to treat the rotation step by using the idea of locally optimal conjugate gradient direction from the eigensolver. The main point is to enlarge the subspace expansion utilizing the vector in previous step therefore to get more accurate approximation of the softest mode without additional calculation of the forces. The new method was tested by the Baker test system. The results show that our method greatly increases the efficiency of the rotation move by a factor of 3.

II. LOCALLY OPTIMAL CONJUGATE GRADIENT FOR FINDING THE SOFTEST MODE

As described above, locating the softest mode in the rotation step is equivalent to solve the smallest eigenpair. Different methods have been be considered to choose the subspace of rotation, including conjugate gradient,[23,25] Lanczos,[17] Broyden,[29] and L-BFGS,[26,33,34] etc.

Among these optimization-based methods, the Broyden and L-BFGS algorithms are more easily trapped by the higher eigenpairs. If we regard the rotation and geometry steps as an inner-outer iteration scheme, the conjugate-gradient type method is more preferable than other methods because the previous eigenpair is naturally a good initial guess in the next iteration, while other methods much less depend on the initial guess. This is similar to the self-consistent field in the electronic structure calculations, where the conjugate-gradient method is widely used as the eigensolver.

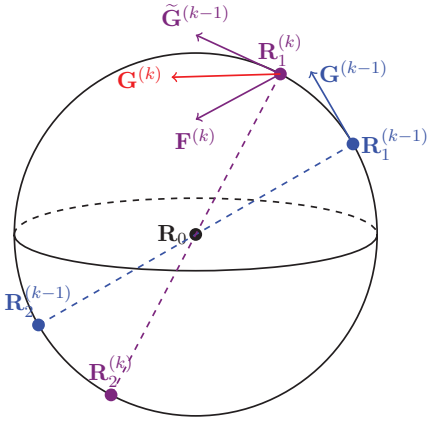

FIG. 2. Rotation via conjugate direction.

For this reason, we look for more efficient conjugate-gradient type method for the rotation step. First we briefly introduce the plane of rotation in dimer conjugate gradient method which is closely related to the proposed approach.

A. Dimer conjugate gradient method

First we define $\mathbf{F} = \mathbf{F}_2^{\perp} - \mathbf{F}_1^{\perp}$ which is "parallel" to the gradient.

In dimer conjugate gradient method, the conjugate direction is a linear combination of the current gradient and the previous conjugate direction, see Figure 2. More precisely, we use $\mathbf{G}^{(k)}$ to denote the conjugate direction at kth iteration, then

$$\mathbf{G}^{(k)} = \mathbf{F}^{(k)} + \gamma_k \widetilde{\mathbf{G}}^{(k-1)},$$

where

$$\gamma_k = \frac{(\mathbf{F}^{(k)} - \mathbf{F}^{(k-1)})^T \mathbf{F}^{(k)}}{\mathbf{F}^{(k)T} \mathbf{F}^{(k)}}$$

and $\widetilde{\mathbf{G}}^{(k-1)}$ is the rotation of $\mathbf{G}^{(k-1)}$ from $\mathbf{R}_1^{(k-1)}$ to $\mathbf{R}_1^{(k)}$.

Let $\mathbf{T}^{(k)} = \mathbf{G}^{(k)}/\|\mathbf{G}^{(k)}\|$ be the unit vector along the conjugate direction. Then the conjugate gradient method finds the minimum curvature in the rotation plane spanned by $\{\mathbf{N}^{(k)}, \mathbf{T}^{(k)}\}$. Namely, the best rotation direction is to be determined by

$$\min_{a^2+b^2=1} (a\ b) \begin{pmatrix} \mathbf{N}^{(k)T} \\ \mathbf{T}^{(k)T} \end{pmatrix} \mathbf{H} (\mathbf{N}^{(k)} \quad \mathbf{T}^{(k)}) \begin{pmatrix} a \\ b \end{pmatrix}, \tag{2}$$

when $\mathbf{N}^{(k)}, \mathbf{T}^{(k)}$ are known. Let $(a^{(k)}\ b^{(k)})$ be the minimizer of (2), then the next direction is updated as $\mathbf{N}^{(k+1)} = (\mathbf{N}^{(k)}\ \mathbf{T}^{(k)}) \begin{pmatrix} a \\ b \end{pmatrix}$.

B. Locally optimal conjugate gradient method

Following the idea of LOBPCG method[35] for solving symmetric sparse eigenproblem, we impose one additional vector $\mathbf{P}^{(k)}$. Instead of minimizing the 2×2 sub-eigenproblem in (2), we minimize a 3×3 sub-eigenproblem in each

iteration (except the first iteration) as follows

$$\begin{aligned}\left(a^{(k)}\ b^{(k)}\ c^{(k)}\right) = \arg \min_{\|a\mathbf{N}^{(k)}+b\Phi^{(k)}+c\mathbf{P}^{(k)}\|=1} \left(a\ b\ c\right) \\ \times \begin{pmatrix} \mathbf{N}^{(k)^T} \\ \Phi^{(k)^T} \\ \mathbf{P}^{(k)^T} \end{pmatrix} \mathbf{H} \begin{pmatrix} \mathbf{N}^{(k)} & \Phi^{(k)} & \mathbf{P}^{(k)} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix},\end{aligned} \tag{3}$$

where $\Phi^{(k)} = \mathbf{F}^{(k)}/\|\mathbf{F}^{(k)}\|$. $\mathbf{P}^{(k+1)}$ is updated by

$$\mathbf{P}^{(k+1)} = \frac{b^{(k)}\Phi^{(k)} + c^{(k)}\mathbf{P}^{(k)}}{\|b^{(k)}\Phi^{(k)} + c^{(k)}\mathbf{P}^{(k)}\|}$$

and $\mathbf{P}^{(1)} = 0$. As a result, the new rotational direction is obtained by

$$\mathbf{N}^{(k+1)} = \begin{pmatrix} \mathbf{N}^{(k)} & \Phi^{(k)} & \mathbf{P}^{(k)} \end{pmatrix} \begin{pmatrix} a^{(k)} \\ b^{(k)} \\ c^{(k)} \end{pmatrix}.$$

Solving (3) is equivalent to solving the smallest eigenvalue λ corresponding to the following generalized eigenvalue problem

$$\begin{aligned}&\begin{pmatrix} \mathbf{N}^{(k)^T} \\ \Phi^{(k)^T} \\ \mathbf{P}^{(k)^T} \end{pmatrix} \mathbf{H} \begin{pmatrix} \mathbf{N}^{(k)} & \Phi^{(k)} & \mathbf{P}^{(k)} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} \\ &= \lambda \begin{pmatrix} \mathbf{N}^{(k)^T} \\ \Phi^{(k)^T} \\ \mathbf{P}^{(k)^T} \end{pmatrix} \begin{pmatrix} \mathbf{N}^{(k)} & \Phi^{(k)} & \mathbf{P}^{(k)} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}.\end{aligned}$$

The overlapping matrix in the right side is not identical in general. Since $\mathbf{N}^{(k)}$ and $\mathbf{\Phi}^{(k)}$ are orthonormal and $\mathbf{P}^{(k)}$ is normalized, we look into the structure of the overlapping matrix which is of the form

$$\begin{pmatrix} 1 & 0 & \mathbf{N}^{(k)^T}\mathbf{P}^{(k)} \\ 0 & 1 & \Phi^{(k)^T}\mathbf{P}^{(k)} \\ \mathbf{P}^{(k)^T}\mathbf{N}^{(k)} & \mathbf{P}^{(k)^T}\Phi^{(k)} & 1 \end{pmatrix}.$$

Then a simplified Cholesky factorization can be adopted to convert the generalized eigenvalue problem to the standard one.

In principle, we can keep all historical vectors and optimize globally. However, it has been observed[35] that the approximation of the first eigenpair using the locally optimal idea is about the same quality as those by global optimization. Moreover, the locally optimal method is more stable when the Hessian matrix is ill-conditioned. And in Ref. 36, locally optimal rotation (LOR) is considered as a Jacobi conjugation scheme

$$\mathbf{G}^{(k)} = \mathbf{F}^{(k)} + \gamma_k \tilde{\mathbf{G}}^{(k-1)}, \quad \gamma_k = -\frac{\mathbf{F}^{(k)^T}\mathbf{J}\left(C_{\mathbf{N}}^{(k)}, \mathbf{N}^{(k)}\right)\mathbf{G}^{(k-1)}}{\mathbf{G}^{(k-1)^T}\mathbf{J}\left(C_{\mathbf{N}}^{(k)}, \mathbf{N}^{(k)}\right)\mathbf{G}^{(\mathbf{k-1})}},$$

$$\mathbf{J}(C_{\mathbf{N}}^{(k)}, \mathbf{N}^{(\mathbf{k})}) = (\mathbf{I} - \mathbf{N}^{(k)}\mathbf{N}^{(k)^T})\left(\mathbf{H} - C_{\mathbf{N}}^{(k)}\mathbf{I}\right)(\mathbf{I} - \mathbf{N}^{(k)}\mathbf{N}^{(k)^T}),$$

where $\mathbf{J}(C, \mathbf{N})$ is the Jacobi correction operator and $\mathbf{I}$ is the identity matrix. LOR implicitly employs $C_{\mathbf{N}}^{(k+1)}$ instead of $C_{\mathbf{N}}^{(k)}$ in the Jacobi correction operator.

C. Translations of the forces

Given a matrix, the important characterization of an eigensolver is that only one matrix-vector multiplication is carried out in each iteration. The situation is slightly more tricky because of the fact that the potential energy surface is usually not exactly quadratic.

In the original dimer method, $\mathbf{HN}^{(k+1)}$ is approximated by calculating one extra force at $\mathbf{R}_0 + \epsilon\mathbf{N}^{(k+1)}$ for some small ϵ. Then

$$\mathbf{HN}^{(k+1)} \approx -\frac{\mathbf{F}(\mathbf{R}_0 + \epsilon\mathbf{N}^{(k+1)}) - \mathbf{F}(\mathbf{R}_0)}{\epsilon}.$$

As we already have

$$\mathbf{N}^{(k+1)} = a^{(k)}\mathbf{N}^{(k)} + b^{(k)}\Phi^{(k)} + c^{(k)}\mathbf{P}^{(k)},$$

it is naturally to translate the existing forces to the new direction

$$\mathbf{HN}^{(k+1)} = a^{(k)}\mathbf{HN}^{(k)} + b^{(k)}\mathbf{H}\Phi^{(k)} + c^{(k)}\mathbf{HP}^{(k)}.$$

And similarly, $\mathbf{HP}^{(k+1)}$ can be translated from $\mathbf{H}\Phi^{(k)}$ and $\mathbf{HP}^{(k)}$. Therefore, we only carry out $\mathbf{H}\Phi^{(k+1)}$ in the next iteration. A similar idea has been used in L-BFGS method in Ref. 26. However, our treatment does not use the extrapolation technique, as a result, behaviors more stable.

Now the question is whether the errors in the force calculations will be accumulated during the translations. It is interesting that it indeed improves the numerical stability. The reason is that when we do the translation, we implicitly use the same Hessian matrix on different directions. While the direct approximation of the forces involves high order perturbations especially when $\mathbf{N}^{(k+1)}$ almost converges. The detailed analysis will be presented in a separate paper.

In particularly, the calculated curvature at the $k+1$st iteration is guaranteed to be less than or equal to the one at kth iteration. It is easily seen from the 3×3 sub-problem by setting $(a^{(k)}, b^{(k)}, c^{(k)}) = (1, 0, 0)$. This phenomenon is verified by the experiments in Sec. III.

D. Overall algorithm

Combining the techniques above, we obtain the LOR algorithm stated in Algorithm I. The iteration stops when the residual norm is small enough, i.e., equivalently $\|\mathbf{F}^{(k)}\| < \tau$ for some k, where τ is a given tolerance or when the curvature stalls.

We emphasize that there is only one force call in each iteration in Algorithm I.

III. RESULTS AND DISCUSSION

To test the efficiency of our approach, we chose Baker reaction system[37] as the testing examples, which contains 25 different chemical reactions as listed in Table I. The same system has been utilized to test the modified dimer method. Columns (11) and (12) in Table I are the numbers of iterations and forces by using CG algorithm in original dimer method. Columns (21) and (22) are the numbers of iterations and forces by using LOR. For comparison purpose,

ALGORITHM I. Locally Optimal Rotation (LOR)

procedure LOR ($\mathbf{R}_0^{(1)}$, $\mathbf{N}^{(1)}$) ▷ Initial structure and mode

$\mathbf{F}^{(1)} \leftarrow (\mathbf{I} - \mathbf{N}^{(1)}\mathbf{N}^{(1)^T})\,\mathbf{HN}^{(1)}$ ▷ The force orthogonal to $\mathbf{N}^{(1)}$ at $\mathbf{R}_0^{(1)}$

if $\|\mathbf{F}^{(1)}\| < \tau$, then stop ▷ Convergence check

$\Phi^{(1)} \leftarrow \mathbf{F}^{(1)}/\|\mathbf{F}^{(1)}\|$

$$\min_{a^2+b^2=1} (a\ b)\begin{pmatrix}\mathbf{N}^{(1)^T}\\ \Phi^{(1)^T}\end{pmatrix}\mathbf{H}(\mathbf{N}^{(1)}\quad\Phi^{(1)})\begin{pmatrix}a\\ b\end{pmatrix}$$ ▷ 2×2 eigenproblem

$\mathbf{N}^{(2)} \leftarrow a\mathbf{N}^{(1)} + b\Phi^{(1)}$

$\mathbf{P}^{(2)} \leftarrow \Phi^{(1)}$

$\mathbf{F}^{(2)} \leftarrow (\mathbf{I} - \mathbf{N}^{(2)}\mathbf{N}^{(2)^T})\mathbf{HN}^{(2)}$

for $k = 2$ to rotmax **do**

if $\|\mathbf{F}^{(k)}\| < \tau$, then stop

$\Phi^{(k)} \leftarrow \mathbf{F}^{(k)}/\|\mathbf{F}^{(k)}\|$

$$\min_{\|a\mathbf{N}^{(k)}+b\Phi^{(k)}+c\mathbf{P}^{(k)}\|=1} (a\ b\ c)\begin{pmatrix}\mathbf{N}^{(k)^T}\\ \Phi^{(k)^T}\\ \mathbf{P}^{(k)^T}\end{pmatrix}\mathbf{H}(\mathbf{N}^{(k)}\quad\Phi^{(k)}\quad\mathbf{P}^{(k)})\begin{pmatrix}a\\ b\\ c\end{pmatrix}$$ ▷ 3×3 eigenproblem

$$\mathbf{N}^{(k+1)} \leftarrow \frac{a\mathbf{N}^{(k)} + b\Phi^{(k)} + c\mathbf{P}^{(k)}}{\|a\mathbf{N}^{(k)} + b\Phi^{(k)} + c\mathbf{P}^{(k)}\|}$$

$$\mathbf{P}^{(k+1)} \leftarrow \frac{b\Phi^{(k)} + c\mathbf{P}^{(k)}}{\|b\Phi^{(k)} + c\mathbf{P}^{(k)}\|}$$

$\mathbf{F}^{(k+1)} \leftarrow (\mathbf{I} - \mathbf{N}^{(k+1)}\mathbf{N}^{(k+1)^T})\mathbf{HN}^{(k+1)}$

end for

end procedure

TABLE I. The energy and gradient calculation steps from original dimer method, the LOR method and the L-BFGS method in the TS location of Baker reactions with 10 maximum rotations each geometry step. Columns (11) and (12) are the numbers of iterations and forces by using CG algorithm in original dimer method. Columns (21) and (22) are the numbers of iterations and forces by using LOR. Columns (31) and (32) are the numbers of iterations and forces by using L-BFGS, respectively.

	System	Dimension	(11)	(12)	(21)	(22)	Ratio (%)	(31)	(32)
1	$CH_3O \rightarrow CH_2OH$	15	9	114	10	40	35	9	52
2	Claisen rearrangement	42	20	410	24	104	25	29	117
3	$H_3CCH_3 \rightarrow H_2CCH_2 + H_2$	24	14	236	15	74	31	14	59
4	Parent Diels-Alder	48	28	594	28	175	29	39	257
5	β-(formyloxy) ethyl	30	16	286	16	74	26	17	59
6	Bicyclo110 butane TS1	30	32	542	22	104	19	27	131
7	Bicyclo110 butane TS2	30	35	748	29	126	17	30	133
8	$H_2CCHOH \rightarrow H_3CCHO$	21	28	468	26	128	27	24	118
9	$H_2CNH \rightarrow HCNH_2$	15	20	216	18	83	38	27	118
10	$H_2CO \rightarrow H_2 + CO$	12	17	237	16	66	28	13	71
11	$H_2O + PO_3^- \rightarrow H_2PO_4$	28	18	347	21	108	31	27	150
12	$H_3CCH_2F \rightarrow H_2CCH_2 + HF$	24	20	334	24	111	33	32	211
13	$HCCH \rightarrow CCH_2$	12	15	161	18	101	63	25	174
14	$HCNH_2 \rightarrow HCN + H_2$	15	11	87	11	54	62	14	64
15	$HCN \rightarrow HNC$	9	10	115	10	45	39	9	59
16	$HCOCl \rightarrow HCl + CO$	12	30	530	23	164	31	##	##
17	$HCONH_3^+ \rightarrow NH_4^+ + CO$	21	19	357	27	149	42	##	##
18	$HCONHOH \rightarrow HCOHNHO$	21	14	286	20	86	30	##	##
19	$HNCCS \rightarrow HNC + CS$	15	18	332	13	85	26	##	##
20	$HNC + H_2 \rightarrow H_2CNH$	15	14	112	14	59	53	15	63
21	Ring-opening cyclopropyl	24	##	##	19	88	#	19	89
22	Rotational TS in acrolein	24	44	775	19	76	10	##	##
23	Rotational TS in butadiene	30	41	645	30	154	24	##	##
24	Silylene insertion	33	23	446	25	117	26	18	99
25	s-Tetrazine $\rightarrow 2HCN + N_2$	24	15	278	19	128	46	19	92

TABLE II. The iteration number and average rotation steps per iteration of 25 reactions in the Baker test system. The "##" sign in the table represents that the corresponding method either produces the wrong TS or diverges after 1000 forces. Reaction 21 is not counted when calculating the average **Itr** and **Av. Rot** for the original dimer method. Reactions 16, 17, 18, 19, 22, and 23 are not counted when calculating the average **Itr** and **Av. Rot** for the L-BFGS method.

System		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
Itr	CG	9	20	14	28	16	32	35	28	20	17	18	20	15
	LOR	10	24	15	28	16	22	29	26	18	16	21	24	18
	L-BFGS	9	29	14	39	17	27	30	24	27	13	27	32	25
Av. Rot	CG	10.7	18.5	14.9	19.2	15.9	14.9	19.4	14.7	8.8	11.9	18.8	14.7	8.7
	LOR	2.0	2.3	2.9	4.3	2.6	2.7	2.3	2.9	2.6	2.1	3.1	2.6	3.6
	L-BFGS	3.8	2.0	2.2	4.6	1.5	2.9	2.4	2.9	2.4	3.5	3.6	4.6	5.0
System		**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**Av.**
Itr	CG	11	10	30	19	14	18	14	##	44	41	23	15	**21.3**
	LOR	11	10	23	27	20	13	14	19	19	30	25	19	**19.9**
	L-BFGS	14	9	##	##	##	##	15	19	##	##	18	19	**21.4**
Av. Rot	CG	5.9	9.5	15.7	16.8	18.4	16.4	6.0	##	15.6	13.7	17.4	16.5	**14.30**
	LOR	2.9	2.5	5.1	3.5	2.3	4.5	2.2	2.6	2.0	3.1	2.7	4.7	**2.98**
	L-BFGS	2.6	4.6	##	##	##	##	2.2	2.7	##	##	3.5	2.8	**3.14**

we also implemented the L-BFGS rotational algorithm used in the DL-FIND package.[27] The numbers of iterations and forces by using the L-BFGS algorithm are listed in columns (31) and (32), respectively. For all the reactions studied, we started from the same guess structure ($\mathbf{R}_{GS}$) with at least one negative eigenvalue as suggested by Baker.[15] The initial mode for the dimer was set as $\mathbf{N}_{ini} = \mathbf{R}_{GS} - \mathbf{R}_{IS}$. All calculations were performed using VASP package[38] with ultrasoft pseudo potentials in which the GGA-PBE exchange-correlation functional was utilized. The rotation stops if the $\|\mathbf{F}\|$ is lower than 0.1 eV/Å or 10 rotational moves (20 steps of energy and force calculations) have been made. The L-BFGS algorithm[26] also stops when the rotation angle is less than 5°. The TS searching is converged if the maximum force on each freedom (max $|\mathbf{F}|$) is below 0.05 eV/Å. All the determined TSs have been checked with the literature structure[15] to ensure that the correct TS is identified. In Table I, the "##" sign in the table represents that the corresponding method either produces the wrong TS or diverges after 1000 forces.

From Table I, we found that the average number of steps by using original method is 361 for the 24 reactions with located correct TS, and the method fails in one reaction. By using the LOR method, the average number of steps is reduced to 100 for 25 reactions with located correct TS, and the method succeeds in all reactions. The overall efficiency is increased by about 261%. With our implementation, the L-BFGS method failed in 6 reactions. And it is still a little worse than the LOR method for other reactions. Table II listed the iteration numbers (**Itr**) and the average rotation steps (**Av. Rot**) per iteration of each reaction in Baker system (reaction 21 is not counted). We can see that on average, the LOR method

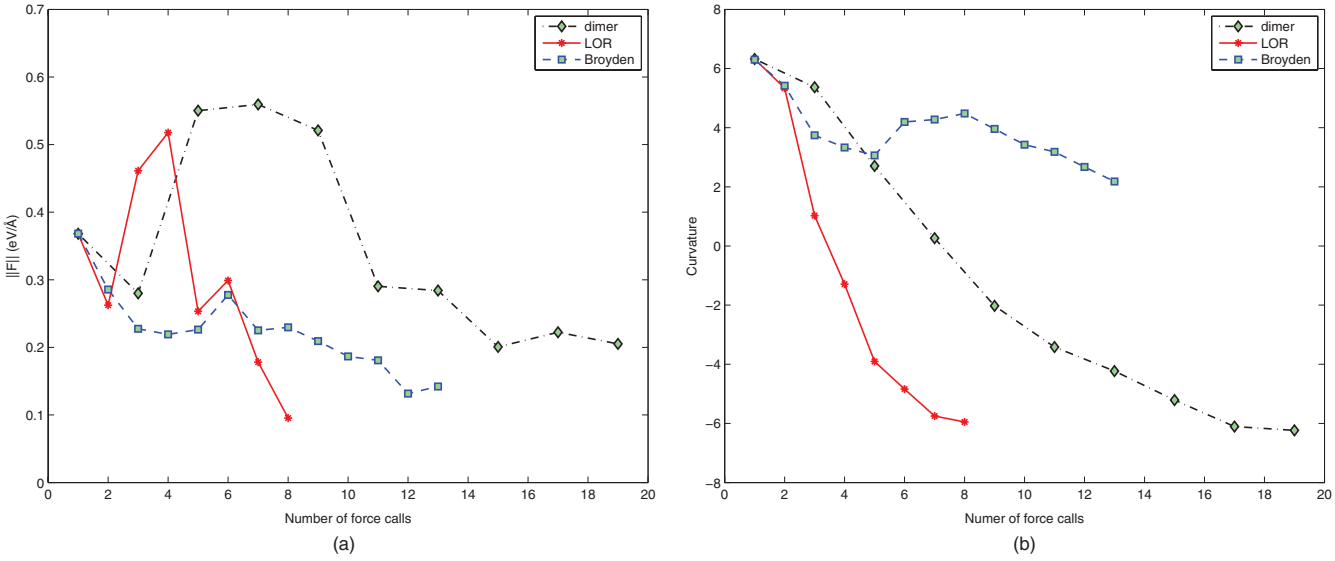

FIG. 3. The $\|\mathbf{F}\|$ trajectory (a) and curvature trajectory (b) of rotation move. $\mathbf{R}_0$ is determined as the initial structure of the reaction parent Diels-Alder. The initial mode is determined the same as that in Baker test.

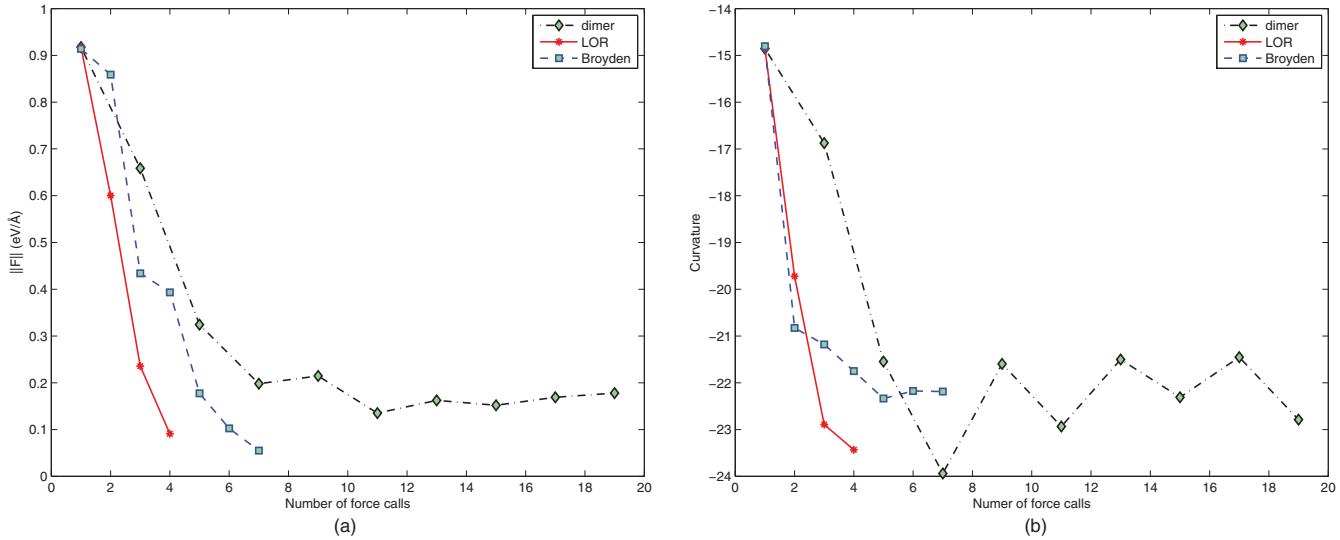

FIG. 4. The $\|\mathbf{F}\|$ trajectory (a) and curvature trajectory (b) of rotation move. $\mathbf{R}_0$ is determined as the initial structure of the reaction β-(formyloxy) ethyl. The initial mode is determined the same as that in Baker test.

does not save the iteration number to locate the transition state. It is reasonable because the iteration number is mainly determined by the translation move that is the same for these methods. However, the **Av. Rot** is decreased from 14.2 in original dimer to 3.0 in our method which means the efficiency of rotation is increased by 373%. Moreover, the average **Itr** and **Av. Rot** of the L-BFGS method for the 19 reactions are 21.4 and 3.14 which are very close to those of LOR.

We have further investigated the convergence behavior of LOR by utilizing the reaction parent Diels-Alder, the reaction β-(formyloxy) ethyl, and the reaction $H_2CO \rightarrow H_2 + CO$ as examples because the **Av. Rot** of these reactions are close to the average value of the system. The initial structures were set as $\mathbf{R}_0$, that is, only the first rotation steps were compared. Both methods start from the initial modes and converge to $\|\mathbf{F}\| < 0.1$ eV/Å. The results are shown in Figures 3–5, respectively. For comparison, the results of the Broyden method[29] is also shown because it uses the same initial modes in the first rotation steps. We can see that the trajectories of the LOR method are smoother than those of CG algorithm in original dimer method and the Broyden method, and the LOR method always converges faster. The curvatures calculated by the LOR method monotonically decrease since the 2nd iteration while CG algorithm and the Broyden method do not. It seems that the Broyden method did not converge to the softest mode in the reaction parent Diels-Alder. It is worth noticing that this would happen for the Broyden method as stated in Ref. 29.

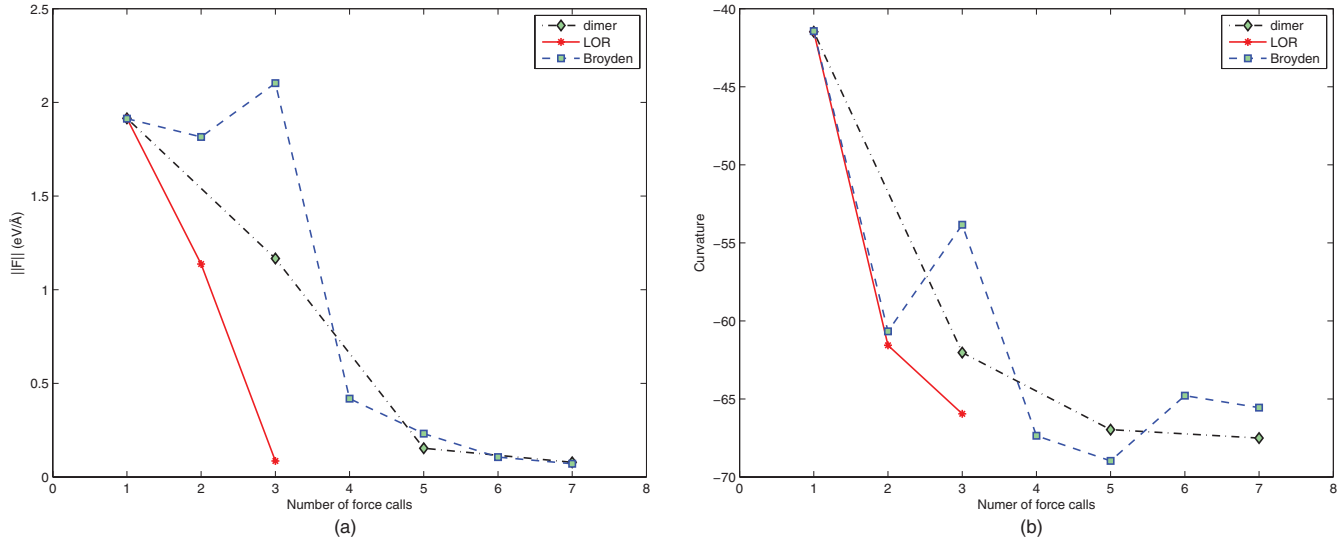

FIG. 5. The $\|\mathbf{F}\|$ trajectory (a) and curvature trajectory (b) of rotation move. $\mathbf{R}_0$ is determined as the initial structure of the reaction $H_2CO \rightarrow H_2 + CO$. The initial mode is determined the same as that in Baker test.

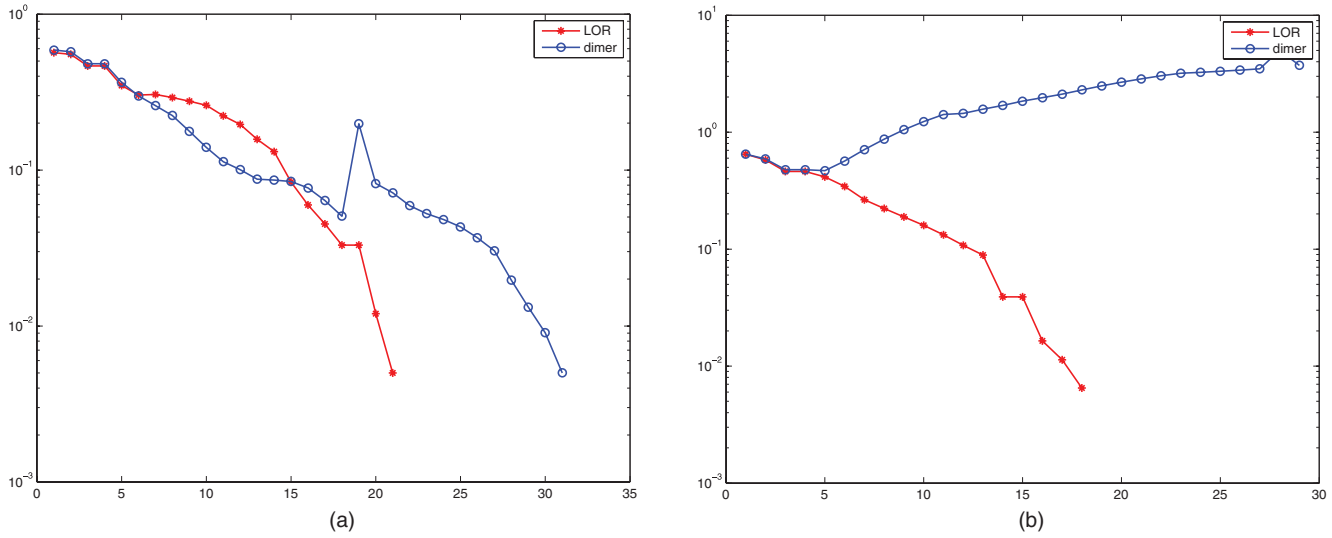

FIG. 6. The distances between each structure in the iteration and the find saddle point structure for (a) the reaction bicyclo110 butane TS1 and (b) the reaction ring-opening cyclopropyl.

We measured the distances between each structure in the iteration and the find saddle point structure (for the example that the dimer method fails, we use the saddle point found by LOR instead). Since we essentially use the same step length formula as the dimer methods, the geometry steps behavior similarly for most examples. However, we can still see from Figure 6 that for the reaction bicyclo110 butane TS1 and the reaction ring-opening cyclopropyl, the dimer method hovered over or escaped away from the saddle point due to the inaccurate directions. This confirms the stability of the LOR method.

We are also interested in seeing the effects of reducing the maximum number of rotations. So we changed it from 10 to 5 (up to 10 force calculations in each iteration). The results are shown in Table III. Not surprisingly, the behaviors of CG

TABLE III. The energy and gradient calculation steps from original dimer method, the LOR method and the L-BFGS method in the TS location of Baker reactions with 5 maximum rotations each geometry step. Columns (11) and (12) are the numbers of iterations and forces by using CG algorithm in original dimer method. Columns (21) and (22) are the numbers of iterations and forces by using LOR. Columns (31) and (32) are the numbers of iterations and forces by using L-BFGS, respectively.

	System	Dimension	(11)	(12)	(21)	(22)	Ratio (%)	(31)	(32)
1	$CH_3O \rightarrow CH_2OH$	15	15	81	9	38	47	9	41
2	Claisen rearrangement	42	22	234	20	96	41	21	96
3	$H_3CCH_3 \rightarrow H_2CCH_2 + H_2$	24	15	162	13	53	33	14	61
4	Parent Diels-Alder	48	29	322	27	163	51	##	##
5	β-(formyloxy) ethyl	30	15	168	16	74	44	17	59
6	Bicyclo110 butane TS1	30	30	288	22	104	36	29	132
7	Bicyclo110 butane TS2	30	40	468	31	132	28	28	139
8	$H_2CCHOH \rightarrow H_3CCHO$	21	23	233	20	98	42	39	199
9	$H_2CNH \rightarrow HCNH_2$	15	21	150	18	83	55	31	131
10	$H_2CO \rightarrow H_2 + CO$	12	13	131	16	66	50	12	51
11	$H_2O + PO_3^- \rightarrow H_2PO_4$	28	19	217	21	108	50	17	88
12	$H_3CCH_2F \rightarrow H_2CCH_2 + HF$	24	16	163	24	111	68	25	118
13	$HCCH \rightarrow CCH_2$	12	13	124	23	112	90	##	##
14	$HCNH_2 \rightarrow HCN + H_2$	15	11	85	11	51	60	15	73
15	$HCN \rightarrow HNC$	9	10	101	10	45	45	14	68
16	$HCOCl \rightarrow HCl + CO$	12	39	383	25	140	37	##	##
17	$HCONH_3^+ \rightarrow NH_4^+ + CO$	21	44	477	49	332	70	##	##
18	$HCONHOH \rightarrow HCOHNHO$	21	##	##	##	##	#	##	##
19	$HNCCS \rightarrow HNC + CS$	15	22	256	13	78	30	##	##
20	$HNC + H_2 \rightarrow H_2CNH$	15	14	111	14	59	53	14	60
21	Ring-opening cyclopropyl	24	##	##	18	83	#	19	85
22	Rotational TS in acrolein	24	17	182	18	79	43	##	##
23	Rotational TS in butadiene	30	##	##	36	187	#	30	185
24	Silylene insertion	33	18	182	25	116	64	23	119
25	s-Tetrazine $\rightarrow 2HCN + N_2$	24	17	193	16	83	43	30	150

algorithm changed a lot while the LOR method did not. This is because CG algorithm did not achieve the softest modes in many geometry steps. Although we see the improvements on the force calls for most reactions, it failed in 3 reactions. And the L-BFGS failed on two additional reactions 4 and 13.

Finally, the geometry step may be further improved by using the continuous translational move scheme proposed previously[29] and an accurate mode direction can certainly help to stabilize the continuous move. This will be the future direction to explore.

IV. CONCLUSIONS

We proposed a locally optimal rotation strategy in the calculation of TST states in this paper. This new strategy not only accelerates the rotations, but also improves the convergence numerically. Experiments with the Baker test system demonstrates the efficiency of the proposed method.

ACKNOWLEDGMENTS

The work of W. Gao is supported by the National Natural Science Foundation of China under Grant No. 11071047 and the 111 Project. The work of Z. Liu is supported by National Natural Science Foundation of China (Grant Nos. 20825311 and 21173051) and 973 program (Grant No. 2011CB808500). The authors are grateful to Weinan E for helpful discussions and the anonymous referees for valuable suggestions.

[1]G. Henkelman and H. Jónsson, J. Chem. Phys. **113**, 9978 (2000).
[2]G. Henkelman, B. P. Uberuaga, and H. Jónsson, J. Chem. Phys. **113**, 9901 (2000).
[3]G. Mills and H. Jónsson, Phys. Rev. Lett. **72**, 1124 (1994).
[4]D. Sheppard, R. Terrell, and G. Henkelman, J. Chem. Phys. **128**, 134106 (2008).
[5]R. Elber and M. Karplus, Chem. Phys. Lett. **139**, 375 (1987).
[6]S. A. Trygubenko and D. J. Wales, J. Chem. Phys. **120**, 2082 (2004).
[7]S. A. Trygubenko and D. J. Wales, J. Chem. Phys. **120**, 7820 (2004).
[8]E. F. Koslover and D. J. Wales, J. Chem. Phys. **127**, 134102 (2007).
[9]J. M. Carr, S. A. Trygubenko, and D. J. Wales, J. Chem. Phys. **122**, 234903 (2005).
[10]W. E, W. Ren and E. Vanden-Eijnden, Phys. Rev. B **66**, 052301 (2002).
[11]B. Peters, A. Heyden, A. T. Bell, and A. Chakraborty, J. Chem. Phys. **120**, 7877 (2004).
[12]J. Simons, P. Jorgensen, H. Taylor, and J. Ozment, J. Phys. Chem. **87**, 2745 (1983).
[13]Y. G. Khait and Y. V. Puzanov, J. Mol. Struct.: THEOCHEM **398–399**, 101 (1997).
[14]C. J. Cerjan and W. H. Miller, J. Chem. Phys. **75**, 2800 (1981).
[15]J. Baker, J. Comput. Chem. **7**, 385 (1986).
[16]G. T. Barkema and N. Mousseau, Phys. Rev. Lett. **77**, 4358 (1996).
[17]R. Malek and N. Mousseau, Phys. Rev. E **62**, 7723 (2000).
[18]D. J. Wales, J. Chem. Soc., Faraday Trans. **86**, 3505 (1990).
[19]D. J. Wales, Mol. Phys. **74**, 1 (1991).
[20]L. J. Munro and D. J. Wales, Phys. Rev. B **59**, 3969 (1999).
[21]Y. Kumeda, D. J. Wales, and L. J. Munro, Chem. Phys. Lett. **341**, 185 (2001).
[22]D. J. Wales, *Energy Landscapes* (Cambridge University Press, Cambridge, UK/New York, 2003).
[23]G. Henkelman and H. Jónsson, J. Chem. Phys. **111**, 7010 (1999).
[24]R. A. Olsen, G. J. Kroes, G. Henkelman, A. Arnaldsson, and H. Jónsson, J. Chem. Phys. **121**, 9776 (2004).
[25]A. Heyden, A. T. Bell, and F. J. Keil, J. Chem. Phys. **123**, 224101 (2005).
[26]J. Käestner and P. Sherwood, J. Chem. Phys. **128**, 014106 (2008).
[27]J. Kästner *et al.*, J. Phys. Chem. A **113**, 11856 (2009).
[28]H.-F. Wang and Z.-P. Liu, J. Am. Chem. Soc. **130**, 10996 (2008).
[29]C. Shang and Z.-P. Liu, J. Chem. Theory Comput. **6**, 1136 (2010).
[30]W. E and X. Zhou, Nonlinearity **24**, 1831 (2011).
[31]A. Samanta and W. E, J. Chem. Phys. **136**, 124104 (2012).
[32]J. Y. Zhang and Q. Du, J. Comput. Phys. **231**, 4745 (2012).
[33]J. M. Carr and D. J. Wales, J. Chem. Phys. **123**, 234901 (2005).
[34]T. James and D. J. Wales, J. Chem. Phys. **122**, 134306 (2005).
[35]A. V. Knyazev, SIAM J. Sci. Comput. **23**, 517 (2001).
[36]E. E. Ovtchnnkov, SIAM J. Numer. Anal. **46**, 2567 (2008).
[37]J. Baker and F. Chan, J. Comput. Chem. **17**, 888 (1996).
[38]G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).