Journal of Computational Physics 309 (2016) 69–87

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Iterative minimization algorithm for efficient calculations of transition states

Weiguo Gao [a,b], Jing Leng [a], Xiang Zhou [c,*]

[a] School of Mathematical Sciences, Fudan University, Shanghai, 200433, China
[b] MOE Key Laboratory of Computational Physical Sciences, Fudan University, Shanghai, 200433, China
[c] Department of Mathematics, City University of Hong Kong, Tat Chee Ave, Kowloon, Hong Kong

ARTICLE INFO

Article history:
Received 13 January 2015
Received in revised form 25 September 2015
Accepted 29 December 2015
Available online 4 January 2016

Keywords:
Saddle point
Transition states
Energy landscape
Eigenvector-following
Iterative minimization

ABSTRACT

This paper presents an efficient algorithmic implementation of the iterative minimization formulation (IMF) for fast local search of transition state on potential energy surface. The IMF is a second order iterative scheme providing a general and rigorous description for the eigenvector-following (min-mode following) methodology. We offer a unified interpretation in numerics via the IMF for existing eigenvector-following methods, such as the gentlest ascent dynamics, the dimer method and many other variants. We then propose our new algorithm based on the IMF. The main feature of our algorithm is that the translation step is replaced by solving an optimization subproblem associated with an auxiliary objective function which is constructed from the min-mode information. We show that using an efficient scheme for the inexact solver and enforcing an adaptive stopping criterion for this subproblem, the overall computational cost will be effectively reduced and a super-linear rate between the accuracy and the computational cost can be achieved. A series of numerical tests demonstrate the significant improvement in the computational efficiency for the new algorithm.

1. Introduction

Many phenomena in physics, material sciences, chemistry and biology can be abstractly formulated as a system that navigates over a complex potential energy landscape of high or infinite dimensions. Rare events, which correspond to infrequently hops between different local minima of the potential function, have attracted extensive research work to understand quite many important physical processes in natural sciences and engineering. The examples include chemical reactions, phase transitions of condensed matter, etc. To study rare event, one of the fundamental questions is to find energy barriers and transition states. The barrier is the energy difference between a local minimum and its transition state. The transition state is a critical point on the potential energy surface with exactly one negative Hessian eigenvalue. We call this type of saddle points as index-1 saddle points. Furthermore, finding saddle points provides valuable information for many bifurcation problems, especially for subcritical bifurcations, since the unstable manifold of index-1 saddle point connects two locally stable solutions.

While locating potential energy minima can routinely be done, at least for local search, thanks to the significant progress of nonlinear optimization, finding saddle points can be extremely difficult and remains one of the major challenges for large

* Corresponding author.
E-mail address: xiang.zhou@cityu.edu.hk (X. Zhou).

http://dx.doi.org/10.1016/j.jcp.2015.12.056
0021-9991/

systems. Several methods exist to locate the transition states. They can be divided into surface-walking algorithms, which use only local quantities around one point on the potential energy surface, and the chain-of-states methods that connect different points on the surface. The nudged elastic band method [1] is an example of the chain-of-state methods and the string method [2–4] presents an elegant dynamical description and a convenient numerical technique in the path space for such chain-of-states methods. For the surface-walking algorithms, the early developments include quasi-Newton methods introduced by Cerjan and Miller [5] and later modified by Simons and co-workers and Wales [6]. One of these methods widely used for *ab initio* molecular calculations is the partitioned rational function optimization algorithm of Baker [7], which utilizes either an approximate or exact Hessian matrix. For large systems exact Hessians have to be calculated regularly. If the Hessian matrix cannot be determined analytically, the second derivative matrix has to be determined numerically. For large systems this step often becomes prohibitively expensive.

There has been an idea of "eigenvector-following" methodology which uses only the min-mode of the Hessian matrix and moves the system uphill along this min-mode direction. The min-mode is the eigenvector corresponding to the minimal Hessian eigenvalue. This idea of gently ascending potential energy surface following such eigenvectors can be dated back to Crippen and Scheraga in 1971 [8]. There have been significant developments of new algorithms based on eigenvector-following methodology, for instance, the dimer method [9] and its many variants [10–12], the activation–relaxation techniques and the variants [13–15]. Instead of calculating the full Hessian matrix these algorithms calculate only the lowest eigenvalue and the corresponding eigenvector by finite difference scheme or Lanczos methods. Numerous applications for practical problems have proven that these eigenvector-following methods generally run faster and converge better than the previous Newton–Raphson root-finding methods.

Along with the substantial progress in algorithmic developments and applications of this popular eigenvector-following methodology, a first rigorous mathematical analysis for the local convergence to index-1 saddle point is established in [16] by formulating eigenvector-following methodology as a coupled dynamical system, with the name "gentlest ascent dynamics" (GAD). Index-1 saddle point on the potential energy surface becomes a locally stable equilibrium point in the GAD, thus the local (linear) convergence to index-1 saddle points is guaranteed. The applications of the GAD include [17] and [18]. A different but similar dynamical system named shrinkage dimer dynamics is also pursued in [19] by introducing one more dimer length variable. The method in [20] and the application in [21] combined a modified string method technique in which one end of the string follows a modified dynamics and then the inexact Newton method.

Based on the work of the GAD, we recently proposed a new description of the eigenvector-following' methodology in [22]. This new model is an iterative mapping and is named as "iterative minimization formulation", or IMF in short. In each iteration of the IMF, an auxiliary function is constructed as a new objective function, based on the local quantities around the current point on the potential energy. Then a local minimizer of this objective function is assigned as the new position for the next iteration. Theoretically, this iterative scheme can be locally described by a continuously differentiable mapping near the saddle point. We proved that the iterative mapping defined in such ways has quadratic convergence rate. This rate is the best rate for all numeral schemes based on the eigenvector-following methodology by using only the min-mode.

In the work of the IMF [22], we did not specify how to solve the subproblem of minimizing auxiliary objective function at each iteration, which is a crucial step in computations. This article is to address this practical issue of algorithm design for the IMF. It is already observed in [22] that the GAD is equivalent to solving the minimization subproblem in the IMF by using a single steepest descent step. We continue in this article to reveal the connections of other eigenvector-following based methods to the traditional optimization schemes for the subproblem in the IMF. See the detailed discussion in the Appendix A. So, the IMF is a quite general and rigorous mathematical description and offers a unified description of various algorithms related to the eigenvector-following methodology.

The contribution of this paper is to present an efficient and easy-to-implement algorithm for the IMF. To alleviate the computational bottleneck for the IMF, we propose an adaptive stopping rule for solving the subproblem of minimizing the new objective function at each iteration. Meanwhile, we propose an efficient proposal for the calculation of the gradient of the new objective function, which requires the minimal number of force calculations of the original potential function. Therefore, with our stopping rule and gradient calculation scheme, any gradient-based optimization solver can be exploited in principle as a good inexact solver for the subproblem. We call our proposed algorithm as "iterative minimization algorithm" (IMA). The main advantage of the IMA is its efficiency: on the one side, it explores the quadratic iterative rate of the IMF and on the other side, it minimizes the computational cost for the subproblem solver. As a result, the IMA not only needs less number of the computation-intensive rotation steps, but also reduces the overall computational cost to achieve a prescribed accuracy.

The rest of this paper is organized as follows. In Section 2, we review the iterative minimization formulation Our main work of the new algorithm is presented in Section 3. Section 4 includes several numerical examples. Section 5 is the concluding summary. The Appendix A is about the analysis of a series of existing algorithms by the perspective of the IMF.

2. Iterative minimization formulation

2.1. The problem

Let M be a Hilbert space with norm $|\cdot|$ and inner product $\langle\cdot,\cdot\rangle$ in its tangent space T_M. Given a differentiable (and sufficiently smooth) potential function $V : M \to \mathbb{R}$. The vector field ∇V generates the gradient descent flow on M,