Curved-line search algorithm for ab initio atomic structure relaxation

Zhanghui Chen, Jingbo Li, Shushen Li, and Lin-Wang Wang

Phys. Rev. B **96**, 115141 — Published 21 September 2017

DOI: 10.1103/PhysRevB.96.115141

A curved line search algorithm for *ab initio* atomic structure relaxation

Zhanghui Chen,[1] Jingbo Li,[2] Shushen Li,[2] and Lin-Wang Wang[1,*]

[1] *Materials Sciences Division, Lawrence Berkeley National Laboratory, One Cyclotron Road, Mail Stop 50F, Berkeley, California 94720, United States*

[2] *State Key Laboratory of Superlattices and Microstructures, Institute of Semiconductors, Chinese Academy of Sciences, P.O. Box 912, Beijing 100083, People's Republic of China*

(Dated: September 8, 2017)

Ab initio atomic relaxations often take large numbers of steps and long time to converge, especially when the initial atomic configurations are far from the local minimum or there are curved and narrow valleys in the multi-dimensional potentials. An atomic relaxation method based on on-the-flight force learning and a corresponding new curved line search algorithm is presented to accelerate this process. Results demonstrate the superior performance of this method for metal clusters when compared with the conventional conjugate-gradient method.

One major usage of *ab initio* density functional theory (DFT) in material science simulation is to determine the ground state atomic configuration for a given system [1, 2]. Overall, such applications probably take most of the DFT simulation time. There are two types of ground structure searching. The first is to find global minimum among many local minima [3, 4]. This has become an intensely studied topic in material design projects [5–9]. Various types of evolutionary algorithms [6–10] or simulated annealing [11] schemes have been developed, as well as the minimum hopping methods [5]. The second type is the conventional local minimum optimization, which is the concern of the current study. The related calculation is dominated by the conjugated gradient (CG) method [12–14] and the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method [15–17]. Although these methods guarantee to converge into a local minimum, the convergence rate could be agonizingly slow, e.g. with hundreds of steps, thus a faster method will be extremely helpful. This local minimum problem also presents itself in the global minimum search since each global minimum search step usually deploys one or more local minimizations [5–9]. One reason for the slow convergence of the local minimization steps is the possible narrow and curved energy valley leading to the minimum, which prevents the efficient execution of the conventional CG or BFGS methods. Imaging a rotation of a molecule on the surface of a substrate. Such rotation cannot be described by a straight line in cartesian coordinates which is used under CG or BFGS methods. In higher dimension, the situation can be more complicated, making it impossible to find the natural degree of freedom (e.g., the rotation angle). One such example is a metal cluster [4, 18, 19] (which will be studied in this paper), where hundreds of steps might be needed to relax a structure while there is no obvious natural (or say internal) degree of freedom to speed up the convergence. To overcome these problems, one needs to do the minimization steps along guided curved lines following the energy valleys. We will call such algorithms the guided curved-line-search (CLS) algorithms.

The issue is how to find such guided curved lines. In this work, we will show that such guided curved line can be provided by model surrogate potentials with their parameters provided by on-the-flight fitting (OTFF) to the *ab initio* atomic forces [20–22]. We will demonstrate the efficiency of our CLS algorithm on metal clusters. Overall, we have found that the CLS method can speed up the traditional CG method by a factor of 2 to 4 for both the number of steps and wall clock times, in problems (Pt, Co, CuAu clusters) with initial configurations far from minima or with narrowly curved energy valley. CLS as well as our previous modified pre-conditioned CG algorithms for problems with ill-conditioned Hessian matrix [23], also demonstrated that OTFF can be effectively used to speed up the atomic relaxations, not just molecular dynamics as it has been usually used [20–22].

As mentioned before, the guided curved line will be provided by a surrogate potential. One possible option is to carry out *ab initio* line minimization along the steepest descent line (SDL) of this surrogate potential. We will use on-the-flight fitting (OTFF) to ensure that the atomic forces of this surrogate potential at the beginning of each step equal that of the *ab initio* forces. When the system approaching the final minimum point, the curved line will become straight in the small scale, then the curved line search will go back to the conventional straight line search. In practice, we found that the SDL can be warped with sharp twists in high dimensions. Besides, using SDL will miss the conjugated gradient feature between different line searches. To overcome these shortcomings, we will use the surrogate potential conjugate gradient descent line (SP-CGDL). To construct SP-CGDL, the conventional CG formalism is applied to the initial atomic force direction to yield the CG search direction. Then a straight line minimum search based on the surrogate potential is carried out. From the new line minimum point of the current surrogate function, subsequent CG straight lines are carried out. Thus, our SP-CGDL curved search line is consisted with many straight lines segments of the surrogate potential CG path. The *ab initio* line minimization will be carried out along this SP-CGDL on DFT energy landscape. One might worry that the *ab initio* energy function along this segmented line might not be smooth enough to carry out *ab initio* line minimization. But in practice, we found that one can effectively use the Brents algorithm [24] to search for the *ab initio* line minimum along this SP-CGDL, and such line search often finds the minimum at a few segments down the road along the SP-CGDL. Typically two *ab initio* calculations are needed in the Brents algorithm to search for the minimum along the SP-CGDL [25], much like the conventional line minimization calculation. After the *ab initi*o line minimization are done, we call this one step, and the algorithm will repeat itself (from OTFF to construction of SP-CGDL, then *ab initio* line minimization). The overall flow chart of this CLS algorithm is shown in Figure 1. Note, the above procedure maintains the feature of CG, when close to the minimum.

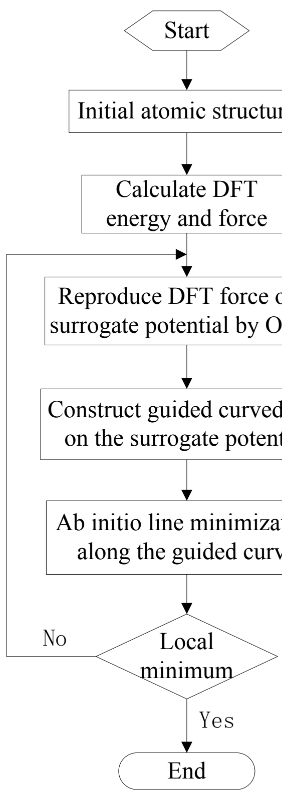

FIG. 1. The flow chart of the curved-line-search (CLS) algorithm. OTFF indicates on-the-flight fitting.

We will use metal cluster [4, 18, 19] to demonstrate our CLS algorithm. The metal cluster potential is intrinsically high dimensional due to their long range atom-atom interaction. As a result, it is often difficult to reach their local minima. The metal cluster is an important subfield related to catalysts [4, 10, 18]. A lot of works have been done in searching of the optimal cluster structures, and the density of local minima in energy [4, 5, 10, 18, 19, 26–28]. For metallic systems, we found that the N-body Gupta force field [29] is a very good general potential. It has been used to model various types of metal clusters. The potential is a special case of the embedded atom potential [30] based on the second moment approximation of the tight binding theory and it has the following form:

$$E_N = \sum_{i=1}^{N} \left\{ \sum_{j=1(j\neq i)}^{N} A_{ij} \exp\left(-p_{ij}\left(\frac{r_{ij}}{r_{ij}^0}-1\right)\right) - \left[\sum_{j=1(j\neq i)}^{N} \xi_{ij}^2 \exp\left(-2q_{ij}\left(\frac{r_{ij}}{r_{ij}^0}-1\right)\right)\right]^{1/2} \right\}$$

where r_{ij} represents the distance between the atom i and j in the cluster. The five groups of parameters A_{ij}, ξ_{ij}, p_{ij}, q_{ij}, r_{ij}^0 are allowed to vary independently to match the forces from DFT calculations. We restrict $A_{ij} = A_{ji}$, $\xi_{ij} = \xi_{ji}$, $p_{ij} = p_{ji}$, $q_{ij} = q_{ji}$ and $r_{ij}^0 = r_{ji}^0$, and set them to zero when the r_{ij} is larger than a cut-off distance to limit the number of the variables (see Supplemental Material [31]). Parameters are also restricted to vary within a physically meaningful range.

Because the analytic expression for atomic forces of this model is a non-linear function of these parameters, in order to have an accurate force fitting, we have used a parallel differential evolutional algorithm [10] to globally minimize the force error. The resulting best solution is further optimized by a CG local minimization algorithm for these parameters. This approach enable us to always fit the atomic forces with an error less than 0.005 eV/Å, which is a few times lower than the typical *ab initio* minimization stoping criterion. Although the fitting procedure (at the beginning of every *ab initio* line minimization step) might sound complicated, its computational cost is negligible, about 5% of the *ab initio* computational time.

To show the quality of atomic force fitting, we present the atomic force error in Figure 2 for a Pt_{100} cluster (with 100 Pt atoms). To begin with, we use the Gupta parameters from Ref. 29, 32, and 33 which have the parameters for almost all the major metallic elements. The atomic force error compared to *ab initio* calculation using these original parameters without fitting is about 1 eV/Å. After the parameter fitting, they becomes about 10^{-3} eV/Å. This improvement on the force is at no cost of degradation of other properties of this potential. For example, Figure 2(c),(d) compare the atomic force changes between the Gupta and DFT results when the atomic positions have been randomly displaced. The original Gupta result is already rather good, and it has been improved after force fitting. Such data to some extent shows the second-order differential information of energy landscape.

To demonstrate the speedup of the CLS method, we first test five random Pt_{20} clusters [10] with different initial structures and corresponding different initial energy.

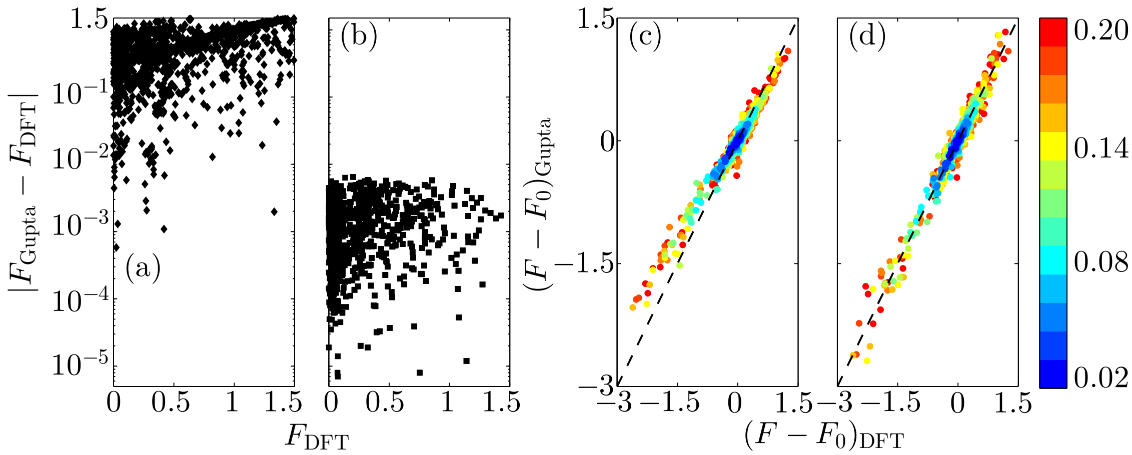

FIG. 2. (a-b) Deviation of atomic forces for Pt_{100} between DFT calculations and approximate models: (a) conventional un-fitted Gupta potential; (b) force-fitted Gupta potential. Their x-axis is DFT force $F_{\rm DFT}$ while y-axis is the deviation $|F_{\rm Gupta} - F_{\rm DFT}|$ (in unit of eV/Å). (c-d) Comparisons of the values of force difference $F - F_0$ between DFT and approximate models: (c) un-fitted potential; (d) force-fitted potential, where F_0 is the force at the atomic structures (R_0) used for force fitting and F is the one at randomly displaced structures (R) around R_0. Their x-axis is DFT force difference $(F - F_0)_{\rm DFT}$ while y-axis is the approximate one (in unit of eV/Å). Color bar for (c) and (d) indicates the distance (in unit of Å) of R from R_0. Note that the plot contains several different groups of R_0.

The convergence results are shown in Figure 3 in comparison with the conventional CG results. We see that, the initial energy is more than 10 eV higher than their local minima energy, indicating these initial structures are far from stable minima. Such situation is very common in structure searches. Their CG relaxations need more than 200 *ab initio* force evaluations. However, by using the new CLS algorithm, a factor of 2-4 speedup are achieved for most cases. Especially, a factor of 3-6 speedup are achieved in the initial relaxation steps. Note, the wall clock time of other operations is much less than the one of DFT force evaluations. In our test, one force fitting costs about 5% time of one DFT force evaluation, and all the other lines of codes including construction of SP-CGDL cost less than 1% since these operations are performed on classical force field. Each *ab initio* line minimization usually calls about two DFT calculations, thus the total computational time in CLS is about 3% higher than the one in CG at the same number of DFT calculations. This demonstrate that CLS can speedup the relaxation for both the number of steps and wall clock time.

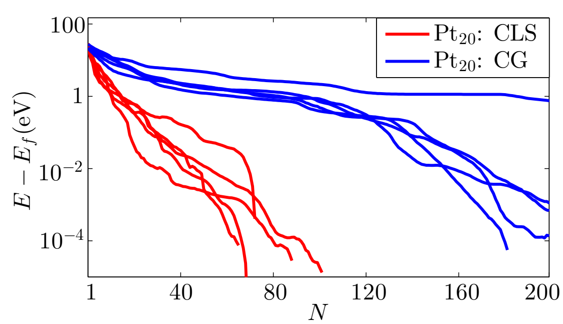

FIG. 3. The relaxation process of five random Pt_{20} clusters with different initial structures and corresponding different initial energies by the CLS and CG method, respectively. The x-axis is the number (N) of *ab initio* force evaluation while the y-axis is $E - E_f$ (in unit of eV), where E is the energy of the current step and E_f is the energy of the finally sought structure.

In actual work, one often uses Gupta to pre-relax the initial atomic cluster to a Gupta local minimum, then uses conventional *ab initio* CG relaxation to further relax the total energy of the system. We will call such scheme pre-CG. One can also start with the Gupta relaxed minimum, then use our CLS method, we will call such method pre-CLS. Their results for these Pt_{20} clusters and a bigger-size Pt_{100} cluster are shown in Figure 4. We can see that, due to the good approximation of Gupta to DFT energy, the initial energy after pre-relaxation is closer to the local minima, compared to the one of Figure 3. Such pre-relaxation does provide a good initial speedup for *ab initio* energy minimization, e.g., the $E - E_f$ converges faster to 10^{-1} level. However, the subsequent relaxation is as slow as the one without pre-relaxation. This is due to the issue mentioned earlier. There are no straight lines connecting these configurations to their local minima. The relaxation path could be twisted or curved in the energy landscape, making the CG method very inefficient. We see that, by using CLS method, the relaxation is much faster, and pre-CLS out performs pre-CG by a factor of 2 to 4. For the global search problems, or to search for local minima density, one issue is that the pre-CG or pre-CLS method tends to mislead the system to the same local minimum near the

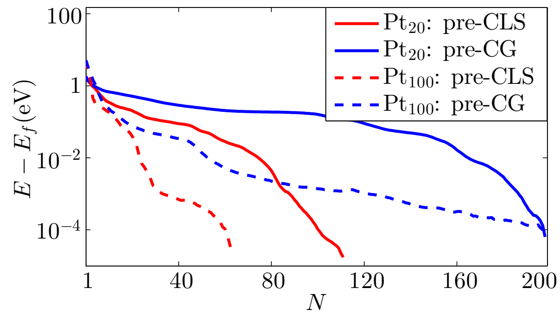

FIG. 4. The mean energy evolution in the *ab initio* relaxation of five pre-relaxed Pt_{20} clusters, and the relaxation of Pt_{100} cluster. The x-axis is the number (N) of *ab initio* force evaluation while the y-axis is $E - E_f$ (in unit of eV), where E is the energy of the current step and E_f is the energy of the finally sought structure.

Gupta potential basins for different initial configurations. This reduces the diversity of the global search and might miss the true DFT global minimum. On the other hand, the CLS method without pre-relaxation is exempted from such a problem. Since the cost of CLS is only modestly increased compared to pre-CLS, we suggest CLS in both local relaxation and global search problems.

Finally, we performed CLS relation on other types of atomic clusters. The results for Co_{120} and $Cu_{20}Au_{18}$ alloy clusters are shown in Figure 5. The initial geometries for the two clusters are from the global minimum of conventional un-fitted Gupta force field. We can see that a factor of 2 to 4 speedup is achieved, demonstrating the generality of CLS algorithm for metallic cluster systems.

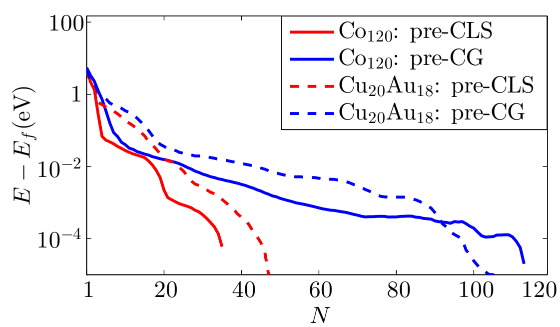

FIG. 5. The relaxation process of Co_{120} and $Cu_{20}Au_{18}$ clusters with the initial geometry from the global minimum of conventional un-fitted Gupta force field. The x-axis is the number (N) of *ab initio* force evaluation while the y-axis is $E - E_f$ (in unit of eV), where E is the energy of the current step and E_f is the energy of the finally sought structure.

We have also considered another two methods, i.e., BFGS [15, 34] and Fast Inertial Relaxation Engine (FIRE) [35], which have been used in the acceleration of atomic relaxations. BFGS updates the Hessian matrix as a pre-condition in the quasi-Newton optimization. FIRE makes use of inertia of molecular dynamics in the Newton optimization. We have found that both BFGS and FIRE are not significantly better than CG in the relaxations of Pt_{100}, Co_{120} and $Cu_{20}Au_{18}$ clusters and they are both inferior when compared to our CLS method (see Supplemental Material [36]).

In summary, we have presented a curved line search (CLS) algorithm to speed up *ab initio* atomic structure relaxation. This CLS uses a classical potential to provide the curved line on which *ab initio* line minimization is carried out. The parameters of this classical potential are fitted on-the-flight at every step to the *ab initio* atomic forces. We tested this approach using metal clusters with Gupta force field as the classical potential and we expect similar approaches can be applied to other systems. Compared to the traditional methods, we found CLS can speed up by a factor of 2-4. The CLS method is expected to be useful for general optimization problems.

This work was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division, of the under Contract No. DE-AC02-05-CH11231 within the Non-Equilibrium Magnetic Materials program (MSMAG). This research used the resources of the National Energy Research Scientific Computing Center (NERSC) and Oak Ridge Leadership Computing Facility (OLCF) that are supported by the Office of Science of the U.S. Department of Energy, with the computational time allocated by the Innovative and Novel Computational Impact on Theory and Experiment (INCITEE) project.

* lwwang@lbl.gov

[1] N. Schuch and F. Verstraete, Nat. Phys. **5**, 732 (2009).
[2] T. L. Beck, Rev. Mod. Phys. **72**, 1041 (2000).
[3] D. J. Wales, Science **293**, 2067 (2001).
[4] D. J. Wales and H. A. Scheraga, Science **285**, 1368 (1999).
[5] S. Goedecker, W. Hellmann, and T. Lenosky, Phys. Rev. Lett. **95**, 055501 (2005).
[6] L. Zhang, J.-W. Luo, A. Saraiva, B. Koiller, and A. Zunger, Nat. Commun. **4** (2013).
[7] L. B. Vilhelmsen and B. Hammer, Phys. Rev. Lett. **108**, 126101 (2012).
[8] Q. Li, D. Zhou, W. Zheng, Y. Ma, and C. Chen, Phys. Rev. Lett. **110**, 136403 (2013).
[9] C.-H. Hu, A. R. Oganov, Q. Zhu, G.-R. Qian, G. Frapper, A. O. Lyakhov, and H.-Y. Zhou, Phys. Rev. Lett. **110**, 165504 (2013).
[10] Z. Chen, X. Jiang, J. Li, S. Li, and L. Wang, J. Comp. Chem. **34**, 1046 (2013).
[11] K. Doll, J. C. Schön, and M. Jansen, Phys. Rev. B **78**, 144110 (2008).
[12] J. R. Shewchuk, *An Introduction to the Conjugate Gradient Method Without the Agonizing Pain*, Tech. Rep. (Pittsburgh, PA, USA, 1994).
[13] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).

[14] M. D. Segall, P. J. D. Lindan, M. J. Probert, C. J. Pickard, P. J. Hasnip, S. J. Clark, and M. C. Payne, J. Phys.: Condens. Matter. **14**, 2717 (2002).
[15] D. C. Liu and J. Nocedal, Math. Prog. **45**, 503 (1989).
[16] X. Gonze, J.-M. Beuken, R. Caracas, F. Detraux, and M. Fuchs, Comp. Mater. Sci. **25**, 478 (2002).
[17] P. Giannozzi, S. Baroni, and N. Bonini, J. Phys.: Condens. Matter. **21**, 395502 (2009).
[18] X. Lai, R. Xu, and W. Huang, J. Chem. Phys. **135**, 164109 (2011).
[19] L. X. Zhan, J. Z. Y. Chen, W. K. Liu, and S. K. Lai, J. Chem. Phys. **122**, 244707 (2005).
[20] G. Csányi, T. Albaret, M. C. Payne, and A. De Vita, Phys. Rev. Lett. **93**, 175503 (2004).
[21] Y. Lee and G. S. Hwang, Phys. Rev. B **85**, 125204 (2012).
[22] X. Zhang, Q. Peng, and G. Lu, Phys. Rev. B **82**, 134120 (2010).
[23] Z. Chen, J. Li, S. Li, and L. Wang, Phys. Rev. B **89**, 144110 (2014).
[24] R. P. Brent, *Algorithms for minimization without derivatives* (Courier Dover Publications, 2013).
[25] A. Wächter and L. T. Biegler, Math. Prog. **106**, 25 (2006).
[26] V. Kumar and Y. Kawazoe, Phys. Rev. B **77**, 205418 (2008).
[27] L. Xiao and L. Wang, J. Phys. Chem. A **108**, 8605 (2004).
[28] X. Wang and D. Tian, Comput. Mater. Sci. **46**, 239 (2009).
[29] K. Michaelian, N. Rendón, and I. L. Garzón, Phys. Rev. B **60**, 2000 (1999).
[30] M. S. Daw and M. I. Baskes, Phys. Rev. B **29**, 6443 (1984).
[31] See Supplemental Material at [URL will be inserted by publisher] for the details of cut-off distance. ().
[32] R. Ismail, Theoretical studies of free and supported nanoalloy clusters, University of Birmingham (2013).
[33] F. Cleri and V. Rosato, Phys. Rev. B **48**, 22 (1993).
[34] J. Sherman and W. J. Morrison, "Adjustment of an inverse matrix corresponding to a change in one element of a given matrix," (1950).
[35] E. Bitzek, P. Koskinen, F. Gähler, M. Moseler, and P. Gumbsch, Phys. Rev. Lett. **97**, 170201 (2006).
[36] See Supplemental Material at [URL will be inserted by publisher] for the results and discussions of BFGS and FIRE relaxations. ().