Towards Higher Scalability than Embarrassing Parallelism: Some Thoughts on Ab Initio Semiclassical Molecular Dynamics Simulations on Exascale Machines

Yu Zhuang[1], Michele Ceotto[2], William L. Hase[3], and Wibe A. de Jong[4]
[1] Computer Science Department, Texas Tech University, Lubbock, Texas 79409
[2] Dipartimento di Chimica, Università degli Studi di Milano, 20133 Milano, Italy
[3] Chemistry & Biochemistry Department, Texas Tech University, Lubbock, Texas 79409
[4] Environmental Molecular Science Laboratory, Pacific Northwest National Laboratory, Richland, WA 99352

Introduction

Semiclassical molecular dynamics simulations [1-6] study the atomic level motions, and provide a generally applicable tool for capturing quantum dynamics effects, making them essential for a broad range of studies in chemical and materials science [6-20]. A semiclassical theory that has proven to be one of the most successful in describing quantum effects is the coherent states implementation of the semicalssical-initial value representation (SC-IVR) [21-26], which goes under the name of Heller-Herman-Kluk-Kay (HHKK) SC-IVR approximation. The HHKK SC-IVR approximation of the quantum mechanical evolution operator needs the monodromy matrix [26], whose integration requires the Hessian at each time step. Thus, the trajectory calculation for a semiclasscial simulation needs both the gradient and Hessian at each time step. SC-IVR dynamics depends only on the local potential, i.e. it can be performed "on-the-fly" as a direct ab initio molecular dynamics method. This aspect is of fundamental relevance, as a major challenge in quantum dynamics is the simulation of complex systems where the development of an analytic potential energy surface may be a formidable task. Ab initio SC-IVR dynamics is very promising for the quantum dynamics of complex systems but at a higher computational cost than SC-IVR on fitted surfaces [27-33].

Enormous Computation Cost

A quantum mechanics electronic structure calculation is required to obtain the gradient and Hessian of the potential energy. This calculation involves solving an eigen-related problem of the Schrodinger equation, and has an enormous computation cost (see the SIAM Review survey paper [34] by Saad et al., with description of the two approaches that won the Nobel Prize in Chemistry in 1998). With minutes to hours of execution time for an electronic structure calculation for a few atoms on a multi-core processor, a trajectory that has hundreds to thousands of time steps takes hours to months of computation time on a multi-core processor. In addition, a dynamics study usually requires calculating tens to thousands of trajectories.

Approaches to Reduction of Computation Time

There have been substantial, multidisciplinary efforts for reducing the computation time of these dynamics simulations. Semi empirical methods [35,36], frozen-core approximations [37], etc. [38] are chemical physics theories that target the reduction of the computation cost of electronic structure calculations. The computation time may be further reduced by parallelization of the electronic structure calculation. These approaches have been implemented in DOE-funded electronic structure calculation packages [39-41], including NWChem [39].

In addition to the aforementioned approaches, another obvious way for further reducing the computation time of ab initio dynamics simulations is to utilize the parallelism in the calculations of multiple trajectories, i.e. to let multiple trajectories be calculated on different and mutually disjoint sets of compute nodes, while the electronic structure calculation at each time step of a single trajectory is computed in parallel by all cores of the set of compute nodes assigned to compute the trajectory.

Thus, a dynamics simulation study can have two levels of parallelism: the coarse-grain parallelism offered by the parallelizable computations of multiple trajectories, and the parallelism in each electronic structure calculation that is performed at each time step of a trajectory. For existing simulation methods, these two levels of parallelism are almost all the parallelism that can be utilized by a parallel machine for an ab initio dynamics simulation. To our knowledge, little additional parallelization can be done for further measureable

computation time reduction. That is, we believe that the vast majority of the parallelism in the calculation of a single trajectory is in each electronic structure calculation (ESC). We believe this since (i) different electronic structure calculations are at different time steps and are inherently not parallel, and (ii) the cost of other computations in a trajectory calculation is negligible compared with that of ESCs.

Can We Do Better?

While the aforementioned two levels of parallelism is almost all the utilizable parallelism, further reduction of computation time without sacrificing the simulation quality calls for unconventional, or revolutionary, approaches. An approach we consider is one that destroys the embarrassing parallelism of multiple trajectory calculations by creating data sharing among the calculation of the trajectories.

The SC-IVR approximation of the quantum mechanical evolution operator needs the gradient and Hessian of the ab initio potential energy at each time step, where the Hessian calculation comprises a substantial fraction of the computation cost. Recently, this team of researchers invented a class of high-accuracy approximation methods for high dimensional data [42], and applied them to the Hessian approximation for semi-classical dynamics simulations [43,44]. With our Hessian approximation method, for every K time steps, the ab initio Hessian is calculated in the first step followed by K-1 steps of the Hessian approximation. Our on-going work for 10-atom glycine shows that the Hessian approximation with K = 10 produces simulation results almost identical to those obtained with a complete ab initio Hessian simulation. A glycine trajectory with the B3LYP/6-31G** method takes 11.5 days on two quad-core processors for 5000 time steps without the Hessian approximation, and the Hessian approximation with K = 10 reduces the computation time by over 6 times to less than 2 days.

Then the question is may the Hessian approximation be extended to exascale computing to enable simulations of larger ensembles of atoms. The Hessian approximation needs ab initio data that are adequately close so the approximation is accurate. With the independent and embarrassingly parallel computing of multiple trajectories, only previously calculated ab initio data for the same trajectory is used for each Hessian approximation. Sharing ab inito data between multiple trajectories destroys the embarrassing parallelism, but will lead to more and closer data points for a better approximation, offering the prospect of reducing the cost of the Hessian calculations to a level measurably lower than when each trajectory is calculated independently.

But critical issues remain in order to determine if the benefits of such data sharing exceed the incurred overheads. Issue 1 is that using ab initio data from multiple trajectories will improve the approximation accuracy, but will incur much higher communication cost if a large amount of data from the trajectories needs to be sent among processors. A solution to be investigated is to not send the data but send the trajectory calculation, which however calls for careful planning of the partition/distribution of ab initio data among different processors and the use of a good partition/distribution strategy based on the high-dimensional space of the nuclear coordinates.

Issue 2 is that by including more data from multiple trajectories, it takes more time to find data points for the approximation. The search index is the high-dimensional nuclear coordinates, so an efficient algorithm is needed to search such a high-dimensional data set, much larger than that for a single trajectory. Efficient search algorithms exist for one-dimensional data, e.g. binary search of 1-D sorted data. High-dimensional search algorithms also exist. Adaptation of existing, or invention of new, search algorithms need to be investigated to take care of the particular needs of the problem.

In addition to the above two overhead-related issues, another obvious one is the following. The approach we used for the Hessian approximation is Hessian updating, which approximates the Hessian in a fashion that one direction (or one dimension) of the Hessian is approximated with the closest available data while all other directions remain almost unchanged. This is a good high-dimensional data approximation when there is only sparse data available for the approximation. There are existing meshless methods for high dimensional data [45-48], and to the best of our knowledge none of them have been used for the Hessian approximation. While our analysis shows updating methods are more suitable for approximating high-order differentiation data than meshless methods, it is tempting to investigate if better Hessian updating methods can be invented that incorporate some of the good features of other high-dimensional data approximation methods.

References

[1] W. H. Miller, Classical S Matrix: Numerical Application to Inelastic Collisions, *J. Chem. Phys.*, Vol. 53, 3578-3587 (1970).

[2] W. H. Miller, Classical-Limit Quantum Mechanics and the Theory of Molecular Collisions, *Adv. Chem. Phys.*, Vol. 25, 69-177 (1974).

[3] W. H. Miller, Spiers Memorial Lecture, Quantum and semiclassical theory of chemical reaction rates, Faraday Discuss. Vol. 110, 1-21 (1998).

[4] W. H. Miller, The Semiclassical Initial Value Representation: A Potentially Practical Way for Adding Quantum Effects to Classical Molecular Dynamics Simulations, J. Phys. Chem. A, Vol.105, 2942-2955 (2001).

[5] Paranjothy, M., Sun, R., Zhuang, Y., Hase, W. (2012). "Direct chemical dynamics simulations: coupling of classical and quasiclassical trajectories with electronic structure theory". *Wiley Interdisciplinary Reviews (WIREs) Computational Molecular Science*. doi: 10.1002/wcms.1132.

[6] N. J. Wright and N. Makri, Forward–backward semiclassical dynamics for condensed phase time correlation functions, *J. Chem. Phys*. 119, 1634 (2003).

[7] S. Garashchuk, F. Grossmann and D. Tannor, Semiclassical approach to the hydrogen-exchange reaction and transition-state dynamics, J. Chem. Soc., Faraday Trans., Vol. 93 (5), 781-789 (1997).

[8] S. Garashchuk and D.J. Tannor, Correlation function formulation for the state selected total reaction probability, Chem. Phys. Vol. 109 (8), 3028-3026 (1998).

[9] A. R. Walton, D. E. Manolopoulos, Application of the frozen Gaussian approximation to the photodissociation of CO_2,Chem. Phys. Lett., Vol. 244 (5), 448-455 (1995).

[10] X. Sun and W. H. Miller, Semiclassical initial value representation for rotational degrees of freedom: The tunneling dynamics of HCl dimer, J. Chem. Phys., Vol. 108(21), 8870-8877 (1998).

[11] V. Guallar, V.S. Batista, and W.H. Miller. Semiclassical molecular dynamics simulations of excited state double-proton transfer in 7-azaindole dimers, J. Chem. Phys., Vol.110 (20), 9922-9936 (1999).

[12] E. A. Coronado, V. S. Batista, and W.H. Miller, Nonadiabatic photodissociation dynamics of ICN in the continuum: A semiclassical initial value representation study, J. Chem. Phys., Vol. 112(13), 5566-5575 (2000).

[13] Michael Thoss, William H. Miller, and Gerhard Stock, Semiclassical description of nonadiabatic quantum dynamics: Application to the S1–S2 conical intersection in pyrazine, J. Chem. Phys., Vol. 112 (23), 10282-10292 (2000).

[14] Michael F. Herman and David F. Coker, Classical mechanics and the spreading of localized wave packets in condensed phase molecular systems, *J. Chem. Phys.*, Vol. 111 (5), 1801-1807 (1999).

[15] C. J. Margulis, D. A. Horner, S. Bonella, and D. F. Coker, Vibrational dynamics of the I3 radical: A semiempirical potential surface, and semiclassical calculation of the anion photoelectron spectrum, *J. Phys. Chem. A*, Vol. 103 (47), 9552–9563 (1999).

[16] B. G. Alberding, M. H. Chisholm1, J. C. Gallucci, Y. Ghosh, and T. L. Gustafson, Electron delocalization in the S1 and T1 metal-to-ligand charge transfer states of trans-substituted metal quadruply bonded complexes, *PNAS*, Vol. 108 (2011), no.20, 8152–8156.

[17] L.J. Leger and J.T. Visentine. A consideration of atomic oxygen interactions with the space station. *J. Spacecraft & Rockets*, 23 (1986), 505–511.

[18] R.H. Notter. *Lung surfactants: basic science and clinical applications.* Macel Dekker, New York, 2000.

[19] A. Podgorski, T.R. Sosnowski, an L. Gradon. Deactivation of the pulmonary surfactant dynamics by toxic aerosols and gases. *J. Aerosol. Med.*, Vol. 14 (2001), 455.

[20] C.W. Spicer, R.A. Plastridge, B.J. Finlayson-Pitts, A.M. Grannas, and P.B. Shepson. Molecular halogens before and during ozone depletion events in the Arctic at polar sunrise: concentrations and sources. *Atmospheric Environment,* Vol. 36 (2002), pp. 2721–2731.

[21] E. J. Heller, Time-dependent approach to semiclassical dynamics, J. Chem. Phys., Vol. 62, 1544-1555 (1975).

[22] E. J. Heller, Frozen Gaussians: A very simple semiclassical approximation, J. Chem. Phys., Vol.75, 2923-2931 (1981).

[23] M. F. Herman and E. Kluk, A semiclasical justification for the use of non-spreading wavepackets in dynamics calculations, Chem. Phys., Vol. 91,27-34 (1984).

[24] K. G. Kay, Integral expressions for the semiclassical time-dependent propagator,J. Chem. Phys.,100,4377-4392 (1994).

[25] K. G. Kay, Numerical study of semiclassical initial value methods for dynamics, J. Chem. Phys.,100,4432-4445 (1994).

[26] R. Gelabert, X. Giménez, M. Thoss, H. Wang, and W.H. Miller, A Log-Derivative Formulation of the Prefactor for the Semiclassical Herman-Kluk Propagator, *J. Phys. Chem. A,* Vol. 104, 10321-10327 (2000).

[27] M. Ceotto, S. Atahan, G.F. Tantardini, and A. Aspuru-Guzik, Multiple coherent states for first-principles semiclassical initial value representation molecular dynamics, *J. Chem. Phys*. 130, 234113, (2009).

[28] M. Ceotto, D. dell'Angelo, and G.F. Tantardini, Multiple coherent states semiclassical initial value representation spectra calculations of lateral interactions for CO on Cu(100), *J. Chem. Phys*. 133, 054701 (2010).

[29] M. Ceotto, S. Valleau, G.F. Tantardini, and A. Aspuru-Guzik, First principles semiclassical calculations of vibrational eigenfunctions, *J. Chem Phys*. 134, 234103 (2011)

[30] M. Ceotto, G.F. Tantardini, and A. Aspuru-Guzik, Fighting the curse of dimensionality in first-principles semiclassical calculations: Non-local reference states for large number of dimensions, *J. Chem. Phys*. 135, 214108 (2011).

[31] M. Ceotto, S. Atahan, S. Shim, G.F. Tantardini, and A. Aspuru-Guzik, First-principles semiclassical initial value representation molecular dynamics, *Phys. Chem. Chem. Phys*. 11, 3861 (2009).

[32] J. Tatchen and E. Pollak, Semiclassical on-the-fly computation of the S0→S1 absorption spectrum of formaldehyde, *J. Chem. Phys*., 2009, 130, 041103.

[33] Wong, S.Y.Y.; Benoit, D.M.; Lewerenz, M.; Brown, A.; Roy, P.-N., Determination of molecular vibrational state energies using the ab initio semiclassical initial value representation: Application to formaldehyde, *J. Chem. Phys*. 2011, 134, 094110.

[34] Y. Saad, J. Chelikowsky, and S. Shontz, Numerical methods for electronic structure calculations of materials, *SIAM Review*, Vol. 52 (2010), no. 1, pp. 3-54.

[35] Michael J. S. Dewar and Walter Thiel (1977). Ground states of molecules. 38. The MNDO method. Approximations and parameters. *The Journal of the American Chemical Society* (ACS Publications), Vol.99 (15): 4899–4907.

[36] James J. P. Stewart (1989). Optimization of parameters for semiempirical methods I. Method. *The Journal of Computational Chemistry* (Wiley InterScience) 10 (2): 209–220. doi:10.1002/jcc.540100208.

[37] U. von Barth and C. D. Gelatt, Validity of the frozen-core approximation and pseudopotential theory for cohesive energy calculations, Phys. Rev. B, Vol. 21, (1980), 2222-2228.

[38] A. Kahros and B.J. Schwartz, Going beyond the frozen core approximation: Development of coordinate-dependent pseudopotential and application to Na^+. J. Chem. Phys., Vol. 138 (2013), 054110.

[39] M. Valiev, E.J. Bylaska, N. Govind, K. Kowalski, T.P. Straatsma, H.J.J. van Dam, D. Wang, J. Nieplocha, E. Apra, T.L. Windus, W.A. de Jong, NWChem: a comprehensive and scalable open-source solution for large scale molecular simulations, *Comput. Phys. Commun*. Vol. 181, 1477 (2010).

[40] M.W.Schmidt, K.K.Baldridge, J.A.Boatz, S.T.Elbert, M.S.Gordon, J.H.Jensen, S.Koseki, N.Matsunaga, K.A.Nguyen, S.J.Su, T.L.Windus, M.Dupuis, J.A.Montgomery, General Atomic and Molecular Electronic Structure System (GAMESS), *J.Comput.Chem*. Vol. 14, 1347-1363(1993), doi:10.1002/jcc.540141112.

[41] C.L. Janssen, I.M.B. Nielsen, M.L. Leininger, E.F. Valeev, J. Kenny, and E.T. Seidl, The Massively Parallel Quantum Chemistry program (MPQC), version 3.0.0-alpha. Sandia National Laboratories, Livermore, CA, 2007, http://www.mpqc.org.

[42] H. Wu, M. Rahman, J. Wang, U. Lourderaj, W. L. Hase, and Y. Zhuang, Higher-accuracy schemes for approximating the Hessian from electronic structure calculations in chemical dynamics simulations, *J. Chem. Phys*. 133, 074101-074110 (2010).

[43] Zhuang, Y., Ceotto, M., Siebert, M., Hase, W., Kay, K. (2012). Evaluating the Accuracy of Hessian Approximations for Direct Dynamics Simulations. *Journal of Chemical Theory and Computation,* Vol. 9, 54-64.

[44] M. Ceotto, W.L. Hase, Y. Zhuang, Accelerated ab initio Semiclassical molecular dynamics using a compact finite difference Hessian scheme, *Journal of Chemical Physics*, Vol. 138, 054116 (2013).

[45] R. Dawes, D. L. Thompson, A. F. Wagner, and M. Minkoff, Interpolating moving least-squares methods for fitting potential energy surfaces: A strategy for efficient automatic data point placement in high dimensions, *J. Chem. Phys.*, Vol.128 (2008), 084107.

[46] Y. Guo, I. Tokmakov, D. L. Thompson, A. F. Wagner, and M. Minkoff, Interpolating Moving Least-Squares Methods for Fitting Potential Energy Surfaces: Improving Efficiency via Local Approximants, *J. Chem. Phys*., Vol. 127 (2007), 214106.

[47] W.L. Thacker, J. Zhang, L.T. Watson, J.B. Birch, M.A. Iyer and M.W. Berry, Algorithm 905: SHEPPACK: Modified Shepard algorithm for interpolation of scattered multivariate data, ACM Transactions on Mathematical Software (TOMS), Vol. 37, No. 3, Sep. 2010.

[48] K.C. Thompson, M.J.T. Jordan, and M.A. Collins, Polyatomic molecular potential energy surfaces by interpolation in local internal coordinates, *J. Chem. Phys.*, Vol. 108, No. 20, May 1998, 8302–8316.