GA

BRUECKNER–GOLDSTONE QUANTUM MONTE CARLO

Allocation: Blue Waters Professor/200 Knh
PI: So Hirata[1]
Collaborators: Cole M. Johnson[1], Alexander E. Doran[1]

[1]University of Illinois at Urbana-Champaign

EXECUTIVE SUMMARY

We introduced and fully developed novel scalable algorithms and software for predictively accurate (*ab initio*) electronic-structure calculations for large molecules and solids, which are not easily subjected to fast calculations by fragmentation. We transform the usual, nonscalable sum-of-products expressions of many-body perturbation and Green's function theories in the complete-basis-set limit into a few high-dimensional integrals, which are then evaluated by a highly scalable Metropolis Monte Carlo algorithm. They can compute energy differences (including quasiparticle energy bands) directly without a sign problem at an operation cost whose size dependence is one or two ranks lower than their deterministic counterparts. They execute efficiently on many CPUs or many GPUs, easily achieving an unprecedented speedup (for an ab initio electron-correlation calculation) by a factor of 31,000 (on 256 GPUs) relative to a serial calculation.

RESEARCH CHALLENGE

Existing algorithms of predictive computational chemistry are not scalable with respect to either system size or computer size. For example, the memory and arithmetic operation costs of the simplest *ab initio* electron-correlated theory, i.e., the second-order Møller–Plesset perturbation (MP2) theory, increase as the fourth and fifth power of system size, respectively, and its computational kernel is poorly parallelizable. The situation is even worse for higher-order MP and other *ab initio* electron-correlation methods such as coupled-cluster theory. This is compounded with the extremely slow convergence of their results with the size of an expansion basis set, further driving up the cost to achieve the complete-basis-set (CBS) limit. This project seeks to invent and fully develop fundamentally scalable algorithms for predictive computational chemistry (i.e., without sacrificing accuracy by introducing arbitrary approximations with uncontrollable errors) with respect to both system and computer sizes by combining quantum-Monte-Carlo-like stochastic algorithms and *ab initio* electron-correlated theory. We will then deploy such algorithms on Blue Waters in chemistry/solid-state-physics applications with unprecedented accuracy and problem sizes.

METHODS & CODES

We mathematically transformed the usual sum-of-products expressions of MP2, second-order Green's function (GF2) theory, and their CBS corrections by explicitly correlated (F12) ansätze into single high-dimensional integrals by a Laplace transform. These integrals are then evaluated by a Metropolis Monte Carlo method with judiciously chosen weight functions. The resulting stochastic (Brueckner–Goldstone quantum Monte Carlo) methods [1—Monte Carlo MP2 (MC-MP2) [2], Monte Carlo GF2 (MC-GF2) [3], Monte Carlo explicitly correlated MP2 (MC-MP2-F12) [4,5], and Monte Carlo explicitly correlated GF2 (MC-GF2-F12) [6]—can compute energy differences (correlation energies and electron detachment/attachment energies) directly without a sign problem in a scalable manner with respect to both computer size (on thousands of CPUs or hundreds of GPUs) and system size (the operation cost is linear scaling per MC step and cubic to quartic scaling to achieve given relative accuracy and the memory cost is negligible) [7]. They can also calculate quasiparticle energy bands of a solid for the entire Brillouin zone as nearly continuous curves of a wave vector [8] and have been extended to third-order MP (MP3) [9] using an expedient interpretation of Brueckner–Goldstone diagrams as well as a convergence-acceleration scheme (redundant-walker algorithm) [10].

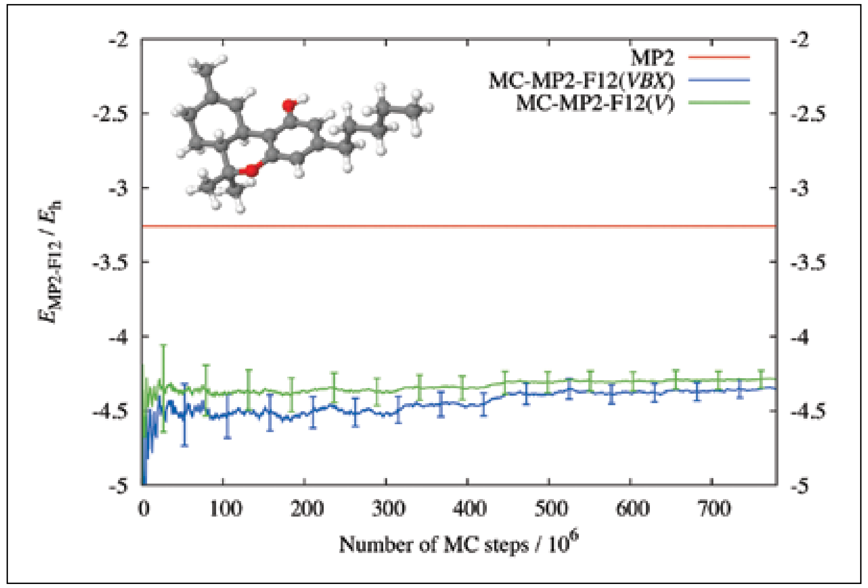

Figure 1: Monte Carlo explicitly correlated MP2 energies of tetrahydrocannabinol (472 basis functions).

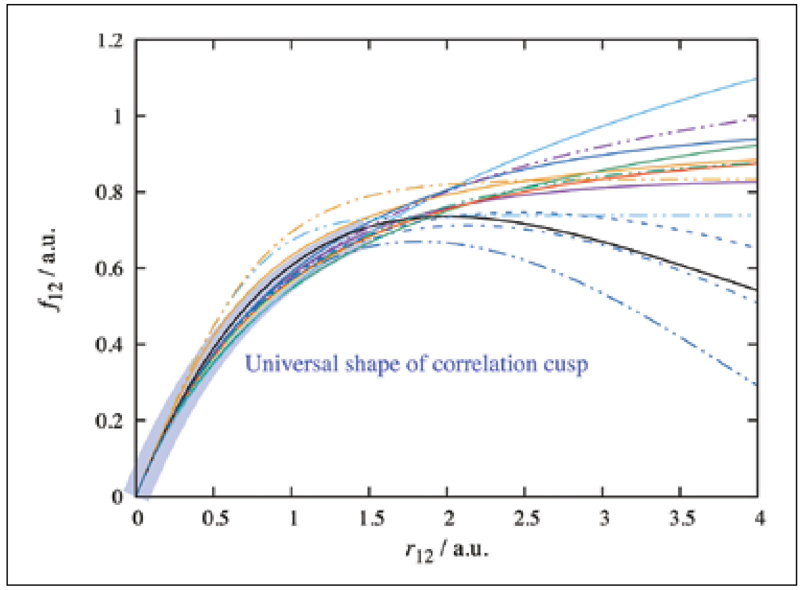

Figure 2: The universal size (the radius of 0.8 Ångstrom) and concave shape of a correlation hole uncovered by Monte Carlo MP2 calculations with 17 different correlation factors.

RESULTS & IMPACT

The MC-MP2-F12 method enabled an exact (CBS-limit) MP2 energy calculation of tetrahydrocannabinol (472 basis functions) without a local-correlation scheme (Fig. 1). Exploiting the extraordinary flexibility of this algorithm in using virtually any explicitly correlated factor, we numerically characterized the performance of 17 such factors. We observed that highly performing factors share the same short-range behavior within the radius of 1.5 Bohr, while differing greatly in the long-range behavior. This result reveals fundamental electron-correlation physics that a correlation hole of a pair of electrons has a universal size (1.5 Bohr) and concave shape (dictated by Kato's cusp condition) regardless of its molecular environment or energy (Fig. 2). We have completed the development of the MC-GF2-F12 method, which can compute electron detachment/attachment energies directly in the CBS limit. With this, we computed exact GF2 electron affinities with a statistical uncertainty of 0.03 eV for C_{60} and C_{70}, which play important roles in heterojunction solar cells as an electron acceptor but resist a local-correlation scheme for fast calculations.

These calculations were based on the redundant-walker algorithm, which propagates more walkers than minimally necessary and permutes them in all possible ways when being substituted into the integrand, thereby multiplying the sampling efficiency. We introduced [7] a two-level parallelism in which dense matrix multiplications for many walkers are fine-grained on a GPU and a Monte Carlo integration itself is coarse-grained across multiple CPU-GPUs. In this way, not only did we observe a speedup by a factor of 31,000 on 256 GPUs relative to a serial execution, but we also found that the saturation point of the acceleration is significantly delayed to a much greater number of walkers. This is a rare instance in which the parallel architecture (GPU) and algorithm (the redundant-walker algorithm) mutually enhance each other.

WHY BLUE WATERS

The stability and ease of use (OS, compilers, libraries, and NCSA expertise) as well as the balanced deployment of CPUs and GPUs are all essential for rapid coding/profiling of new scalable algorithms from scratch and their capacity testing.

PUBLICATIONS AND DATA SETS

Johnson, C. M., S. Hirata, and S. Ten-no, Explicit correlation factors. *Chem. Phys. Lett.*, published online (2017): https://doi.org/10.1016/j.cplett.2017.02.072.

Johnson, C. M., A. E. Doran, J. Zhang, E. F. Valeev, and S. Hirata, Monte Carlo explicitly correlated second-order many-body perturbation theory. *J. Chem. Phys.*, 145 (2016), p. 154115.

Doran, A. E. and S. Hirata, Monte Carlo MP2 on many graphical processing units. *J. Chem. Theory Comput.*, 12 (2016), pp. 4821–4832.