3E09

Analysis Tools for Linear Scaling Quantum Chemistry Calculations

○William Dawson[1], Luigi Genovese[2], Takahito Nakajima[1]
[1] RIKEN Center for Computational Science, Japan
[2] French Alternative Energies and Atomic Energy Comission, France

[Abstract] The development of efficient linear-scaling algorithms alongside steady increases in computational power have made ab-initio calculations with thousands of atoms more and more common place. However, compared to calculations on small systems, there still remains a deficiency in the kinds of tools available to post-process and analyze the results. In this talk, we will discuss new algorithms and computational techniques we have been developing to help interpret the results of calculations in the linear-scaling regime. One key approach to understanding large systems is the ability to partition a system into smaller building blocks. In this talk, we will consider how to automatically decompose a system into building blocks using the information in the one particle density matrix. We will demonstrate the utility of these tools by presenting an analysis of various biological systems in realistic environments.

[Introduction] Commonly used computational techniques for computing the properties of molecular systems from first-principles are often limited to systems of fewer than one thousand atoms due to computational costs that grow cubically or worse with the size of the system. In order to overcome these bottlenecks, linear-scaling methods both for the construction of the Hamiltonian and for computing the density matrix have been developed[1]. However, while these techniques can enable calculations on larger molecules in more realistic environments, this increase in system size also leads to an increased challenge in interpreting calculation results. This challenge is made even more difficult by deficiencies in analysis tools that can efficiently be applied to large systems.

[Methods] One of the main bottlenecks that needs to be avoided when performing linear-scaling calculations is the computation of the density matrix through the eigendecomposition of the Hamiltonian. Linear-scaling methods directly compute the density matrix using methods such as the Fermi Operator Expansion, density matrix minimization, density matrix purification, etc, which can effectively exploit the sparsity of the density matrix. While these methods allow for the efficient calculation of the ground state density, information about the eigenspectrum is lost.

The basis of these diagonalization free methods is the calculation of a projector on to the occupied subspace of a matrix. By changing the target subspace, these methods can also be employed to slice up the eigenspectrum into arbitrary patches, which might then be recombined to recover the full eigenspectrum[2]. In many cases, the projectors associated with these patches of the spectrum also have significant sparsity, allowing for an efficient divide and conquer approach to computing the eigenspectrum. Once a local energy envelope has been chosen, we might also investigate the locality in space using the sparsity of the density matrix. Mohr et al.[3] has shown that a good measure for the locality of the region is whether the sub-density matrix associated with a given region is idempotent.

[Results and Discussion] In Fig. 1, we plot the eigenspectrum of a small protein (1L2Y [4]) in a salt water solution computed using the linear scaling version of BigDFT[5]. The density matrix is moderately sparse, with about 14% of its entries being nonzero. We have also computed the density matrices associated with only certain subsets of the energy spectrum. While arbitrary energy regions can lead to significant fill in, if these subset regions are chosen well, a significant amount of sparsity remains. This locality in space and energy allows for efficient calculation of the molecular orbitals of this system. This information, combined with the purity indicator metric, allows for the partitioning of large systems into smaller building blocks.

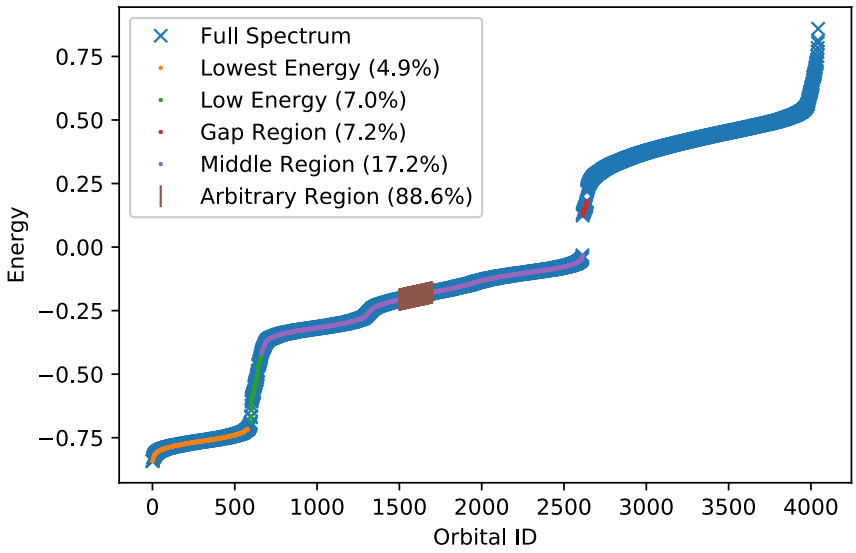

Fig. 1. Density matrix sparsity for different energy regions.

[References]

[1] Bowler, D. R., and Tsuyoshi Miyazaki. "Methods in electronic structure calculations." Reports on Progress in Physics 75, no. 3 (2012): 036503.

[2] Bai, Zhaojun, and James Demmel. "Using the matrix sign function to compute invariant subspaces." SIAM Journal on Matrix Analysis and Applications 19, no. 1 (1998): 205-225.

[3] Mohr, Stephan, Michel Masella, Laura E. Ratcliff, and Luigi Genovese. "Complexity Reduction in Large Quantum Systems: Fragment Identification and Population Analysis via a Local Optimized Minimal Basis." Journal of chemical theory and computation 13, no. 9 (2017): 4079-4088.

[4] Neidigh, Jonathan W., R. Matthew Fesinmeyer, and Niels H. Andersen. "Designing a 20-residue protein." Nature Structural and Molecular Biology 9, no. 6 (2002): 425.

[5] Mohr, Stephan, Laura E. Ratcliff, Paul Boulanger, Luigi Genovese, Damien Caliste, Thierry Deutsch, and Stefan Goedecker. "Daubechies wavelets for linear scaling density functional theory." The Journal of chemical physics 140, no. 20 (2014): 204110.