Application of Compressive Sensing to Simulated Chemical Spectra

Camille Houferak, Joe Kasper, Joseph J. Radler, Shichao Sun

March 12, 2016

Abstract

Abstract

I. Introduction and Overview

Computational chemistry seeks to construct fine-grained models of chemical systems in order to model and predict experimentally observable quantities in support of experimental research. A major implementation of computational modeling is the construction of idealized chemical spectra, which are essentially the 'fingerprints' of a chemical species in a system of interest. Traditionally, spectra have been generated experimentally, but recently have been generated from models of molecular electronic structure which comprises a major field of study in the computational chemistry community. However, generating predictive spectra from model and simulation can be a difficult undertaking, as the number of degrees of freedom for a so-called *full configuration* solution increases exponentially (or factorially!) for each additional electron found in the system. Full configuration calculations are ideal for capturing *all* of the physics of a system, but come at great computational expense as every interaction between each particle must be accounted for in the Hamiltonian and the second-order differential *Schrödinger Equation* solved numerically. If atomic nuclear dynamics are also modeled, this increases the complexity even further as each solution for each time step. Consequently, approximations and calculation schemes that minimize the complexity and computational cost of a model are highly valued in the field. Accordingly, techniques for reducing or removing the necessity of small time-steps and long timescales are also of great utility, as dynamical simulations examining nuclear motion often require picosecond timescales to capture the physics of interest whereas models examining only electronic dynamics require only femtosecond timescales.

Things are not so bleak, however, as there is an abundance of mathematical techniques for reducing the scale of such problems. This study focuses on one such technique called

Compressed Sensing (CS) that is used in data analysis to reduce the necessary sampling rate. CS does this by assuming a high degree of sparsity in a signal and makes use of the L^1 norm minimization technique to locate the vector of signal transform coefficients that is the most sparse. As CS is frequently applied to time-frequency analysis problems, we apply it to modeled chemical spectra which heavily feature Fourier analysis for signal resolution. This particular investigation focuses on *Optical Absorption Spectroscopy* – an electronic process in molecular systems where electrons are excited to higher discrete quantum states by absorption of incident photons.

Modeling absorption spectra computationally presents a unique challenge since the propagation of electronic states through time is the fundamental process behind the generation of experimentally observable spectra. Several schemes exist for performing calculations that capture electron dynamics, however here we focus only on simulated spectra obtained through a *Real-Time Time-Dependent Self-Consistent Field* (RT-TDSCF) scheme which solves electron dynamics by self-consistently solving a set of pseudoeigenvalue equations for each iteration in time while holding the nuclear positions fixed (using the *Born-Oppenheimer approximation*). As one can imagine, this can become prohibitively expensive for large systems with many electrons or large numbers of time steps; therefore we seek to find the best CS sampling rate to reduce the number of necessary time steps to achieve satisfactory optical absorption spectral resolution for molecular systems.

II. Theoretical Background

A. An Introduction to Optical Absorption Spectroscopy

B. Time-Dependent Self-Consistent Field Theory

C. Compressive Sensing

D. Applications to Chemical Spectroscopy

III. Algorithm Implementation and Development

IV. Computational Results

A. Ideal Data Cases

In order to test the validity of our CS algorithm for data reconstruction, we first applied it to idealized data sets represented by known linear combinations of sinusoidal functions. Having knowledge of the exact "signal" *a priori* allows us to determine the sampling rates and discrete transforms that result in the best approximations under ideal conditions.

B. Simulated Optical Absorption Spectra Cases

V. Summary and Conclusions

References

[1]

[2]

Appendix A   MATLAB Functions and Implementations

function subsection

Appendix B   MATLAB Code

```
%Insert Matlab Code Here.  Whitespace is retained within {Verbatim}
```

Appendix C   Custom MATLAB Functions

Custom MATLAB Function 1