Application of compressed sensing to the simulation of atomic systems

Xavier Andrade, Jacob N. Sanders, and Alán Aspuru-Guzik[1]

Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138

Compressed sensing is a method that allows a significant reduction in the number of samples required for accurate measurements in many applications in experimental sciences and engineering. In this work, we show that compressed sensing can also be used to speed up numerical simulations. We apply compressed sensing to extract information from the real-time simulation of atomic and molecular systems, including electronic and nuclear dynamics. We find that, compared to the standard discrete Fourier transform approach, for the calculation of vibrational and optical spectra the total propagation time, and hence the computational cost, can be reduced by approximately a factor of five.

sparse signal reconstruction | molecular dynamics | electron dynamics

A recent development in the field of data analysis is the compressed sensing (CS) (or compressive sampling) method (1, 2). The foundation of the method is the concept of sparsity: A signal expanded in a certain basis is said to be sparse when most of the expansion coefficients are zero. This extra information can be used by the CS method to significantly reduce the number of measurements needed to reconstruct a signal. CS has been successfully applied to data acquisition in many different areas (3), including the improvement of the resolution of medical magnetic resonance imaging (4) and the experimental study of atomic and quantum systems (5–7).

In this article we show that CS can also be an invaluable tool for some numerical simulations with a considerable reduction of the computational cost. We focus on atomistic simulations of nanoscopic systems by using CS to extract frequency-resolved information from real-time methods such as molecular dynamics (MD) and real-time electron dynamics.

MD (8, 9) is one of the most widely used methods to study atomistic systems computationally as it can be used to compute many static and dynamical properties. In MD the trajectory of the atomic nuclei is obtained by integrating their equations of motion either with parametrized force fields or else by explicitly modeling the electrons (10). Given the importance of MD, developing methods that can improve the precision and reduce the computational cost of this method, especially for ab initio MD, can have a large impact in the field of atomistic simulation.

Real-time electron dynamics, in particular real-time time-dependent density functional theory (TDDFT) (11), plays a similarly important role in the study of linear and nonlinear electronic properties (12–15). Because of its scalability and parallelizability, real-time TDDFT is particularly efficient for large electronic systems (16), so an additional reduction in the computational cost can extend the boundaries of the system sizes that can be studied.

Many physical properties are represented by frequency-dependent quantities. To obtain these from real-time information, usually a discrete Fourier transform (FT) is used. Our approach is to replace this FT by a calculation of the Fourier coefficients based on the CS method. To obtain a given frequency resolution, the CS method requires a total propagation time that is several times smaller than that required by the FT.

CS has the potential to provide across the board speedup for many applications involving the computation of sparse spectra. Moreover, this speedup may be obtained without making any changes to the underlying propagation code used in different types of electronic and nuclear calculations; one simply replaces the FT algorithm with the CS method, making the approach quite straightforward to implement. This paper introduces CS and demonstrates its broad utility in computational chemistry and physics by applying it to the calculation of various nuclear and electronic spectra of small molecules. The resulting computer code is available as open source software.

The article is structured as follows. We first introduce the CS method and show how it may be applied to the determination of Fourier coefficients. Next, we apply CS to the calculation of vibrational, optical absorption, and circular dichroism (CD) spectra. We then proceed to a discussion of the numerical methods used in our CS implementation. Finally, we offer conclusions and an outlook.

Compressed Sensing

In this section, we briefly introduce the application of the CS method to the calculation of Fourier coefficients. More details about the method and its origins may be found in refs. 17–19.

For simplicity, we assume that we want to calculate a certain frequency-resolved quantity $g(\omega)$ that is given by the sine transform of a certain time-resolved function $h(t)$

$$g(\omega) = \int_{-\infty}^{\infty} dt \sin(\omega t) h(t) \qquad [1]$$

(the analysis is equally valid for the cosine transform). Because we are interested in numerical solutions, we need to think in terms of discrete quantities. For the numerical problem, we consider that $h(t)$ has been sampled at N_t equidistant times $t_j = \Delta t j$. We want to obtain $\{g_1, g_2, ..., g_{N_\omega}\}$ at N_ω equally spaced frequencies $\omega_j = \Delta \omega j$.

In principle, the g_k can be directly obtained using the discrete FT,

$$g_k = \sum_{j=1}^{N_t} \Delta t \sin(\omega_k t_j) h_j. \qquad [2]$$

However, if we expect that many of the Fourier coefficients are zero, a property known as sparsity, we can use this additional information to obtain more precise results. This property is the basis for the CS scheme.

We start by reformulating the discrete Fourier transform in Eq. 2 as a matrix inversion problem. From this perspective, we are trying to solve the linear equation for $\boldsymbol{g}$,

$$F\boldsymbol{g} = \boldsymbol{h}, \qquad [3]$$

where F is the $N_\omega \times N_t$ Fourier matrix with entries

$$F_{jk} = \frac{2}{\pi} \Delta\omega \sin(\omega_j t_k). \qquad [4]$$

Our objective is to obtain sensible results with N_t as small as possible. Thus, we are interested in the case $N_\omega > N_t$, where the linear system is underdetermined, and there are many solutions for g (in fact, one of them will be given by Eq. 2). From all the solutions of Eq. 3, we select the one that has the largest number of zero coefficients: the sparsest solution. This procedure turns out to be equivalent to the so-called basis pursuit (BP) optimization problem (18)

$$\min_g |g|_1 \quad \text{subject to} \quad Fg = h, \qquad [5]$$

which is what one solves in practice (where $|g|_1 = \sum_k |g_k|$ is the standard 1-norm).

The CS scheme can be generalized to allow for a certain amount of noise in the time-resolved signal. In this case the problem to be solved is known as basis pursuit denoising (BPDN)

$$\min_g |g|_1 \quad \text{subject to} \quad |Fg - h| < \eta, \qquad [6]$$

where η represents the level of noise in the signal. BPDN is the formulation we use in our case, because we expect a certain amount of noise coming from the finite-precision numerical calculations (and possibly other sources).

Vibrational Spectra

MD can be used to obtain information about the vibrational modes of atomic systems. Experimentally, the quantities that usually give access to the vibrational modes are the infrared and Raman spectra that can be obtained from MD as the Fourier components of the electronic polarization and polarizability, respectively. If we are only interested in the vibrational frequencies, from the nuclear velocities $\{v_i\}$ we can calculate the velocity autocorrelation function

$$\gamma(t) = \frac{\left\langle \sum_i v_i(t) \cdot v_i(0) \right\rangle}{\left\langle \sum_i v_i(0) \cdot v_i(0) \right\rangle}, \qquad [7]$$

where the brackets indicate the average over the ensemble. The cosine transform of γ is the vibrational frequency distribution (20)

$$f(\omega) = \int dt \gamma(t) \cos(\omega t). \qquad [8]$$

Because this spectrum is composed of a finite number of frequencies (fewer than three times the number of atoms in the system), the calculation is ideal for the CS method.

To illustrate the properties of the CS method we start with a simple case, the single vibrational frequency of a diatomic molecule, Na_2, which we simulate using ab initio MD. In Fig. 1, we show how the vibrational spectrum depends on the amount of time for which the velocity autocorrelation is calculated, both for CS and for the discrete FT with a standard polynomial damping function. Whereas the FT requires long propagations to resolve the vibrational frequency, the CS method gives a well-defined peak even with less than one oscillation of the molecular vibrational mode. That the peak is well-defined, however, does not imply that the peak position is converged. As seen in Fig. 2, the peak position oscillates with the total sampling time until it converges to the true value after a few periods are sampled. For the FT, on the other hand, the peak position, obtained by fitting the
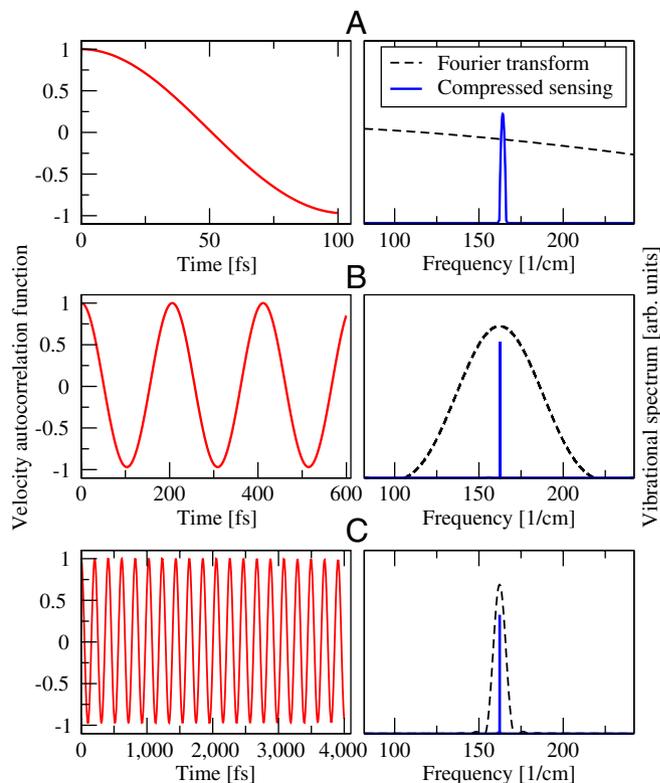
Fig. 1. Frequency distribution spectrum of Na_2 calculated using FT and CS for different total propagation times: (A) 100 fs, (B) 600 fs, and (C) 4,000 fs. (Left) The velocity autocorrelation function and (Right) the frequency spectrum.

spectrum to a Gaussian function, is well-defined when more than one period is sampled, but the width of the peak converges very slowly with the sampling time. This difference in behavior of CS and FT has been reported previously, for example, in ref. 21. We remark that the CS process does not use any additional information about the signal beyond assuming it is sparse.

The advantage of the CS approach is further illustrated when several closely spaced frequencies must be resolved. To demonstrate this advantage, we calculate the vibrational spectrum for a benzene molecule from an ab initio MD simulation in Fig. 3. We can see that the CS approach with 1,000 fs gives a spectrum that is better resolved than the FT results for 5,000 fs, which directly
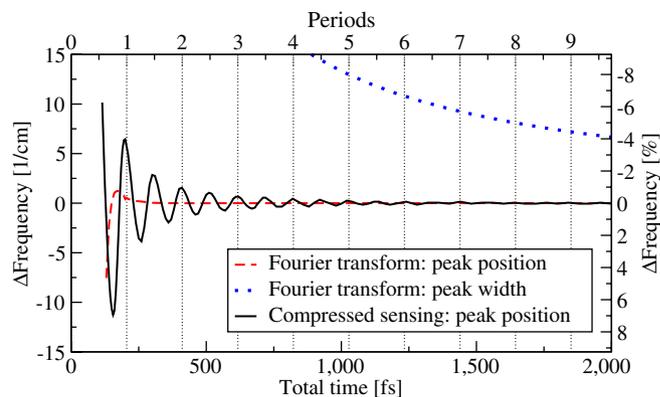
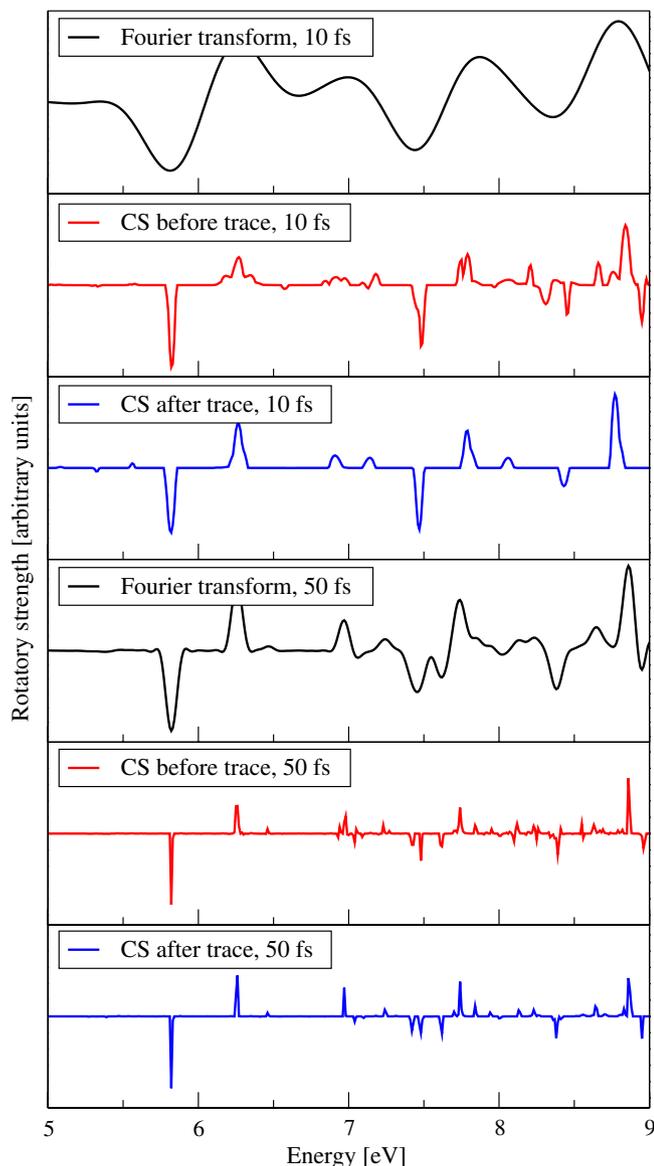**Fig. 2.** Error in the vibrational frequency of Na_2 computed by CS with respect to total time. For comparison we plot the positions and width of the peak obtained by using a discrete FT. For the FT, the position, ω_0, and the width, σ, are fit by assuming the peak has a Gaussian form $A \exp[-(\omega - \omega_0)^2/(2\sigma^2)]$. For CS, the peak width corresponds to the energy spacing, 0.01 1/cm.

Fig. 3. Vibrational frequency distribution spectrum of benzene. Comparison of a CS calculation with 1,000 fs and FT with 5,000 fs.

translates into a reduction of the computational time by five times or more. It is reasonable to expect that equivalent gains can be obtained for the computer simulation of other vibrational spectroscopies such as infrared and Raman.

Optical Absorption Spectra

Optical absorption is an electronic process. Although it can be calculated from a linear response framework (22, 23), it can also be obtained from real-time electron dynamics (12). To obtain the spectrum from real-time dynamics the electronic system is propagated under the effect of an electric field of the form $E(r, t) = \kappa\delta(t)$. From the propagation the time-dependent dipole moment $p(t)$ is obtained, and from the dipole the frequency-dependent polarizability can be obtained as (atomic units are used in the next two sections)

$$\alpha_{ij}(\omega) = \frac{1}{\kappa_i} \int_0^\infty dt e^{-i\omega t}[p_j(t) - p_j(0)]. \qquad [9]$$

To obtain the full α tensor three propagations are required (with κ in different directions).

The absorption cross-section is related to the trace of the imaginary part of the polarizability tensor

$$\sigma(\omega) = \frac{4\pi\omega}{3c} \mathfrak{Im} \sum_{i=1}^{3} \alpha_{ii}(\omega). \qquad [10]$$

The optical absorption spectrum is an ideal candidate for the application of CS. For a molecule, the electronic transitions between bound states produce a discrete spectrum in the low energy region. At higher energies, the transitions to unbound states produce a continuous spectrum. In general, electronic structure calculations cannot capture this continuous spectrum and approximate it as a sequence of discrete excitations, so in practice the whole spectrum is sparse.

In Fig. 4, we show the optical absorption spectrum for benzene calculated via real-time TDDFT. There we illustrate the effect of the propagation time on the spectrum for CS and FT. From the figure, it is clear that the CS method is capable of resolving the spectrum much better and with a shorter propagation time than a discrete FT. Because the two methods have different convergence properties, as discussed in the previous section, it is difficult to rigorously quantify the speedup afforded by CS. Nevertheless, by comparing the FT at 25 fs with CS at 5 fs, we can see that the FT requires approximately five times the propagation time as CS to achieve a similar resolution. Note that for a sampling time of 1 fs, the CS spectrum is not particularly sparse, perhaps due to the fact

Fig. 4. Optical absorption spectra for benzene computed from real-time TDDFT with different propagation times. Comparison between CS and discrete FT. Experimental results from ref. 41.

that the 1-norm minimization is not a good measure of sparsity in the very short-time regime (1).

CD Spectra

Another property that can be calculated from real-time electron dynamics is CD spectra (24, 25). A CD spectrum measures the difference in a chiral molecule's response to left and right circularly polarized light. The real-time calculation is performed in the same manner as the optical absorption case, but the key quantity to be calculated is now the time-dependent orbital magnetization $m(t)$. The frequency-dependent electric-magnetic cross-response tensor may be obtained as

$$\beta_{ij}(\omega) = \frac{ic}{\omega\kappa_i} \int_0^\infty dt e^{-i\omega t}[m_j(t) - m_j(0)]. \qquad [11]$$

The rotatory strength, which is the quantity typically plotted in CD spectra, is related to the trace of the imaginary part of $\beta(\omega)$ according to

$$R(\omega) = \frac{\omega}{\pi c} \mathfrak{Im} \sum_{i=1}^{3} \beta_{ii}(\omega).$$ [12]

The rotatory strength $R(\omega)$ is suitable to the CS scheme because it is sparse in frequency space. In fact, the peaks in a CD spectrum are located at the same positions as in an absorption spectrum. However, the CD spectrum contains both positive and negative peaks.

Fig. 5 compares the CD spectrum for (R)-methyloxirane as computed by FT and CS for two different propagation times (10 and 50 fs). As can be seen from the figure, for a given propagation time, the CS method provides better spectral resolution than the discrete FT. In fact, just as with linear absorption, FT requires a propagation time approximately five times as long as CS to obtain a comparable spectral resolution (as can be seen by comparing the 50 fs FT with the 10 fs CS).

Fig. 5 also illustrates another feature of CS: Unlike a discrete FT, the CS method is nonlinear. Adding together time-resolved signals and then applying CS generally gives different results from applying CS first and then adding together the results in the frequency domain, particularly if not all the peaks are well-resolved. In other words, the use of CS to convert time-resolved data into the frequency domain, as in Eq. 11, and the calculation of the trace, as in Eq. 12, do not commute. Hence, there are two approaches to obtain the CD spectrum: We can perform CS for each propagation direction and then compute the trace, or we can compute the trace in the time domain and then perform CS. Both approaches are shown in Fig. 5; at 10 fs, they give similar but not identical results. Performing the CS after the trace appears to give a cleaner spectrum with fewer and better resolved peaks, as expected from the sparsity requirement. For longer propagation times (50 fs), all of the peaks are more fully resolved and the two approaches to CS provide much improved resolution over a FT for a given propagation time. This effect is also present in the absorption spectra, but to a much lesser degree, probably because the spectra are nonnegative.

Numerical Methods

Numerically, to find a spectrum using the CS method we need to solve Eq. 6. This problem is not trivial, so we rely on the spectral projected gradient for 1-norm minimization (SPGL1) algorithm developed by van den Berg and Friedlander (26). To avoid numerical stability issues we work with a normalized BPDN problem, where the prefactor $2\Delta\omega/\pi$ of the F matrix, Eq. 4, is left out and h is normalized. These prefactors are then included in g after the solution is found. This strategy has the additional advantage of making the noise parameter η of Eq. 6 dimensionless.

Because we do not have an a priori estimate for η, we do not set it directly. As the SPGL1 algorithm finds a sequence of approximate solutions with decreasing values of η, we set the target value to zero. We assume that the calculation is converged when the value of η falls below a certain threshold (10^{-7}) or the active space of the system, the set of nonzero coefficients, has not changed for a certain number of iterations (50). In the former case we consider that a solution of the BP problem, Eq. 5, has been found. For all the calculations presented here $\eta < 10^{-3}$.

CS is much more costly numerically than the discrete FT approach, as it usually involves several hundred matrix multiplications. However, this increased computational cost is not a problem for the applications we are proposing because this process normally only takes a few minutes, much less than the computation time required to simulate the real-time dynamics of large atomic systems.

All the calculations presented in this article are performed using the OCTOPUS code (16, 27) at the (time-dependent) density functional theory level with the PBE exchange correlation functional (28). The adiabatic MD calculations are performed from first principles using the modified Ehrenfest method (29, 30) with an ionic timescale factor (μ) of 30 for Na_2 and 5 for benzene. The systems are given initial velocities equivalent to 300 K and the MD is performed at constant energy.

All calculations use norm-conserving pseudo-potentials with a real-space grid discretization. The shape of the grid is a union of boxes around each atom. We choose the grid parameters to ensure proper convergence of the results to an accuracy of 1%. For Na_2 we use a spacing of 0.375 a.u. with a sphere radius of 12 a.u., and the MD time step is 0.057 fs. For benzene, the grid spacing is 0.35 a.u., the radius is 14 a.u., and the time step is 0.0085 fs for MD and 0.0017 fs for real-time TDDFT. For (R)-methyloxirane, the spacing is 0.378 a.u., the sphere radius is 15.1 a.u., and the time step is 0.0008 fs for real-time TDDFT. For the vibrational spectrum calculation we use a time step 10 times the one of the MD, the energy step is 0.01 1/cm, and the maximum spectrum

Fig. 5. CD spectrum computed for (R)-methyloxirane from real-time TDDFT with different propagation times. Comparison between discrete FT and CS. Because the CS process is nonlinear we compute the spectra in two ways. CS before trace: The spectra is calculated for each direction using CS and then the trace, Eq. 12, is computed. CS after trace: The trace is calculated in the time domain and then the CS process is used over this averaged signal.

energy is 5,000 1/cm. For the benzene optical absorption spectra, we use a time step of 0.0017 fs, the energy step is 0.027 eV, and the maximum spectrum energy is 820 eV. For the (R)-methyloxirane CD spectra, the time step is 0.0008 fs, the energy step is 0.01 eV, and the maximum spectrum energy is 330 eV. The structure of benzene was taken from ref. 31 and the structure of (R)-methyloxirane was taken from ref. 32.

All discrete FTs are performed using third-order polynomial damping: Each signal at time t is multiplied by $p(t) = 1 - 3(t/T)^2 + 2(t/T)^3$ prior to Fourier transform, where T is the time length of the signal (33).

The SPGL1 method used for CS was implemented into OCTOPUS based on a Fortran translation of the original MatLab code of van den Berg and Friedlander (26). We plan to release this implementation as a stand-alone tool in the near future (for the moment the code can be obtained from the OCTOPUS repository).

Conclusions

We have shown that the CS method can be applied to the numerical calculation of different kinds of atomic and electronic spectra. These results give a significant reduction of the computational time required for the numerical simulations. For the systems tested in this work, we find that it is possible to obtain a speedup of five for CS as compared to the discrete FT. Because the number of samples required for CS depends on the sparsity of the signal (18), this ratio could change for other systems.

The effect of this reduction is to increase the size of the systems that are currently accessible to numerical simulations, and to make possible simulations with more precise, but more costly, methods. It also means that other types of simulations become more affordable from a real-time perspective, for example the combined dynamics of nuclei and electrons that are constrained to short simulation times by the fast dynamics of the electron.

In this work, we have shown the application of CS to the calculation of a few types of spectra, but the method most likely can be applied to other quantities as well, such as nonlinear optical response (14, 15), magnetic CD (34), semiclassical nuclear dynamics (35), and two-dimensional spectroscopy (36, 37). Of course, the method is not limited to atomistic simulations and could be applied to simulations in all scientific fields.

Probably the main advantage of CS is that it is a simple and universal approach that works as a drop-in replacement for the FT for many applications where discrete or sparse spectra need to be calculated. The main limitation of the CS approach is that it may not be as beneficial for quantities that are not sparse.

There are some cases where the sparsity requirement might be circumvented. For example, though the real part of the polarizability tensor is not sparse, it could be computed from the imaginary part by using the Kramers–Kronig relation. Another example is crystalline systems (38), where there is a continuum of excitation energies. In this case it might be possible to apply the CS scheme to each k-point separately.

Another important issue is the performance of CS in resolving spectral features such as peak broadening. These features do not appear in the spectra calculated in this article, but might appear when internal structure or environmental effects are considered. This issue will be the focus of future work. Another interesting continuation of our work would be to compare CS with other approaches for calculating spectral properties such as the maximum entropy method (39). This method is related to CS but may better reproduce broadened features in spectra (40).

Nevertheless, we want to emphasize that for the sparse spectra obtained in atomistic simulations, compressed sensing is a promising method. We expect that compressed sensing will become widely used in the scientific computing community once its advantageous properties become more widely known. The main difficulty in the adoption of CS is that it is more complex to implement than a discrete FT. This problem can be solved by providing libraries and utilities that can be used by researchers. However, the CS method also has features researchers will need experience to understand. For example, the method is nonlinear as discussed and the peak width is not always related to the convergence of the spectrum.

We believe that the direct application of the compressed sensing methodology to numerical simulation opens the path for more challenging applications. The principles of sparsity could be used to design algorithms for numerical simulations that have a reduced computational cost not only in the number of operations, but also in memory and data transfer bandwidth requirements; we could call this idea compressed computing.

1. Candes E, Romberg J, Tao T (2006) Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans Inf Theory* 52:489–509.
2. Donoho D (2006) Compressed sensing. *IEEE Trans Inf Theory* 52:1289–1306.
3. Baraniuk R, Candes E, Elad M, Ma Y (2010) Applications of sparse representation and compressive sensing. *Proc IEEE* 98:906–909.
4. Lustig M, Donoho D, Pauly JM (2007) Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn Reson Med* 58:1182–1195.
5. Shabani A, et al. (2011) Efficient measurement of quantum dynamics via compressive sensing. *Phys Rev Lett* 106:100401.
6. Al Quraishi M, McAdams HH (2011) Direct inference of protein DNA interactions using compressed sensing methods. *Proc Natl Acad Sci USA* 108:14819–14824.
7. Katz O, Levitt JM, Silberberg Y (2010) Compressive Fourier transform spectroscopy. *Frontiers in Optics* (Optical Society of America, Washington, DC), Available at arXiv:1006.2553, p FTuE3.
8. Rapaport DC (1995) *The Art of Molecular Dynamics Simulation* (Cambridge University Press, Cambridge, UK).
9. Allen P, Tildesley D (1989) *Computer Simulation of Liquids* (Oxford Science Publications, Clarendon, Oxford).
10. Marx D, Hutter J (2009) *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods* (Cambridge University Press, Cambridge, UK).
11. Runge E, Gross EKU (1984) Density-functional theory for time-dependent systems. *Phys Rev Lett* 52:997–1000.
12. Yabana K, Bertsch GF (1996) Time-dependent local-density approximation in real time. *Phys Rev B* 54:4484–4487.
13. Castro A, Marques MAL, Alonso JA, Rubio A (2004) Optical properties of nanostructures from time-dependent density functional theory. *J Comput Theor Nanosci* 1:231–255.
14. Castro A, Marques MAL, Alonso JA, Bertsch GF, Rubio A (2004) Excited states dynamics in time-dependent density functional theory. *Eur Phys J D* 28:211–218.
15. Takimoto Y, Vila FD, Rehr JJ (2007) Real-time time-dependent density functional theory approach for frequency-dependent nonlinear optical response in photonic molecules. *J Chem Phys* 127:154114.
16. Andrade X, et al. (2012) Time-dependent density-functional theory in massively parallel computer architectures: The octopus project. *J Phys Condens Matter* 24:233202.
17. Baraniuk R (2007) Compressive sensing. *IEEE Signal Proc Mag* 24:118–121.
18. Candes E, Wakin M (2008) An introduction to compressive sampling. *IEEE Signal Proc Mag* 25:21–30.
19. Chartrand R, Baraniuk RG, Eldar YC, Figueiredo MAT, Tanner J (2010) Introduction to the issue on compressive sensing. *IEEE J Sel Topics Signal Process* 4:241–243.
20. Dickey JM, Paskin A (1969) Computer simulation of the lattice dynamics of solids. *Phys Rev* 188:1407–1418.
21. Zhu L, Zhang W, Elnatan D, Huang B (2012) Faster STORM using compressed sensing. *Nat Methods* 9:721–723.
22. Casida M (1996) Time-dependent density functional response theory of molecular systems: Theory, computational methods, and functionals. *Recent Developments and Applications of Modern Density Functional Theory* (Elsevier, Amsterdam) p 391.

23. Andrade X, Botti S, Marques MAL, Rubio A (2007) Time-dependent density functional theory scheme for efficient calculations of dynamic (hyper)polarizabilities. *J Chem Phys* 126:184106.

24. Yabana K, Bertsch GF (1999) Application of the time-dependent local density approximation to optical activity. *Phys Rev A* 60:1271–1279.

25. Varsano D, et al. (2009) Towards a gauge invariant method for molecular chiroptical properties in TDDFT. *Phys Chem Chem Phys* 11:4481–4489.

26. van den Berg E, Friedlander MP (2009) Probing the Pareto frontier for basis pursuit solutions. *SIAM J Sci Comput* 31:890–912.

27. Castro A, et al. (2006) Octopus: A tool for the application of time-dependent density functional theory. *Phys Status Solidi B* 243:2465–2488.

28. Perdew JP, Burke K, Ernzerhof M (1996) Generalized gradient approximation made simple. *Phys Rev Lett* 77:3865–3868.

29. Alonso JL, et al. (2008) Efficient formalism for large-scale ab initio molecular dynamics based on time-dependent density functional theory. *Phys Rev Lett* 101:096403.

30. Andrade X, et al. (2009) Modified Ehrenfest formalism for efficient large-scale ab initio molecular dynamics. *J Chem Theory Comput* 5:728–742.

31. Curtiss LA, Raghavachari K, Redfern PC, Pople JA (1997) Assessment of Gaussian-2 and density functional theories for the computation of enthalpies of formation. *J Chem Phys* 106:1063–1079.

32. Carnell M, et al. (1991) Experimental and quantum-theoretical investigation of the circular dichroism spectrum of R-methyloxirane. *Chem Phys Lett* 180:477–481.

33. Yabana K, Nakatsukasa T, Iwata JI, Bertsch GF (2006) Real-time, real-space implementation of the linear response time-dependent density-functional theory. *Phys Status Solidi B* 243:1121–1138.

34. Lee KM, Yabana K, Bertsch GF (2011) Magnetic circular dichroism in real-time time-dependent density functional theory. *J Chem Phys* 134:144106.

35. Ceotto M, Atahan S, Tantardini GF, Aspuru-Guzik A (2009) Multiple coherent states for first-principles semiclassical initial value representation molecular dynamics. *J Chem Phys* 130:234113.

36. Mančal T, Pisliakov AV, Fleming GR (2006) Two-dimensional optical three-pulse photon echo spectroscopy. I. Nonperturbative approach to the calculation of spectra. *J Chem Phys* 124:234504.

37. Yuen-Zhou J, Aspuru-Guzik A (2011) Quantum process tomography of excitonic dimers from two-dimensional electronic spectroscopy. I. General theory and application to homodimers. *J Chem Phys* 134:134505.

38. Bertsch G, Iwata JI, Rubio A, Yabana K (2000) Real-space, real-time method for the dielectric function. *Phys Rev B* 62:7998–8002.

39. Gull S, Daniell G (1978) Image reconstruction from incomplete and noisy data. *Nature* 272:686–690.

40. Stern AS, Donoho DL, Hoch JC (2007) NMR data processing using iterative thresholding and minimum l1-norm reconstruction. *J Magn Reson* 188:295–300.

41. Koch E, Otto A (1972) Optical absorption of benzene vapour for photon energies from 6 eV to 35 eV. *Chem Phys Lett* 12:476–480.

CHEMISTRY