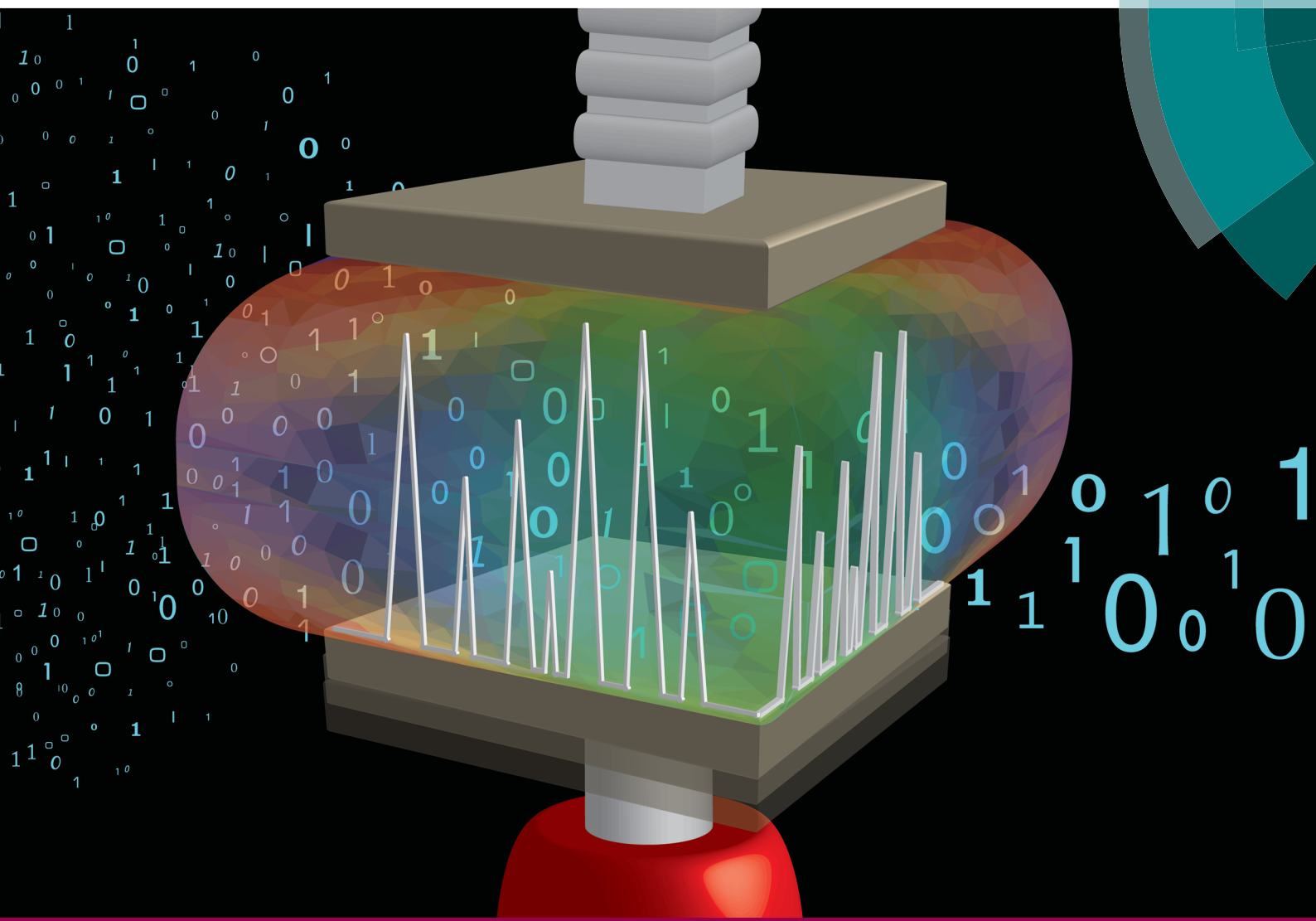
Volume 31 | Number 11 | November 2016 | Pages 2103–2348

# JAAS

Journal of Analytical Atomic Spectrometry
www.rsc.org/jaas



## Themed issue: 2016 International Glow Discharge Spectroscopy Symposium

ISSN 0267-9477

ROYAL SOCIETY OF CHEMISTRY

TUTORIAL REVIEW
Gerardo Gamez
Compressed sensing in spectroscopy for chemical analysis

175 YEARS

JAAS

TUTORIAL REVIEW

View Article Online
View Journal | View Issue



Cite this: *J. Anal. At. Spectrom.*, 2016, **31**, 2165

# Compressed sensing in spectroscopy for chemical analysis

Gerardo Gamez

Analytical chemists are instilled with the Shannon–Nyquist theorem for measuring practices: the sampling frequency must be greater than two times the frequency of the signal of interest to avoid misrepresentation. Furthermore, chemical analysis techniques keep evolving to yield increasing amounts of information resulting in greater demands in data collection, data analysis and computer memory. This leads to typical practices of performing software data compression. In addition, techniques that require increases in spatial frequency of array detectors typically lead to expensive hardware solutions. Over the past ten years, compressed sensing has presented a sampling paradigm shift. The main idea is to perform compression during data acquisition which leads to several advantages, including faster analysis time and availability of cost effective alternatives to array detectors. In this short review, a summary of the main concepts of compressed sensing are presented. In addition, selected examples of compressed sensing applications in spectrochemical analysis are used to showcase its advantages and potential.

Received 19th July 2016
Accepted 5th September 2016

DOI: 10.1039/c6ja00262e

www.rsc.org/jaas

## Introduction

Spectroscopists, and metrologists in general, are tasked with measuring analog signals. On the other hand, computational systems are digital and this means that one must sample the analog signal to transform it into digital form for storage and processing. The trick is to sample in such a way as to be able to faithfully reconstruct the signal of interest. Here enters the Shannon–Nyquist sampling theorem which states that the minimum sampling rate should be higher than twice the frequency of the signal of interest.[1] This has been the dogma for quite some time now.

There are several issues associated with such practice. The first issue is that the amount of data collected can be quite daunting, especially when several dimensions are monitored. Such is the case for hyperspectral imaging, where the signal intensity is followed along the spatial $x$, $y$, $z$ dimensions, the spectral $\lambda$ dimension, and even the temporal dimension.[2] Such file sizes pose a burden in terms of available computer memory. The second issue is that of array detectors and cost. The tendency is to try to acquire an image with a wide field-of-view and at a high resolution which means that more elements are needed in an array detector. This is typically translated into prohibitive hardware costs. However, even when one has enough computer memory and hardware resources to meet the first two issues mentioned above, there is a third issue: time. Such highly dimensional signals take an increasing amount of time for data collection, processing and analysis.

The issues of computer memory burden as well as data processing and data analysis time can be improved through data compression practices. Data compression can be undertaken because not all elements of such data contain relevant information. An example of data compression is when a picture taken with a consumer point-and-shoot camera is saved in JPEG format instead of as a RAW data file.[3] The difference is hardly discernible to the naked eye, yet the file sizes are much smaller. In this case, the data is said to be sparse, and if not perfectly sparse then compressible, which means that most of the significant information is contained in a few elements.

*Texas Tech University, Department of Chemistry and Biochemistry, Lubbock, TX 79409-41061, USA. E-mail: Gerardo.gamez@ttu.edu*

*Gerardo Gamez obtained his B.Sc. and M.S. in Chemistry at the University of Texas at El Paso with Prof. Gardea-Torresdey. He obtained his PhD in Analytical Chemistry at Indiana University-Bloomington with Prof. Hieftje. He performed postdoctoral work at ETH Zurich with Prof. Zenobi and later became a Scientist at EMPA Thun. In 2013 he joined the faculty at Texas Tech University's Department of Chemistry and Biochemistry. His research is focused on plasma-based high-throughput elemental mapping and ambient desorption/ionization MS sources, as well as plasma fundamental studies. He has co-authored >50 publications, received the Payling award and JAAS Emerging Investigator Lectureship.*

However, if not all data elements contain relevant information, would it not be better to only collect those elements with the important part of the signal? This is where compressed sensing (CS) comes in. In CS, the compression is performed during the signal acquisition, thus alleviating the issues mentioned above.

Several decades before compressed sensing came to be known as such, radio astronomers[4] and seismic explorers[5] were able to beat the limitation imposed by the Shannon–Nyquist theorem. In the case of the geological experiments it involved sending a seismic pulse from the surface to the sub-surface and measuring the reflected pulse generated at the underground layer boundaries with sensors at the surface. In principle, the convoluted observed information is not sufficient to recreate the shape of the layers in a unique way. Nevertheless, geologists realized that subsurface layers are mostly homogeneous with very few interfaces where the reflected pulses are generated, thus a sparse spike train. It was this awareness of the signal sparsity that allowed undersampled reconstruction. In a similar way, in the mid-2000s, researchers showed they could reconstruct magnetic resonance images uniquely from undersampled data.[6] This work inspired Candes, Romberg and Tao,[7,8] as well as Donoho,[9] to publish the papers that have made compressed sensing such a hot topic.

Only a decade has passed since these CS framework manuscripts were published but compressed sensing has exploded into many different fields, including geology, medicine, astronomy, *etc.* This review is not meant to be comprehensive, but only a tutorial introduction to the subject through a conceptual view followed by a presentation of some applications in spectroscopy relevant to chemical analysis. Several excellent reviews and materials are available for further reading,[6–15] some of which offer a more rigorous mathematical description of the subject[7–9,16] and describe applications in other fields.

## Compressed sensing: a conceptual view

### Group sampling and the case of the counterfeit gold coin

A quick look at group sampling will allow us to get a more intuitive understanding of how can one obtain the information of interest without being bound by the Shannon–Nyquist limits. In this case there is a collection of eight gold coins but one of them is a counterfeit. In principle, we would have to weigh each coin to identify the bad one. However, the task is to identify the bad one through its weight difference while making as few weighings as possible. This can be achieved by weighing the coins in groups because we know how much a genuine gold coin should weigh. The coins can be weighed in groups of four and, with the right combinations, only three weighings are required. In the first measurement coins 1, 2, 3, and 4 are weighed simultaneously, in the second measurement, coins 1, 2, 5, and 6, and in the third measurement, coins 1, 3, 5, and 7. If none of the measurements give the expected, or correct, weight then we know coin 1 is counterfeit because it was the only coin included in all three measurements. Conversely, if all measurements yield the correct weight we can conclude coin 8 is the counterfeit coin because it was not included in any measurement. If only the first measurement is incorrect then coin 4 is counterfeit because it was only included in that measurement. The same logic applies to the second or third measurements to identify coin 6 or coin 7, respectively, as counterfeit. If the first and second measurements are incorrect but not the third it means that coin 2 is counterfeit because it was the only coin included in the first two measurements. A similar concept applies to coin 3 with measurements one and three, as well as coin 5 with measurements two and three.

One might argue that three group measurements compared to eight individual ones (probably less if the counterfeit coin is weighed early) is not a significant improvement. Nevertheless, what if there were 1000 coins or more? Then the advantages are very clear. Something else to note is that in the above situation it was known beforehand that there was one counterfeit coin, which allowed the design of the measurement combinations, or matrix. What if there were several counterfeit coins? In this case, the best way to decide which coin will be included in what measurement is randomly and independently. This would be analogous to flipping each coin to decide if it is going to be included in that measurement. The importance of such practice is that it gives equal probability to each component of being measured.

### Reconstructing the signal

In the 8 coin example it is intuitive to see how to identify the counterfeit coin from the measurement matrix. Nevertheless, with a thousand coins it is not so clear. Typically, one is used to solving problems that are determined, that is, systems with the same number of independent equations and unknowns. When the problem is determined it is certain to have a unique solution. This would be the case when each coin is measured separately in the above example or when obtaining a $256 \times 256$ picture image by measuring each pixel individually from the 65 536 array.

On the other hand, compressed sensing systems are by nature underdetermined which means there are less equations in comparison to the number of unknowns. After all, the purpose is to get the relevant information from less measurements. At first glance, however, the problem seems insurmountable, if there are not enough measurements then there are not enough restrictions on the system which results in an infinite number of solutions. How, then, was it possible to solve the 8 coin example above? The answer is sparsity. In the 8 coin example it was known *a priori* that there was only one counterfeit coin, which allowed inferring, or reconstructing, the original signal from the measurement matrix. This can be better illustrated by writing out the measurement matrix ($\phi$) and applying it to the sample, or coin, matrix ($X$) to obtain a resulting weights matrix ($b$), or $b = \phi X$. The $\phi$ matrix is known from the experiment design and the $b$ matrix is known from the measurements performed. The task is to reconstruct the original signal, or in this case the $X$ matrix. It is clear from Fig. 1 that the sparsity restriction of only one counterfeit coin leads to

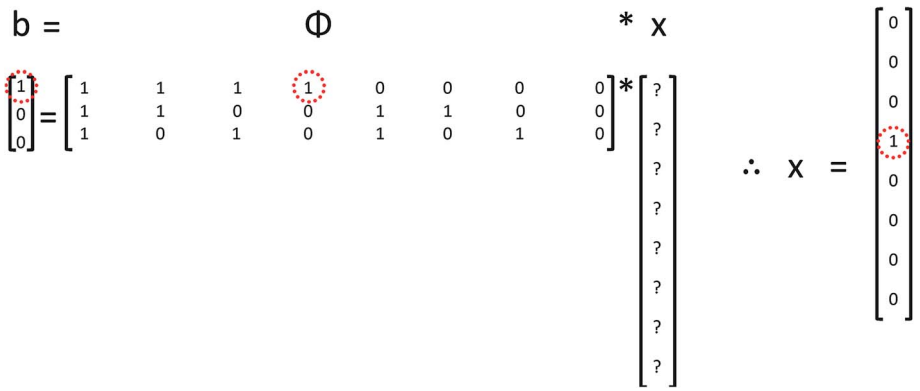



Fig. 1 Matrix representation for finding a single counterfeit gold coin among 8 gold coins. Each row of the measurement matrix $\phi$ represents a combination of coins to be weighed. Each element of a row in $\phi$ determines if the 1st coin, 2nd coin, 3rd coin, *etc.*, will be included (1) in that measurement or not (0). Each row of the resulting weights matrix $b$ denotes the weight of the corresponding combination of coins, or row, in $\phi$. $X$ is the sample matrix and each row stands for a different coin. In the sample matrix, a deviation from the nominal gold coin weight is denoted with a 1 and the absence of a deviation with a 0. In this case, the resulting weights matrix $b$ shows a deviation only in the first measurement, thus, it is clear that the 4th coin is the counterfeit one because it is only included in the first measurement and we know *a priori* that there is only one counterfeit gold coin.

a unique solution of the $X$ matrix. Thus, sparsity is a very powerful restriction for obtaining a unique solution to underdetermined systems.

Then again, not all signals of interest are originally sparse. Take for example a complex wave signal which is very dense when plotted as a function of time. Nevertheless, such signals can be represented in a different basis, or domain, where they are sparse (Fig. 2). Thus, upon subjecting the wave signal in question to a Fourier transform it becomes clear that the signal is sparse in the frequency domain. The same can be said for images which may appear very dense originally but most of the information can be contained within a few elements upon processing with, for example, a discrete cosine transform. Selecting only those elements for reconstructing the image is at the heart of jpeg file compression (Fig. 2).

### $l_1$-Norm minimization, unit circles & sparsity

The unique solution to the $X$ matrix in Fig. 1 is intuitive at a glance but when one has many more coins, or pixels to measure, then an algorithm to reach the unique solution has to be followed. There are several ways to do it, for example, one can take a brute force approach and attempt all potential answers

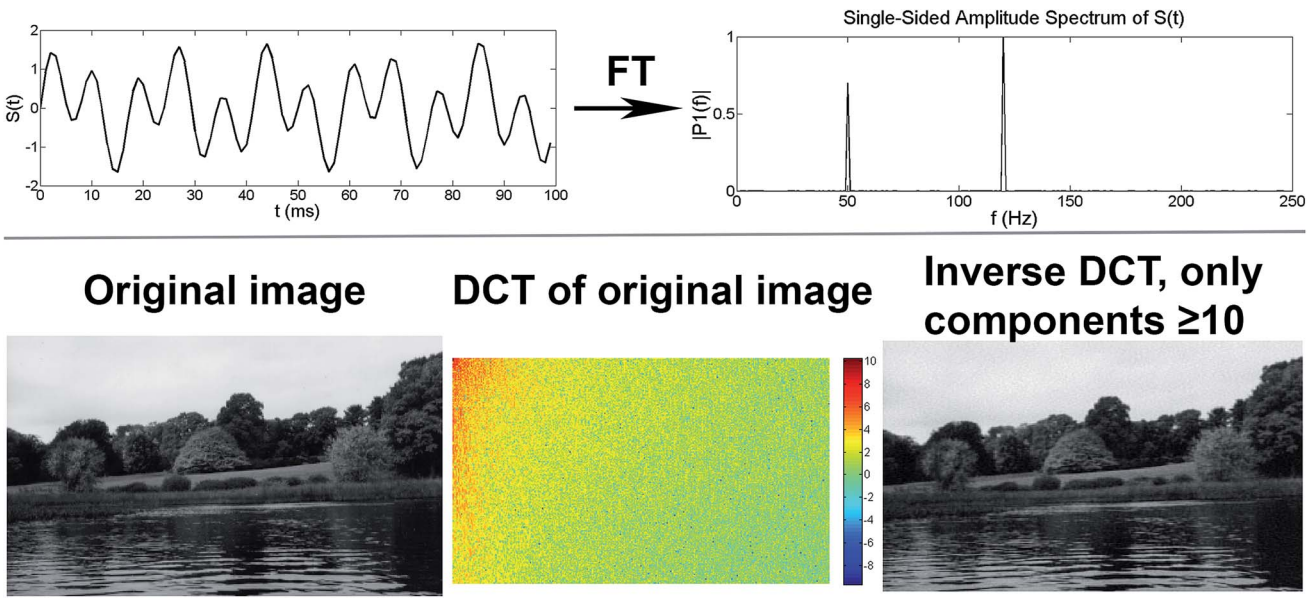



Fig. 2 Signals which are not sparse in their original basis can be sparse when represented in another basis. (Top) A complex signal waveform that varies as a function of time is sparse in the frequency domain. (Bottom) An originally dense image can be represented in a different basis through a discrete cosine transform (DCT). A very accurate reconstruction of the image by inverse DCT with only the components ≥10 shows that the image is highly compressible. Both examples were completed with Matlab software and the included sample autumn.tif image.

until the correct one is found. This approach, however, is very impractical and consumes an exceedingly long time. For example, if someone says "I am thinking of a number between 1 and 100, what is it?" It would take an average of 50 tries ($N/2$) if one were to guess individual numbers. A better way to do it is with a game of "20 questions" where one can ask if the number is greater than 50 and depending on the answer one can adapt the next question until the correct number is found. This approach would take $\log_2 N$ questions, or ~7 for our 100 number example above.[10] Here, the approach to find the unique solution that will be briefly introduced is called $l_1$-norm minimization which has virtually become analogous with compressed sensing. It is similar to the "20 questions" approach but it is non-adaptive to make it relevant in practice.

A norm is a quantity that describes the size of a vector, or matrix, in vector space. The $l_p$-norm of $x$ is:

$$\|x_p\| = \sqrt[p]{\sum_{i=1}^{N} {x_i}^p}$$

The $l_0$-norm is of interest because it will yield the total number of non-zero elements in $x$ given that any non-zero element to the zeroth power will result in one. Technically, the $l_0$-norm is not a norm but that is outside the scope of this review so for now the nomenclature will be kept for clarity purposes. Going back to our 8 coin example in Fig. 1 it is clear that the unique solution is found by minimizing the $l_0$-norm of $X$, *i.e.* its total number of non-zero elements, again because of the sparsity restriction. The drawback of using the $l_0$-norm is that its minimization is regarded as computationally intractable, or NP-hard, if the system is large, which basically means it is too complex to solve.[16,17] On the other side of the spectrum is the $l_2$-norm minimization, which is the very well established least squares minimization approach. In this case, minimization can be performed very efficiently, unfortunately it rarely leads to the sparsest, or in this case correct, solution.[12,14,16] One can get a conceptual view of this by looking at the geometric representation of the line including all the possible solutions to $b = \phi X$ and how it intersects with the two-dimensional "$l_p$-ball", or for $X$ with only two components (Fig. 3). The "$l_p$-ball" contains all equivalent $l_p$-norms which are equal or less than the "ball" radius. The task is to have the $b = \phi X$ line intersect with the "$l_p$-ball" at the minimum $l_p$-norm which is at the axis, where one of the two $x$ components is zero. This is clearly the case with the $l_0$-norm but, again, minimization becomes a computationally intractable problem. On the other hand, it is also clear that the $b = \phi X$ line does not intersect the "$l_2$-ball" at the axis which means it will not give the sparsest $l_2$-norm. Consequently, one can turn to the $l_1$-norm.[12,14,16,17] The "$l_1$-ball" is anisotropic in a similar way to the "$l_0$-ball" such that the vertices lie at the axis. Thus, there is a very high probability that $l_1$-norm minimization will yield the same solution as the $l_0$-norm minimization. This is very powerful because there are several algorithms to minimize the $l_1$-norm and, even though this is not as efficient as a least squares minimization, they have gotten to the point where it is practical to work with thousands of measurements.[18] In addition, such recovery algorithms are able to give very good approximations even when the signal is not perfectly sparse.[8] Furthermore, even when the signal is perturbed by error there are means to recover the signal such that the recovery error is comparable to the error in the measurement.[19]

## Challenges in compressed sensing

One of the challenges in compressed sensing is that of dynamic range. Going back to the counterfeit gold coin example, a smaller dynamic range is needed when weighing the coins individually as opposed to weighing combinations of four coins. This may not be such a problem in this case but it quickly becomes an issue when trying to obtain a megapixel spectral image through compressed sensing approaches. Most measurements of random combinations will give values close to an average but some will give much higher or much lower values. If all the random combinations are to be included within the dynamic range of the detector then the bit-depth will limit the ability to distinguish the differences between the values around the average. One way to get around this is to zoom-in around the average such that the much higher and lower values are outside the measuring window (*i.e.* saturated). Then, one can either reject the saturated measurements or factor them into the recovery algorithm.[20] One of the reasons this works is because each measurement carries approximately equal amount of information.

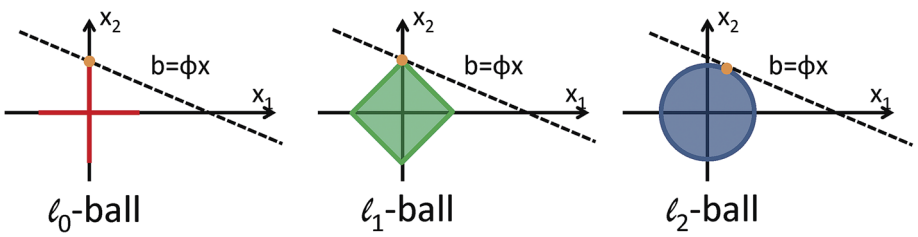



Fig. 3 Simplified geometric representation of the $l_p$-norm minimization where the dotted line including all the possible solutions to $b = \phi X$ intersects with different two-dimensional $l_p$-balls. It is clear that the $l_0$-norm minimization would yield the sparsest and correct solution as shown by the intersection at the axis (minimum number of non-zero elements) but this is an intractable computational problem. The $l_2$-norm minimization (ball shape obtained from Euclidean geometry) is very efficient computationally but does not yield the sparsest solution as shown by the intersection away from the axis. Nevertheless, the $l_1$-ball does have its apex along the axis (ball shape obtained from taxi-cab geometry) thus $l_1$-norm minimization will yield the sparsest solution that can be computed.

Another challenge is when the signal is shot-noise, or Poisson noise, limited. For example, it has been shown that in this case the error bound decreases as the number of sensors goes down for an image of fixed intensity.[21] This may be counter intuitive at first but no so much considering the analogous situation of multiplex advantage turning into a disadvantage in shot-noise limited situations.

## Selected compressed sensing designs for spectral imaging

### Single pixel cameras

One of the most prominent designs for compressed sensing spectral imaging is the single pixel camera.[22,23] Here, an object is imaged onto a digital micromirror device (DMD). The DMD can then sequentially reflect pseudorandom combinations of bundles of light from the image onto a detector consisting of a single sensor (Fig. 4a). Finally, recovery of the original image can be performed by using the known measurement matrix and data. The advantage is that one can circumvent the need for prohibitively expensive, or non-existing in some spectral regions, array detectors while minimizing the time penalties of pixel-by-pixel scanning, or whisker broom, approaches. Nevertheless, a sequence of measurements is still needed for reconstruction such that some time sensitive applications may be outside of the capabilities of this design. Several embodiments now include not just spatial encoding, as described above, but also spectral encoding and/or a one dimensional sensor array for multi-spectral imaging applications.

### Coded apertures

Another approach to compressed sensing spectral imaging is through the use of coded apertures. A coded aperture snapshot spectral imager (CASSI) spatially encodes the image before it passes through a dispersive element onto an array detector.[15,24,25] Basically, the coded aperture replaces the entrance slit on a spectrograph (Fig. 4b). The result will be a "stack" of spatially encoded images which are shifted by wavelength according to the dispersion factor of the spectrograph and measured simultaneously (or integrated) by the array detector. The spectral images can then be recovered through CS algorithms. The main advantage is that several spectral images can be obtained with a single camera shot, although objects with abundant spectral and spatial features need multiple shots for a faithful reconstruction. Nevertheless, it is still more amenable to time sensitive situations compared to the single-pixel approach. On the other hand, it requires an array detector which may be prohibitively expensive in some situations.

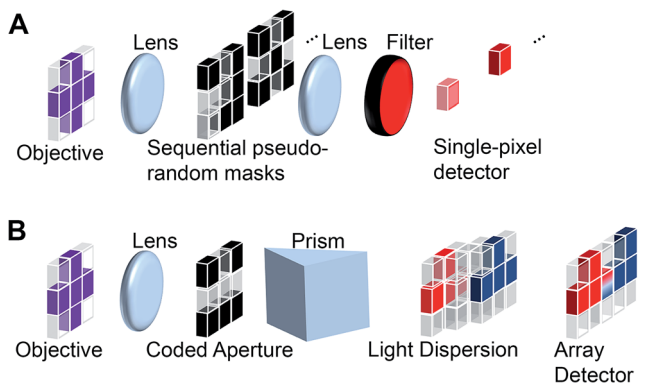



Fig. 4 Typical architectures for compressed sensing spectral imaging. In the single-pixel camera (A) a sequence of pseudo-random binary masks encode the image such that a series of combinations of light bundles arrive at the single-pixel detector. In the coded aperture approach (B) a single aperture replaces the entrance slit on a spectrograph.

## Selected examples of CS in analytical spectrometry

Ever since the seminal paper by Candes, Romberg, and Tao in 2006,[7] there have been a lot of applications reported on compressed sensing regarding imaging and spectroscopy. Table 1 shows many of the different techniques that have benefited from CS approaches. It is evident that there are few studies where atomic spectroscopy or elemental analysis has taken advantage of CS but they have demonstrated that the potential gains are superb. Here, several of these studies will be described, together with some examples from molecular spectroscopy, with the aim of giving the reader a sense of the possibilities and advantages but this is not meant to be an exhaustive list.

The advantages of CS are uniquely harnessed by multidimensional spectroscopy for chemical analysis, even when these are not multiple spatial dimensions. A great example of CS in atomic spectroscopy would be the work of Aspuru-Guzik *et al.*[36] with phase-modulation two-dimensional fluorescence of $^{87}$Rb vapor. In this 2D ultrafast spectroscopy technique the sample is excited using a series of laser pulses separated by coherence times, population times, and measurement times. The authors

Table 1 Selected examples of imaging and spectroscopy techniques benefiting from compressed sensing

| |
|---|
| X-ray diffraction[26,27] |
| X-ray fluorescence spectroscopy[28] |
| Laser induced breakdown spectroscopy[29,30] |
| Fluorescence spectroscopy[31–36] |
| Bioluminescence spectroscopy[37] |
| Raman spectroscopy[38–43] |
| Infrared spectroscopy[44–51] |
| Sum frequency generation spectroscopy[52,53] |
| Terahertz spectroscopy[54–57] |
| Multidimensional nuclear magnetic resonance[58,59] |
| Magnetic resonance imaging[60–66] |
| Electron microscopy[67–71] |
| Electron energy loss spectroscopy[72,73] |
| Atomic force microscopy[74–79] |
| Helium atom scattering[80] |
| Super-resolution imaging[81–88] |
| Confocal microscopy[89–91] |
| Mass spectrometry[92–96] |

showed how using the same sample data they were able to obtain an order-of-magnitude better spectral resolution through uniform grid sampling CS (Fig. 5). Conversely, they also show that it is possible to get similar spectral resolution using less than 5% of the sampled data but with random sampling in the time domain. Kubarych and co-workers also applied CS to multi-dimensional spectroscopy but in this case 2D-IR.[45] They show that the typical features of a 2D-IR system can be reproduced with the CS approach in less than 1/16$^{th}$ of the time. It is worth noting that for many FT spectroscopy techniques the Nyquist sampling theorem applied to time-domain digitization determines maximum observable frequency rather than resolution. A more recent example of compressed sensing applied to multidimensional spectroscopy is the single-point array reconstruction by spatial encoding (SPARSE) approach by Harel *et al.*[97]

Another example of CS in elemental analysis is the work of Bals *et al.*[68] where they achieve 3D elemental mapping at the atomic scale. The authors implement high-angle annular dark-field (HAADF) scanning transmission electron microscopy (STEM) where the projected image intensity depends on the material thickness and the atomic number. Taking advantage of compressed sensing tomography and statistical parameter estimation, they were able to reconstruct the 3D elemental composition of an Au/Ag core shell nanoparticle (Fig. 6) with as little as 5 orthogonal projections.

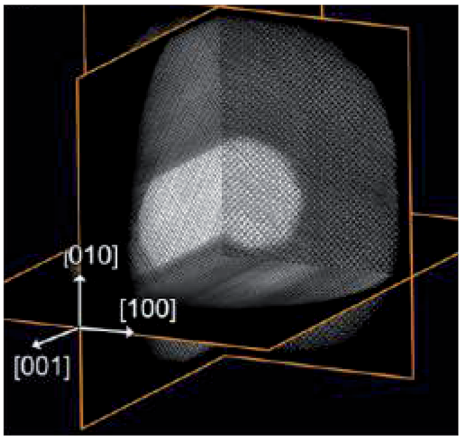



Fig. 6 3D elemental map of an Au core/Ag shell nanoparticle obtained *via* HAAD STEM compressed sensing tomography. Only five projections were needed to reconstruct the image. The nanoparticles had an average length of (35.3 ± 4.6) nm and width of (20.4 ± 1.9) nm. Adapted with permission from Bals *et al.*[68] Copyright (2016) American Chemical Society.

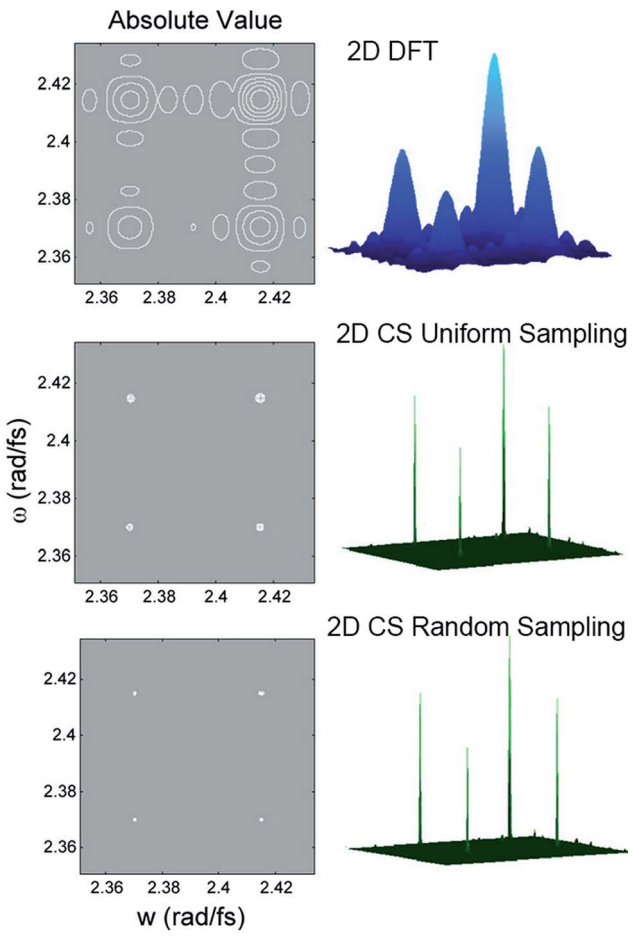



Fig. 5 Phase-modulation two-dimensional fluorescence of $^{87}$Rb vapor *via* 2D discrete Fourier transform, 2D compressed sensing with uniform grid sampling, and 2D compressed sensing with random sampling. The better spectral resolution with CS techniques is evident. Adapted with permission from Aspuru-Guzik *et al.*[36] Copyright (2012) American Chemical Society.

Imaging of laser induced breakdown plasmas has also benefited from compressed sensing approaches. Mochizuki and co-workers[30] demonstrate the use of an ultra-high-speed multi-aperture CMOS compressed sensing imager on LIBS plasmas. This imaging architecture is interesting to examine, basically, it features a 5 × 3 aperture array where each aperture consists of a lens which projects an image onto a CMOS sub-region. The binary modulation of each aperture is implemented in time, such that the captured images are temporally-multiplexed. This approach allows single-shot burst-readout image acquisition rates of 200 Mfps. The authors observe the evolution of a laser induced breakdown plasma in air and show the development of a primary and secondary plasma under their operating conditions (Fig. 7).

In 2006, Brady *et al.* implemented a coded aperture optical spectrograph.[42] The binary (open/closed) 2D pattern on the aperture based on Hadamard *S*-matrices allowed reconstruction from the spectral projections on the focal plane array detector. The increased light throughput due to the size of the aperture (1.73 mm × 4.32 mm) results in a signal-to-noise increase of $(N/2)^{0.5}$ where $N$ refers to the Hadamard matrix order. The large coded aperture also allows for better coupling of large etendue, *i.e.* geometric extent, sources. They show a proof-of-principle spectra with the atomic emission of an argon lamp and move on to perform a Raman spectroscopy quantitative study of ethanol in tissue. In another interesting example, Ben-Amotz and co-authors[40] designed and implemented a multivariate hyperspectral Raman imager. In their design, the collected Raman scattering passes through a volume-phase holographic (VPH)

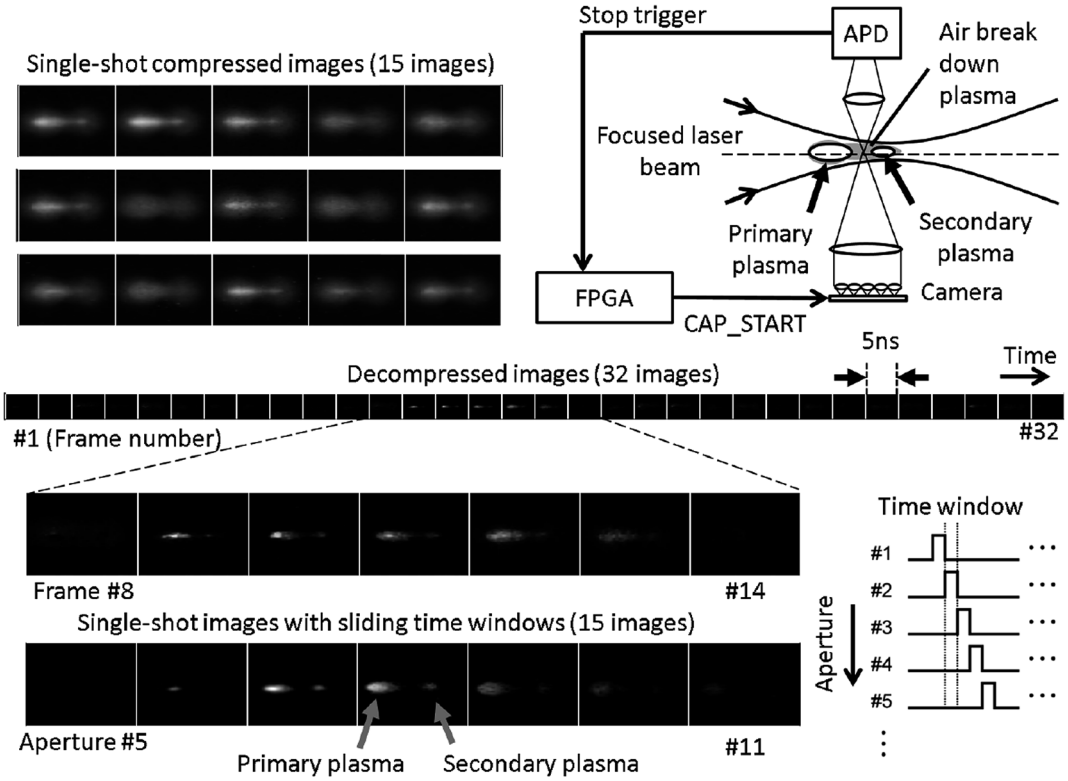



Fig. 7 Ultra-high-speed multi-aperture CMOS compressed sensing imager on LIBS plasmas. An image frame rate of 200 Mfps was achieved with the CS approach. Reprinted with permission from Mochizuki *et al.*[30] Copyright (2015) IEEE.

grating onto a liquid-crystal based reflective spatial light modulator (SLM) such that the reflected spatial/spectral encoded light comes back through the VPH grating and is directed by a beam splitter to single-pixel sensor. The programmable optical filter can then be used to circumvent expensive array detectors or to pre-train filter functions to look for specific analytes which makes the chemical imaging much faster (>1 ms per pixel image).

Baldelli and co-workers[52] have used compressed sensing for sum frequency generation hyperspectral microscopy. In SFG the non-linear optical interaction between a visible laser and an infrared laser on a sample surface is utilized. Advantages include the ability to obtain the surface chemical information provided by IR but collected at visible wavelengths. In their work, the authors encode the visible laser with a sequence of pseudorandom patterns *via* a digital micromirror device. The signal is then collected with a spectrograph and the image recovered through CS algorithms. The authors report hyperspectral images recovered with acquisition in a fifth of the time compared to typical methods.

Glass *et al.* have reported the use of 1D[93] and 2D[96] coded apertures for mass spectrometry. In their 2D *S*-matrix coded aperture work, they use a 90-degree magnetic sector mass spectrometer with a 0.45 T magnet and an array detector comprised of a micro channel plate, phosphor screen and camera. The mass spectrometer features a large uniform ion flux and a magnetic sector with a wide gap and better able to preserve field homogeneity. They show proof-of-principle electron impact ionization mass spectra of ethanol, acetone, and argon (Fig. 8). An increase in throughput by 3.5 times was observed with the 2D coded aperture accompanied with a 1.3× to 1.4× loss in resolution. On the other hand, a similar throughput increase (3.5×) with a typical slit would come with an equivalent loss in resolution (3.5×).

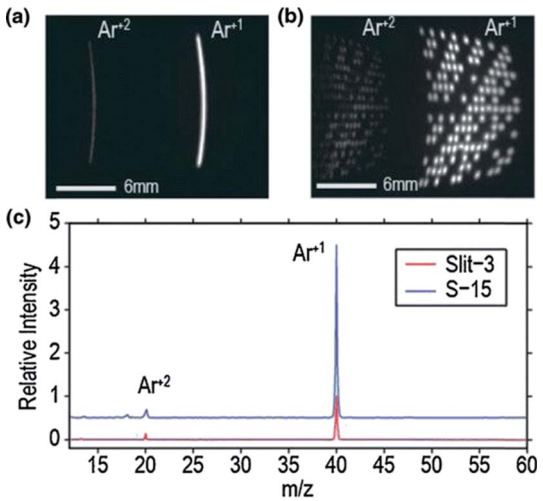



Fig. 8 Ion images from magnetic sector spectrograph with a typical entrance slit (a) compared to a 2D coded aperture (b). The corresponding recovered spectra (c) shows the better S/N for the coded aperture (S – 15) *vs.* the normal slit (slit – 3). Reprinted from Glass *et al.*[96] with permission of Springer.

Furthermore, CS has been shown to be advantageous with regards to data mining. For example, Griffin *et al.*[29] considered the benefits of using compressed sensing for spectroscopic classification of laser induced breakdown spectroscopy data. They compare CS *versus* other approaches of LIBS materials identification such as principal component analysis or partial least squares. One of their main observations is that the signal obtained *via* CS techniques is inherently classified, thus one does not need to reconstruct the signal prior to classification. Conversely, Conrad and co-authors[98] developed a CS approach to pick features and classify proteomics mass spectral data and improve bio-medical interpretation. The sparse proteomics analysis algorithm was reported to perform better than other popular algorithms for the cases studied.

## Conclusions and perspective

The conceptual view of compressed sensing given here has served to develop an intuition of how it works. Furthermore, selected architectures and applications have served to demonstrate the far-reaching and ongoing impact of compressed sensing techniques. The CS benefits of better spatial/spectral resolution with fully sampled data sets, faster acquisition times with undersampled data sets, or higher throughput have been harnessed by many techniques (see Table 1). In addition, novel CS instrument designs and implementations are constantly being reported, like a camera with an order-of-magnitude thinner optics,[49] a lens-less camera,[99] or the use of a CMOS camera's electronics to perform the encoding instead of a physical mask.[100] Nevertheless, there are only a few implementations of CS in atomic spectroscopy, or elemental analysis, thus there are many potential benefits to be realized here. This is illustrated in a paper describing a CS spectral imaging system implemented on plasma optical emission spectroscopy that is concurrently submitted to this issue.[101] It is clear that compressed sensing will keep growing in the field of spectrochemical analysis and will continue permeating into different areas of study, accelerated by further research into its current challenges.

## Acknowledgements

G. Gamez would like to acknowledge funding from the National Science Foundation under CHE – 1610849.

## References

1 R. J. Marks II, *Introduction to Shannon Sampling and Interpolation Theory*, Springer-Verlag, New York, 1991.
2 S. Lopez, T. Vladimirova, C. Gonzalez, J. Resano, D. Mozos and A. Plaza, *Proc. IEEE*, 2013, **101**, 698–722.
3 P. Buonora and F. Liberati, *D-Lib Magazine*, 2008, **14**, DOI: 10.1045/july2008-buonora.
4 J. Högbom, *Astron. Astrophys., Suppl. Ser.*, 1974, **15**, 417–426.
5 J. F. Claerbout and F. Muir, *Geophysics*, 1973, **38**, 826–844.
6 B. Hayes, *Am. Sci.*, 2009, **97**, 276–280.
7 E. J. Candes, J. Romberg and T. Tao, *IEEE Trans. Inf. Theory*, 2006, **52**, 489–509.
8 E. J. Candes, J. K. Romberg and T. Tao, *Communications on Pure and Applied Mathematics*, 2006, **59**, 1207–1223.
9 D. L. Donoho, *IEEE Trans. Inf. Theory*, 2006, **52**, 1289–1306.
10 D. Mackenzie, *What's Happening in the Mathematical Sciences*, 2009, vol. 7, pp. 114–127.
11 D. J. Holland and L. F. Gladden, *Angew. Chem., Int. Ed.*, 2014, **53**, 13330–13340.
12 J. Romberg, *IEEE Signal Process. Mag.*, 2008, **25**, 14–20.
13 R. M. Willett, R. F. Marcia and J. M. Nichols, *Opt. Eng.*, 2011, **50**, 072601.
14 R. G. Baraniuk, *IEEE Signal Process. Mag.*, 2007, **24**, 118–121.
15 G. R. Arce, D. J. Brady, L. Carin, H. Arguello and D. S. Kittle, *IEEE Signal Process. Mag.*, 2014, **31**, 105–115.
16 K. Bryan and T. Leise, *SIAM Rev.*, 2013, **55**, 547–566.
17 D. L. Donoho, *Communications on Pure and Applied Mathematics*, 2006, **59**, 797–829.
18 J. A. Tropp and S. J. Wright, *Proc. IEEE*, 2010, **98**, 948–958.
19 J. A. Tropp, *IEEE Trans. Inf. Theory*, 2006, **52**, 1030–1051.
20 J. N. Laska, P. T. Boufounos, M. A. Davenport and R. G. Baraniuk, *Appl. Comput. Harmon. Anal.*, 2011, **31**, 429–443.
21 M. Raginsky, R. M. Willett, Z. T. Harmany and R. F. Marcia, *IEEE Trans Sig. Process.*, 2010, **58**, 3990–4002.
22 M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly and R. G. Baraniuk, *IEEE Signal Process. Mag.*, 2008, **25**, 83–91.
23 D. Takhar, J. N. Laska, M. B. Wakin, M. E. Duarte, D. Baron, S. Sarvotham, K. E. Kelly and R. G. Baraniuk, in *Computational Imaging IV*, ed. C. A. Bouman, E. L. Miller and I. Pollak, 2006, vol. 6065, pp. 6509–6509.
24 D. J. Brady and M. E. Gehm, *Proceedings of SPIE - The International Society for Optical Engineering*, Kissimmee, FL, 2006.
25 R. F. Marcia, Z. T. Harmany and R. M. Willett, in *Computational Imaging Vii*, ed. C. A. Bouman, E. L. Miller and I. Pollak, 2009, vol. 7246.
26 J. A. Greenberg and D. J. Brady, in *Computational Imaging Xii*, ed. C. A. Bouman and K. D. Sauer, 2014, vol. 9020.
27 J. Greenberg, K. Krishnamurthy and D. Brady, *Opt. Lett.*, 2014, **39**, 111–114.
28 C. Hall, R. G. Acres, A. Winnett and F. Wang, *J. Instrum.*, 2016, **11**, C03048.
29 S. T. Griffin, E. Jacobs and O. Furxhi, *Proceedings of SPIE - The International Society for Optical Engineering*, Orlando, FL, 2011.
30 F. Mochizuki, K. Kagawa, S. I. Okihara, M. W. Seo, B. Zhang, T. Takasawa, K. Yasutomi and S. Kawahito, *Digest of Technical Papers - IEEE International Solid-State Circuits Conference*, 2015.
31 M. Dahan, *Imaging Systems and Applications, ISA 2012*, Monterey, CA, 2012.
32 Q. Pian, R. Yao and X. Intes, *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*, 2016.
33 V. Studera, J. Bobin, M. Chahida, H. S. Mousavia, E. Candes and M. Dahane, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, E1679–E1687.

34 J. Wang, C. Kuang, Y. Wang and X. Liu, *Zhongguo Jiguang*, 2013, **40**, 1204003.
35 W. Zou and X. Pan, *BioMedical Engineering OnLine*, 2014, **13**, 119.
36 J. N. Sanders, S. K. Saikin, S. Mostame, X. Andrade, J. R. Widom, A. H. Marcus and A. Aspuru-Guzik, *J. Phys. Chem. Lett.*, 2012, **3**, 2697–2702.
37 Y. Lu, X. Zhang, A. Douraghy, D. Stout, J. Tian, T. F. Chan and A. F. Chatziioannou, *Opt. Express*, 2009, **17**, 8062–8080.
38 D. G. Carreno and H. A. Fuentes, *Symposium of Signals, Images and Artificial Vision - 2013, STSIVA 2013*, Bogota, 2013.
39 A. Cocking, N. Mehta, K. Shi and Z. Liu, *Opt. Express*, 2015, **23**, 24991–24996.
40 B. M. Davis, A. J. Hemphill, D. C. Maltas, M. A. Zipper, P. Wang and D. Ben-Amotz, *Anal. Chem.*, 2011, **83**, 5086–5092.
41 C. Jenila and A. S. Raja, *Opt. Quantum Electron.*, 2015, **47**, 3855–3862.
42 S. T. McCain, M. E. Gehm, Y. Wang, N. P. Pitsianis and D. J. Brady, *Appl. Spectrosc.*, 2006, **60**, 663–671.
43 J. Monsalve, H. Vargas and H. Arguello, *2015 20th Symposium on Signal Processing, Images and Computer Vision, STSIVA 2015-Conference Proceedings*, 2015.
44 R. Bhargava, *Appl. Spectrosc.*, 2012, **66**, 1091–1120.
45 J. A. Dunbar, D. G. Osborne, J. M. Anna and K. J. Kubarych, *J. Phys. Chem. Lett.*, 2013, **4**, 2489–2492.
46 J. R. Dupuis, M. Kirby and B. R. Cosofret, *Proceedings of SPIE - The International Society for Optical Engineering*, 2015.
47 G. Frigo, S. Brigadoi, G. Giorgi, G. Sparacino and C. Narduzzi, *2015 IEEE International Symposium on Medical Measurements and Applications, MeMeA 2015-Proceedings*, 2015.
48 G. Frigo, S. Brigadoi, G. Giorgi, G. Sparacino and C. Narduzzi, *IEEE Trans. Instrum. Meas.*, 2016, **65**, 1310–1318.
49 M. Shankar, R. Willett, N. Pitsianis, T. Schulz, R. Gibbons, R. T. Kolste, J. Carriere, C. Chen, D. Prather and D. Brady, *Appl. Opt.*, 2008, **47**, B1–B10.
50 Y. Wu, G. R. Arce and D. W. Prather, *Applied Industrial Optics: Spectroscopy, Imaging and Metrology, AIO 2012*, Monterey, CA, 2012.
51 Z. J. Zhang, L. Liu, X. R. Li, L. Gan, Y. Huang and Y. C. Shen, *IET Seminar Digest*, 2016.
52 D. Zheng, L. Lu, Y. Li, K. F. Kelly and S. Baldelli, *J. Phys. Chem. Lett.*, 2016, **7**, 1781–1787.
53 X. Cai, B. Hu, T. Sun, K. F. Kelly and S. Baldelli, *J. Chem. Phys.*, 2011, **135**, 194202.
54 S. Augustin, J. Hieronymus, P. Jung and H. W. Hübers, *J. Infrared, Millimeter, Terahertz Waves*, 2015, **36**, 496–512.
55 D. Coltuc, *Proceedings of SPIE - The International Society for Optical Engineering*, 2015.
56 P. Duan, Y. Wang, D. Xu, C. Yan, Z. Yang, W. Xu, W. Shi and J. Yao, *Appl. Opt.*, 2016, **55**, 3670–3675.
57 C. C. Nadell, C. M. Watts, J. A. Montoya, S. Krishna and W. J. Padilla, *Adv. Opt. Mater.*, 2016, **4**, 66–69.
58 M. J. Bostock, D. J. Holland and D. Nietlispach, *J. Biomol. NMR*, 2012, **54**, 15–32.
59 D. J. Holland, M. J. Bostock, L. F. Gladden and D. Nietlispach, *Angew. Chem., Int. Ed.*, 2011, **50**, 6548–6551.
60 M. Lustig, D. L. Donoho, J. M. Santos and J. M. Pauly, *IEEE Signal Process. Mag.*, 2008, **25**, 72–82.
61 B. Deka and S. Datta, *ACM International Conference Proceeding Series*, 2014.
62 S. Geethanath, R. Reddy, A. S. Konar, S. Imam, R. Sundaresan, D. R. Ramesh Babu and R. Venkatesan, *Crit. Rev. Biomed. Eng.*, 2013, **41**, 183–204.
63 S. G. Lingala and M. Jacob, *IEEE Trans. Med. Imag.*, 2013, **32**, 1132–1145.
64 Y. Nan, Z. Yi and C. Bingxia, *Proceedings - 2015 7th International Conference on Information Technology in Medicine and Education, ITME 2015*, 2015.
65 D. S. Smith, X. Li, R. G. Abramson, C. C. Quarles, T. E. Yankeelov and E. B. Welch, *Canc. Imag.*, 2013, **13**, 633–644.
66 J. Trzasko and A. Manduca, *IEEE Trans. Med. Imag.*, 2009, **28**, 106–121.
67 H. S. Anderson, J. Ilic-Helms, B. Rohrer, J. Wheeler and K. Larson, *Proceedings of SPIE - The International Society for Optical Engineering*, Burlingame, CA, 2013.
68 B. Goris, A. De Backer, S. Van Aert, S. Gómez-Graña, L. M. Liz-Marzán, G. Van Tendeloo and S. Bals, *Nano Lett.*, 2013, **13**, 4236–4241.
69 R. Leary, Z. Saghi, P. A. Midgley and D. J. Holland, *Ultramicroscopy*, 2013, **131**, 70–91.
70 A. Al-Afeef, W. P. Cockshott, I. MacLaren and S. McVitie, *Scanning*, 2016, **38**, 251–276.
71 M. D. Guay, W. Czaja, M. A. Aronova and R. D. Leapman, *Sci. Rep.*, 2016, **6**, 27614.
72 A. Hörl, A. Trügler and U. Hohenester, *ACS Photonics*, 2015, **2**, 1429–1435.
73 O. Nicoletti, F. De La Peña, R. K. Leary, D. J. Holland, C. Ducati and P. A. Midgley, *Nature*, 2013, **502**, 80–84.
74 Y. Luo and S. B. Andersson, *Nanotechnology*, 2015, **26**, 505703.
75 S. B. Andersson and L. Y. Pao, *Proceedings of the American Control Conference*, Montreal, QC, 2012.
76 Y. Luo and S. B. Andersson, *Proceedings of the American Control Conference*, 2015.
77 B. D. Maxwell and S. B. Andersson, *Proceedings of the American Control Conference*, Portland, OR, 2014.
78 P. S. Pedersen, J. Ostergaard and T. Larsen, *2015 IEEE Global Conference on Signal and Information Processing, GlobalSIP 2015*, 2015.
79 B. Song, N. Xi, R. Yang, K. W. C. Lai and C. Qu, *Proceedings of the IEEE Conference on Nanotechnology*, Portland, OR, 2011.
80 A. Jones, A. Tamtögl, I. Calvo-Almazán and A. Hansen, *Sci. Rep.*, 2016, **6**, 27776.
81 S. Gazit, A. Szameit, Y. C. Eldar and M. Segev, *Opt. Express*, 2009, **17**, 23920–23946.
82 W. AlSaafin, S. Villena, M. Vega, R. Molina and A. K. Katsaggelos, *Digit. Signal Process.*, 2016, **50**, 180–190.

83 Y. Mao, Y. Wang, J. Zhou and H. Jia, *Infrared Phys. Technol.*, 2016, **76**, 735–739.
84 R. F. Marcia and R. M. Willett, *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, Las Vegas, NV, 2008.
85 Z. Pan, J. Yu, H. Huang, S. Hu, A. Zhang, H. Ma and W. Sun, *IEEE Transactions on Geoscience and Remote Sensing*, 2013, **51**, 4864–4876.
86 Y. Sun, G. Gu, X. Sui and Y. Liu, *IEEE Photonics J.*, 2016, **8**, 6900112.
87 Y. Sun, X. Sui, G. Gu, Y. Liu and S. Xu, *IEEE Photonics J.*, 2016, **8**, 6900508.
88 S. Zhang, G. Dong and G. Kuang, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2016, **9**, 2184–2196.
89 R. Horisaki, Y. Tampa and J. Tanida, *Computational Optical Sensing and Imaging, COSI 2012*, Monterey, CA, 2012.
90 P. Ye, J. L. Paredes, G. R. Arce, Y. Wu, C. Chen and D. W. Prather, *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, Taipei, 2009.
91 P. Ye, J. L. Paredes, Y. Wu, C. Chen, G. R. Arce and D. W. Prather, *Proceedings of SPIE - The International Society for Optical Engineering*, San Jose, CA, 2009.
92 B. Andreas, D. Patrick, T. Dennis, A. Theodore and M. Peter, *Inverse Problems*, 2013, **29**, 125015.
93 E. X. Chen, Z. E. Russell, J. J. Amsden, S. D. Wolter, R. M. Danell, C. B. Parker, B. R. Stoner, M. E. Gehm, J. T. Glass and D. J. Brady, *J. Am. Soc. Mass Spectrom.*, 2015, **26**, 1633–1640.
94 J. X. Liu and Q. S. Sun, *Proceedings - 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2012*, Chongqing, 2012.
95 J. X. Liu and Q. S. Sun, *IET Signal Processing*, 2013, **7**, 201–209.
96 Z. E. Russell, E. X. Chen, J. J. Amsden, S. D. Wolter, R. M. Danell, C. B. Parker, B. R. Stoner, M. E. Gehm, D. J. Brady and J. T. Glass, *J. Am. Soc. Mass Spectrom.*, 2015, **26**, 248–256.
97 A. P. Spencer, B. Spokoyny, S. Ray, F. Sarvari and E. Harel, *Nat. Commun.*, 2016, **7**, 10434.
98 T. Conrad, M. Genzel, N. Cvetkovic, N. Wulkow, A. Leichtle, J. Vybiral, G. Kutyniok and C. Schutte, arXiv:1506.03620, 2015.
99 G. Huang, H. Jiang, K. Matthews and P. Wilford, in *2013 20th IEEE International Conference on Image Processing*, 2013, pp. 2101–2105.
100 R. Robucci, J. D. Gray, L. K. Chiu, J. Romberg and P. Hasler, *Proc. IEEE*, 2010, **98**, 1089–1101.
101 J. Usala, A. Maag, T. Nelis and G. Gamez, Compressed Sensing Spectral Imaging for Plasma Optical Emission Spectroscopy, *J. Anal. At. Spectrom.*(*submitted*).