Measure What Should be Measured: Progress and Challenges in Compressive Sensing

Thomas Strohmer

***Abstract*—Is compressive sensing overrated? Or can it live up to our expectations? What will come after compressive sensing and sparsity? And what has Galileo Galilei got to do with it? Compressive sensing has taken the signal processing community by storm. A large corpus of research devoted to the theory and numerics of compressive sensing has been published in the last few years. Moreover, compressive sensing has inspired and initiated intriguing new research directions, such as matrix completion. Potential new applications emerge at a dazzling rate. Yet some important theoretical questions remain open, and seemingly obvious applications keep escaping the grip of compressive sensing. In this paper[1] I discuss some of the recent progress in compressive sensing and point out key challenges and opportunities as the area of compressive sensing and sparse representations keeps evolving. I also attempt to assess the long-term impact of compressive sensing.**

I. Introduction

"Measure what can be measured", this quote often attributed to Galileo Galilei, has become a paradigm for scientific discovery that seems to be more dominant nowadays than ever before[2]. However, in light of the data deluge we are facing today, it is perhaps time to modify this principle to *"Measure what should be measured"*. Of course the problem is that a priori we often do not know what we should measure and what not. What is important and what can be safely ignored?

A typical example is a digital camera, which acquires in the order of a million measurements each time a picture is taken, only to dump a good portion of the data soon after the acquisition through the application of an image compression algorithm. In contrast, *compressive sensing* operates under the premise that signal acquisition and data compression can be carried out simultaneously: *"Measure what should be measured!"*

On the one end of the spectrum of scientific endeavour, the concept of compressive sensing has led to the development of new data acquisition devices. On the other end, the beauty of the underlying mathematical theory has attracted even pure mathematicians. And "in between", scientists from physics, astronomy, engineering, biology, medical image processing, etc. explore the possibilities of sparse representations and the opportunities of compressive sensing. It is therefore only natural for such a timely journal as the IEEE Signal Processing Letters, that compressive sensing and sparsity are now incorporated into the new EDICS categories.

At the mathematical heart of compressive sensing lies the discovery that it is possible to reconstruct a sparse signal exactly from an underdetermined linear system of equations *and* that this can be done in a computationally efficient manner via convex programming. To fix ideas and notation, consider $\mathbf{Ax} = \mathbf{y}$, where $\mathbf{A}$ is an $m \times n$ matrix of rank m with $m < n$. Here, $\mathbf{A}$ models the measurement (or sensing) process, $\mathbf{y} \in \mathbb{C}^m$ is the vector of observations and $\mathbf{x} \in \mathbb{C}^n$ is the signal of interest. Conventional linear algebra wisdom tells us that in principle the number of measurements m has to be at least as large as the signal length n, otherwise the system would be underdetermined and there would be infinitely many solutions. Most data acquisition devices of current technology obey this principle in one way or another (for instance, devices that follow Shannon's Sampling Theorem which states that the sampling rate must be at least twice the maximum frequency present in the signal).

Now assume that $\mathbf{x}$ is sparse, i.e., $\mathbf{x}$ satisfies $s := \|\mathbf{x}\|_0 \ll n$ (where $\|\mathbf{x}\|_0 := \#\{k : x_k \neq 0\}$), but we do *not* know the locations of the non-zero entries of $\mathbf{x}$. Due to the sparsity of $\mathbf{x}$ one could try to compute $\mathbf{x}$ by solving the optimization problem

$$\min_{\mathbf{z}} \|\mathbf{z}\|_0 \quad \text{s.t.} \;\; \mathbf{Az} = \mathbf{y}. \tag{1}$$

However solving (1) is an NP-hard problem and thus practically not feasible. Instead we consider its convex relaxation

$$\min_{\mathbf{z}} \|\mathbf{z}\|_1 \quad \text{s.t.} \;\; \mathbf{Az} = \mathbf{y}, \tag{2}$$

which can be solved efficiently via linear or quadratic programming techniques. It is by now well-known that under certain conditions on the matrix $\mathbf{A}$ and the sparsity of $\mathbf{x}$, both (1) and (2) have the same unique solution [2]–[4]. The *Restricted Isometry Property* (RIP) and the *coherence* of a matrix are to date the most widely used tools to derive such conditions. Indeed, for a properly chosen $\mathbf{A}$ about $m = s \log n$ measurements suffice to uniquely recover $\mathbf{x}$ from (2). In other words, a sparse signal can be sampled at almost its "information rate" using non-adaptive linear measurements.

Compressive sensing took the signal processing community by storm. As the graph in [5] shows, the number of publications dealing with sparse representations and compressive sensing has grown rapidly over the last couple of years. Admittedly, we were in a somewhat lucky situation when compressive sensing arrived on the scene: Researchers in

T. Strohmer is with the Department of Mathematics, University of California at Davis, Davis CA-95616, USA. This work was supported in part by the National Science Foundation via grant DTRA-DMS 1042939 and by DARPA via grant N66001-11-1-4090.

[1] This is not a regular IEEE-SPL paper, but rather an invited contribution offering a vision for key advances in emerging fields.

[2] The full quote says *"Measure what can be measured and make measurable what cannot be measured"*, but it is disputed whether Galilei ever said or wrote these words [1]. Nevertheless, the quote is widely accepted as a very fitting characterization of the leitmotif of Galilei's work with respect to the central role of the *experiment* in the *Nuova Scienza*.

signal processing, applied harmonic analysis, imaging sciences, and information theory had already fostered a culture of close collaboration and interaction over the last two decades or so, laying the foundation for a strong willingness from engineers, statisticians, and mathematicians to cooperate and learn from each other. This fact definitely contributed to the very quick adoption of compressive sensing by the various research communities.

Is compressive sensing overrated? Will compressive sensing revolutionize data acquisition? Can compressive sensing live up to our (admittedly, rather high) expectations? What are the most promising applications? Are there still interesting open mathematical problems? And what will come after compressive sensing and sparse representations? While this article may not be able to provide satisfactory answers to all these questions, it is nevertheless strongly motivated by them. I will discuss open problems and challenges, and while doing so, shed light on some recent progress. I will also attempt to evaluate the impact of compressive sensing in the context of future scientific developments.

I also want to draw the reader's attention to the enlightening article *"Sparse and Redundant Representation Modeling — What Next?"* by Michael Elad in the very same issue of this journal. I have tried to keep the topics discussed in my article somewhat complementary to his, but, naturally, our two articles do overlap at places, which was in part not avoidable, since we were writing them at about the same time. The reader, who wonders why I did not mention the one or the other important open problem, may likely find it very eloquently described in Elad's paper. I want to stress at this point that the areas of compressive sensing and sparse representations obviously have a strong overlap, but one should not conflate them completely.

I assume that the reader is familiar with the basics of compressive sensing and sparse representations. Excellent introductions to compressive sensing are the review articles [6], [7], the soon-to-be-published book [8], and of course the original research papers [2]–[4]. A great source for sparse and redundant representations is [9]. The reader who wants to get an overview of recent developments in these areas should also check out Igor Carron's informative blog *Nuit Blanche* (http://nuit-blanche.blogspot.com).

II. Progress and Challenges

In this section I will discuss some problems which I consider important future research directions in compressive sensing. They range from very concrete to quite abstract/conceptual, from very theoretical to quite applied. In some of the problems mentioned below we already have seen significant progress over the last few years, others are still in their infancy. The ordering of the problems does not reflect their importance, but is chosen to best aid the narrative of the paper. The list is by no means exhaustive, moreover it is subjective and biased towards the author's background, taste, and viewpoints. To highlight the connection with the new EDICS related to sparsity and compressive sensing, I am listing the EDICS most relevant for each subsection: Subsection II-A: MLAS-SPARS, SAM-SPARCS; Subsection II-B: DSP-SPARSE; Subsection II-C: MLAS-SPARS, IMD-SPAR, SAM-SPARCS; Subsection II-D: MLAS-SPARS, IMD-SPAR, SAM-SPARCS; Subsection II-E: DSP-SPARSE, SAM-SPARCS; Subsection II-F: DSP-SPARSE; Subsection II-G: DSP-SPARSE, SAM-SPARCS.

A. Structured sensing matrices

Much of the theory concerning explicit performance bounds for compressive sensing revolves around Gaussian and other random matrices. These results have immense value as they show us, in principle, the possibilities of compressive sensing. However, in reality we usually do not have the luxury to choose $\mathbf{A}$ as we please. Instead the sensing matrix is often dictated by the physical properties of the sensing process (e.g., the laws of wave propagation) as well as by constraints related to its practical implementability. Furthermore, sensing matrices with a specific structure can give rise to fast algorithms for matrix-vector multiplication, which will significantly speed up recovery algorithms. Thus the typical sensing matrix in practice is not Gaussian or Bernoulli, but one with a very specific structure, e.g. see [10]–[13]. This includes deterministic sensing matrices as well as matrices whose entries are random variables which are coupled across rows and columns in a peculiar way. This can make it highly nontrivial to apply standard proof techniques from the compressive sensing literature.

Over the last few years researchers have developed a fairly good understanding of how to derive compressive sensing theory for a variety of structured sensing matrices that arise in applications, see for instance the survey article [14] for many examples and references as well as the work of Rauhut [15]. Despite this admirable progress, the derived bounds obtained so far are not as strong as those for Gaussian-type random matrices. One either needs to collect more measurements or enforce more restrictive bounds on the signal sparsity compared to Gaussian matrices, or one has to sacrifice *universality*. Here, universality means that a fixed (random) sensing matrix guarantees recovery of *all* sufficiently sparse signals. In comparison, to obtain competitive theoretical bounds using structured sensing matrices we may have to assume that the locations and/or the signs of the non-zero entries of the signal are randomly chosen [15]–[17]. As a consequence the performance guarantees obtained are not universal, as they "only" hold for *most* signals.

So far involved and cumbersome combinatorial arguments, which need to be carefully adapted to the algebraic structure of the matrix for each individual case, often provide the best theoretical performance bounds for structured matrices – and yet, as mentioned before, these bounds still fall short of those for Gaussian matrices. Can we overcome these limitations of the existing theory by developing a collection of tools that allows us to build a compressive sensing theory for structured matrices that is (almost) on par with that for random matrices?

Now let us change our viewpoint somewhat. Assume that we *do* have the freedom to design the sensing matrix. The only condition we impose is that we want deterministic (explicit) constructions with the goal to establish performance bounds that are comparable to those of random matrices, for instance by establishing appropriate RIP bounds. Most bounds to date on the RIP for deterministic matrix constructions are based on the coherence, which in turn causes the number of required samples to scale quadratically with the signal sparsity. In [18] the authors use extremely sophisticated and delicate arguments to achieve an ε-improvement in this scaling behavior of the bounds.

This poses the question, whether we can come up with deterministic matrices which satisfy the RIP in the optimal range of parameters. It may well be that the so constructed matrices will have little use in practice. But if we succeed in this enterprise, I expect the mathematical techniques developed for this purpose to have impact far beyond compressive sensing.

B. Caught between two worlds: The gridding error

With a few exceptions, the development of compressive sensing until recently has focused on signals having a sparse representation in discrete, finite-dimensional dictionaries. However, signals arising in applications such as radar, sonar, and remote sensing are typically determined by a few parameters in a continuous domain.

A common approach to make the recovery problem amenable to the compressive sensing framework, is to discretize the continuous domain. This will result in what is often called the *gridding error* or basis mismatch [19]. By trying to mitigate the gridding error, we quickly find ourselves caught between two worlds, the continuous and the discrete world. The issue is best illustrated with a concrete example. Suppose the signal of interest is a multitone signal of the form

$$y(t) = \sum_{k=1}^{s} c_k e^{j2\pi t f_k}, \tag{3}$$

with unknown amplitudes $\{c_k\}$ and unknown frequencies $\{f_k\} \subset [-W, W]$. Assume we sample y at the time points $\{t_l\}_{l=1}^{m} \subset [0, 1)$, the goal is to find $\{f_k\}_{k=1}^{s}$ and $\{c_k\}_{k=1}^{s}$ given $\mathbf{y} := \{y(t_1), \ldots y(t_m)\}$. This is the well-known spectral estimation problem, and numerous methods have been proposed for its solution. But the keep in mind that I chose (3) only for illustration purposes, in truth we are interested in much more general sparse signals. We choose a regular grid $\mathcal{G} = \{\frac{\Delta i}{2W}\}_{i=-N}^{N} \subset [-W, W]$, where Δ is the stepsize. Let the sensing matrix be

$$\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_n], \qquad \text{with } \mathbf{a}_i = \frac{1}{\sqrt{m}} \{e^{j2\pi t_l \Delta i/(2W)}\}_{i=-N}^{N}.$$

(An approximation to) the spectral estimation problem can now be expressed as

$$\mathbf{A}\mathbf{x} = \mathbf{y} + \mathbf{e}, \tag{4}$$

with $\mathbf{e} = \mathbf{n} + \mathbf{d}$ being the error vector, where $\mathbf{n}$ models additive measurement noise and $\mathbf{d}$ represents noise due to the discretization or gridding error[3]. Assuming that the frequencies f_k fall on the grid points in $\mathcal{G}$, the vector $\mathbf{x}$ will have exactly s non-zero entries with coefficients $\{c_k\}_{k=1}^{s}$. In general however the frequencies will not lie on the grid $\mathcal{G}$, resulting in a large gridding error, which creates a rather unfavorable situation for sparse recovery. To guarantee that (4) is a good approximation to the true spectral estimation problem, we need to ensure a small gridding error. For each f_k to be close to some grid point in $\mathcal{G}$, we may have to choose Δ to be very small. However, this has two major drawbacks: (i) the number of columns of $\mathbf{A}$ will be large, which will negatively impact the numerical efficiency of potential recovery algorithms. (ii) The coherence of $\mathbf{A}$ will be close to 1, which implies extremely bad theoretical bounds when applying standard coherence-based estimates.

Thus we are caught in a conundrum: Choosing a smaller discretization step on the one hand reduces the gridding error, but on the other hand increases the coherence as well as the computational complexity. This problem begs for a clever solution.

The *finite rate of innovation* concept [20] might be useful in this context, but that concept by itself does not lead to stable and fast algorithms or a framework that can handle signals that are only approximately sparse.

Promising approaches to mitigate the gridding problem can be found in [21], [22]. Both of the proposed approaches have their benefits, but also some drawbacks. Since the purpose of this paper is to point out open problems, let me focus on the drawbacks here, but I want to stress that I find the simplicity of [21] and the ingenuity of [22] very appealing. The theoretical assumptions on the signal sparsity and the dynamic range of the coefficients in [21] are much more restrictive than those of the best results we have for standard compressive sensing. Moreover, only approximate, but not exact, support recovery is guaranteed. The framework of [22], based on an intriguing approach to superresolution in [23], does not require a discretization step, but it is currently limited to very specific classes of signals. Also, the proposed numerical algorithm lacks some of the simplicity of ℓ_1-minimization.

Can we develop a rigorous compressive sensing framework for signals which are sparse in a continuous domain, that is applicable to a large class of signals, and comes with simple, efficient numerical algorithms that preserve as much as possible the simplicity and power of the standard compressive sensing approach? Can we derive theoretical guarantees about the superresolution capabilities of compressive sensing based methods? In this context we refer the reader to [24], where an infinite-dimensional framework for compressive sensing is proposed.

C. Structured sparsity and other prior information

The work of Lustig and collaborators in MRI [10] has shown that a careful utilization of the distribution of the large wavelet coefficients across scales can lead to substantial improvements in the practical performance of compressive sensing in MRI. “Classical” compressive sensing theory does not assume any structure or other prior information about the locations of the non-zero entries of the signal. How can we best take advantage of the knowledge that all sparsity patterns

[3]We could also have captured the gridding error as a perturbation $\mathbf{E}$ of the sensing matrix, $\tilde{\mathbf{A}} := \mathbf{A} + \mathbf{E}$, but it would not change the gist of the story.

may not be equally likely in a signal? This question is a topic of active research, e.g. see [25]–[27] as well as many more references in [14].

Structured sparsity is only one of many kinds of prior information we may have about the signal or image. Besides obvious constraints such as non-negativity of the signal coefficients, there is also application-specific prior information, such as the likelihood of certain molecule configurations or a minimum distance between sparse coefficients due to some repelling force. In particular in the low SNR regime the proper utilization of available prior information can have a big impact on the quality of the recovered signal. The aim is to develop frameworks that can incorporate various kinds of prior information both at the theoretical and the algorithmic level of compressive sensing.

D. Beyond sparsity and compressive sensing

A very intriguing extension of compressive sensing is the problem of recovering a low-rank matrix from incomplete information, also known as the problem of *matrix completion* or *matrix recovery* [28], [29]. Let $\mathbf{X}$ be an $n \times n$ matrix. We do not require that $\mathbf{X}$ is a sparse matrix, but instead we assume that most of its singular values are zero, i.e., the rank of $\mathbf{X}$ is small compared to n. Suppose we are given a linear map $\mathcal{A} : \mathbb{C}^{n\times n} \to \mathbb{C}^m$ and measurements $\mathbf{y} = \mathcal{A}(\mathbf{X})$. Can we recover $\mathbf{X}$? Trying to find $\mathbf{X}$ by minimizing the rank of $\mathbf{Z}$ subject to $\mathcal{A}(\mathbf{Z}) = y$ would be natural but is computationally not feasible. Inspired by concepts of compressive sensing we are led to consider the *nuclear norm* minimization problem

$$\min \|\mathbf{Z}\|_* \quad \text{subject to } \mathcal{A}(\mathbf{Z}) = y,$$

where $\|\mathbf{Z}\|_*$ denotes the sum of the singular values of $\mathbf{Z}$. A large body of literature has been published on the topic of matrix completion, covering conditions and algorithms under which the nuclear norm minimization (or variations thereof) can indeed recover $\mathbf{X}$. Interestingly, the paper [30] derives a framework that allows us to translate (some) recovery conditions from compressive sensing to the setting of matrix completion.

Many high-dimensional data structures are not just sparse in some basis, but in addition are highly correlated across some coordinate axes. For instance spectral signatures in a hyperspectral data cube are often highly correlated across wavelength. Suppose now $\mathbf{X}$ is a hyperspectral data matrix whose columns represent hyperspectral images and the column index corresponds to wavelength. We would like to acquire the information represented by $\mathbf{X}$ with very few measurements only. We take measurements of the form $y = \mathcal{A}(\mathbf{X})$, where $\mathcal{A}$ is a properly designed sensing operator. Following ideas in [31] and [32], it is intuitively appealing to combine the powers of compressive sensing and matrix completion and consider the following optimization problem

$$\text{minimize } \|\mathbf{Z}\|_* + \lambda \mathcal{S}(\mathbf{Z}) \qquad \text{subject to } \mathcal{A}(\mathbf{Z}) = y \qquad (5)$$

in order to recover $\mathbf{X}$. Here the functional $\mathcal{S}$ is chosen to exploit the sparsity inherent in $\mathbf{X}$. For instance we may choose

$$\mathcal{S}(\mathbf{X}) = \sum_k \|X_k\|_{TV},$$

where X_k is the k-th column of $\mathbf{X}$, see [31]. Or we could set

$$\mathcal{S}(\mathbf{X}) = \|\mathbf{U}\mathbf{X}\mathbf{V}^*\|_1,$$

where $\mathbf{U}$ and $\mathbf{V}$ are transforms designed such that $\mathbf{X}$ is sparse with respect to the tensor basis $\mathbf{U} \otimes \mathbf{V}$, see [32]. Clearly, many variations of the theme are possible, cf. [32], [33] for further discussion and examples. Optimization problems of this kind have significant potential in a wide range of applications, such as dynamic MRI, hyperspectral imaging, or target tracking.

All this leads us naturally to the quite ambitious task of constructing a unifying framework that allows to make statements about the recovery conditions of mathematical objects that obey some minimal complexity measure via methods from convex optimization. An interesting step along this lines is taken in the paper [34]. Such an undertaking must incorporate a further investigation of the connection between compressive sensing and information theory. A Shannon-information theoretic analog of compressive sensing was recently introduced by Wu and Verdú, see [35]. Further exciting results in this direction can be found in [36], [37].

E. Nonlinear compressive sensing

So far we have assumed that the observations we are collecting can be modeled as *linear* functionals of the form $\langle \mathbf{x}, \mathbf{a}_k \rangle, k = 1, \dots, m$, where $\mathbf{a}_k^*$ is a sensing vector representing a row of $\mathbf{A}$. However in many applications we can only take *nonlinear* measurements. An important example is the case where we observe signal intensities, i.e., the measurements are of the form $|\langle \mathbf{x}, \mathbf{a}_k \rangle|^2$, the phase information is missing. The problem is then to reconstruct $\mathbf{x}$ from intensity measurements only. A classical example is the problem of recovering a signal or image from the intensity measurements of its Fourier transform. Problems of this kind, known as *phase retrieval* arise in numerous applications, including X-ray crystallography, diffraction imaging, astronomy, and quantum tomography [38].

Concepts from compressive sensing and matrix completion have recently inspired a new approach to phase retrieval called *PhaseLift* [39]. It has been shown that if the vectors $\mathbf{a}_k$ are sampled independently and uniformly at random on the unit sphere, then the signal $\mathbf{x}$ can be recovered exactly (up to a global phase factor) from quadratic measurements by solving a trace-norm minimization problem provided that m is on the order of $n \log n$ measurements[4]. PhaseLift does not assume that the signal is sparse. It is natural to ask if we can extend the compressive sensing theory to the recovery of sparse signals from intensity measurements. Some initial results can be found in [41], [42], but it is clear that this development is still in its infancy and much more remains to be done. For instance, it would be very useful for a variety of applications to know how many measurements are required to recover an s-sparse signal $\mathbf{x} \in \mathbb{C}^n$ from Fourier-type intensity measurements.

Another type of nonlinear measurements is the case of quantized samples, and in the extreme case, 1-bit measurements [43], [44] (which is in a sense the opposite of intensity

[4] We know meanwhile that in the order of n measurements suffice, see [40].

measurements). But what about more general nonlinear measurements? For which types of nonlinear measurements can we build an interesting and relevant compressive sensing theory? I expect such a potential framework to have wide impact in disciplines like biology, where we often encounter all kinds of nonlinear processes driven by a few parameters.

F. Numerical algorithms

In recent years we have seen a large variety of numerical algorithms being developed to solve various versions of the compressive sensing problem. While the user of compressive sensing now has a plethora of algorithms to choose from, a comparison of the advantages and disadvantages of individual algorithms is difficult. Some algorithms provide guaranteed recovery of *all* sufficiently sparse signals, others succeed only for many or most signals. Some algorithms claim to be numerically efficient, yet are only so, when very specific sensing matrices are used or certain assumptions are fulfilled. Other algorithms are fast, but in order to succeed they require more measurements than competing methods. Fortunately the number of researchers who have made implementations of their algorithms available is large (much larger than in many other areas where numerical algorithms play a key role), making it fairly easy to test many of the published algorithms in a variety of scenarios.

Compressive sensing and matrix completion have stimulated the development of a variety of efficient algorithms for ℓ_1-minimization and semidefinite programming, see for instance [45]–[48]. Many of these algorithms come with rigorous theoretical guarantees. Based on heuristic considerations, some of these algorithms have been extended to solve non-convex problems, such as ℓ_p-minimization with $p < 1$. To what extent can we support these promising empirical results for non-convex optimization with theoretical convergence guarantees?

Iterative thresholding algorithms have been proposed as numerically efficient alternatives to convex programming for large-scale problems [49]–[51]. But until recently, known thresholding algorithms have offered substantially worse sparsity-undersampling tradeoffs than convex optimization. *Message passing algorithms* are a breakthrough in this regard [52]. Approximate Message Passing (AMP) algorithms proposed by Donoho, Montanari and their coworkers, are low-complexity iterative thresholding algorithms which can achieve optimal performance in terms of sparsity-undersampling tradeoff [36]. These AMP algorithms are also able to utilize block sparsity for instance. Interestingly, in the message passing framework of Donoho and Montanari we observe a shift from sparsity to the (Rényi) information dimension, which in turn leads us right to the discussion at the end of Subsection II-D. There are also intriguing connections to statistical physics.

However, it remains a major challenge to extend the theory underlying AMP from (Gaussian or band-diagonal) random matrices to those structured sensing matrices that are encountered in practice. Initial investigations have been carried out by Schniter and collaborators, see e.g. [53].

G. Hardware design

The concept of compressive sensing has inspired the development of new data acquisition hardware. By now we have seen compressive sensing "in action" in a variety of applications, such as MRI, astronomy, and analog-to-digital conversion, see Igor Carron's list of compressive sensing hardware [54]. Yet, the construction of compressive sensing-based hardware is still a great challenge.

But the process of developing compressive sensing hardware is not the job of the domain scientist alone. The knowledge gained during this process feeds back into the "production cycle" of compressive sensing, as theoreticians (have to) learn how to adapt their theory to more realistic scenarios, and in turn may then be able to provide the practitioner with better insight into performance bounds and improved design guidelines. Noise is a major limiting factor. Calibration remains a big problem. An efficient feedback loop between the different scientists working on theory, algorithms, and hardware design will be key to ensure further breakthroughs in this area.

III. The Future of Compressive Sensing

The surest way for a scientist to make a fool of himself/herself is by attempting to predict the future. But let me try anyway. Is compressive sensing here to stay? How important will it be in the future? And how will it evolve?

Where tremendous hope and a lot of enthusiasm meet, there is naturally the danger of a hype and thus the possibility of dramatic failure. Will the roadmap of compressive sensing be *from hope to hype to history*? It is clear that when we look back, say ten years from now, there will be areas where the concept of compressive sensing was not successful. One reason for such a failure may be that compressive sensing seemed a promising solution as long as we looked at an isolated subproblem. Yet, once we consider the subproblem in the context of the bigger problem from which it was extracted, the efficiencies gained via compressive sensing may have diminished.

However, I will not attempt to predict in which areas compressive sensing may not fulfill its promise. After all, if there will not be any crushed hopes, then we simply did not aim high enough. Or, in the words of Mario Andretti: *"If everything seems under control, you're just not going fast enough!"*. Instead let me sketch some areas, where I believe that compressive sensing will have (and in part already has had) a major impact.

There is a growing gap between the amount of data we generate and the amount of data we are able to store, communicate, and process. As Richard Baraniuk points out, in the year 2011 we produced already twice as many data as could be stored [55]. And the gap keeps widening. As long as this development continues there is an urgent need for novel data acquisition concepts like compressive sensing.

There is an obvious intellectual achievement, in which compressive sensing and sparse representations play a key role: Advanced probability theory and (in particular) random matrix theory, convex optimization, and applied harmonic analysis will become and already have become standard ingredients of the toolbox of many engineers. At the same time, mathematicians will have gained a much deeper understanding of how to confront real-world applications. Compressive sensing teaches us (or forces us?) to work across disciplines, but not in form of an alibi collaboration whose main purpose is to convince program directors and proposal reviewers to fund our next "interdisciplinary" project. No, it creates interdisciplinary collaborations for the only sensible reason: because some important problems simply cannot be solved otherwise! Furthermore, compressive sensing has advanced the development of ℓ_1-minimization algorithms, and more generally of nonsmooth optimization. These algorithms find wide-spread use in many disciplines, including physics, biology, and economics.

There will be "conceptual" achievements. For example, analogously to how wavelet theory has taught us how to think about multiscale and sparsity (despite the fact that wavelets could not live up to many of our expectations), compressive sensing will teach us how to think properly about minimal complexity and how to exploit it in a computationally efficient manner, and it may even be instrumental in developing a rigorous information theory framework for various areas such as molecular biology.

To revolutionize technology we will need to develop hardware and algorithms via an integrated, transdisciplinary approach. Hence, in the future when we design sensors, processors, and other devices, we may no longer speak only about *hardware* and *software*, where each of these two components is developed essentially separately. Instead, we may have to add a third category, which we could call *hybridware* or *mathematical sensing*[5], where the physical device and the mathematical algorithm are completely intertwined and co-designed right from the beginning.

Hence, looking further into the future, maybe the most important legacy of compressive sensing will be that it has forced us to think about information, complexity, hardware, and algorithms in a truly integrated manner.

Acknowledgement

The author would like to thank Emmanuel, Candès, Michael Elad, Holger Rauhut, and Anna Scaglione for their useful feedback on an earlier version of this paper.

References

[1] A. Kleinert, "Der messende Luchs. Zwei verbreitete Fehler in der Galilei-Literatur. (The Measuring Lynx. Two Widespread Mistakes in the Galileo Literature)," *N.T.M. Zeitschrift für Geschichte der Wissenschaften, Technik und Medizin*, vol. 17, pp. 199–206, 2009.

[2] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertaitny principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[3] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Trans. on Information Theory*, vol. 52, pp. 5406–5425, 2006.

[4] D. L. Donoho, "Compressed sensing," *IEEE Trans. on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[5] M. Elad, "Sparse and Redundation Representation Modeling – What Next?," *IEEE Signal Processing Letters*, 2012.

[5] *"Mathematical Sensing , Dennis Healy's Vision"*, Presentation by Ronald Coifman given at the Minisymposium "In Memory of Dennis Healy and His Scientific Vision", 2010 SIAM Annual Meeting.

[6] R. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, 2007.

[7] E. Candès and M. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[8] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. Springer, to appear.

[9] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Springer, 2010.

[10] M. Lustig, D. Donoho, J. Santos, and J. Pauly, "Compressed sensing MRI," *IEEE Signal Processing Magazine*, vol. 27, pp. 72–82, 2008.

[11] M. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Trans. on Signal Processing*, vol. 57, no. 6, pp. 2275–2284, 2009.

[12] J. Haupt, W. Bajwa, G. Raz, and R. Nowak, "Toeplitz compressed sensing matrices with applications to sparse channel estimation," *IEEE Trans. Inform. Theory*, vol. 56, no. 11, pp. 5862–5875, 2010.

[13] G. Pfander, H. Rauhut, and J. Tropp, "The restricted isometry property for time-frequency structured random matrices," *Preprint*, 2011.

[14] M. Duarte and Y. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Trans. Sig. Proc.*, to appear.

[15] H. Rauhut, "Compressive sensing and structured random matrices," in *Theoretical Foundations and Numerical Methods for Sparse Recovery* (M. Fornasier, ed.), vol. 9 of *Radon Series Comp. Appl. Math.*, pp. 1–92, deGruyter, 2010.

[16] J. A. Tropp, "On the conditioning of random subdictionaries," *Appl. Comput. Harmon. Anal.*, vol. 25, pp. 1–24, 2008.

[17] E. Candès and Y. Plan, "Near-ideal model selection by ℓ_1 minimization," *Annals of Statistics*, vol. 37, no. 5A, pp. 2145–2177, 2009.

[18] J. Bourgain, S. Dilworth, K. Ford, S. Konyagin, and D. Kutzarova, "Explicit constructions of RIP matrices and related problems," *Duke Math. J.*, vol. 159, no. 1, pp. 145–185, 2011.

[19] Y. Chi, A. Pezeshki, L. Scharf, and R. Calderbank, "Sensitivity to basis mismatch in compressed sensing," in *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, (Dallas, Texas), 2010.

[20] P. Dragotti, M. Vetterli, and T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 1741–1757, 2007.

[21] A. Fannjiang and W. Liao, "Coherence pattern-guided compressive sensing with unresolved grids," *SIAM J. Imaging Sci.*, vol. 5, pp. 179–202, 2012.

[22] G. Tang, B. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *Preprint, [arvix:1207.6053]*, 2012.

[23] E. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Comm. Pure Appl. Math.*, to appear.

[24] B. Adcock and A. Hansen, "Generalized sampling and infinite-dimensional compressed sensing," *DAMTP Tech. Rep. 2011/NA12*, 2011.

[25] R. Baraniuk, V. Cevher, M. Duarte, and C. Hedge, "Model-based compresssive sensing," *IEEE Trans. Information Theory*, vol. 56, no. 4, pp. 1982–2001, 2010.

[26] Y. Eldar, P. Kuppinger, and H. Bölcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Trans. on Signal Processing*, vol. 58, no. 6, pp. 3042–3054, 2010.

[27] G. Yu, G. Sapiro, and S. Mallat, "Solving inverse problems with piecewise linear estimators: From Gaussian mixture models to structured sparsity," *IEEE Trans. on Image Processing*, vol. 21, no. 5, pp. 2481–2499, 2012.

[28] B. Recht, M. Fazel, and P. Parrilo, "Guaranteed minimum rank solutions of matrix equations via nuclear norm minimization," *SIAM Review*, to appear.

[29] E. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717–772, 2009.

[30] S. Oymak, K. Mohan, M. Fazel, and B. Hassibi, "A simplified approach to recovery conditions for low rank matrices," *Preprint, [arXiv:1103.1178]*, 2011.

[31] M. Golbabaee and P. Vandergheynst, "Joint trace/TV norm minimization: a new efficient approach for spectral compressive imaging," in *IEEE International Conference on Image Processing (ICIP)*, 2012.

[32] E. Candès and M. Soltanolkotabi, "Discussion of "Latent Variable Graphical Model Selection via Convex Optimization"," *Annals of Statistics*, to appear.

[33] A. Waters, A. Sankaranarayanan, and R. Baraniuk, "SpaRCS: Recovering low-rank and sparse matrices from compressive measurements," in *Neural Information Processing Systems (NIPS)*, (Granada, Spain), 2011.

[34] V. Chandrasekaran, B. Recht, P. Parrilo, and A. Willsky, "The convex geometry of linear inverse problems," *Foundations of Computational Mathematics*, to appear.

[35] Y. Wu and S. Verdú, "Rényi information dimension: Fundamental limits of almost lossless analog compression," *IEEE Trans. on Inform. Theory*, vol. 56, pp. 3721–3748, 2010.

[36] D. Donoho, A. Javanmard, and A. Montanari, "Information-theoretically optimal compressed sensing via spatial coupling and approximate message passing," *Preprint, [arXiv:1112.0708]*, 2011.

[37] Y. Wu and S. Verdú, "Optimal phase transitions in compressed sensing," *Preprint, [arXiv:1111.6822]*, 2012.

[38] M. Klibanov, P. Sacks, and A. Tikhonravov, "The phase retrieval problem," *Inverse problems*, vol. 11, p. 1995, 1–28.

[39] E. Candès, T. Strohmer, and V. Voroninski, "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming," *Comm. Pure Appl. Math.*, to appear.

[40] E. Candès and X. Li, "Solving quadratic equations via PhaseLift when there are about as many equations as unknowns," *Preprint, [arXiv:1208.6247]*, 2012.

[41] W. Chan, M. Moravec, R. Baraniuk, and D. Mittleman, "Terahertz imaging with compressed sensing and phase retrieval," *Optics Letters*, vol. 33, no. 9, pp. 974–976, 2008.

[42] H. Ohlsson, A. Yang, R. Dong, and S. Sastry, "Compressive phase retrieval from squared output measurements via semidefinite programming," *Preprint, [arxiv:1111.6323v3]*, 2012.

[43] L. Jacques, J. Laska, P. Boufounos, and R. Baraniuk, "Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors," *Preprint*, 2011.

[44] Y. Plan and R. Vershynin, "One-bit compressed sensing by linear programming," *Preprint*, 2011.

[45] M. A. T. Figueiredo, R. D. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Sel. Top. Signal Proces.*, vol. 1, no. 4, pp. 586–598, 2007.

[46] E. Hale, W. Yin, and Y. Zhang, "Fixed-point continuation for ℓ_1-minimization: methodology and convergence," *SIAM J. Optim.*, vol. 19, no. 3, pp. 1107–1130, 2008.

[47] M. Burger, M. Moeller, M. Benning, and S. Osher, "An adaptive inverse scale space method for compressed sensing," *Technical Report 11-08, UCLA*, 2011.

[48] S. Ma, D. Goldfarb, and L. Chen, "Fixed point and bregman iterative methods for matrix rank minimization," *Math. Program.*, vol. 128, no. 1-2, pp. 321–353, 2011.

[49] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Comm. Pure Appl. Math.*, vol. 57, no. 11, pp. 1413–1457, 2004.

[50] M. Fornasier and H. Rauhut, "Iterative thresholding algorithms," *Appl. Comput. Harmon. Anal.*, vol. 25, no. 2, pp. 187–208, 2008.

[51] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sci.*, vol. 2, no. 1, pp. 183–202, 2009.

[52] D. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. National Academy of Sciences (PNAS)*, vol. 106, no. 45, pp. 18914–18919, 2009.

[53] S. Som and P. Schniter, "Compressive imaging using approximate message passing and a Markov-tree prior," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3439–3448, 2012.

[54] https://sites.google.com/site/igorcarron2/compressedsensinghardware.

[55] R. Baraniuk, "More is Less. Signal Processing and the Data Deluge," *Science*, vol. 331, p. 717, 2011.