Sparse and Low Rank Approximation (11w5036)

Holger Rauhut (University of Bonn),
Gitta Kutyniok (University of Osnabrück),
Joel Tropp (California Institute of Technology),
Özgür Yılmaz (University of British Columbia)

March 6 – 11, 2011

1 Overview of the Field

Digital computers and their efficiency at processing data play a central role in our modern technology. Today, we use digital hardware in every aspect of our daily lives. Cell phones, digital cameras, MP3 players and DVD players are only a few examples where signals of interest, which are inherently analog, are acquired, converted to digital bit streams, stored in compressed form, and transmitted over noisy channels. In all these applications, the current demand is towards "higher-resolution" and "faster". In turn, the amounts of data we need to deal with is getting exceedingly large.

Recently, a new signal processing paradigm based on "sparsity" has emerged. This new paradigm exploits an empirical observation: many types of signals, e.g., audio, natural images, and video, can be well represented by "very few" elements of a suitable representation system. This fact has revolutionized data processing, and led to a rethinking about how to acquire various types of signals using a very limited amount of linear measurements. Consequently, this new paradigm has the potential to change the way we collect information about many types of signals.

One exciting development is the introduction of the method of Compressed Sensing [16, 7, 22, 36] which, for instance, led to the design of a single-pixel digital camera built at the Rice University, see `http://dsp.rice.edu/cscamera`. This research area is very new and hence various key research questions are still under investigation, and also the application of these techniques to, for instance, astronomical image and signal processing, radar imaging, wireless communication, and seismology is just in its beginning stage. Advancing research on these questions as well as studying the potential for diverse applications was a primary motivation for this workshop.

2 Mathematical Foundations

The field of sparse and low-rank approximation is based on several mathematical ideas, and it has a variety of striking applications.

2.1 Sparsity

A function is *sparse* or *compressible* if it can be approximated by a short linear combination of basis functions drawn from a fixed collection. "Short" means that the number of basis functions needed is somewhat smaller than the ambient dimension of the function space. Compressible functions have less information than their dimension suggests. In many problems, the family of basis functions should describe the type of structures one hopes to find in the class of functions of interest; it may or may not form an orthogonal system.

Classical example A smooth function is well approximated by its lowest-order frequency components.

Less classical example A piecewise smooth function is well approximated by the most significant terms in its wavelet expansion.

Modern example Low-rank matrices can be expressed using a sum of a few rank-one matrices.

Sparsity is the main idea behind transform coding, the concept at the bottom of the JPEG scheme for image compression. Roughly, we transform an image into the discrete cosine domain; we retain only the terms in the expansion that have large coefficients; we store the few remaining coefficients. This operation achieves compression. To decompress, we apply the inverse discrete cosine transform to the compressed representation.

Another important perspective on sparsity is framed in terms of a linear inverse problem. Consider the system of linear equations

$$f = Ax$$

where f is the function of interest and A is a fixed matrix. The columns of A are the basis functions. We say that f has a sparse representation when the variable x is a sparse vector. In practice, of course, the equality is only approximate! But this situation arises often enough to base a field of research on it.

2.2 Algorithms that enforce sparsity

Given an observation of a function f, we would like to find a sparse (approximate) solution x to the linear inverse problem $f = Ax$. In general, this problem is computationally hard because it might require a combinatorial search. Remarkably, in certain situations, we can still accomplish this goal.

One approach is to solve a regularized inverse problem. Given an observed function f and a matrix A, solve [11, 16]

$$\min_x ||x||_1 \quad \text{subject to} \quad f = Ax.$$

In other words, instead of the classical ℓ_2 regularization, we use an ℓ_1 regularization instead, which has the effect of promoting sparsity in the solution.

The ℓ_1-regularized optimization problem has a (trivial) closed-form solution when A is invertible. More generally, it can be solved using special-purpose convex optimization algorithms.

When the columns of A are roughly orthogonal, the solution to the ℓ_1-minimization problem is close to the ”true” sparse x^* for which $f = Ax^*$. We can quantify “roughly orthogonal” in several different ways, but the technical details are not that important here.

2.3 Compressed sensing

Consider a binary 0–1 vector x^* of length d with at most s entries equal to one. It is very inefficient to express this vector in the standard coordinate basis because it contains so many zeros. In fact, it only takes about $s \log(d/s)$ bits to encode this vector (using run-length coding!).

The vector x^* is sparse in the sense that I described before because it can be represented efficiently in the standard basis. Is there some type of measurement that automatically sieves out the locations of the unit entries? In fact, the inner product between x^* and a random (say, Gaussian) vectors provides a lot of information about x^*.

Suppose that we could take measurements

$$f_i = \langle x^*, g_i \rangle, \quad i = 1, 2, ..., m$$

where g_i are independent, standard Gaussian random vectors. It turns out that, provided $m > 2s \log(d/s)$, the real-valued measurements $f_1, ..., f_m$ contain all the information we need to (stably) reconstruct x^* [8, 17, 10].

How? Use tractable algorithms for sparse solution of linear inverse problems. We just need to find a sparse solution to $f = Ax$ where the rows of A are the vectors $g_1, ..., g_m$. We can do this using ℓ_1 minimization. It turns out that the Gaussian matrix A has the property that the columns are "sufficiently orthogonal" to ensure that the true solution x^* is the unique min-ℓ_1 solution of $f = Ax$.

3 Recent Developments and Open Problems

While Gaussian random measurements are known to be optimal for compressive sensing they have the drawback of lacking a fast matrix vector multiply. Structured random measurement maps may overcome this computational bottleneck. Partial random Fourier matrices are known since the advent of compressive sensing to provide near optimal recovery guarantees. Recently, further types of structured random matrices were investigated. These include partial random circulant matrices [36, 37] (mentioned in the talk by Romberg), and time-frequency structured random matrices [34] (talk by Pfander). For both matrices the optimal expected recovery guaranties are so far open.

The extension of compressive sensing techniques to the recovery of low rank matrices is another recent development. A number of efficient recovery algorithms for this task have been developed over the last two years, including nuclear norm minimization [20, 39] (reported in the talk by Maryam Fazel), singular value thresholding [6], iteratively reweighted least squares [23], max norm minimization [30] (as presented by Nathan Srebro), and more. Also a number of theoretical results for low rank matrix recovery is available by now [39, 9, 38, 32] (talks by Maryam Fazel, Ben Recht), but many more developments can be expected, such as extensions to structured random measurement maps (initial insights were reported by Justin Romberg).

Ideas and techniques from compressive sensing were recently applied in new areas, which have not been the focus of research in compressive sensing in the beginning. The use of compressive sensing for quantization is investigated in the initial contributions [4, 27], and was reported at the workshop. Many open questions in this context remain, e.g., whether the use of efficient structured measurement maps is possible.

The Johnson-Lindenstrauss embedding represents a fundamental tool for dimensionality reduction, and is usually realized via random matrices. It was an early insight [3] that the fundamental concentration inequality leading to the Johnson-Lindenstrauss map, can also be used for proving the restricted isometry property, a fundamental concept in compressed sensing. It came as a recent surprise that also the converse is true: A matrix satisfying the restricted isometry property provides a Johnson-Lindenstrauss embedding after randomizing the column signs [29]. Eldar and Needell developed a variant of the Kaczmarz method for solving large linear systems, which exploits Johnson-Lindenstrauss maps [18]. In [21] a method for numerically solving dynamical systems in high dimensions using the Johnson-Lindenstrauss embedding was developed.

Manifold learning and processing is an increasingly important task with applications to machine learning, signal processing and more. In some cases, these manifolds are embedded into a very high dimensional space. Recently, investigations on the use of compressive measurement maps (such as Johnson-Lindenstrauss maps) in this context have been initiated (talks by Herman and Iwen). First results are promising, and many more development in this direction can be expected.

Recently, sparse approximation techniques were also applied for the numerical analysis of partial differential equations (PDEs) with random coefficients, as well as in parametric PDEs [13]. Here it can be shown that under certain conditions the solution to the parametric PDE as a function of the physical variable and parameter can be well-approximated by a sparse expansion in tensor product Legendre polynomials in the parameter, and with coefficient functions depending on the physical variable. The application of sparse approximation schemes seems promising, and moreover, this result leads to decay rates for greedy approaches in reduced basis methods [5]. Many more contributions in this direction are expected in the near future.

4 Presentation Highlights

Reflecting the high-level of activity in the area, the workshop brought together world-experts from various areas of mathematics and engineering along with young researchers. Accordingly, nine talks were given by junior researchers (one of them was a graduate student, the remaining eight were postdocs). Below, we categorize the presentations according to their specific focus.

4.1 Compressed Sensing

There were a large number of presentations that were on compressed sensing and its variations as well as implications of certain techniques developed in CS to other areas of mathematics.

4.1.1 Theory and algorithms for sparse recovery

P. Boufounos presented recent work in sparsity constrained optimization of arbitrary cost functions—as opposed to the quadratic cost functions of the usual sparse recovery algorithms— where he generalizes the Restricted Isometry Property (RIP) to a stability condition on the Hessian of the cost function. Furth ermore, Boufounos discussed the Gradient Support Pursuit (GraSP), an algorithm that generalizes known CS algorithms to perform sparsity constrained minimization of arbitrary cost functions.

V. Temlyakov considered greedy algorithms for sparse recovery, which constitute a viable alternative to ℓ_1 minimization in this context. He discussed so-called Lebesgue-type inequalities for greedy algorithms in Hilbert and Banach spaces.

H. Mansour studied weighted ℓ_1 minimization for signal construction from CS measurements when partial support information is available [26]. He proved that if the support estimate is more than % 50 accurate, then this alternative algorithm recovers the underlying sparse or compressible signal under weaker sufficient conditions than those for ℓ_1 minimization.

V. Goyal emphasized that most theoretical guarantees in sparse approximation are uniform over deterministic signal classes and consequently pessimistic. He proposed a Bayesian formulation based on the "replica method" from statistical physics [35]. Using this formulation, he presented an exact asymptotic analysis of many commonly-used decoders such as basis pursuit, lasso, linear estimation with thresholding, and zero-norm regularized estimation, which gives results that are more encouraging for compressed sensing than the usual uniform guarantees.

4.1.2 Mathematical and practical challenges in implementing compressed sensing

R. Saab showed that one can successfully employ "noise-shaping" $\Sigma\Delta$ quantizers for compressed sensing. He proved that, by using appropriate Sobolev dual frames in the reconstruction, $\Sigma\Delta$ quantizers utilize the inherent redundancy more efficiently than "any" round-off type quantization algorithm, at least in the case of Gaussian measurement matrices. Consequently, $\Sigma\Delta$ quantizers outperform the round-off quantization schemes (generally referred to as memoryless scalar quantizers) in the compressed sensing setting [27].

R. Baraniuk considered 1-bit quantization for compressed sensing. He proposed the notion of a "Binary Stable Embedding", a property that ensures stable reconstruction from 1-bit measurements. Using this property, he presented some recovery guarantees for 1-bit compressed sensing [4]. In addition, he discussed algorithmic issues regarding 1-bit compressed sensing.

Y. Eldar reviewed her earlier work on "analog compressed sensing" which she calls "xampling". The goal in xampling is to build sub-Nyquist analog-to-digital converters. She then discussed how sub-Nyquist sampling of pulse streams can be used to identify linear time-varying systems and showed that sufficiently-underspread parametric linear systems, described by a finite set of delays and Doppler-shifts, are identifiable from a single observation as long as the time-bandwidth product of the input signal is proportional to the square of the total number of delay-Doppler pairs in the system [2].

R. Calderbank addressed two practical challenges: The first is a wireless uplink where mobile users in some geographic area need to register with a base station. The second is the distribution of information in a wireless sensor network. Inspired from compressed sensing and Reed-Muller codes, he presented a new deterministic framework for communication in wireless networks which can successfully address the above-mentioned challenges.

T. Strohmer used sparse MIMO Radar as a case study to analyze the potential benefits and pitfalls of compressive sensing, He noted that compressed sensing "finds itself caught between two extremes - the parametric and the non-parametric world". He noted that even though at first glance the benefits of Sparse MIMO Radar seem obvious (robust with respect to noise, can detect weak targets), a careful analysis raises some challenging questions. He concluded that compressed sensing combines some of the strengths as well as some of the weaknesses of both the parametric and non-parametric methods.

A. Pezeshki discussed the sensitivity of compressed sensing to "basis mismatch", i.e., mismatch between the assumed and the actual models for sparsity. He presented some numerical examples, based on Fourier imaging, that demonstrate a significant dependence on the correct choice of the associated sparsity basis. He also presented theoretical results on how basis mismatch effects the "recovery error" that apply when a wide class of sparse recovery algorithms, including basis pursuit, are used [12].

4.1.3 Generalizations of compressed sensing

M. Iwen considered the problem of approximating low-dimensional manifolds in high-dimension using a small number of linear measurements, i.e., compressive measurements. He provided a simple reconstruction algorithm. He also obtained theoretical recovery guarantees.

B. Recht discussed a framework for "extending the catalog of objects and structures that can be recovered from partial information" [10]. He presented a family of algorithms—that are obtained in a convex optimization framework—for obtaining "sparse" approximations of these various objects and discussed general recovery guarantees and implementation schemes for these algorithms. While Recht's approach encompasses ℓ_1 minimization (compressed sensing) and nuclear-norm minimization (low rank matrix completion), it also shows that in various different problems an analogous approach can be adopted.

4.1.4 Mathematical results inspired by compressed sensing

F. Krahmer established a new connection between the Johnson-Lindenstrauss lemma (which is a central dimension-reduction result used in various computer science applications) and the restricted isometry property (RIP), a central concept in the theory of sparse recovery. Specifically, he proved that $m \times N$ matrices satisfying the RIP of optimal order provide (via randomized column signs) optimal Johnson-Lindenstrauss embeddings up to a logarithmic factor in N. Krahmer's results yield the best known bounds on the necessary embedding dimension m for a wide class of structured random matrices, including partial Fourier and partial Hadamard matrices [29].

D. Needell considered the randomized Kaczmarz method that was recently proposed by Strohmer and Vershynin [40], who also shoved that this methods converges exponentially (in expectation). In her talk, Needell presented a modified version of the randomized Kaczmarz method (where at each iteration the optimal projection from a randomly chosen set is selected). She showed that this modified rate generally improves the convergence rate significantly. Furthermore, using dimension reduction methods based on Johnson-Lindenstrauss lemma, the run-time of the modified algorithm is on the same order as the run-time of the original version [18].

S. Foucart focused on two important problems on the geometry of finite-dimensional ℓ_1 spaces: estimation of Gelfand widths of their unit balls, and Kashin's orthogonal decomposition theorem. Using the theory of compressive sensing, he extended the known results to finite-dimensional ℓ_p spaces ($0 < p \leq 1$) [25]. Furthermore, using compressed sensing methods, he obtained the lower estimate for the Gelfand width of ℓ_1^N in ℓ_2^N.

4.2 Sparse approximation

I. Daubechies focused on a new mathematical theory for the empirical mode decomposition (EMD) algorithm. EMD is a heuristic technique that aims to decompose functions that have sparse expansions in terms of building blocks that can be viewed as locally harmonic functions with slowly varying amplitudes and phases that are well-separated in the time-frequency plane. Daubechies introduced a decomposition that "captures the flavor and philosophy of the EMD approach", however her method of constructing the components is different and uses the so-called "synchrosqueezed wavelet transform". She presented both theoretical results and applications of the algorithm [14].

J. Vybìral considered the problem of approximating a function $f : \mathbb{R}^d \mapsto \mathbb{R}$ that is of the form $f(x) = g(Ax)$ where A is a $k \times d$ matrix with $k \ll d$ from only a few values of f. Vybìral presented a randomized algorithm and proved that (under some conditions on g and A) the algorithm produces a uniform approximation to f with high probability [24].

A. Cohen also discussed mathematical problems that involve functions of a very large number of variables. He focused on such problems in partial differential equations depending on parametric or stochastic variables where numerical difficulties arise due to the so-called ”curse of dimensionality”. Cohen explained how these difficulties may be handled in various contexts using the concepts of variable reduction and sparse approximation.

S. Kunis emphasized that sparse approximation is a powerful approach for handling high-dimensional problems, however he noted that it is critical that algorithms such as FFT need also be customized as to allow the use of sparsity to reduce the computational cost. To that end, he considered two generalizations of the FFT: a hyperbolic-cross FFT [15] and the so-called butterfly sparse FFT.

4.3 Low rank approximation

M. Fazel focused on theoretical recovery conditions for low rank matrices from (possibly noisy) linear measurements. She presented a framework that can be used to extend robust recovery results from vectors (as in the case of sparse recovery) to low-rank matrix recovery. She noted that this methodology is simple and leads to the tightest known sufficient conditions (e.g., in terms of the RIP of the measurement map) for low rank matrix recovery [32].

A. Agarwal discussed matrix decomposition problems where the goal is to recover from noisy observations matrices that can be decomposed as a sum of an (approximately) low rank matrix and a second matrix with some low dimensional structure. Agarwal proved a general theorem bounding the approximation error addressing the approximation obtained via a convex optimization problem that included nuclear norm minimization and an appropriate regularizer. Agarwal also used his results to study some special cases in the context of robust PCA [1].

N. Srebro's talk was also on matrix recovery. Srebro focused on two forms of matrix regularization which constrain the norm of the factorization: the trace-norm (i.e., nuclear-norm) and the so-called max-norm (i.e., $\gamma_2 : \ell_1 \to \ell_\infty$ norm). Srebro discussed how these two norms relate to the rank and showed that simple low rank matrix completion guarantees can be obtained using these norms. Srebro also argued that the max-norm may be a better surrogate for the rank than the nuclear norm [30].

J. Romberg presented several architectures that use simple analog building blocks (vector-matrix multiply, modulators, filters, and ADCs) to implement different types of measurement schemes with “structured randomness”. These sampling schemes allow us to take advantage of the (a priori unknown) correlation structure of the ensemble, which leads to a low rank matrix to be determined. Exploiting this fact reduces the total number of observations required to reconstruct the collection of signals.

4.4 Applications in various disciplines

M. Fornasier talked about a recent project where he combines various results on compressed sensing theory and algorithms, and learning functions in high-dimension to address problems related to learning, simulation, and control of particle systems, kinetic equations, and fluid dynamics models of interacting agents in high-dimension. He proposed an approach for the simulation of dynamical systems governed by functions of adjacency matrices in high-dimension by random projections via Johnson-Lindenstrauss embeddings, and recovery by compressed sensing techniques [21].

M. Herman presented results on denoising point cloud data using non-local techniques. His approach is grid-free, and thus preserves subtle geometric information in the data and thereby permits the identification of 3-dimensional structures in the point cloud. Herman successfully applied the techniques to LIDAR data to denoise objects of codimension 1 and 2 simultaneously.

A. Maleki's talk focused on a theoretical analysis of the nonlocal means algorithm, an image denoising method that has been successful in applications. Maleki presented results on the asymptotic risk analysis of this algorithm for images that are piecewise constant with a sharp edge discontinuity. He proved that the associated mean-square risk is suboptimal (and close to that of wavelet thresholding).

F. Herrmann addressed applications of compressed sensing to the full-waveform inversion problem in exploration seismology. He explained that the demand for higher quality imaging results in very large problem sizes which is a fundamental obstacle. To address this problem, he proposed a dimension-reduction strategy based on compressed sensing and stochastic optimization [31].

M. Sacchi's presentation was also on applications in exploration seismology focusing on industrial applications where the goal is to “reconstruct” 5-dimensional seismic data cubes from the observations (that are "incomplete" and noisy). He discussed reconstruction of large seismic volumes via rank reduction methods by posing the problem as a tensor completion problem.

J.L. Starck noted that the European Space Agency's PLANCK mission is designed to deliver full-sky coverage, low-noise level, high resolution temperature and polarisations maps. He reviewed some of the key problems of the PLANCK data analysis and presented how sparsity can be used to analyze such data set [33].

5 Scientific Progress Made and Outcome of the Meeting

The meeting was very successful. Many participants mentioned to us that they greatly enjoyed the excellent talks, the fruitful discussions, the inspiring atmosphere at BIRS, and the neverending support of the BIRS staff.

Let us now briefly summarize the impact of our meeting.

- *Initiation of communications between practitioners and mathematicians.*

  One main goal of our workshop was to invite people from very different areas, in particular, including practitioners who utilize methodologies from sparse and low rank approximation. This included researchers in the ares of astronomical image and signal processing, radar imaging, wireless communication, and seismology. This BIRS workshop was a unique opportunity to initiate a fertile discussion between researchers from these areas and mathematicians. The talks by the practitioners indeed led to vivid debates; here we mention, in particular, the talk by Jean-Luc Starck on "Sparse analysis of the PLANCK cosmic microwave background data" and the talk by Richard Baraniuk on "1-bit Compressive Sensing and Binary Stable Embeddings"; as references see [33] and [4], respectively. We can report that these discussions led to several new collaborations between these two groups.

- *Intensification of new directions in the field.*

  Various new directions both theoretical as well as applied were presented during talks, and vividly discussed afterwards. One main new direction in the field is matrix identification, which was the main focus of three talks. The general problem is to minimize the rank of a matrix subject to some general linear constraint, which has applications in various areas of high dimensional data processing. This direction is currently in its beginning stage and far more developments can be expected in the near future. As further directions, which are even more at the beginning stage, we exemplarily mention the high level analysis of linear inverse problems via convex geometry (Ben Recht) [10]. This workshop was hence a unique opportunity to discuss the most recent results and stimulate these directions.

- *Discussion of methodologies.*

  Several talks made the audience sensitive to careful utilization of theoretically derived results in applications. One example was the discussion of the problem of basis mismatch in the talk by Ali Pezeshki (cf. [12]), in which he described the effect when the wrong sparsifying basis is chosen. Another talk which should be mentioned in this regard was given by Rayan Saab, see [27]. He discussed the sensitivity of Compressed Sensing to quantization, and provided one solution for this problem. Finally, the talk by Thomas Strohmer discussed various obstacles which have to be overcome to use sparse recovery methods for radar analysis; we refer to [28, 19]. These talks were very beneficial for the audience, in particular, for those not familiar with applications, making them sensitive to the needs of the practitioners.

- *Introduction of young scientists.*

  Several of our participants were young, very promising scientists such as Felix Krahmer, Arian Maleki, and Rayan Saab. This workshop gave them an exceptional chance to present themselves and get in contact with the leading researchers in this area, and also to broaden their horizon. After the workshop, we received excited feedbacks from this group, voicing the general opinion that this was a unique opportunity. Testimonial opinions are put on the BIRS webpage at `http://www.birs.ca/events/2011/5-day-workshops/11w5036/testimonials`.

- *Manifestation of the future direction of the field.*

  Since this workshop brought together the main leaders in the broad area of sparse and low rank approximations, it presented the possibility to debate and manifest the future directions of this field. Many intense discussions already took place right after most talks; in addition, we scheduled a general discussion session on the last day of the conference. Interesting open problems and possible future research directions came up. We therefore expect this workshop to have a signal effect which will significantly influence the research in this field in the future. As particular topics that came up in the discussion session we mention the following. In some practical recovery problems noise is added on the vector to be recovered. This poses severe problems in the application of compressed sensing, where in previous analysis the noise is assumed to be added to the measurement vector. It seems that methods have to be developed that are less sensitive to such type of noise (although it is presently not clear, whether this is possible at all). Furthermore, in the theoretical analysis of matrix completion more realistic probability models of selecting entries should be investigated. For instance, in applications to global positioning, usually only distance information on small distances is available. The near-optimal construction of deterministic compressed sensing matrices was discussed as well. While this remains a fundamental and important problem, no route to its solution could be identified in the discussion, which is certainly due to the hardness of the problem.

References

[1] A. Agarwal, S. Negahban, and M. Wainwright, *Noisy matrix decomposition via convex relaxation: Optimal rates in high dimensions*, *Preprint*, 2010.

[2] W. U. Bajwa, K. Gedalyahu, and Y. C. Eldar, *Identification of Parametric Underspread Linear Systems and Super-Resolution Radar*, IEEE Transactions on Signal Processing, 59(6):2548–2561, 2011.

[3] R. G. Baraniuk, M. Davenport, R.A. DeVore, and M. Wakin. A simple proof of the restricted isometry property for random matrices. *Constr. Approx.*, 28(3):253–263, 2008.

[4] L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, *Robust 1-Bit Compressive Sensing via Binary Stable Embeddings of Sparse Vectors*, preprint (2011).

[5] P. Binev, A. Cohen, Wolfgang Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk. Convergence rates for greedy algorithms in reduced basis methods. *Preprint*, 2010.

[6] J.F. Cai, E.J. Candès, and Z. Shen. A singular value thresholding algorithm for matrix completion. *SIAM J. Optim.*, 20(4):1956–1982, 2010.

[7] E.J. Candès, J., Terence Tao, and J. Romberg. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory*, 52(2):489–509, 2006.

[8] E.J. Candès and T. Tao. Near optimal signal recovery from random projections: universal encoding strategies? *IEEE Trans. Inform. Theory*, 52(12):5406–5425, 2006.

[9] E.J. Candès and B. Recht. Exact matrix completion via convex optimization. *Found. Comput. Math.,*, 9:717–772, 2009.

[10] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. Willsky, *The Convex Geometry of Linear Inverse Problems.* preprint (2010).

[11] S.S. Chen, D.L. Donoho, and M.A. Saunders. Atomic decomposition by Basis Pursuit. *SIAM J. Sci. Comput.*, 20(1):33–61, 1999.

[12] Y. Chi, A. Pezeshki, L. L. Scharf, and R. Calderbank, *Sensitivity to basis mismatch of compressed sensing*, IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), Dallas, TX, Mar. 14-19, 2010.

[13] A. Cohen, R. DeVore, and Christoph Schwab. Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs. *Anal. Appl.*, 2011.

[14] I. Daubechies, J. Lu, and H.T. Wu, *Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool*, Applied and Computational Harmonic Analysis, 30(2):243–261, 2011.

[15] M. Döhler, S. Kunis, and D. Potts, *Nonequispaced hyperbolic cross fast Fourier transform*, SIAM J. Numer. Anal. 47:4415-4428, 2010.

[16] D.L. Donoho. Compressed sensing. *IEEE Trans. Inform. Theory*, 52(4):1289–1306, 2006.

[17] D.L. Donoho and J.Tanner. Counting faces of randomly-projected polytopes when the projection radically lowers dimension. *J. Amer. Math. Soc.*, 22(1):1–53, 2009.

[18] Y. Eldar and D. Needell. Acceleration of randomized Kaczmarz method via the Johnson-Lindenstrauss lemma. *Numer. Algorithms*, to appear.

[19] A. Fannjiang, P. Yan, and T. Strohmer, *Compressed Remote Sensing of Sparse Objects*, SIAM J. Imag. Sci., to appear.

[20] M. Fazel. *Matrix rank minimization with applications*. PhD thesis, 2002.

[21] M. Fornasier, J. Haskovec, and Jan Vybiral. Particle systems and kinetic equations modeling interacting agents in high dimension. *preprint*, 2011.

[22] M. Fornasier and H. Rauhut. Compressive sensing. In O. Scherzer, editor, *Handbook of Mathematical Methods in Imaging*, pages 187–228. Springer, 2011.

[23] M. Fornasier, H. Rauhut, and R. Ward. Low-rank matrix recovery via iteratively reweighted least squares minimization. *Preprint*, 2010.

[24] M. Fornasier, K. Schnass, and J. Vybìral, *Learning Functions of Few Arbitrary Linear Parameters in High Dimensions*, *Preprint*, 2010.

[25] S. Foucart, A. Pajor, H. Rauhut, and T. Ullrich, *The Gelfand widths of ℓ_p-balls for* $0 < p \leq 1$. Journal of Complexity, 26(6):629–640, 2010.

[26] M. Friedlander, H. Mansour, R. Saab, and Ö. Yılmaz, *Recovering Compressively Sampled Signals Using Partial Support Information*. *Preprint*, 2010.

[27] S. Güntürk, M. Lammers, A. Powell, R. Saab, and Ö. Yılmaz, *Sobolev duals for random frames and Sigma-Delta quantization of compressed sensing measurements*, *Preprint*, 2010.

[28] M. Herman and T. Strohmer, *High Resolution Radar via Compressed Sensing*, IEEE Trans. Signal Proc. **57** (2009), 2275–2284.

[29] F. Krahmer and R. Ward. New and improved Johnson-Lindenstrauss embeddings via the Restricted Isometry Property. *SIAM J. Math. Anal.*, 43(3):1269–1281, 2011.

[30] J. Lee, Benjamin Recht, R. Salakhutdinov, N. Srebro, and J.A. Tropp. Practical large-scale optimization for max-norm regularization. In *Advances in Neural Information Processing Systems 23 (NIPS)*, pages 1297–1305, December 2010.

[31] X. Li, A. Aravkin, T. van Leeuwen, F. Herrmann, *Full-waveform inversion with randomized L1 recovery for the model updates*, EAGE 2011, Vienna.

[32] S. Oymak, K. Mohan, M. Fazel, and B. Hassibi. A simplified approach to recovery conditions for low-rank matrices. *preprint*, 2011.

[33] P. Paykari and J.-L. Starck. *Cosmic Microwave Background Data Analysis*, in A dvances in Machine Learning and Data Mining for Astronomy, Chapman and Hall, Edtitors M. Way, J. Scargle, K. Ali, A. Srivastava, to appear.

[34] G.E. Pfander and H. Rauhut. Sparsity in time-frequency representations. *J. Fourier Anal. Appl.*, 16(2):233–260, 2010.

[35] S. Rangan, A. Fletcher, and V. Goyal. *Asymptotic Analysis of MAP Estimation via the Replica Method and Applications to Compressed Sensing*, *Preprint*, 2009.

[36] H. Rauhut. Compressive Sensing and Structured Random Matrices. In M. Fornasier, editor, *Theoretical Foundations and Numerical Methods for Sparse Recovery*, volume 9 of *Radon Series Comp. Appl. Math.*, pages 1–92. deGruyter, 2010.

[37] H. Rauhut, J. Romberg, and J. Tropp. Restricted isometries for partial random circulant matrices. *Appl. Comput. Harmon. Anal.*, to appear.

[38] B. Recht. A simpler approach to matrix completion. *J. Mach. Learn. Res.*, to appear.

[39] B. Recht, M. Fazel, and P. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.*, 52(3):471–501, 2010.

[40] T. Strohmer and R. Vershynin. *A randomized Kaczmarz algorithm with exponential convergence*, Journal of Fourier Analysis and Applications, 15(1):262–278, 2009.