Robustness to unknown error in sparse regularization

Simone Brugiapaglia*, Ben Adcock†
Department of Mathematics
Simon Fraser University
Burnaby, BC V5A 1S6, Canada
Email: *simone_brugiapaglia@sfu.ca, †ben_adcock@sfu.ca

Richard K. Archiblad
Computer Science and Mathematics Division
Oak Ridge National Laboratory
Oak Ridge, TN 37831, USA
Email: archibaldrk@ornl.gov

***Abstract*—From a numerical analysis perspective, assessing the robustness of ℓ^1-minimization is a fundamental issue in compressed sensing and sparse regularization. Yet, the recovery guarantees available in the literature usually depend on *a priori* estimates of the noise, which can be very hard to obtain in practice. In this work, we study the performance of ℓ^1-minimization when these estimates are not available, providing robust recovery guarantees for quadratically constrained basis pursuit and random sampling in bounded orthonormal systems. Applications of this work include approximation of high-dimensional functions, infinite-dimensional sparse regularization for inverse problems, and fast algorithms for non-Cartesian Magnetic Resonance Imaging.**

I. Introduction

In *Compressed Sensing* (CS) and sparse representations we deal with underdetermined linear systems of equations [7], [9]

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{A} \in \mathbb{C}^{m\times N}$, with $m \ll N$, is the sensing matrix, $\mathbf{x} \in \mathbb{C}^N$ is an unknown signal, and $\mathbf{y} \in \mathbb{C}^m$ is the vector of measurements perturbed by noise $\mathbf{n} \in \mathbb{C}^m$. This corruption could be due to physical noise produced by the measuring device, to approximation errors in the model, or to numerical factors. Some examples are model error in inverse problems such as MRI [13], [16], the so-called "inverse crime" committed in infinite-dimensional CS when truncating the signal to its finite dimensional representation [1], [4], or the quadrature error involved in the evaluation of the bilinear form associated with a PDE [5], [6].

A standard tool to regularize the inverse problem (1) and recover a good approximation $\hat{\mathbf{x}}$ to the solution $\mathbf{x}$ (assumed to be sparse or compressible) is the *Quadratically Constrained Basis Pursuit* (QCBP) optimization program

$$\hat{\mathbf{x}} := \arg\min \|\mathbf{z}\|_1, \quad s.t. \quad \|\mathbf{A}\mathbf{z} - \mathbf{y}\|_2 \leq \eta, \tag{2}$$

also called *Basis Pursuit* (BP) when $\eta = 0$.

II. The crucial role of η

Usually, in order to study the recovery guarantees of (2), the parameter η is assumed to control the noise magnitude, i.e.,

$$\|\mathbf{n}\|_2 \leq \eta. \tag{3}$$

Indeed, under the regime (3) and with suitable hypotheses on the sensing matrix $\mathbf{A}$ (e.g., based on the restricted isometry property), the following type of recovery error estimate holds with high probability

$$\|\mathbf{x} - \hat{\mathbf{x}}\|_2 \lesssim \frac{\sigma_s(\mathbf{x})_1}{\sqrt{s}} + \eta, \tag{4}$$

where $\sigma_s(\mathbf{x})_1$ is the best s-term approximation error of $\mathbf{x}$ with respect to the ℓ^1-norm [8], [11].

Unfortunately, *a priori* estimates of the noise of the form (3) may not be available in real applications of CS. Moreover, since the recovery error estimate (4) is sensitive to η, the choice of this parameter is crucial. In practice, one could resort to cross-validation in order to tune this parameter, but this technique could be time-consuming or inaccurate and it is not properly understood from a theoretical perspective [10].

In order to show the importance of η, let us consider a simple example. In Figure 1, we plot the recovery error of BP ($\eta = 0$) and of QCBP with $\eta = 0.01$ as a function of m/N for Fourier and Gaussian random measurements.

III. Robust recovery guarantees

The goal of this work is to establish robust recovery guarantees for QCBP (and BP) under the regime

$$\|\mathbf{n}\|_2 \geq \eta. \tag{5}$$

In this scenario, recovery estimates analogous to (4)–where η is replaced by $\|\mathbf{n}\|_2$–hold for BP [11]. They are based on the so-called quotient property, which is known to be fulfilled only by random Gaussian matrices [20] and by Weibull matrices [12], under suitable restrictions on the number of measurements m.

In this work, we prove robust recovery error estimates for QCBP (and BP) when the matrix $\mathbf{A}$ is built by random sampling in *bounded orthonormal systems* [14] (this framework includes, e.g., the partial discrete Fourier transform, the nonharmonic Fourier transform, and subsampled isometries). In particular, under suitable hypotheses involving the restricted isometry constants and the singular values of $\mathbf{A}$, we provide recovery error estimates in probability of the form

$$\|\mathbf{x} - \hat{\mathbf{x}}\|_2 \lesssim \frac{\sigma_s(\mathbf{x})_1}{\sqrt{s}} + \eta + \log^2(N) \max\{\|\mathbf{n}\|_2 - \eta, 0\}. \tag{6}$$

The effect of the unknown error $\mathbf{n}$ is encapsulated in the third term on the right-hand side (compare with (4)). As is to be expected, this term approaches zero as the estimation of the model error η improves.

The main tool employed in our analysis is the theory of asymptotic estimates for the singular values of random matrices with heavy-tailed rows [19]. In particular, a key role is played by the following incoherence parameter

$$\mu := \mathbb{E}\Big[\max_{i\in[m]} \sum_{k\in[m]\setminus\{i\}} |\langle \mathbf{a}_i, \mathbf{a}_k\rangle|^2\Big], \tag{7}$$

where $[m] := \{1, \ldots, m\}$ and $\mathbf{a}_i$ are the rows of $\mathbf{A}$.

Finally, we discuss extensions and applications of this analysis. First, the case of weighted ℓ^1-minimization. This is used notably in high-dimensional function approximation and interpolation [2], [15], [17], with applications in uncertainty quantification for parametric PDEs. Second, fast methods for non-Cartesian MRI, where model error arises from gridding non-uniform Fourier data to a uniform grid, and can seriously hamper reconstruction quality [3], [13]. See Figure 2 for an illustration. Third, low-rank matrix recovery.

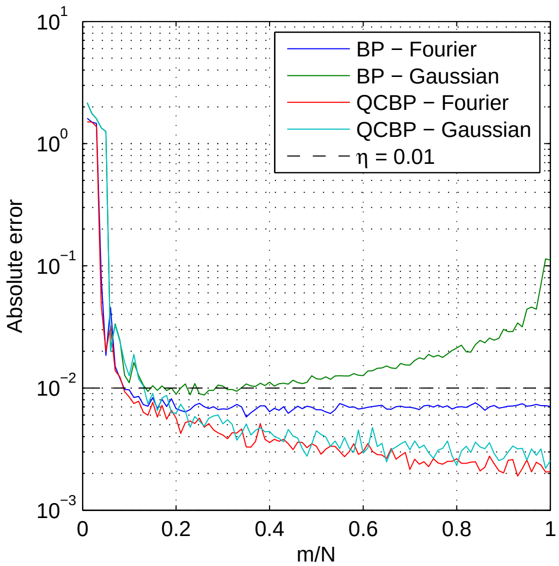

Fig. 1. Numerical assessment of BP ($\eta = 0$) and QCBP (with $\eta = 0.01$) for Fourier and Gaussian measurements corrupted by noise $\mathbf{n}$ of magnitude $\|\mathbf{n}\|_2 = 0.01$. The solution $\mathbf{x}$ is a randomly generated 10-sparse vector in $\mathbb{C}^{1000}$. The absolute error $\|\mathbf{x} - \hat{\mathbf{x}}\|_2$ is plotted as a function of the ratio m/N. The results are produced using the MATLAB package SPGL1 [18]. The QCBP solver, where relation (3) holds, is very robust for both Fourier and Gaussian measurements. On the contrary, for BP, where relation (3) does not hold anymore, the situation is different: the solver's performance highly depends on the type of measurements and on the ratio m/N. Notably, Fourier measurements, coming from randomly subsampling the rows of the DFT matrix, are much more stable in the BP case when $m/N \to 1$ than Gaussian measurements.

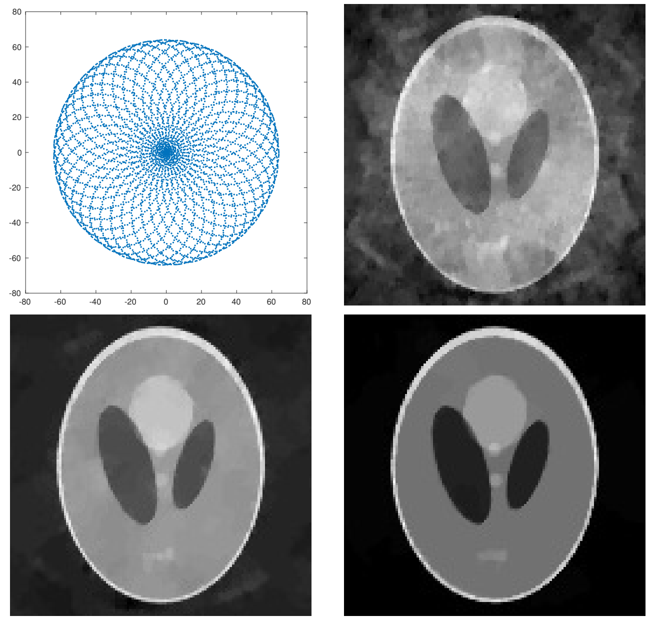

Fig. 2. The effect of gridding error on non-Cartesian MRI. A 128×128 resolution phantom is sampled using a rosette sampling trajectory (top left) giving $m = 5001$ non-Cartesian Fourier measurements. Upfront gridding of the data is performed, and then the images are recovered using total variation minimization. Top right: recovered image using standard nearest neighbor gridding. The signal-to-error ratio is SER = 7.91dB. Bottom row: recovered image using the novel fractional integer nearest neighbor gridding introduced in [3] with parameter $n_{up} = 2$ (left) and $n_{up} = 4$ (right). The signal-to-error ratios are SER = 11.34dB and SER = 12.0dB respectively. Standard gridding leads to $\mathcal{O}(1)$ model error. The fractional nearest neighbor gridding has $\mathcal{O}(1/n_{up})$ model error.

References

[1] B. Adcock, "Infinite-dimensional ℓ^1 minimization and function approximation from pointwise data" *Constr. Approx.*, (to appear), 2016.

[2] B. Adcock, "Infinite-dimensional compressed sensing and function interpolation", *Found. Comput. Math.*, (to appear), 2016.

[3] B. Adcock, R. Archibald, A. Gelb, G. Song, R. B. Platte and E. G. Walsh, "Parameter assessment from time-dependent MR signals using sequential imaging", preprint, 2016.

[4] B. Adcock and A.C. Hansen, "Generalized sampling and infinite-dimensional compressed sensing," *Found. Comput. Math.*, vol. 16, no. 5, pp. 1263–1323, 2016.

[5] S. Brugiapaglia, S. Micheletti, and S. Perotto, "Compressed solving: A numerical approximation technique for elliptic PDEs based on Compressed Sensing," *Comput. Math. Appl.*, vol. 70, no. 6, pp. 1306–1335, 2015.

[6] S. Brugiapaglia, F. Nobile, S. Micheletti, and S. Perotto, "A theoretical study of COmpRessed SolvING for advection-diffusion-reaction problems", *Math. Comp.*, (to appear), 2016.

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

[8] A. Cohen, W. Dahmen, and R.A. DeVore, "Compressed sensing and best k-term approximation," *J. Amer. Math. Soc.*, vol. 22, no. 1, pp. 211–231, 2009.

[9] D. L. Donoho, "Compressed Sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[10] A. Doostan and H. Owhadi, "A non-adapted sparse approximation of PDEs with stochastic inputs," *J. Comput. Phys*, vol. 230, no. 8, pp. 3015–3034, 2011.

[11] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. New York: Springer Science+Business Media, 2013.

[12] S. Foucart, "Stability and robustness of ℓ^1-minimizations with Weibull matrices and redundant dictionaries," *Linear Algebra Appl.*, vol. 441, pp. 4–21, 2014.

[13] M. Guerquin-Kern, L. Lejeune, K.P. Pruessmann, and M. Unser, "Realistic analytical phantoms for parallel magnetic resonance imaging," *IEEE Trans. Med. Imag.*, vol. 31, no. 3, pp. 626–636, 2012.

[14] H. Rauhut, "Compressive sensing and structured random matrices," in *Theoretical Foundations and Numerical Methods for Sparse Recovery*, Radon Series on Computational and Applied Mathematics, vol. 9., M. Fornasier, Ed. Berlin: de Gruyter, 2010.

[15] H. Rauhut and R. Ward, "Interpolation via weighted ℓ_1 minimization", *Appl. Comput. Harmon. Anal.*, vol. 40, no. 2, pp. 321–351, 2016.

[16] B. Roman, A. C. Hansen and B. Adcock, "On asymptotic structure in compressed sensing", arXiv:1406.4178, 2016.

[17] H. Tran, C. Webster and G. Zhang, "Analysis of quasi-optimal polynomial approximations for parameterized PDEs with deterministic and stochastic coefficients", Technical Report ORNL/TM-2015/497, Oak Ridge National Laboratory, 2015.

[18] E. van den Berg and M.P. Friedlander, "Probing the Pareto frontier for basis pursuit solutions," *SIAM J. Sci. Comput.*, vol. 31, no. 2, pp. 890–912, 2008/09.

[19] R. Vershynin, "Introduction to the non-asymptotic analysis of random matrices" in *Compressed Sensing: Theory and Applications*, Y. Eldar and G. Kutyniok, Ed. Cambridge: Cambridge University Press, 2012.

[20] P. Wojtaszczyk, "Stability and instance optimality for Gaussian measurements in compressed sensing," *Found. Comput. Math.*, vol. 10, no.1, pp. 1–13, 2010.