# Sensitivity Considerations in Compressed Sensing

Louis L. Scharf,<sup>1</sup> Edwin K. P. Chong,<sup>1,2</sup> Ali Pezeshki,<sup>2</sup> and J. Rockey Luo<sup>2</sup>

<sup>1</sup>Department of Mathematics, Colorado State University

Fort Collins, CO 80523, USA

<sup>2</sup>Department of Electrical and Computer Engineering, Colorado State University

Fort Collins, CO 80523, USA

Abstract-In [1]-[4], we considered the question of basis mismatch in compressive sensing. Our motivation was to study the effect of mismatch between the mathematical basis (or frame) in which a signal was assumed to be sparse and the physical basis in which the signal was actually sparse. We were motivated by the problem of inverting a complex space-time radar image for the field of complex scatterers that produced the image. In this case there is no apriori known basis in which the image is actually sparse, as radar scatterers do not usually agree to place their ranges and Dopplers on any apriori agreed sampling grid. The consequence is that sparsity in the physical basis is not maintained in the mathematical basis, and a sparse inversion in the mathematical basis or frame does not match up with an inversion for the field in the physical basis. In [1]-[3], this effect was quantified with theorem statements about sensitivity to basis mismatch and with numerical examples for inverting time series records for their sparse set of damped complex exponential modes. These inversions were compared unfavorably to inversions using fancy linear prediction. In [4] and this paper, we continue these investigations by comparing the performance of sparse inversions of sparse images, using apriori selected frames that are mismatched to the physical basis, and by computing the Fisher information matrix for compressions of images that are sparse in a physical basis.

## I. INTRODUCTION

There are many ways that a *sparse inversion* can arise in signal processing. For example:

1) An over-determined separable linear model is replaced by an under-determined linear model. In the separable linear model, parameters such as complex scattering coefficients compose a small number of modes that are nonlinearly modulated by mode parameters such as frequency, wavenumber, and delay. In the approximating linear model, scattering coefficients compose a large number of modes whose frequencies, wavenumbers, and delays are chosen apriori. In the over-determined separable linear model, the problem is to estimate the mode parameters and the scattering coefficients. In the under-determined linear model, the mode parameters are assumed to be determined already at an agreed grid spacing, so the only problem is to estimate the complex scattering coefficients for the few modes that are active and call the parameter estimates the parameters belonging to these active modes.

The question that arises is, what is the sensitivity to mismatch between the unknown apriori basis in the overdetermined problem and the assumed apriori frame in the under-determined problem? How does the performance of the sparse inversion compare with the Cramer-Rao bound (CRB)?

2) A recorded radar image can be compressed with random filters and beamformers, to be subsequently inverted for ranges, Dopplers, and DOAs of the scattering field. The question that arises is, how well does the compressed version of the space-time image preserve information about the field of scatterers? How is Fisher information preserved, or equivalently, how do CRBs for the compressed image compare with CRBs for the original field?

In this paper, we address these questions with results from our continuing investigations. Our results suggest that a compressed set of recordings, processed according to the principles of sparse inversion, will suffer performance losses with respect to classical methods. Our continuing aim is to manage these losses by exploiting apriori knowledge to tailor the compression to the problem.

## II. FROM AN OVER-DETERMINED LINEAR SEPARABLE MODEL TO AN UNDER-DETERMINED LINEAR MODEL

Begin with the separable linear model

$$\mathbf{y} = \mathbf{x} + \mathbf{n}; \ \mathbf{x} = \mathbf{A}(\boldsymbol{\phi})\boldsymbol{\theta}$$

where  $\mathbf{x}, \mathbf{n}, \mathbf{y}$  are elements of the set  $\mathbb{C}^n$ . The parameters  $\boldsymbol{\theta} \in \mathbb{C}^q$  are parameters that linearly compose the signal **x**, and the parameters  $\boldsymbol{\phi} = [\phi_1, \dots, \phi_p]^T \in \mathbb{C}^p$ , with  $\phi_i \in \Phi$  for  $i = 1, \dots, p$ , are parameters that nonlinearly modulate the modes of the model matrix  $A(\phi)$ . An interesting special case is when p = q and  $\mathbf{A}(\boldsymbol{\phi}) \in \mathbb{C}^{n \times q}$ is structured as  $\mathbf{A}(\boldsymbol{\phi}) = [\mathbf{a}(\phi_1), \mathbf{a}(\phi_2), \cdots, \mathbf{a}(\phi_p)]$ , where  $\mathbf{a}(\phi_i) = [1, z(\phi_i), z^2(\phi_i), z^{n-1}(\phi_i)]^T$ , with  $z(\phi_i) \in \mathbb{C}$ . This models DOA modulation in a linear array or linear angle modulation in a time series. There is a great number of methods of modal analysis in the published literature for estimating the parameters  $(\theta, \phi)$ , and their performance is well understood. There are general formulas for the Fisher information matrix and the CRB on the error covariance matrix (see, e.g., [5]), and in specific cases where the matrix  $A(\phi)$ is composed of damped complex exponential modes, there are special formulas that have been numerically evaluated for

This paper was presented at the 2011 Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, Nov. 6-9, 2011. The work reported here is supported by DARPA under contract N66001-11-C-4023 and by NSF under grant CCF-1018472.

concentration ellipses [6]. Generally, the Fisher information matrix in the model  $\mathbf{y} : \mathcal{CN}_n[\mathbf{A}(\phi)\boldsymbol{\theta}, \sigma^2 \mathbf{I}]$  may be written as [7]

$$\mathbf{J}_{\mathbf{y}} = \frac{1}{\sigma^2} \mathbf{G}^H \mathbf{G}$$

where G is a matrix of sensitivities:

$$\mathbf{G} = [\mathbf{G}_{\phi}, \mathbf{G}_{\theta}],$$
  
$$\mathbf{G}_{\phi} = [\mathbf{g}_{\phi}[1], \mathbf{g}_{\phi}[2], \dots, \mathbf{g}_{\phi}[p]]$$
  
$$\mathbf{G}_{\theta} = [\mathbf{g}_{\theta}[1], \mathbf{g}_{\theta}[2], \dots, \mathbf{g}_{\theta}[q]].$$

The sensitivity vectors  $\mathbf{g}$  are partial derivatives of  $\mathbf{A}(\phi)\boldsymbol{\theta}$  with respect to the elements of  $\phi$  and  $\boldsymbol{\theta}$ . Thus the Fisher information matrix  $\mathbf{J} = \frac{1}{\sigma^2} \mathbf{G}^H \mathbf{G}$  is a Grammian of sensitivities. There is a nice geometrical interpretation [7], showing that Fisher information is poor when a one-dimensional subspace  $\langle \mathbf{g}\phi[i] \rangle$  lies near to the multidimensional subspace spanned by all other sensitivity vectors. This explains the difficulty in resolving closely-spaced modes.

Suppose the over-determined linear separable model is replaced by the under-determined linear model

$$\mathbf{y} = \mathbf{x} + \mathbf{n}; \ \mathbf{x} = \mathbf{B}\mathbf{c}$$

where the nonlinearly-modulated mode matrix  $\mathbf{A}(\phi) = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p]$ , with  $\mathbf{a}_i = \mathbf{a}(\phi_i)$  for  $i = 1, 2, \dots, p$ , has been replaced by the fixed frame  $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_N]$ , N > n > p, with the  $\mathbf{b}_i$  a gridding of the set  $\{\mathbf{a}(\phi), \phi \in \Phi\}$ . Whether or not the measurement  $\mathbf{y}$  is compressed, there is the question of inverting each of these models for their mode parameters  $\phi$ . Generally, a *sparse inversion* inverts for sparse  $\mathbf{c}$  in the model  $\mathbf{x} = \mathbf{B}\mathbf{c}$  under a constraint on the fit of  $\mathbf{B}\mathbf{c}$  to  $\mathbf{y}$ . For example, one may solve the following  $\ell_1$  minimization problem:

$$\min \|\mathbf{c}\|_1 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{B}\mathbf{c}\|_2^2 \le \epsilon^2$$

Once sparse c is returned, the estimate for  $\phi$  is extracted from the support set of the identified c. As the gridding of the mode space never produces an exact match between the modes  $\mathbf{a}_i$ and the gridded modes  $\mathbf{b}_i$ , there is a question of sensitivity to mismatch. These questions have been answered analytically and experimentally in [1]–[3], but no Monte Carlo simulation of estimator errors has been undertaken to date. In [4] and this paper, we present a small subset of our simulations for estimators and estimator errors using sparse inversions based on a second order cone program for  $\ell_1$  minimization and sparse inversion based on orthogonal matching pursuit (OMP).

In Figs. 1(a)–(c), we present experimental mean-squared estimation errors for estimating the frequency of an  $f_1 = 0.5$  Hz, unit amplitude, complex exponential from a 25 sample measurement of this complex exponential, plus an interfering  $f_2 = 0.52$  Hz, unit-amplitude, complex exponential, plus complex proper Gaussian noise of variance  $\sigma^2$ . The 0.2 Hz separation between the two tones is a separation of half the Rayleigh limit of (1/25) Hz. For reference, the mean-squared error (MSE) of errors uniformly distributed over this Rayleigh limit is -34 dB. SNR(dB) is defined to be  $10 \log_{10}(1/\sigma^2)$ . Mean-squared error is plotted as  $10 \log_{10}(MSE)$ . The linear dotted line is the CRB. The parameter L denotes the ratio N/n, so it may be called an *expansion factor*. At expansion factor L, the number of modes in the frame B is nL, corresponding to complex exponentials spaced at 1/nL Hz. The resolution of the frame may be said to be 1/nL Hz. These modes are placed so that there is always a half-cell mismatch between the actual frequencies and the mode frequencies. The width of the half-cell (1/2nL) is indicated in Figs. 1(a)–(c) for the various L values as asterisks on the right side of the plot.

Fig. 1(a) demonstrates that at a low expansion factor of L = 2the  $\ell_1$  inversions are noise-defeated below 5 dB and resolutionlimited above 5 dB, meaning they are limited by the resolution of the frame. That is, below 5 dB, mean-squared error is biassquared plus variance, while above 5 dB, mean-squared error is bias-squared due to coarse-grained resolution of the frame. At L = 4, the inversions are noise-defeated below 5 dB, noise-limited from 5 to 10 dB, and resolution-limited above 10 dB. As the expansion factor is increased, corresponding to finer- and finer-grained resolution in frequency, the frame loses its incoherence, meaning the dimension of the null space increases so much that there are many sparse inversions that meet a fitting constraint. As a consequence, we anticipate that for larger values of L the mean-squared error never reaches its resolution limit, as variance overtakes bias-squared to produce mean-squared errors that are larger than those of inversions that used smaller expansion factors. This suggests that there is a clear limit to how much bias-squared can be reduced with frame expansion, before variance overtakes bias-squared to produce degraded  $\ell_1$  inversions. Figs. 1(b) and (c) make these points for OMP inversions. The interesting thing about these results is that OMP inversions extend the threshold behavior of the inversions, they track the CRB more closely in the noise-limited region, and they reach their resolution limit for larger values of L before reaching their null-space limit at high SNRs. For example, at L = 8 the null-space limit has not yet been reached at SNR= 30 dB, whereas for L = 14, the null-space limit is reached before the resolution limit can be reached.

In all of these experiments, the fitting error is matched to the noise variance. With mismatch between fitting error and noise variance the results are more pessimistic. The results suggest that the replacement of an over-determined separable linear model with an under-determined linear model, which is then sparsely inverted, must be very carefully managed to find an appropriate expansion factor and fitting parameter, if performance loss with respect to classical methods is to be managed.

*Remark.* The results reported in Fig. 1 are actually too optimistic, as the mode amplitude are equal. For a weak mode in the presence of a strong interfering mode, the results are much worse.



3



Beam space processing replaces a set of n spatial measurements  $\mathbf{y} \in \mathbb{C}^n$  with a set of  $m \leq n$  beam space measurements  $\mathbf{z} = \mathbf{T}\mathbf{y}, \mathbf{T} \in \mathbb{C}^{m \times n}$ , where the rows of the beamforming matrix  $\mathbf{T}$  are beam-steering vectors. Conventionally these rows are designed to form spatial beams, and typically the beam centers are spaced at m equally-spaced resolutions of electrical angle between  $(-\pi, \pi]$ . So the question is, how does the Fisher information matrix for the DOA parameters of a sparse scattering or radiating field, computed from the original measurements  $\mathbf{y}$ , compare with the Fisher information matrix for these and parameters, computed from the beam space measurements  $\mathbf{z} = \mathbf{T}\mathbf{y}$ ? And how do these results compare for a conventional choice of the beamforming matrix and for a random choice? We offer answers to both of these questions.

For any transformation **T**, the effect on the Fisher information matrix in the model  $\mathbf{y} : C\mathcal{N}_n[\mathbf{A}(\boldsymbol{\phi})\boldsymbol{\theta}, \sigma^2 \mathbf{I}]$  is [7]

$$\mathbf{J}_{\mathbf{T}\mathbf{y}} = \frac{1}{\sigma^2} \mathbf{G}^H \mathbf{P}_{\mathbf{T}^H} \mathbf{G}$$

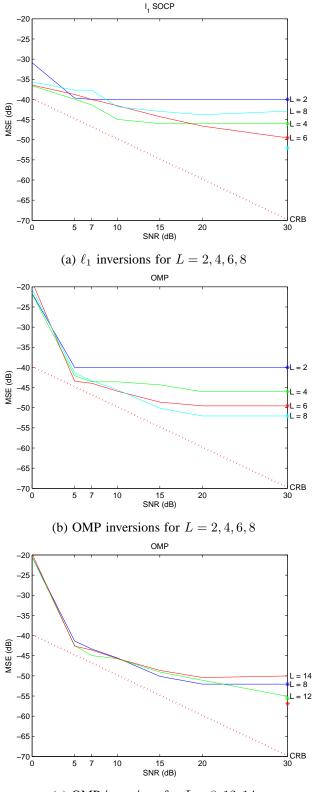
where  $\mathbf{P}_{\mathbf{T}^{H}} = \mathbf{T}^{H} (\mathbf{T}\mathbf{T}^{H})^{-1}\mathbf{T}$  is a projection onto the range space of  $\mathbf{T}^{H}$ , and  $\mathbf{G}$  is a matrix of sensitivities. It is the effect of the transformation  $\mathbf{T}$  on the sensitivity vectors,  $\mathbf{T}\mathbf{G}$ , and on the noise covariance,  $\mathbf{T}\mathbf{T}^{H}$ , that determines the impact of the transformation on Fisher information.

*Remark.* If the parameters  $\phi$  modulate a covariance matrix  $\mathbf{R}(\phi)$  in the second-order model  $\mathbf{y} : \mathcal{CN}_N[\mathbf{0}, \mathbf{R}(\phi)]$ , then the effect of transformation  $\mathbf{T}$  is to produce Fisher information

$$\left(\mathbf{J}_{\mathbf{Ty}}\right)_{ij} = \operatorname{tr}\left\{\left(\mathbf{TRT}^{H}\right)^{-1}\mathbf{T}\frac{\partial\mathbf{R}}{\partial\phi_{i}}\mathbf{T}^{H}(\mathbf{TRT}^{H})^{-1}\mathbf{T}\frac{\partial\mathbf{R}}{\partial\phi_{j}}\mathbf{T}^{H}\right\}.$$

Figs. 2(a)–(d) demonstrate the results of random beamsteering with a transformation **T**. The experiment is the following: two plane waves propagate across the face of an n = 64 element linear array or two linearly modulated complex exponentials populate a time series. One of the plane waves arrives on boresight and the other arrives at electrical angle  $\phi$ , with  $\phi$ swept over  $(-3\frac{2\pi}{n}, 3\frac{2\pi}{n}]$  with a step size much finer than the Rayleigh limit of  $\frac{2\pi}{n}$  radians. For each such angle, the CRB for estimating the boresight angle is computed and plotted, with and without beamsteering with compressive beamformer **T**. In our experiments, **T** is chosen to be a random complex Gaussian matrix, a Slepian matrix of approximate beamwidth  $2\pi/m$  (see, e.g., [8] and [9]), or a monopulse beamsteering consisting of a sum beam and a difference beam (see, e.g., [10]). The SNR is normalized to 0 dB, as it is only the relative values of the CRB that interest us.

Moving from Fig. 2(a) to Fig. 2(d), the *compression fac*tor n/m is increased from 1 to 8 for a random complex Gaussian compressor. In Fig. 2(a), the compression factor is n/m = 1. The random beamformers build a nonsingular matrix, with probability one. So the CRB is the CRB of the original measurements, without beamforming. In Fig. 2(d), the compression factor is n/m = 8. The band of CRBs



(c) OMP inversions for L = 8, 12, 14

Fig. 1. Experimental mean-squared estimation errors for estimating the frequency of an  $f_1 = 0.5$  Hz, unit amplitude, complex exponential from a 25 sample measurement of this complex exponential, plus an interfering  $f_2 = 0.52$  Hz, unit-amplitude, complex exponential, plus complex proper Gaussian noise of variance  $\sigma^2$ . Experimental mean-squared errors are shown for  $\ell_1$  and OMP inversions for different expansion factors L. The asterisks indicate the width of the half-cell (1/2nL) for the various L values.

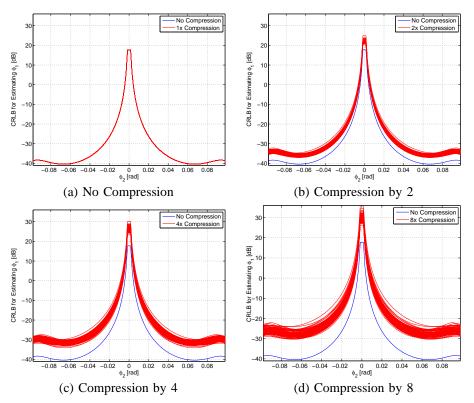


Fig. 2. Cramer-Rao lower bound for estimating the electrical angle of source at boresight, as the electrical angle  $\phi$  of an interfering source varies from  $-3\frac{2\pi}{n}$  to  $3\frac{2pi}{n}$ . Each red curve in a subfigure corresponds to one realization of a random complex Gaussian beamforming matrix **T** of size *m* by *n*. A total of 100 realizations is superimposed in each subfigure. (a) No Compression (m = n = 64), (b) Compression by 2 (m = n/2 = 32), (c) Compression by 4 (m = n/4 = 16), and (d) Compression by 8 (m = n/8 = 8).

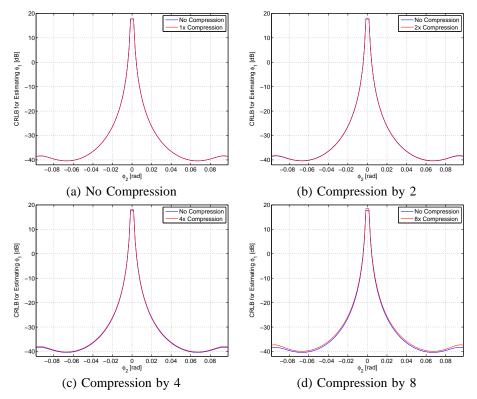
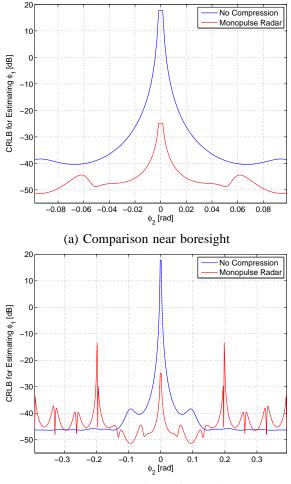


Fig. 3. Cramer-Rao lower bound for estimating the electrical angle of source at boresight, as the electrical angle  $\phi$  of an interfering source varies from  $-3\frac{2\pi}{n}$  to  $3\frac{2pi}{n}$ , from compressive measurements made by a Slepian beamformer of approximate beamwidth  $2\pi/m$ . (a) No Compression (m = n = 64), (b) Compression by 2 (m = n/2 = 32), (c) Compression by 4 (m = n/4 = 16), and (d) Compression by 8 (m = n/8 = 8).

shows CRBs for many random choices of **T**. The variation between good and bad random choices of **T** is 6 dB, and the best choice has a loss of 9 dB with respect to the CRB for the original measurements. This represents an increase in variance in estimating the boresight angle by a factor of at least 9 radians-squared, or an increase in standard deviation by a factor of at least  $\sqrt{9}$  radians. These results quantify the performance loss of compressed random beamforming, providing a basis for selecting a compression factor that meets performance specifications.

Figs. 3(a)–(d) illustrate the same results for Slepian beamforming and Figs. 4(a) and (b) illustrate the results for monopulse beamsteering. Within the main beam defined by the Rayleigh limit, the increase in CRB for the Slepian beamformer is negligible, while the monopulse beamformer outperforms the uncompressed array processing by filtering out broad wavenumber noise. This shows that if the vicinity of closely-spaced sources is known, this knowledge can be exploited to tailor the compression to improve resolution.



(b) Comparison in a wider region

Fig. 4. Cramer-Rao lower bound for estimating the electrical angle of source at boresight, as the electrical angle  $\phi$  of an interfering source varies from  $-3\frac{2\pi}{n}$  to  $3\frac{2\pi}{n}$ , from compressive measurements made by a Monopulse beamformer.

In all of these experiments, it is the so-called deterministic Fisher information  $\mathbf{J}_{\mathbf{T}\mathbf{y}} = \frac{1}{\sigma^2} \mathbf{G}^H \mathbf{P}_{\mathbf{T}^H} \mathbf{G}$  that is computed and inverted for the CRB on the variance of an estimator of boresight angle.

#### **IV. CONCLUSIONS**

We have reviewed two problems in signal processing that may be approximated and re-framed in such a way that a sparse inversion of compressed measurements for modal parameters can be considered. But in each case there are consequences arising from model mismatch and from loss of Fisher information. Our results illuminate these consequences and suggest remedies based on apriori knowledge.

#### REFERENCES

- [1] Y. Chi, L. L. Scharf, A. Pezeshki, and R. Calderbank, "Sensitivity to basis mismatch of compressed sensing for spectrum analysis and beamforming," in *Proc. 6th U.S./Australia Joint Workshop on Defence Applications of Signal Processing (DASP)*, Lihue, HI, Sep. 26-28 2009.
- [2] Y. Chi, A. Pezeshki, L. L. Scharf, and R. Calderbank, "Sensitivity to basis mismatch in compressed sensing," in *Proc. 2010 IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Dallas, TX, Mar. 14-19 2010, pp. 3930–3933.
- [3] Y. Chi, L. L. Scharf, A. Pezeshki, and R. Calderbank, "Sensitivity to basis mismatch in compressed sensing," *IEEE Trans. Signal Processing*, vol. 59, no. 5, pp. 2182–2195, May 2011.
- [4] L. L. Scharf, E. K. P. Chong, A. Pezeshki, and J. Luo, "Compressive sensing and sparse inversion in signal processing: Cautionary notes," in *Proc. 7th U.S./Australia Joint Workshop on Defence Applications of Signal Processing (DASP)*, Coolum, Queensland, Australia, Jul. 9-14 2011.
- [5] H. L. V. Trees and K. L. Bell, Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking. IEEE Press, 2007.
- [6] L. T. McWhorter and L. L. Scharf, "Cramer-Rao bounds for deterministic modal analysis," *IEEE Trans. Signal Process.*, vol. 41, no. 5, pp. 1847–1862, May 1993.
- [7] L. L. Scharf and L. M. Whorter, "Geometry of the Cramer-Rao bound," Signal Processing, vol. 31, no. 3, pp. 301–311, Apr. 1993.
- [8] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis and uncertainty - V: The discrete case," *Bell Syst. Tech. J.*, pp. 1371–1430, 1978.
- [9] C. T. Mullis and L. L. Scharf, "Quadratic estimators of the power spectrum," in Advances in Spectrum Estimation, S. Haykin, Ed. Prentice Hall, 1990, vol. 1, ch. 1, pp. 1–57.
- [10] E. Mosca, "Angle estimation in amplitude comparison monopulse system," *IEEE Trans. Aerospace and Electronic Systems*, vol. AES-5, no. 2, pp. 205–212, Mar. 1969.