

Compressed sensing with uncertainty. The Bayesian estimation perspective

Stéphanie Bernhardt, Remy Boyer, Sylvie Marcos, Pascal Larzabal

► **To cite this version:**

Stéphanie Bernhardt, Remy Boyer, Sylvie Marcos, Pascal Larzabal. Compressed sensing with uncertainty. The Bayesian estimation perspective. IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP 2015), Dec 2015, Cancùn, Mexico. hal-01245392

HAL Id: hal-01245392

<https://hal-supelec.archives-ouvertes.fr/hal-01245392>

Submitted on 3 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Compressed Sensing with Uncertainty - the Bayesian Estimation Perspective

Stéphanie Bernhardt, Rémy Boyer and Sylvie Marcos

L2S

Université Paris-Sud (UPS), CNRS, CentraleSupélec

Gif-Sur-Yvette, France

{stephanie.bernhardt, remy.boyer, sylvie.marcos}@lss.supelec.com

Pascal Larzabal

SATIE

ENS Cachan

Cachan, France

pascal.larzabal@satie.ens-cachan.fr

Abstract—The Compressed Sensing (CS) framework outperforms the sampling rate limits given by Shannon’s theory. This gap is possible since it is assumed that the signal of interest admits a linear decomposition of few vectors in a given sparsifying Basis (Fourier, Wavelet, ...). Unfortunately in realistic operating systems, uncertain knowledge of the CS model is inevitable and must be evaluated. Typically, this uncertainty drastically degrades the estimation performance of sparse-based estimators in the low noise variance regime. In this work, the Off-Grid (OG) and Basis Mismatch (BM) problems are compared in a Bayesian estimation perspective. At first glance, we are tempted to think that these two acronyms stand for the same problem. However, by comparing their Bayesian Cramèr-Rao Bounds (BCRB) for the estimation of a L -sparse amplitude vector based on N measurements, it is shown that the BM problem has a lower BCRB than the CS one in a general context. To go further into the analysis we provide for *i.i.d.* Gaussian amplitudes and in the low noise variance regime an interesting closed-form expression of a normalized 2-norm criterion of the difference of the two BCRB matrices. Based on the analysis of this closed-form expression, we obtain two conclusions. Firstly, the two uncertainty problems cannot be confused for a non-zero mismatch error variance and with finite N and L . Secondly, the two problems turn to be similar for any mismatch error variance in the large system regime, *i.e.*, for $N, L \rightarrow \infty$ with constant aspect ratio $N/L \rightarrow \rho$.

Keywords—Compressed Sensing with Uncertainty, Off-Grid effect and Basis Mismatch, Bayesian estimation

I. INTRODUCTION

Compressed Sensing (CS) has been a hot research domain in the last decade [1], [2]. CS potentially outperforms the limit sampling rate predicted by classical sampling theory. This is done by exploiting the a priori knowledge that many natural measurement signals admit a sparse representation in a given redundant dictionary. In today’s world that has to face more and more data, this technique has opened many perspectives on a wide panel of application domains as for instance direction of arrival estimation, sparse radar, etc. The CS framework has thus driven the design of a plethora of sparse-based estimation algorithms as for instance OMP [3], CoSAMP, LASSO or BPDN [4], etc.

In the context of non-ideal CS, there exists two types of uncertainty. The first-one is the well-known Off-Grid (OG)

problem [5], [6] which occurs in sparse estimation. In this case, the continuous estimation parameters cannot match with high probability the pre-fixed discretization of the parameter set. Consider a redundant dictionary obtained from the pre-fixed discretization: as the parameters of interest do not belong to the grid, it appears a ”quantization error”-type [7]. To take into account this effect the measurement vector, denoted by \tilde{y} , is given by a linear combination of few vectors extracted from the initial dictionary corrupted by an additive perturbation while the initial dictionary is known. The Basis Mismatched (BM) problem [8], [6] is described in the following way. The measurement vector, denoted by y , is given by a linear combination of few vectors extracted from the initial dictionary but we dispose of an uncertain knowledge of the initial dictionary or equivalently it is assumed the knowledge of a corrupted dictionary. As a consequence, the sparse-based estimators exhibit very poor estimation accuracies [7] even if the noise variance is low and the support is perfectly estimated (or *a priori* known). Note that the interest of the CRB in the context of noisy CS with deterministic amplitudes and with a perfect knowledge of the support has been demonstrated in references [9], [10], [11]. In particular it is demonstrated that the CRB can be reached by some sparse-based estimators knowing only the support cardinality L and in the low variance regime. We think that this context is well adapted to the BM and OG problems. Based on these arguments, the main concern of this work is to compare these two types of uncertainty in the Bayesian estimation point of view, *i.e.* based on the derivation and the analysis of each BCRB associated with the OG and BM problems.

In the second section we present both uncertainty models, and derive the corresponding BCRB in section 3. First this BCRB is derived for the general case, then we consider the case of an *i.i.d.* gaussian amplitude vector in the low noise variance regime and finally we compare both bounds. In the fourth section we show by simulation the difference of both bounds.

II. COMPRESSED SENSING MODELS WITH UNCERTAINTY

Assume that the dictionary is corrupted according to $\tilde{\mathbf{H}} = \mathbf{H} + \mathbf{E}$ where \mathbf{H} is a random but time-invariant matrix [10] whose entries are *i.i.d.* zero-mean and of variance $1/N$ where N is the number of measurements *i.e.* the number of lines of \mathbf{H} . Matrix \mathbf{E} modelizes the mismatched error following a Gaussian pdf of zero-mean and of variance $\sigma_e^2 \mathbf{I}$. In addition,

This work was supported by the European Commission in the framework of the FP7 Network of Excellence in Wireless COMmunications NEWCOM# (Grant agreement no. 318306), and by the two following projects: ANR MAGELLAN and MI-TITAN.

we define the two following observations

BM problem : $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} = \tilde{\mathbf{H}}\mathbf{x} - \mathbf{E}\mathbf{x} + \mathbf{n}$
 known by the estimator: $\tilde{\mathbf{H}}$
 whose entries are stochastic processes

OG problem : $\tilde{\mathbf{y}} = \tilde{\mathbf{H}}\mathbf{x} + \mathbf{n} = \mathbf{H}\mathbf{x} + \mathbf{E}\mathbf{x} + \mathbf{n}$,
 known by the estimator: \mathbf{H}
 whose entries are random but non-stochastic

The two models are described on Fig. 1.

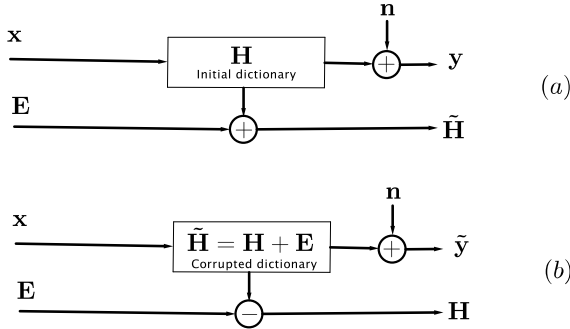


Fig. 1. (a) BM problem, (b) OG problem

In the two above models, the noise term $\mathbf{E}\mathbf{x} + \mathbf{n}$ is data-dependent due to the mismatched error \mathbf{E} . At first glance the two above observations seem to describe the same formalism and intuitively, we are tempted to conclude that the two considered frameworks are related to the same estimation problem. We will show in the sequel that this is a wrong intuition.

III. BCRB WITH UNCERTAINTY

The VanTrees' Bayesian bound [12] is a benchmark against which any Bayesian estimator can be compared. More precisely, let $\hat{\mathbf{x}}$ be an estimate of \mathbf{x} , the BMSEs for each the two contexts of interest are

$$\text{BMSE}_{\text{OG}} = \mathbb{E}_{\tilde{\mathbf{y}}, \mathbf{x} | \mathbf{H}} \|\mathbf{x} - \hat{\mathbf{x}}(\tilde{\mathbf{y}})\|^2 \geq \text{Tr} \mathbf{B}_{\text{OG}} \quad (1)$$

$$\text{BMSE}_{\text{BM}} = \mathbb{E}_{\mathbf{y}, \tilde{\mathbf{H}}, \mathbf{x}} \|\mathbf{x} - \hat{\mathbf{x}}(\mathbf{y}, \tilde{\mathbf{H}})\|^2 \geq \text{Tr} \mathbf{B}_{\text{BM}} \quad (2)$$

where for $1 \leq i, j \leq L$,

$$[\mathbf{B}_{\text{OG}}^{-1}]_{ij} = \left[\text{Var} \left(\frac{\partial \log p(\mathbf{y}, \mathbf{x} | \mathbf{H})}{\partial \mathbf{x}} \right) \right]_{ij}, \quad (3)$$

$$[\mathbf{B}_{\text{BM}}^{-1}]_{ij} = \left[\text{Var} \left(\frac{\partial \log p(\mathbf{y}, \tilde{\mathbf{H}}, \mathbf{x})}{\partial \mathbf{x}} \right) \right]_{ij}. \quad (4)$$

where $\text{Var}(x)$ is the variance of x .

A. BCRB in case of the off-grid problem

The off-grid problem is described according to the following scenario: Estimate the amplitude vector \mathbf{x} based on the observation $\tilde{\mathbf{y}}$ knowing the uncorrupted dictionary \mathbf{H} . So, the conditional observation is

$$\tilde{\mathbf{y}} | \mathbf{H}, \mathbf{x} \sim \mathcal{N}(\mathbf{H}\mathbf{x}, \mathbf{R} = (\sigma_e^2 \|\mathbf{x}\|^2 + \sigma^2) \mathbf{I}). \quad (5)$$

Using the Slepian-Bang formula [13], we obtain the following Fisher Information Matrix (FIM):

$$\mathbf{F}_{\text{OG}} = \frac{\mathbf{H}^T \mathbf{H}}{\sigma_e^2 \|\mathbf{x}\|^2 + \sigma^2} + \frac{2\sigma_e^4 N \mathbf{x} \mathbf{x}^T}{(\sigma_e^2 \|\mathbf{x}\|^2 + \sigma^2)^2} \quad (6)$$

and thus the BCRB takes the following expression

$$\mathbf{B}_{\text{OG}} = (\mathbf{H}^T \mathbf{H} \mathcal{P}(\sigma_e^2, \sigma^2) + \boldsymbol{\Sigma}(\sigma_e^2, \sigma^2) + \mathbf{P}_x)^{-1} \quad (7)$$

where $\mathbf{P}_x = \text{Var} \left(\frac{\partial \log p(\mathbf{x})}{\partial \mathbf{x}} \right)$ and

$$\mathcal{P}(\sigma_e^2, \sigma^2) = \mathbb{E}_x \frac{1}{\sigma_e^2 \|\mathbf{x}\|^2 + \sigma^2}, \quad (8)$$

$$\boldsymbol{\Sigma}(\sigma_e^2, \sigma^2) = \mathbb{E}_x \frac{2\sigma_e^4 N \mathbf{x} \mathbf{x}^T}{(\sigma_e^2 \|\mathbf{x}\|^2 + \sigma^2)^2}. \quad (9)$$

B. BCRB in case of the Basis Mismatched (BM) problem

The Basis Mismatched (BM) problem is described according to the following scenario: Estimate the amplitude vector \mathbf{x} based on the observation \mathbf{y} knowing the corrupted dictionary $\tilde{\mathbf{H}}$. The log-joint pdf is given by

$$\log p(\mathbf{y}, \tilde{\mathbf{H}}, \mathbf{x}) = \log p(\mathbf{y} | \tilde{\mathbf{H}}, \mathbf{x}) + \log p(\tilde{\mathbf{H}}) + \log p(\mathbf{x}) \quad (10)$$

where the conditional observation is given by

$$\mathbf{y} | \tilde{\mathbf{H}}, \mathbf{x} \sim \mathcal{N}(\tilde{\mathbf{H}}\mathbf{x}, (\sigma_e^2 \|\mathbf{x}\|^2 + \sigma^2) \mathbf{I}). \quad (11)$$

Observe that $\partial \log p(\tilde{\mathbf{H}}) / \partial \mathbf{x}$ vanishes [14] and thus the BCRB matrix is given by

$$\mathbf{B}_{\text{BM}} = \left(\mathbb{E}_{\mathbf{x}, \tilde{\mathbf{H}}}(\mathbf{F}_{\text{BM}}) + \mathbf{P}_x \right)^{-1} \quad (12)$$

where the FIM \mathbf{F}_{BM} is given by the Slepian-Bang formula according to

$$\mathbf{F}_{\text{BM}} = \frac{\tilde{\mathbf{H}}^T \tilde{\mathbf{H}}}{\sigma_e^2 \|\mathbf{x}\|^2 + \sigma^2} + \frac{2\sigma_e^4 N \mathbf{x} \mathbf{x}^T}{(\sigma_e^2 \|\mathbf{x}\|^2 + \sigma^2)^2}. \quad (13)$$

Noting that \mathbf{x} and $\tilde{\mathbf{H}}$ are two multidimensional independent processes, we have

$$\mathbb{E}_{\mathbf{x}, \tilde{\mathbf{H}}}(\mathbf{F}_{\text{BM}}) = \mathbb{E}_x \frac{\mathbb{E}_{\tilde{\mathbf{H}}}(\tilde{\mathbf{H}}^T \tilde{\mathbf{H}})}{\sigma_e^2 \|\mathbf{x}\|^2 + \sigma^2} + \mathbb{E}_x \frac{2\sigma_e^4 N \mathbf{x} \mathbf{x}^T}{(\sigma_e^2 \|\mathbf{x}\|^2 + \sigma^2)^2} \quad (14)$$

where $\mathbb{E}_{\tilde{\mathbf{H}}}(\tilde{\mathbf{H}}^T \tilde{\mathbf{H}}) = \mathbf{H}^T \mathbf{H} + \sigma_e^2 \mathbf{I}$. Finally, we obtain

$$\mathbf{B}_{\text{BM}} = ((\mathbf{H}^T \mathbf{H} + \sigma_e^2 \mathbf{I}) \mathcal{P}(\sigma_e^2, \sigma^2) + \boldsymbol{\Sigma}(\sigma_e^2, \sigma^2) + \mathbf{P}_x)^{-1}. \quad (15)$$

C. Comparison

Result 3.1: From expressions (7) and (15), we deduce

$$\mathbf{B}_{\text{BM}} < \mathbf{B}_{\text{OG}} \quad (16)$$

for $\sigma_e^2 \neq 0$.

Proof: The proof is direct considering that all the matrices involved in the two bounds are positive and definite. ■

Remark that we verify the trivial property that if $\sigma_e^2 \rightarrow 0$, then $\mathbf{B}_{\text{BM}} = \mathbf{B}_{\text{OG}}$. The above result shows that for a given noise power, the BM problem exhibits a better accuracy than the OG problem. In the following sections, we provide qualitative analytic expressions to precisely characterize the "distance" between the two studied uncertainty problems. To reach this goal we add some realistic assumptions.

D. Closed-form expressions for i.i.d. Gaussian amplitude vector in the low noise variance regime

In this section,

- A_1 . We consider the low noise variance regime since it is well known that the uncertainty context including the Off-Grid and Basis Mismatched problems appears only in the regime where the noise variance is sufficiently low to be dominated by σ_e^2 . This fact directly implies that any sparse-based estimator in this context cannot be statistically efficient. As this property is a highly desired feature in the context of the estimation theory, the reader can measure the importance of the analysis of this context. At contrary, when the noise variance is high with respect to σ_e^2 , the uncertainty context cannot be measured and thus can be ignored.
- A_2 . The last assumption is to assume that the L amplitudes belonging to the support, *i.e.* taking non-zero values follow an *i.i.d.* centered Normal distribution with variance σ_x^2/L . So, the sparse amplitude vector, denoted by $\tilde{\mathbf{x}}$, is composed by $K - L$ zeros-values and L random non-zero amplitudes. In this case, the prior matrix of the BIM is given by $\mathbf{P}_x = \frac{L}{\sigma_x^2} \mathbf{I}_L$.

Lemma 3.2: In the low noise variance regime, we have

$$\lim_{\sigma^2 \rightarrow 0} \mathcal{P}(\sigma_e^2, \sigma^2) = \frac{L}{(L-2)\sigma_e^2\sigma_x^2}, \quad (17)$$

$$\lim_{\sigma^2 \rightarrow 0} \Sigma(\sigma_e^2, \sigma^2) = \frac{2N}{(L-2)\sigma_x^2} \mathbf{I}_L. \quad (18)$$

Proof: For centered *i.i.d.* Gaussian amplitudes of variance σ_x^2 , we have $1/|\mathbf{x}|^2 \sim \text{Inv} - \chi_L^2$ where $\text{Inv} - \chi_L^2$ stands for an Inverse Chi-squared distribution of L degrees of freedom. Thus, for $L > 2$, $\mathbb{E}_x \frac{1}{|\mathbf{x}|^2} = \frac{L}{\sigma_x^2(L-2)}$. Using the fact that $\lim_{\sigma^2 \rightarrow 0} \mathcal{P}(\sigma_e^2, \sigma^2) = \frac{1}{\sigma_e^2} \mathbb{E}_x \frac{1}{|\mathbf{x}|^2}$ we obtain expression (17).

It is easy to see that $\mathbb{E}_x \frac{\mathbf{x}\mathbf{x}^T}{|\mathbf{x}|^4}$ is a diagonal matrix proportional to the identity matrix, *i.e.* $\mathbb{E}_x \frac{\mathbf{x}\mathbf{x}^T}{|\mathbf{x}|^4} = t \mathbf{I}_L$. Note that $\text{Tr} \mathbb{E}_x \frac{\mathbf{x}\mathbf{x}^T}{|\mathbf{x}|^4} = \mathbb{E}_x \frac{1}{|\mathbf{x}|^2} = \frac{L}{(L-2)\sigma_x^2}$. Thus, we obtain $t = \frac{1}{(L-2)\sigma_x^2}$. Using the fact that $\lim_{\sigma^2 \rightarrow 0} \Sigma(\sigma_e^2, \sigma^2) = 2N \mathbb{E}_x \frac{\mathbf{x}\mathbf{x}^T}{|\mathbf{x}|^4}$ we obtain expression (18). ■

Using the above Lemma, we can give the following result.

Result 3.3: For *i.i.d.* Gaussian amplitude vector and in the low noise variance regime, the traces of the BCRB matrices are given by

$$\text{Tr} \mathbf{B}_{\text{OG}} = \frac{\sigma_x^2(L-2)}{L} \text{Tr} \left[\left(\frac{\mathbf{H}^T \mathbf{H}}{\sigma_e^2} + \mathcal{S}_{N,L} \mathbf{I} \right)^{-1} \right], \quad (19)$$

$$\text{Tr} \mathbf{B}_{\text{BM}} = \frac{\sigma_x^2(L-2)}{L} \text{Tr} \left[\left(\frac{\mathbf{H}^T \mathbf{H}}{\sigma_e^2} + (\mathcal{S}_{N,L} + 1) \mathbf{I} \right)^{-1} \right]. \quad (20)$$

respectively, where $\mathcal{S}_{N,L} = 2\frac{N}{L} - 2 + L$.

E. Comparison of the two bounds

Based on the above Result, we can easily compare the two bounds. Define the following normalized 2-norm criterion:

$$\eta \stackrel{\text{def.}}{=} \|\mathbf{B}_{\text{BM}}^{-1}(\mathbf{B}_{\text{BM}} - \mathbf{B}_{\text{OG}})\|_2. \quad (21)$$

To compare \mathbf{B}_{BM} and \mathbf{B}_{OG} the 2-norm of the difference $\mathbf{B}_{\text{BM}} - \mathbf{B}_{\text{OG}}$ is naturally used. However this 2-norm can be numerically small even for distant bounds due to the small value of \mathbf{B}_{BM} and \mathbf{B}_{OG} . Thus, to improve the robustness of the 2-norm criterion, we normalize the difference by $\mathbf{B}_{\text{BM}}^{-1}$ which takes large values when the bounds are too small to obtain the criterion in (21).

Result 3.4: The normalized 2-norm criterion introduced in (21) takes the following simple form:

$$\eta = \frac{\sigma_e^2}{\lambda_{\min}(\mathbf{H}^T \mathbf{H}) + \sigma_e^2 \mathcal{S}_{N,L}} \quad (22)$$

where $\lambda_{\min}(\cdot)$ stands for the minimal singular values.

Proof: Taking into account $\mathbf{B}_{\text{BM}} = \left(\mathbf{B}_{\text{OG}}^{-1} + \frac{L}{\sigma_x^2(L-2)} \mathbf{I} \right)^{-1}$ and using the definition of the 2-norm involving the maximal singular value, denoted by $\lambda_{\max}(\cdot)$, we obtain

$$\eta = \frac{L}{\sigma_x^2(L-2)} \|\mathbf{B}_{\text{OG}}\|_2 = \lambda_{\max} \left(\left(\frac{1}{\sigma_e^2} \mathbf{H}^T \mathbf{H} + \mathcal{S}_{N,L} \mathbf{I} \right)^{-1} \right).$$

Now using the fact $\lambda_{\max}(\mathbf{Q}^{-1}) = 1/\lambda_{\min}(\mathbf{Q})$ in the above expression, we obtain (22). ■

1) *Analysis of η wrt. σ_e^2 with finite N and L :* As intuitively expected, one can verify $\lim_{\sigma_e^2 \rightarrow 0} \eta = 0$. For growing σ_e^2 , recalling that as $L < N$ we have $\lambda_{\min}(\mathbf{H}^T \mathbf{H}) > 0$, expression (22) is a monotonically increasing function. This result means that the two considered uncertainty problems are not equivalent for sufficient σ_e^2 with finite N and L .

2) *Asymptotic behavior of η in the large system scenario:* In the Random Matrix Theory (RMT) [15], [16], [17], it is assumed that $N, L \rightarrow \infty$ with $N/L \rightarrow \rho$. It is also usual to assume in the CS framework that dictionary \mathbf{H} is a random and time-invariant matrix whose entries follow a given distribution with zero-mean and variance $1/N$. This construction verifies the RIP with high probability [2], [1]. We can now give the following result:

Result 3.5: For any σ_e^2 and if $N, L \rightarrow \infty$ with $N/L \rightarrow \rho$, then $\eta \rightarrow 0$. This means that the OG and BM problems are identical in the Bayesian estimation point of view in the large system scenario.

Proof: It is well-known that $\lambda_{\min}(\mathbf{H}^T \mathbf{H}) \rightarrow \left(1 - \sqrt{\frac{1}{\rho}}\right)^2$ in almost sure convergence [15], [16], [17], meaning that $\lambda_{\min}(\mathbf{H}^T \mathbf{H})$ converges to a finite value. Now, remark that sequence $\mathcal{S}_{N,L} > L$ due to the fact that for avoiding the singularity of the FIM we have $L < N$. So, for infinite L , sequence $\mathcal{S}_{N,L}$ diverges. Thus, criterion given in (22) converges to 0 for any σ_e^2 . ■

IV. NUMERICAL ILLUSTRATIONS

In this section, we assume *i.i.d.* Gaussian amplitude vector and the low noise variance regime. On Fig. 2, it is compared the trace of expressions (7) and (15) obtained thanks to 2000 Monte-Carlo trials to the expressions (20) and (19) obtained under the assumption of low noise variance. We can verify the good agreement of our approximations in the low noise variance regime. It is also possible to numerically check Result (16). In addition, we can see that the two bounds tend to be very close for small σ_e^2 , as intuitively expected.

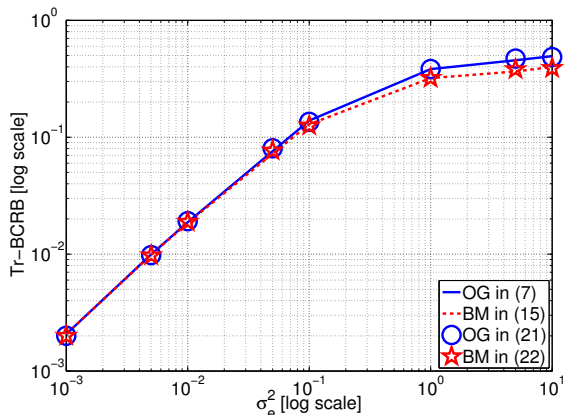


Fig. 2. Trace-BCRB Vs. σ_e^2 with $N = 20$ measurements, $L = 4$ non-zero amplitudes with $\sigma_x^2 = 1$. Expressions (7) and (15) are obtained thanks to 2000 Monte-Carlo trials.

On Fig. 3 are compared expression (21) obtained by means of Monte-Carlo simulations and expression (22). As noticed before, quantity η increases when the error variance σ_e^2 increases. Furthermore, Fig. 3 shows that expression (22) corresponds to the observed Monte Carlo results.

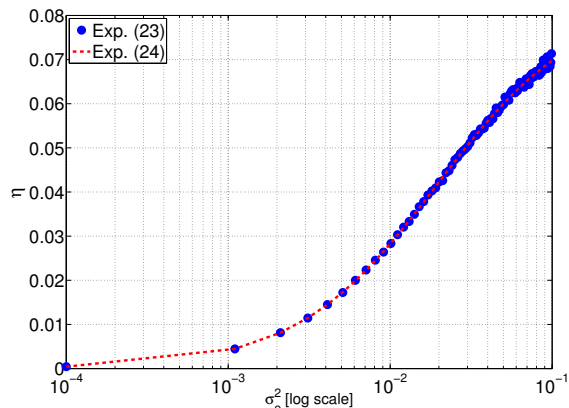


Fig. 3. Criterion η Vs. σ_e^2 .

V. CONCLUSION

In this work, we compare two non-ideal CS frameworks in a Bayesian estimation perspective. More precisely, we focus our study on two well-known uncertainty models known under the acronyms of Off-Grid (OG) and Basis Mismatch (BM) problems. These two problems degrade drastically the estimation accuracy of any sparse-based estimator even if the noise variance is low and the support set of the non-zero amplitudes is perfectly estimated or a priori known. At first glance, the OG and BM problems seem sharing more similarities than

differences. In this work, we first show that the BM problem has a better estimation accuracy than the OG problem in a general context. In a second part, we provide a closed-form expression of a normalized 2-norm for *i.i.d.* Gaussian amplitude vector and in the low noise variance regime. This closed-form expression allows us to precisely quantify the difference between these two problems and show that these two problems cannot be confused even for small mismatch error variance and for a finite number of measurements and support cardinality. Finally, we also show that these two problems turn to be identical in the Bayesian estimation perspective if the number of measurements and support cardinality grow to infinity with a constant ratio.

REFERENCES

- [1] D. L. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [2] E. Candes and M. Wakin, "An introduction to compressive sampling," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 21–30, March 2008.
- [3] Y. C. Pati, R. Rezaeiifar, Y. C. Pati, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," pp. 40–44, 1993.
- [4] S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.
- [5] H. Zhu, G. Leus, and G. Giannakis, "Sparsity-cognizant total least-squares for perturbed compressive sampling," *Signal Processing, IEEE Transactions on*, vol. 59, no. 5, pp. 2002–2016, 2011.
- [6] M. Herman and T. Strohmer, "General deviants: An analysis of perturbations in compressed sensing," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 2, pp. 342–349, 2010.
- [7] S. Bernhardt, R. Boyer, B. Zhang, S. Marcos, and P. Larzabal, "Performance analysis for sparse based biased estimator: Application to line spectra analysis," in *Sensor Array and Multichannel Signal Processing Workshop (SAM), invited paper, 2014 IEEE 8th*. La Coruna, Spain: IEEE, 2014, pp. 365–368.
- [8] Y. Chi, L. Scharf, A. Pezeshki, and A. Calderbank, "Sensitivity to basis mismatch in compressed sensing," *Signal Processing, IEEE Transactions on*, vol. 59, no. 5, pp. 2182–2195, 2011.
- [9] Z. Ben-Haim and Y. Eldar, "The cramer-rao bound for estimating a sparse parameter vector," *Signal Processing, IEEE Transactions on*, vol. 58, no. 6, pp. 3384–3389, June 2010.
- [10] R. Niazadeh, M. Babaie-Zadeh, and C. Jutten, "On the achievability of cramer-rao bound in noisy compressed sensing," *Signal Processing, IEEE Transactions on*, vol. 60, no. 1, pp. 518–526, 2012.
- [11] B. Babadi, N. Kalouptsidis, and V. Tarokh, "Asymptotic achievability of the cramer-rao bound for noisy compressive sampling," *Signal Processing, IEEE Transactions on*, vol. 57, no. 3, pp. 1233–1236, 2009.
- [12] E. L. Lehmann and G. Casella, *Theory of point estimation*. Springer Science & Business Media, 1998, vol. 31.
- [13] P. Stoica and R. L. Moses, *Spectral analysis of signals*. Pearson/Prentice Hall Upper Saddle River, NJ, 2005.
- [14] A. Wiesel, Y. Eldar, and A. Yeredor, "Linear regression with gaussian model uncertainty: Algorithms and bounds," *Signal Processing, IEEE Transactions on*, vol. 56, no. 6, pp. 2194–2205, June 2008.
- [15] R. Couillet and M. Debbah, *Random matrix methods for wireless communications*. Cambridge University Press, 2011.
- [16] W. Hachem, P. Loubaton, X. Mestre, J. Najim, and P. Vallet, "Large information plus noise random matrix models and consistent subspace estimation in large sensor networks," *Random Matrices: Theory and Applications*, vol. 1, no. 02, 2012.
- [17] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Commun. Inf. Theory*, vol. 1, no. 1, pp. 1–182, Jun. 2004.