# Performance Analysis for Sparse Biased Estimator: Application to Line Spectra Analysis

Stéphanie Bernhardt, Remy Boyer, Bo Zhang, Sylvie Marcos, Pascal Larzabal

# Performance Analysis for Sparse Based Biased Estimator: Application to Line Spectra Analysis

Stéphanie Bernhardt, Rémy Boyer, Bo Zhang and Sylvie Marcos
LSS
Université Paris XI (UPS), CNRS
Gif-Sur-Yvette, France
{stephanie.bernhardt, remy.boyer, sylvie.marcos}@lss.supelec.fr

Pascal Larzabal
SATIE
ENS Cachan
Cachan, France
pascal.larzabal@satie.ens-cachan.fr

***Abstract*—Dictionary based sparse estimators are based on the matching of continuous parameters of interest to a discretized sampling grid. Generally, the parameters of interest do not lie on this grid and there exists an estimator bias even at high Signal to Noise Ratio (SNR). This is the off-grid problem. In this work, we propose and study analytical expressions of the Bayesian Mean Square Error (BMSE) of dictionary based biased estimators at high SNR. We also show that this class of estimators is efficient and thus reaches the Bayesian Cramér-Rao Bound (BCRB) at high SNR. The proposed results are illustrated in the context of line spectra analysis and several popular sparse estimators are compared to our closed-form expressions of the BMSE.**

## I. Introduction

Sparse and redundant signal representations have recently drawn much interest in many different applications[1] as for instance source localization, line spectra analysis, range/doppler estimation in radar signal processing etc. The main assumption is that signals of interest can be sparse or compressible in an overcomplete dictionary [1]. An important problem related to this subject is the off-grid problem [2], [3]. Indeed, the parameters of interest are continuous and the dictionary is based on a regularly spaced grid partitioning the parameter set. Thus, there exists an estimation bias (or quantization error) even at high Signal to Noise ratio (SNR) resulting from the fact that the parameters of interest generally do not lie on the grid. As a consequence, the dictionary based estimator is biased even at high SNR.

To decrease the estimation bias, a natural solution is to sample the parameter space more finely at the cost of an increasing computational complexity and a highly coherent dictionary. Several algorithms have been proposed to resolve this problem [3], [4] but to the best of our knowledge, no study gives an analytical expression of the Bayesian Mean Square Error (BMSE) of dictionary based biased estimator at high SNR. Thus, in this work we derive analytical expressions of the BMSE for a regularly spaced dictionary at high SNR for random parameters of interest. We also propose closed form expressions of the Bayesian Cramér-Rao Bound (BCRB) [5] for the considered biased estimator. Finally, we apply our results in the context of the line spectra analysis and compare our bounds with the BMSE of several popular estimators.

## II. Problem formulation

### A. Sparse representation

Consider a known model order $M$, we are interested in the problem of estimating an unknown vector of $M$ parameters $\boldsymbol{\omega} = [\omega_1, \omega_2, \cdots, \omega_M]^T$ given a set of noisy measurements $\{y_t\}_{t=1}^T$. In the dictionary based model estimation approach, the parameter space noted $\Omega$ is discretized into $N$ parameter samples: $\bar{\bar{\omega}} = [\bar{\omega}_1, \bar{\omega}_2, \cdots, \bar{\omega}_N]^T$. The measurement signal can be represented as a linear decomposition in a dictionary [6]:

$$\mathbf{y} = [y_1, \cdots, y_T]^T = \sum_{n=1}^{N} \mathbf{a}(\bar{\omega}_n)\alpha_n + \mathbf{e} \tag{1}$$

where $\mathbf{e}$ is a white Gaussian noise vector of length $T$. $\mathbf{a}(\bar{\omega}_n)$ is a vector representing the parametric function evaluated at $T$ measurement locations [7]. The collection of $\mathbf{a}(\bar{\omega}_n)$ composes the columns of the overcomplete dictionary $\mathbf{A}$ which is a $T \times N$ matrix, with $N >> T$. We consider a regularly sampled dictionary and note $\forall n, r = \bar{\omega}_{n+1} - \bar{\omega}_n$ the grid spacing, the size of the dictionary $N$ is directly related to $r$ by

$$r = \frac{\bar{\omega}_N - \bar{\omega}_1}{N-1}. \tag{2}$$

If all the actual parameters $\omega_m \in \bar{\bar{\boldsymbol{\omega}}}$, then $\alpha_n$ are zeros for all index $n$ except for the index $n_m$ corresponding to $\bar{\omega}_{n_m} = \omega_m$. The estimation problem is then to find the coefficients of the decomposition of the measurements in the known dictionary $\mathbf{A}$. The aim is to estimate the index of the non-zero coefficients, which will correspond to $\omega_m$.

We divide the parameter set $\Omega$ in decision intervals centered at the elements of the dictionary $\bar{\bar{\boldsymbol{\omega}}} = [\bar{\omega}_1, \cdots, \bar{\omega}_N]^T$. We note $\Omega = \Omega_1 \cup_{n=2}^{N-1} \Omega_n \cup \Omega_N$ with $\Omega_1 = \left[\bar{\omega}_1, \bar{\omega}_1 + \frac{r}{2}\right]$, $\Omega_n = \left[\bar{\omega}_n - \frac{r}{2}, \bar{\omega}_n + \frac{r}{2}\right]$ and $\Omega_N = \left[\bar{\omega}_N - \frac{r}{2}, \bar{\omega}_N\right]$.

### B. Dictionary based biased estimator in the high SNR regime

In this work, we assume that the RIP (Resctricted Isometry Property) for sparse estimation is guaranteed [8], this means that the recovery process is stable in the presence of additive noise. We further consider the high SNR regime, in which a relevant assumption to our problem is the following:

$$\forall m, \exists n_m, \forall \omega_m \in \Omega_{n_m}, \quad \hat{\omega}_m(\mathbf{y}) = \bar{\omega}_{n_m} \tag{3}$$

i.e. in high SNR regime, the noise level is low enough so that the estimation is given by the closest value in the dictionary.

Note that in the high SNR regime, $\hat{\omega}_m(\mathbf{y})$ is no longer a random variable since the dependance on the observation $\mathbf{y}$ vanishes and the choice of $\hat{\omega}_m(\mathbf{y})$ is completely driven by the knowledge of the decision interval $\Omega_{n_m}$. Let $\mathbf{b}(\boldsymbol{\omega}) = [b(\omega_1), \cdots, b(\omega_M)]^T$ be the bias vector in which $b(\omega_m) = E(\hat{\omega}_m(\mathbf{y})) - \omega_m = \bar{\omega}_{n_m} - \omega_m$ is the estimation error at high SNR.

### C. Bayesian assumptions on the parameters of interest

*1) Modelization on the dictionary:* In the Bayesian framework, the vector parameter $\boldsymbol{\omega}$ is a vector of random variables equipped with two *a priori* informations: $\boldsymbol{\omega}$ *i)* follows a known prior pdf $p(\boldsymbol{\omega})$ and *ii)* belongs to the parameter product set $\Omega^M = \Omega \times \ldots \times \Omega$. The conditional pdf $p(\boldsymbol{\omega}|\boldsymbol{\omega} \in \Omega^M)$ can be rewritten as a truncated joint pdf according to

$$p(\boldsymbol{\omega}|\boldsymbol{\omega} \in \Omega^M) = \frac{p(\boldsymbol{\omega}) 1_{\Omega^M}(\boldsymbol{\omega})}{\Pr(\boldsymbol{\omega} \in \Omega^M)} \tag{4}$$

[1]This work was supported by the European Commission in the framework of the FP7 Network of Excellence inWireless COMmunications NEWCOM# (Grant agreement no. 318306), PRES UnivSud Paris and the CNRS display/Mastodons project.

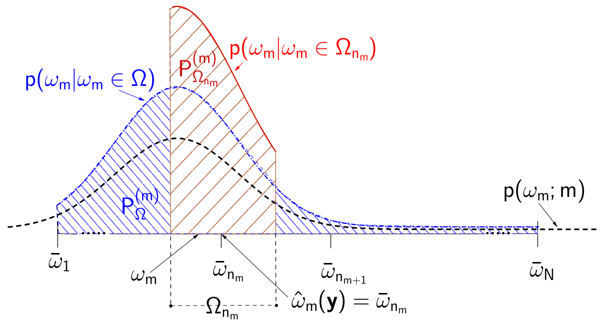


Fig. 1. Truncated pdfs and dictionary based biased estiamator

where $1_{\Omega^M}(\boldsymbol{\omega})$ is the set indicator function and $\Pr(\boldsymbol{\omega} \in \Omega^M) = \int_{\Omega^M} p(\boldsymbol{\omega})d\boldsymbol{\omega}$.

As the parameters $\boldsymbol{\omega}$ are assumed to be independent random variables, we have $\Pr(\boldsymbol{\omega} \in \Omega^M) = \prod_{m=1}^M P_\Omega^{(m)}$ where $P_\Omega^{(m)} = \int_\Omega p(\omega_m; m)d\omega_m$ with $p(\omega_m; m)$ the prior pdf of $\omega_m$ whose parameters can depend on $m$ (see the black dot line on Fig. 1). Using the above expression and the fact that the indicator set function is the product of the indicator functions $1_\Omega(\omega_m)$, we have

$$p(\boldsymbol{\omega}|\boldsymbol{\omega} \in \Omega^M) = \prod_{m=1}^M p(\omega_m|\omega_m \in \Omega) \tag{5}$$

where the distribution of each parameter, knowing the parameter space is the truncated version over the parameter space of $p(\omega_m)$:

$$p(\omega_m|\omega_m \in \Omega) = \frac{p(\omega_m; m)1_\Omega(\omega_m)}{P_\Omega^{(m)}} \tag{6}$$

where $1_\Omega(\omega_m) = 1$ if $\omega_m \in \Omega$ and 0 otherwise. This pdf and its associated normalization factor $P_\Omega^{(m)}$ have been drawn on Fig. 1 with a blue dot line and a blue hatched area, respectively.

*2) Modelization on each decision interval:* Using the law of total probabilities, the pdf of the parameters of interest can be decomposed according to

$$p\left(\omega_m|\omega_m \in \Omega\right) = \sum_{n_m=1}^N P_{\Omega_{n_m}}^{(m)} p\left(\omega_m|\omega_m \in \Omega_{n_m}\right) \tag{7}$$

where $P_{\Omega_{n_m}}^{(m)} = \int_{\Omega_{n_m}} p(\omega_m|\omega_m \in \Omega_{n_m})d\omega_m$ is the probability to be in the decision interval $\Omega_{n_m}$. This pdf and its associated normalization factor $P_{\Omega_{n_m}}^{(m)}$ have been drawn on Fig. 1 with a red line and a red hatched area, respectively.

## III. Analysis of the BMSE and the BCRB

### A. Definition of the BMSE

Conditional to the knowledge of the joint pdf $p(\boldsymbol{\omega})$ and the parameter set $\Omega$, the BMSE is defined by:

$$\mathrm{BMSE}_\Omega = \int_{\mathbb{R}^M} E_\mathbf{y}\left(||\hat{\boldsymbol{\omega}}(\mathbf{y}) - \boldsymbol{\omega}||^2\right) p(\boldsymbol{\omega}|\boldsymbol{\omega} \in \Omega^M)d\boldsymbol{\omega} \tag{8}$$

where $\mathbb{R}^M = \mathbb{R} \times \ldots \times \mathbb{R}$ and $p(\boldsymbol{\omega}|\boldsymbol{\omega} \in \Omega^M)$ is given in (4).

### B. Analytic expressions of the BMSE at high SNR

**Result 1.** *Assume high SNR and independent parameters of interest with known pdf $p(\omega_m)$, the BMSE conditionally to the set $\Omega$ is given by*

$$\mathrm{BMSE}_{\Omega,\mathrm{high}} = \sum_{m=1}^M \sum_{n_m=1}^N P_{\Omega_{n_m}}^{(m)} E_{\omega_m|\omega_m \in \Omega_{n_m}}\left(b(\omega_m)^2\right). \tag{9}$$

*Proof:* Including expression (5) in definition (8) and the law of total probabilities given in (7), we can conclude:

$$\begin{aligned}\mathrm{BMSE}_\Omega &= \sum_{m=1}^M \sum_{n_m=1}^N P_{\Omega_{n_m}}^{(m)} \\ &\quad \int_\mathbb{R} E_\mathbf{y}\left((\hat{\omega}_m(\mathbf{y}) - \omega_m)^2\right) p(\omega_m|\omega_m \in \Omega_{n_m})d\omega_m \\ &\xrightarrow{\text{high SNR}} \sum_{m=1}^M \sum_{n_m=1}^N P_{\Omega_{n_m}}^{(m)} \underbrace{\int_\mathbb{R} b(\omega_m)^2 p(\omega_m|\omega_m \in \Omega_{n_m})d\omega_m}_{E_{\omega_m|\omega_m \in \Omega_{n_m}}(b(\omega_m)^2)}.\end{aligned}$$

In the last expression, we use two properties: *i)* $P_{\Omega_{n_m}}^{(m)}$ is not a function of the parameter $\omega_m$ and can be taken out the integral form and *ii)* the dictionary estimator at high SNR which is given in (3). This property allows us to absorb into the law of total probabilities the error estimation $\hat{\omega}_m - \omega_m$ becoming $\bar{\omega}_{n_m} - \omega_m$ on each interval $\Omega_{n_m}$ and to remove the mathematical expectation on the noise since $\bar{\omega}_{n_m}$ is a deterministic parameter. ■

An equivalent expression of the $\mathrm{BMSE}_{\Omega,\mathrm{high}}$ is given in the following result.

**Result 2.** *Assume high SNR and independent parameters of interest with known pdf $p(\omega_m)$, the BMSE conditionally to the set $\Omega$ is given by*

$$\mathrm{BMSE}_{\Omega,\mathrm{high}} = \sum_{m=1}^M \frac{1}{P_\Omega^{(m)}} \sum_{n_m=1}^N E_{\omega_m}\left(b(\omega_m)^2 1_{\Omega_{n_m}}(\omega_m)\right). \tag{10}$$

*Proof:* Using the definition of the BMSE, expression (6) and the fact that $P_\Omega^{(m)}$ is not a function of the integration parameter $\omega_m$, the $\mathrm{BMSE}_\Omega$ is given by

$$\begin{aligned}\mathrm{BMSE}_\Omega &= \sum_{m=1}^M \frac{1}{P_\Omega^{(m)}} \int_\Omega E_\mathbf{y}\left((\hat{\omega}_m(\mathbf{y}) - \omega_m)^2\right) p(\omega_m; m)d\omega_m \\ &\xrightarrow{\text{high SNR}} \sum_{m=1}^M \frac{1}{P_\Omega^{(m)}} \sum_{n_m=1}^N \underbrace{\int_{\Omega_{n_m}} b(\omega_m)^2 p(\omega_m; m)d\omega_m}_{E_{\omega_m}\left(b(\omega_m)^2 1_{\Omega_{n_m}(\omega_m)}\right)}.\end{aligned}$$

In the last expression, we decompose the integral form defined on the set $\Omega$ into the sum of integrals on the decision intervals $\Omega_{n_m}$, $n_m = 1, \ldots, N$ where the dictionary estimator at high SNR is defined according to expression (3). ■

The BMSE increases proportionally with the number of parameters to estimate, since each one contributes to the error. The parameter $r$ determines the size of $\Omega_{n_m}$ over which the integration is done. When $r$ gets smaller, the BMSE decreases. The influence of $r$ is directly visible when we derive the BMSE for a specific distribution, as for example the uniform distribution.

### C. BMSE for typical priors

*1) Case of the uniform prior:* A natural distribution for $\omega_m$ is the uniform distribution. We consider that each parameter follows a uniform distribution truncated at $\Omega = [\bar{\omega}_1, \bar{\omega}_N]$.

**Result 3.** *The* $\mathrm{BMSE}_{\Omega,\mathrm{high}}$ *for an uniform prior is equal to* $\frac{Mr^2}{12}$.

*Proof:* For $2 \leq n_m \leq N - 1$ we have

$$E_{\omega_m}\left(b(\omega_m)^2 1_{\Omega_{n_m}(\omega_m)}\right) = \int_{\Omega_{n_m}} b(\omega_m)^2 d\omega_m = \frac{r^3}{12}.$$

The border intervals $\Omega_1$ and $\Omega_N$ have a length of $r/2$, therefore $E_{\omega_m}\left(b(\omega_m)^2 1_{\Omega_1(\omega_m)}\right) = E_{\omega_m}\left(b(\omega_m)^2 1_{\Omega_N(\omega_m)}\right) = \frac{r^3}{24}$. Introducing those results in the BMSE formula and secondly introducing (2) gives :

$$\mathrm{BMSE}_{\Omega,\mathrm{high}} = \frac{N-1}{(\bar{\omega}_N - \bar{\omega}_1)} \cdot \frac{r^3}{12} = \frac{Mr^2}{12} \tag{11}$$

■

The dictionary based estimator at high SNR has a similar result as quantization, since it takes the closest value in a discrete regular dictionary of a given continuous parameter.

*2) Case of Gaussian prior:* A Gaussian pdf can modelize some *a priori* knowledge on $\omega_m$. In this scenario, we have the following result.

**Result 4.** *The* $\mathrm{BMSE}_{\Omega,\mathrm{high}}$ *for a prior* $\omega_m \sim \mathcal{N}(\tilde{\omega}_m, \sigma_m^2)$ *is given by:*

$$\mathrm{BMSE}_{\Omega,\mathrm{high}} = \sum_{m=1}^{M} \frac{1}{\operatorname{erf}\left(\frac{\bar{\omega}_N - \tilde{\omega}_m}{\sqrt{2}\sigma_m}\right) - \operatorname{erf}\left(\frac{\bar{\omega}_1 - \tilde{\omega}_m}{\sqrt{2}\sigma_m}\right)} \sum_{n_m=1}^{N} A_{n_m}$$

*where* $\operatorname{erf}(.)$ *is the error function and*

$$\begin{aligned} A_{n_m} = &\left((\tilde{\omega}_m - \bar{\omega}_{n_m})^2 + \sigma_m^2\right)\left(\operatorname{erf}\left(\frac{a + 2(\tilde{\omega}_m - \bar{\omega}_{n_m})}{2\sqrt{2}\sigma_m}\right)\right. \\ &\left. - \operatorname{erf}\left(\frac{b + 2(\tilde{\omega}_m - \bar{\omega}_{n_m})}{2\sqrt{2}\sigma_m}\right)\right) \\ &+ \frac{\sigma_m}{\sqrt{2\pi}}\left(\left(2(\tilde{\omega}_m - \bar{\omega}_{n_m}) - a\right) e^{-\frac{1}{2}\left(\frac{a+2(\tilde{\omega}_m - \bar{\omega}_{n_m})}{2\sigma_m}\right)^2}\right. \\ &\left. - \left(2(\tilde{\omega}_m - \bar{\omega}_{n_m}) - b\right) e^{-\frac{1}{2}\left(\frac{b+2(\tilde{\omega}_m - \bar{\omega}_{n_m})}{2\sigma_m}\right)^2}\right) \end{aligned}$$

*where we have* $a = r, b = -r$ *for* $2 \leq n_m \leq N-1$*;* $a = 0, b = -r$ *for* $n_m = 1$ *and* $a = r, b = 0$ *for* $n_m = N$*.*

*Proof:* We use the change in the variable $b(\omega_m) = \bar{\omega}_{n_m} - \omega_m$ and the formula $\int x^2 \phi(a+bx)dx = b^{-3}\left[\left(a^2+1\right)\Phi(a+bx) + (a-bx)\phi(a+bx)\right] + C$ to derive the mean of the mean square error over each interval for the gaussian distribution. $\phi(x)$ denotes the standard normal distribution and $\Phi(x)$ its repartition function. The denominator in (4) comes from $P_\Omega^{(m)}$ which is the integration of the gaussian distribution over the parameter space $\Omega$. ■

*D. Derivation of the Bayesian Cramér-Rao Bound for biased estimators*

**Result 5.** *The dictionary based biased estimator is statistically efficient in the high SNR regime.*

*Proof:* The Bayesian Cramér Rao Bound (BCRB) for a biased estimator is the mean of the biased version of the biased deterministic CRB as given by [5]

$$\mathrm{BCRB}(\mathbf{b}(\boldsymbol{\omega})) = \int_{\mathbb{R}^M} \mathrm{CRB}(\boldsymbol{\omega}, \mathbf{b}(\boldsymbol{\omega})) p(\boldsymbol{\omega}|\boldsymbol{\omega} \in \Omega^M) d\boldsymbol{\omega}$$

where

$$\mathrm{CRB}(\boldsymbol{\omega}, \mathbf{b}(\boldsymbol{\omega})) \triangleq \mathrm{Tr}\left[\left(\mathbf{I} + \mathbf{D}(\boldsymbol{\omega})\right)\mathbf{J}(\boldsymbol{\omega})^{-1}\left(\mathbf{I} + \mathbf{D}(\boldsymbol{\omega})\right)^T\right] + \|\mathbf{b}(\boldsymbol{\omega})\|^2$$

with $\mathbf{D}(\boldsymbol{\omega}) = \frac{\partial \mathbf{b}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}}$ is the bias gradient matrix and $\mathbf{J}(\boldsymbol{\omega})$ is the Fisher information matrix which is defined, given the measurements $\mathbf{y}$ by $[\mathbf{J}(\boldsymbol{\omega})]_{i,j} = \mathrm{E}_{\mathbf{y}}\left(\frac{\partial \log p(\mathbf{y}|\boldsymbol{\omega})}{\partial \boldsymbol{\omega}_i}\frac{\partial \log p(\mathbf{y}|\boldsymbol{\omega})}{\partial \boldsymbol{\omega}_j}\right)$ where $p(\mathbf{y}|\boldsymbol{\omega})$ is the conditional pdf of the measurement vector.

The BCRB characterizes the smallest achievable variance of any estimator of bias $\mathbf{b}(\omega)$. For the problem of dictionary based estimation, we have $\mathbf{D}(\boldsymbol{\omega}) = -\mathbf{I}$. Therefore $\left(\mathbf{I} + \frac{\partial \mathbf{b}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}}\right) = 0$ and

$$\mathrm{BCRB} = \int_{\mathbb{R}^M} \|\mathbf{b}(\boldsymbol{\omega})\|^2 p(\boldsymbol{\omega}|(\boldsymbol{\omega} \in \Omega^M) d\boldsymbol{\omega} = \mathrm{BMSE}_{\Omega,\mathrm{high}}$$

■

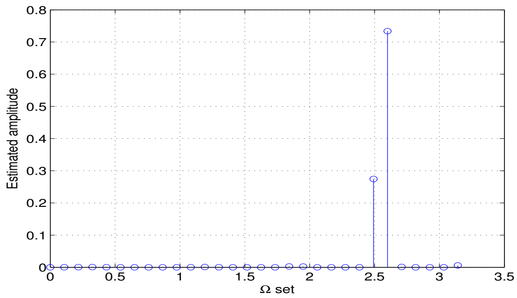


Fig. 2. An example of result of the spgl1 algorithm: the estimated order is 13 here, for amplitude 1, SNR=50dB

## IV. Application to Line Spectra Analysis

*A. Model for dictionary based line spectra analysis*

The parameter to estimate in spectra analysis is the angular frequency $\omega \in \Omega$. Take $T$ time samples, the model of line spectra analysis is [7]:

$$y_t = \sum_{m=1}^{M} \alpha_m e^{i\omega_m(t-1)} + e_t$$

where $t$ is the time sample, $y_t$ is the measured signal, $\alpha_m$ and $\omega_m$ are the amplitude and the angular frequency of the $m$-th spectral line respectively and $e_t$ is the additive noise which is assumed to be white noise.

This model can also be expressed using the sparse representation (1) with a parameter dictionary $\bar{\boldsymbol{\omega}} = [\bar{\omega}_1, \cdots, \bar{\omega}_N]^T$ and $\mathbf{a}(\bar{\omega}_n) = [1, \cdots, e^{i\bar{\omega}_n t}, \cdots, e^{i\bar{\omega}_n(T-1)}]^T$.

*B. Simulations*

We now validate the theoretical BMSE for practical estimators for both gaussian and uniform distributions. The reconstruction algorithm used are Orthogonal Matching Pursuit (OMP) [9], Compressive Sensing Approximate Message Passing (CoSaMP) [10] and basis pursuit [6] using the SPGl1 implementation [11]. OMP and CoSaMP are estimators that need the knowledge of the model order, whereas Basis Pursuit estimate the number of signals. As we know the model order, we consider the $M$ highest peaks of the estimated sparse vector given by the SPGL1 algorithm as the estimated angular frequencies for this algorithm.

Even if we do not consider the model order estimation, it is interesting to note that the SPGl1 algorithm over-estimates the model order: there are several small components that are non-zeros in addition to the actual frequency as shown Fig.2. This results in an under-estimation of the amplitude of the real component. Even at high SNR, the model order is wrongly estimated: Fig.3 plots the average model order over all the iterations as a function of the signal to noise ratio in the case of on and off grid frequencies. The mean order decreases as the SNR increases but reaches a minimum even for the on-grid case: the spg finds the good frequency, and some others with low amplitude. The average order estimate decreases, however we have observed important variations in the model order estimation at a given SNR. Those errors are due to the coherent dictionary which breaks the hypothesis behind BP reconstruction and in the off-grid case, additional error is due to the basis mismatch.

Fig.4 plots the BMSE of the different estimators for $T = 20$ time samples, a dictionary going from 0 to $\pi$ of size $N = 60$ and an uniform distribution. The error decreases as the noise decreases, to finally reach the theoretical BMSE given in (11). The high SNR region is reached for SNR above 30dB. The theoretical BMSE is also validated for a gaussian distribution (see Fig.5) with the same parameter and a normal distribution $p(\omega_m) \sim \mathcal{N}(\pi/3, r^2)$. If a coarser dictionary is used, the BMSE will be higher since $r$ is larger.

The RIP conditions are difficult to prove for an arbitrary matrice, therefore a common measure used to asses the validity of the

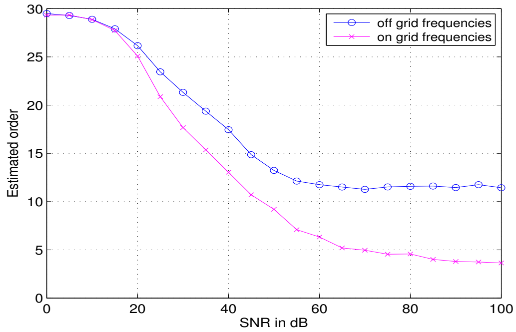


Fig. 3. Mean model order estimate of the BP algorithm as a function of the SNR in dB, $M = 1$

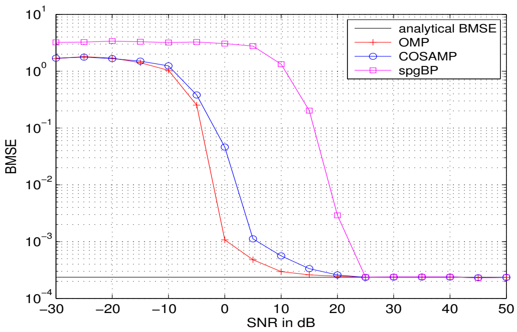


Fig. 4. BMSE in dB for the uniform distribution depending on the SNR in dB, $M = 1$, $T = 20$ and $N = 60$

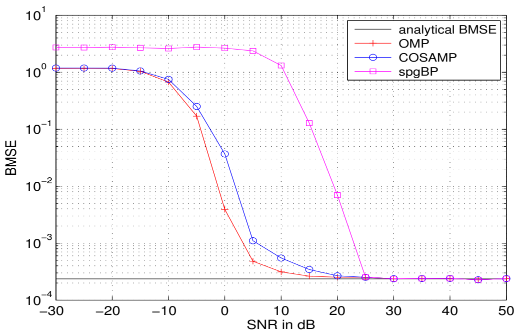


Fig. 5. BMSE for the gaussian distribution depending on the SNR, $M = 1$, $T = 20$ and $N = 60$

compressed sensing estimation is the mutual coherence which is defined as the maximum absolute value of the cross-correlations between the vectors $\mathbf{a}(\bar{\omega}_n)$. The number of spectral lines that can be recovered depends on the coherence $\mu$ [12], it is bounded by $M < \frac{1}{2}\left(1 + \frac{1}{\mu}\right)$. Applying this formula to the above parameters, we get that the maximum number of spectral lines is $M_{\max} = 1$.

Therefore, we use a larger dictionary running from $-\pi$ to $\pi$ and parameters $T = 90$ and $N = 100$ to measure the BMSE with two spectral lines which allows $M_{\max} = 5$. Fig.6 shows the result in the case of two active frequencies, as $M$ increases the BMSE also increases. When the number of sources increases the BMSE and the SNR needed to reach the high SNR region increases. The number of samples $T$ plays also an important role not only on the mutual coherence, but also on the level of noise required to detect a given number of spectral lines.

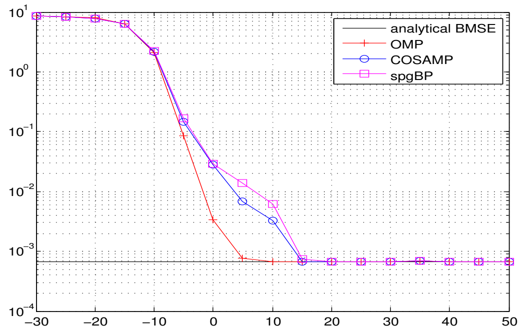


Fig. 6. BMSE for the uniform distribution depending on the SNR, $M = 2$, $T = 90$ and $N = 100$, $\omega \in [-\pi, \pi]$

## V. Conclusion

In this work, we derived and studied the BMSE conditionally to the parameter set for dictionary based biased estimator at high SNR. Two alternative analytical expressions are proposed which we hope will give more perspective on the off-grid problem. As, the parameters of interest are assumed to be random, closed-form expressions of the BMSE are provided for independent uniform and Gaussian pdf. We also show that the dictionary based biased estimator is statistically efficient in the sense that the BMSE reaches the corresponding BCRB. Finally, by means of numerical simulations, we show that our analytical expressions of the BMSE well explain the behavior of several popular sparse estimators. In future work, we will consider the case of non regular and adaptive dictionaries.

## References

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

[2] 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.

[3] 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.

[4] A. Panahi and M. Viberg, "A robust l1 penalized doa estimator," in *Signals, Systems and Computers (ASILOMAR), 2012 Conference Record of the Forty Sixth Asilomar Conference on*, 2012, pp. 2013–2017.

[5] Z. Ben-Haim and Y. Eldar, "A lower bound on the bayesian mse based on the optimal bias function," *Information Theory, IEEE Transactions on*, vol. 55, no. 11, pp. 5179–5196, 2009.

[6] 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. [Online]. Available: http://epubs.siam.org/doi/abs/10.1137/S1064827596304010

[7] C. Austin, J. Ash, and R. Moses, "Dynamic dictionary algorithms for model order and parameter estimation," *Signal Processing, IEEE Transactions on*, vol. 61, no. 20, pp. 5117–5130, 2013.

[8] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *Information Theory, IEEE Transactions on*, vol. 52, no. 2, pp. 489–509, 2006.

[9] Y. C. Pati, R. Rezaiifar, Y. C. P. R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proceedings of the 27 th Annual Asilomar Conference on Signals, Systems, and Computers*, 1993, pp. 40–44.

[10] D. Needell and R. Vershynin, "Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 2, pp. 310–316, 2010.

[11] E. van den Berg and M. P. Friedlander, "SPGL1: A solver for large-scale sparse reconstruction," June 2007, http://www.cs.ubc.ca/labs/scl/spgl1.

[12] M. Duarte and Y. Eldar, "Structured compressed sensing: From theory to applications," *Signal Processing, IEEE Transactions on*, vol. 59, no. 9, pp. 4053–4085, 2011.