Bayesian Sparse Fourier Representation of Off-Grid Targets

Marie Lasserre, Stéphanie Bidon and Olivier Besson
DEOS/ISAE
University of Toulouse
Toulouse, France
Email: firstname.lastname@isae.fr

François Le Chevalier
MS^3
Delft University of Technology
Delft, The Netherlands
Email: F.LeChevalier@tudelft.nl

***Abstract*—We consider the problem of estimating a finite sum of cisoids via the use of a sparsifying Fourier dictionary (problem that may be of use in many radar applications). Numerous signal sparse representation (SSR) techniques can be found in the literature regarding this problem. However, they are usually very sensitive to grid mismatch. In this paper, we present a new Bayesian model robust towards grid mismatch. Synthetic and experimental radar data are used to assess the ability of the proposed approach to robustify the SSR towards grid mismatch.**

I. Introduction

In many radar applications, the received signal is conventionally described by a linear model where the signal of interest, namely the target signal, is represented by a sum of cisoids embedded in additive noise, i.e.,

$$\boldsymbol{y} = \sum_{n=1}^{N} \alpha_n \boldsymbol{a}_n + \boldsymbol{n} \quad \text{with} \quad [\boldsymbol{a}_n]_m = \exp\{j2\pi f_n m\} \tag{1}$$

where

- $\boldsymbol{y} \in \mathbb{C}^M$ is the observation vector and M is the size of the observation space;
- $\alpha_n, \boldsymbol{a}_n$ are respectively the complex amplitude and the steering vector with frequency f_n of the nth target signal;
- $\boldsymbol{n}$ is the noise vector.

Several approaches can be used to estimate the target scene $\{(\alpha_n, f_n)\}$ related to the measurement $\boldsymbol{y}$. These approaches are usually distinguished into two classes according to whether the method assumes, or not, a specific model about the noise covariance matrix. In the latter case the technique is said to be non-parametric (e.g., Fourier transform, Capon's method [1], APES [2]) while in the former case it is called parametric (e.g., subspace method [3], [4], autoregressive model [5]). Over the last few decades, a new estimation paradigm called sparse signal reconstruction (SSR) has emerged. It has been applied to many signal processing applications and, in particular, to the target estimation problem (1), e.g., [6]. SSR aims at describing the signal as a linear combination of a few atoms from a (possibly pre-defined) dictionary. Towards this end and given the problem at hand (1), a natural sparsifying dictionary is the Fourier basis so that (1) can be reformulated as

$$\boldsymbol{y} = \boldsymbol{F}\boldsymbol{x} + \boldsymbol{n} \tag{2}$$

with

- $\boldsymbol{F} \in \mathbb{C}^{M \times \bar{M}}$ the Fourier dictionary of size $M \times \bar{M}$ where usually in SSR $\bar{M} \geq M$;
- $\boldsymbol{x} \in \mathbb{C}^{\bar{M}}$ the sparse vector having ideally exactly N nonzero components.

The literature describing methods solving such SSR problem is nowadays prominent, e.g., ℓ_1 penalized least squares formulations [7] or Bayesian models [8]. However, most of these techniques are very sensitive to grid mismatch [9], [10]. In the context of (2), grid mismatch occurs whenever the frequency of a target signal does not belong to the frequency grid associated with the sparsifying dictionary $\boldsymbol{F}$, i.e., if $f_n \notin \{0, 1/\bar{M}, \ldots, 1 - 1/\bar{M}\}$. Several attempts have been made at robustifying SSR technique towards grid mismatch. The most natural way to deal with grid mismatch is to refine the grid [11]. However, most of the techniques choose to represent grid mismatch via a perturbation matrix $\boldsymbol{E}$ added to $\boldsymbol{F}$ [12]. $\boldsymbol{E}$ usually stems from a first order Taylor expansion [13]–[15]. Another strategy is to consider that $\boldsymbol{F}$ is parameterized by a frequency grid that is estimated jointly with $\boldsymbol{x}$ [16]. In this paper, we adopt the latter strategy and propose a hierarchical Bayesian model which takes into account the possible grid mismatch in a Fourier dictionary. The model assumes a white noise background and is an extension (aside from the dictionary) of a non-robust SSR method [17]. The advantage of the proposed formulation over that of [13], [14], [16] is that it enforces more sparsity via the use of a hierarchical Bernoulli-complex Gaussian prior on $\boldsymbol{x}$. But mostly it is specifically designed to be used, in future work, in other estimation schemes recently developed by the authors.

The remaining of the paper is organized as follows. The proposed Bayesian model and its associated estimation scheme are described respectively in Section II and III. Numerical simulations are conducted in Section IV to assess the benefit of the robustification on both synthetic and experimental radar data.

The work of M. Lasserre, S. Bidon and O. Besson is supported by the Délégation Générale de l'Armement under grant 2012.60.0012.00.470.75.01.

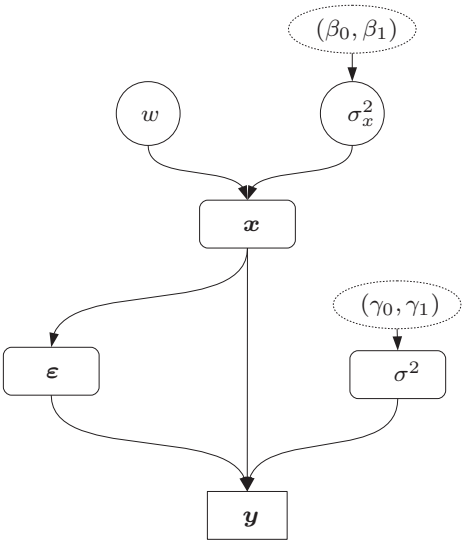


Fig. 1. Directed acyclic graph associated with the hierarchical Bayesian model. The dashed lines indicate that (β_0, β_1) and (γ_0, γ_1) must be set by the operator.

II. Bayesian model

The hierarchical Bayesian model considered in this paper is represented graphically in Fig. 1 and detailed herein. Compared to that of [17], the technical novelty resides into the modeling of the grid mismatch.

A. Observation model

1) Modeling of grid mismatch: We consider the observation model $\boldsymbol{y} = \boldsymbol{F}\boldsymbol{x} + \boldsymbol{n}$ described in (2) with $\bar{M}$ fixed. In absence of grid mismatch, an appropriate choice for the Fourier dictionary $\boldsymbol{F}$ yields

$$[\boldsymbol{f}_{\bar{m}}]_m = 1/\sqrt{M}\exp\{j2\pi m(\bar{m}/\bar{M})\}$$

where $\boldsymbol{f}_{\bar{m}}$ is the $\bar{m}$th column of $\boldsymbol{F}$. We propose to model the possible grid error by introducing a perturbation vector on the frequency axis denoted as

$$\boldsymbol{\varepsilon} = [\varepsilon_0 \quad \dots \quad \varepsilon_{\bar{m}} \quad \dots \quad \varepsilon_{\bar{M}-1}]^T$$

so that the Fourier dictionary is parameterized by $\boldsymbol{\varepsilon}$ as follows

$$\boldsymbol{F} \triangleq \boldsymbol{F}(\boldsymbol{\varepsilon}) = [\boldsymbol{f}_0(\varepsilon_0) \quad \dots \quad \boldsymbol{f}_{\bar{m}}(\varepsilon_{\bar{m}}) \quad \dots \quad \boldsymbol{f}_{\bar{M}-1}(\varepsilon_{\bar{M}-1})]$$

where the $\bar{m}$th column of $\boldsymbol{F}(\boldsymbol{\varepsilon})$ is now expressed as

$$\boxed{[\boldsymbol{f}_{\bar{m}}(\varepsilon_{\bar{m}})]_m = 1/\sqrt{M}\exp\{j2\pi m(\bar{m}+\varepsilon_{\bar{m}})/\bar{M}\}}.$$

To avoid overlapping between the frequency bins of $\boldsymbol{F}$, the amplitude of the grid errors are assumed bounded such that for $\bar{m} = 0, \dots, \bar{M}-1$, $\varepsilon_{\bar{m}} \in [-0.5, 0.5)$.

2) Likelihood: As stated in the introduction, an additive white noise background is considered. More precisely $\boldsymbol{n}$ is assumed to be centered Gaussian with power σ^2, which is denoted as

$$\boldsymbol{n}|\sigma^2 \sim \mathcal{CN}_M\left(\boldsymbol{0}, \sigma^2\boldsymbol{I}\right) \tag{3}$$

where $\boldsymbol{I}$ is the identity matrix. The likelihood function is thus given by

$$\boxed{f(\boldsymbol{y}|\boldsymbol{\varepsilon}, \boldsymbol{x}, \sigma^2) = \frac{1}{\pi^M\sigma^{2M}}\exp\left\{-\frac{\|\boldsymbol{y}-\boldsymbol{F}(\boldsymbol{\varepsilon})\boldsymbol{x}\|_2^2}{\sigma^2}\right\}}. \tag{4}$$

In (4), the vectors of interest are $\boldsymbol{x}$ and $\boldsymbol{\varepsilon}$ while σ^2 is a nuisance parameter. To estimate the target scene $\boldsymbol{x}, \boldsymbol{\varepsilon}$ a Bayesian framework is chosen where each unknown parameter is modeled by a random variable with a given prior probability density function (pdf). In what follows, each prior density is designed to find a convenient balance between 1) mathematical tractability when performing the estimation 2) guarantee to preserve some physical sense to the hierarchical model.

B. Prior pdfs of the parameters

1) Target amplitude vector: Ideally in SSR the vector $\boldsymbol{x}$ introduced in (2) has exactly N nonzero elements whose value represents the post-integration amplitude of the target signal. As in [17] a Bernoulli-complex Gaussian prior is chosen to actually enforce sparsity in $\boldsymbol{x}$. More precisely, the elements $x_{\bar{m}} \triangleq [\boldsymbol{x}]_{\bar{m}}$ of the amplitude vector are assumed independent and identically distributed (iid) according to the following mixed type pdf

$$f(x_{\bar{m}}|w, \sigma_x^2) = (1-w)\delta(|x_{\bar{m}}|) + w\frac{1}{\pi\sigma_x^2}\exp\left\{-\frac{|x_{\bar{m}}|^2}{\sigma_x^2}\right\}. \tag{5}$$

Using the prior (5), denoted as $x_{\bar{m}}|w, \sigma_x^2 \sim \mathcal{B}er\mathcal{CN}\left(w, 0, \sigma_x^2\right)$, leads to considering that a target with power σ_x^2 is present at the $\bar{m}$th frequency with probability w.

2) Grid errors: In this paper, we propose to define the prior pdf of the grid error $\varepsilon_{\bar{m}}$ conditionally to the magnitude of $x_{\bar{m}}$. The idea behind this approach is that it may be unnecessary to estimate a grid error if no target signal is present at the corresponding frequency bin. More specifically, we assume that the $\varepsilon_{\bar{m}}|x_{\bar{m}}$ are iid with pdf

$$f(\varepsilon_{\bar{m}}|x_{\bar{m}} = 0) = \delta(\varepsilon_{\bar{m}}) \tag{6a}$$
$$f(\varepsilon_{\bar{m}}|x_{\bar{m}} \neq 0) = \mathbb{I}_{[-0.5,0.5]}(\varepsilon_{\bar{m}}) \tag{6b}$$

where $\mathbb{I}_A(.)$ is the indicator function of the set A. Note that the uniform distribution (6b) over $[-.5, .5]$ is equivalent to a Beta-like distribution of parameters (1,1) over $[-.5, .5]$, denoted as $\mathcal{B}e_{[-.5,.5]}(1, 1)$.

3) Noise power: A suitable prior for the white noise power σ^2 is an inverse-gamma distribution since it is conjugate to the likelihood (4). The prior pdf of σ^2 can therefore be expressed as

$$f(\sigma^2|\gamma_0, \gamma_1) \propto \frac{e^{-\gamma_1/\sigma^2}}{(\sigma^2)^{\gamma_0+1}}\mathbb{I}_{[0,+\infty)}(\sigma^2) \tag{7}$$

where γ_0, γ_1 are respectively the shape and scale parameters. The distribution (7) is denoted as $\sigma^2|\gamma_0, \gamma_1 \sim \mathcal{IG}(\gamma_0, \gamma_1)$. Note that by tuning adequately the shape and scale parameters (γ_0, γ_1), the prior can be made very informative or on the contrary flat. In radar applications, the thermal noise power is usually well known so that a moderately informative prior can be favored.

C. Prior pdfs of the hyperparameters

Since the probability w and the target signal power σ_x^2 are both unknown, another level is added to the hierachical model.

1) Target signal power: Similarly to the case of σ^2, an inverse-gamma prior is chosen for the target signal power σ_x^2 and is denoted as $\sigma_x^2|\beta_0,\beta_1 \sim \mathcal{IG}(\beta_0,\beta_1)$. Nonetheless note that the shape and scale parameters β_0,β_1 must be chosen to obtain this time a not so informative prior since the target signal amplitudes may vary significantly from one to another.

2) Level of occupancy: If no information is available to the radar operator about the sparsity level of the target scene, a convenient prior is a uniform pdf over the interval $[0,1]$, i.e., $w \sim \mathcal{U}_{[0,1]}$.

III. Bayesian estimation

Herein we propose an estimation scheme of the target scene $\boldsymbol{x},\boldsymbol{\varepsilon}$ according to the Bayesian hierarchical model described in Section II. Particularly in what follows, we propose to study the minimum mean square error (MMSE) estimators

$$\hat{\boldsymbol{x}}_{\text{MMSE}} = \int \boldsymbol{x} f(\boldsymbol{x}|\boldsymbol{y}) d\boldsymbol{x}, \tag{8a}$$

$$\hat{\boldsymbol{\varepsilon}}_{\text{MMSE}} = \int \boldsymbol{\varepsilon} f(\boldsymbol{\varepsilon}|\boldsymbol{y}) d\boldsymbol{\varepsilon}. \tag{8b}$$

The MMSE estimator of $\boldsymbol{x}$ was intractable to derive analytically in [17] when the grid mismatch was not taken into account. Therefore, the analytic calculation of the MMSE estimators (8a) and (8b) seem all the more trying. A Monte-Carlo Markov Chain (MCMC) is thus implemented [18]. The MCMC algorithm simulates iteratively samples $\sigma^{2^{(t)}}, \boldsymbol{\varepsilon}^{(t)}, \boldsymbol{x}^{(t)}, w^{(t)}, \sigma_x^{2^{(t)}}$ according to their conditional posterior distribution $f(\theta_i|\boldsymbol{y},\boldsymbol{\theta}_{-i})$ where $\boldsymbol{\theta} = \left[\sigma^2, \boldsymbol{\varepsilon}^T, \boldsymbol{x}^T, w, \sigma_x^2\right]^T$ and $\boldsymbol{\theta}_{-i}$ is the vector $\boldsymbol{\theta}$ whose ith element has been removed. After a burn-in time N_{bi}, the samples are distributed according to their posterior distribution $f(\theta_i|\boldsymbol{y})$. When a sufficient number of samples N_r is collected, conventional Bayesian estimators can be built empirically

$$\hat{\theta}_{i\text{MMSE}} = N_r^{-1} \sum_{t=1}^{N_r} \theta_i{}^{(t+N_{bi})}. \tag{9}$$

The conditional posterior distributions are obtained from the joint posterior pdf of $\sigma^2,\boldsymbol{\varepsilon},\boldsymbol{x},w,\sigma_x^2|\boldsymbol{y}$

$$\begin{aligned} &f(\sigma^2,\boldsymbol{\varepsilon},\boldsymbol{x},w,\sigma_x^2|\boldsymbol{y}) \\ &\quad\propto f(\boldsymbol{y}|\boldsymbol{\varepsilon},\boldsymbol{x},\sigma^2) f(\boldsymbol{\varepsilon}|\boldsymbol{x}) f(\boldsymbol{x}|w,\sigma_x^2) f(w) f(\sigma_x^2) f(\sigma^2). \end{aligned}$$

In particular, both vectors $\boldsymbol{x}$ and $\boldsymbol{\varepsilon}$ are sampled element-wise. The conditional posterior distributions of $\varepsilon_{\bar{m}}$ and $x_{\bar{m}}$ are derived from their conditional joint posterior distribution

$$\begin{aligned} &f(\varepsilon_{\bar{m}}, x_{\bar{m}}|\boldsymbol{y},\boldsymbol{\varepsilon}_{-\bar{m}},\boldsymbol{x}_{-\bar{m}},\sigma^2,w,\sigma_x^2) \\ &\propto \exp\left\{-\sigma^{-2}\left[|x_{\bar{m}}|^2 - x_{\bar{m}}^* \boldsymbol{f}_{\bar{m}}(\varepsilon_{\bar{m}})^H \boldsymbol{e}_{\bar{m}} - x_{\bar{m}} \boldsymbol{e}_{\bar{m}}^H \boldsymbol{f}_{\bar{m}}(\varepsilon_{\bar{m}})\right]\right\} \\ &\quad\times f(\varepsilon_{\bar{m}}|x_{\bar{m}}) f(x_{\bar{m}}|w,\sigma_x^2) \end{aligned} \tag{10}$$

where $\boldsymbol{e}_{\bar{m}} = \boldsymbol{y} - \sum_{i\neq\bar{m}} \boldsymbol{f}_i(\epsilon_i) x_i$.

A. Sampling of $\boldsymbol{x}$

Following [17], $\boldsymbol{x}$ is sampled element-wise. The $\bar{m}$th element of $\boldsymbol{x}$ follows the distribution $\mathcal{B}er\mathcal{CN}\left(w_{\bar{m}},\mu_{\bar{m}},\eta_{\bar{m}}^2\right)$ with

$$\begin{aligned} \eta_{\bar{m}}^2 &= \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_x^2}\right)^{-1} \\ \mu_{\bar{m}} &= \frac{\eta_{\bar{m}}^2}{\sigma^2} \boldsymbol{f}_{\bar{m}}(\epsilon_{\bar{m}})^H \boldsymbol{e}_{\bar{m}} \\ w_{\bar{m}} &= \frac{w \frac{\eta_{\bar{m}}^2}{\sigma_x^2} \exp\left\{\frac{|\mu_{\bar{m}}|^2}{\eta_{\bar{m}}^2}\right\}}{1 - w + w \frac{\eta_{\bar{m}}^2}{\sigma_x^2} \exp\left\{\frac{|\mu_{\bar{m}}|^2}{\eta_{\bar{m}}^2}\right\}}. \end{aligned}$$

B. Sampling of $\boldsymbol{\varepsilon}$

As for the vector $\boldsymbol{x}$, the parameter $\boldsymbol{\varepsilon}$ is sampled element-wise. Using (10), the conditional posterior distribution of $\varepsilon_{\bar{m}}$ is calculated and rearranged to obtain

$$\begin{aligned} f(\varepsilon_{\bar{m}}|\boldsymbol{y},\boldsymbol{\varepsilon}_{-\bar{m}},\boldsymbol{x},\sigma^2) \propto \exp\left\{\sum_{m=1}^{M-1} \kappa_m \cos\left(2\pi \frac{\varepsilon_{\bar{m}}}{\bar{M}} m - \phi_m\right)\right\} \\ \times f(\varepsilon_{\bar{m}}|x_{\bar{m}}) \end{aligned} \tag{11}$$

where $\kappa_m = \frac{2}{\sigma^2\sqrt{\bar{M}}} \times |b_m|$ and $\phi_m = \angle b_m$. $\boldsymbol{b}$ has been defined as

$$\boldsymbol{b} = x_{\bar{m}}^* \boldsymbol{u}_{\bar{m}}^* \odot \boldsymbol{e}_{\bar{m}}, \tag{12}$$

where $\boldsymbol{u}_{\bar{m}} = \exp\{j2\pi m(\bar{m}/\bar{M})\}$. When $x_{\bar{m}} \neq 0$, $f(\varepsilon_{\bar{m}}|x_{\bar{m}}) = \mathbb{I}_{[-0.5,0.5]}(\varepsilon_{\bar{m}})$. Thus, we recognize from (11) a dilated and truncated generalized von Mises distribution [19]. Such a distribution can be troublesome to sample, so a Metropolis-Hastings (MH) algorithm is used [18]. This algorithm is based on a *proposal distribution* that is easy to simulate from, and is chosen as close as possible from the *target distribution*. The conditional posterior distribution of $\varepsilon_{\bar{m}}$ is represented in Fig.2 for different values of mismatch and post-processing SNR (defined in (16)). This figure suggests that a flat proposal would be appropriate in the case of low power, and a Gaussian distribution in the case of high power. Thus, in our MH algorithm we switch from a flat proposal to a Gaussian proposal (and vice-versa) depending on the estimated target power. This adaptive scheme should be employed only during a burn-in period in order to preserve the convergence properties [18], but in fact it does not damage the performance when used in the whole process.

C. Sampling of w, σ_x^2 and σ^2

Like in [17], the parameter σ^2 and hyperparameters w and σ_x^2 are sampled according to their conditional posterior distribution

$$\sigma^2|\boldsymbol{y},\boldsymbol{x},\boldsymbol{\varepsilon} \sim \mathcal{IG}\left(\gamma_0 + M, \gamma_1 + \| \boldsymbol{y} - \boldsymbol{F}(\boldsymbol{\varepsilon})\boldsymbol{x} \|_2^2\right) \tag{13}$$

$$w|\boldsymbol{x} \sim \mathcal{B}e\left(1 + n_1, 1 + n_0\right) \tag{14}$$

$$\sigma_x^2|\boldsymbol{x} \sim \mathcal{IG}\left(\beta_0 + n_1, \beta_1 + \| \boldsymbol{x} \|_2^2\right) \tag{15}$$

where n_1 is the number of nonzero elements of $\boldsymbol{x}$ and $n_0 = \bar{M} - n_1$.

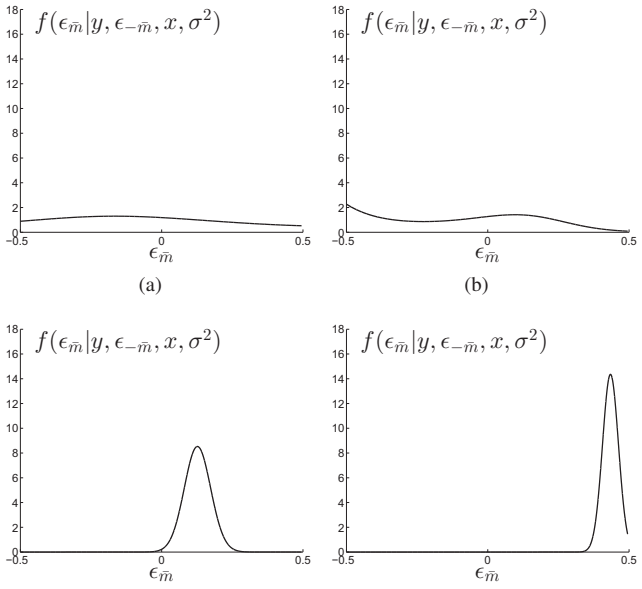


Fig. 2. Shape of the conditional posterior distribution $f(\varepsilon_{\bar{m}}|\boldsymbol{y},\boldsymbol{\varepsilon}_{-\bar{m}},\boldsymbol{x},\sigma^2)$. (a) SNR=5dB, $\varepsilon_{\bar{m}}=.15$. (b) SNR=5dB, $\varepsilon_{\bar{m}}=.45$. (c) SNR=15dB, $\varepsilon_{\bar{m}}=.15$. (d) SNR=15dB, $\varepsilon_{\bar{m}}=.45$.

IV. Numerical simulations

We now present various numerical examples illustrating the performance of the proposed SSR algorithm. Note that for each scenario, the constant hyperparameters (γ_0,γ_1) and (β_0,β_1) are calculated to obtain the desired values of mean and variance of the prior distributions of σ^2 and σ_x^2. Indeed for an inverse gamma distribution $g\sim\mathcal{IG}(\nu_0,\nu_1)$ the mean and variance are respectively

$$\mathrm{m}_g=\frac{\nu_1}{\nu_0-1},\quad \nu_0>1$$
$$\mathrm{var}_g=\frac{\nu_1^2}{(\nu_0-1)^2(\nu_0-2)},\quad \nu_0>2$$

so the parameters are calculated as

$$\nu_0=\frac{\mathrm{m}_g^2}{\mathrm{var}_g}+2$$
$$\nu_1=\mathrm{m}_g\left(\frac{\mathrm{m}_g^2}{\mathrm{var}_g}+1\right).$$

A. Synthetic data

To begin with, synthetic data are generated according to the model described by (1) and (3).

1) Results after one realization: Three targets with possible grid mismatch are simulated in a white noise background with unit power. They all have a post-processing signal-to-noise ratio (SNR) of 25 dB, defined as

$$SNR_n=\frac{\mathcal{E}\{|\alpha_n|^2\}}{\sigma^2}\times M. \tag{16}$$

In Fig. 3 are plotted the post-processing amplitude of the target vector $\hat{\boldsymbol{x}}_{\mathrm{MMSE}}$ and the grid mismatch $\hat{\boldsymbol{\varepsilon}}_{\mathrm{MMSE}}$. These estimates

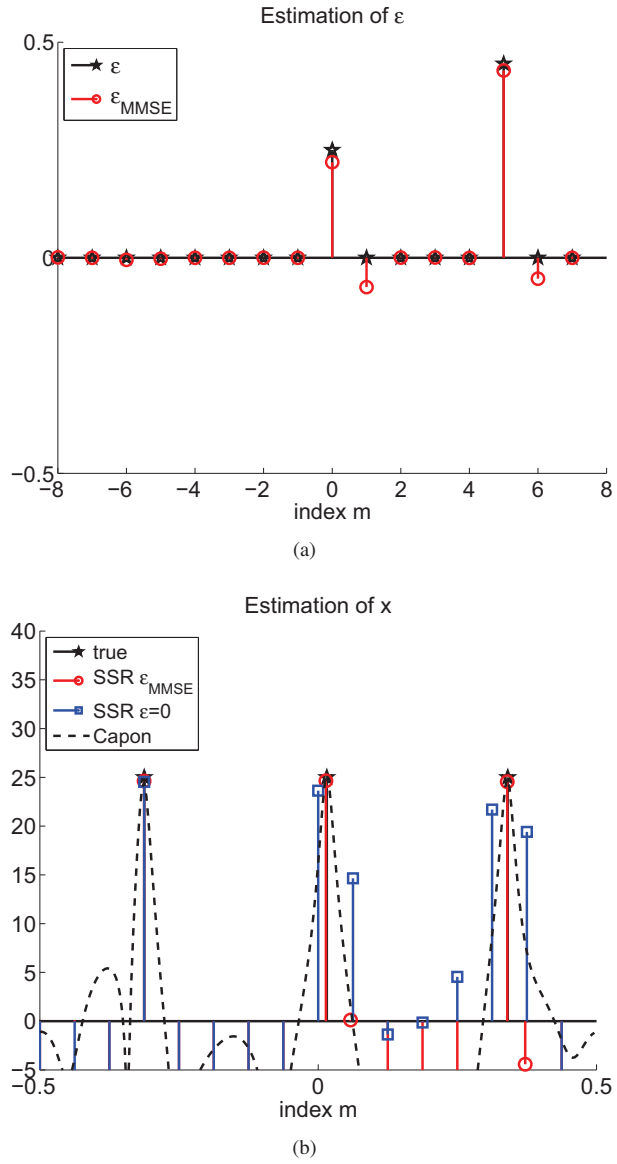


Fig. 3. SSR of synthetic target scene: $M=16$, $N=3$, $\sigma^2=1$, $\bar{M}=M$, $(\mathrm{m}_{\sigma_x^2},\sqrt{\mathrm{var}_{\sigma_x^2}})=M_{dB}+(0,3.5)$ dB where $M_{dB}=10\log_{10}(M)$, $(\mathrm{m}_{\sigma^2},\sqrt{\mathrm{var}_{\sigma^2}})=(0,2.4)$ dB. (a) Grid errors $\boldsymbol{\varepsilon}$. (b) Target amplitude vector $\boldsymbol{x}$.

are compared with the true target scene $\boldsymbol{x},\boldsymbol{\varepsilon}$ and with the target scene estimated by the non-robust method (the proposed SSR technique when the grid mismatch is ignored, i.e., with $\boldsymbol{\varepsilon}=\mathbf{0}$). From Fig. 3, we see that the proposed SSR algorithm estimates correctly both $\boldsymbol{x}$ and $\boldsymbol{\varepsilon}$. On the contrary, if the grid error is ignored, the target energy tends to be spread over the adjacent frequency bins especially when the mismatch is important, i.e., $|\varepsilon_{\bar{m}}|\to.5$. As as result it invalidates the sparsity of the reconstruction (this performance loss is well known).

2) Performance: A simple scenario is considered with a single target at zero velocity for different values of SNR and mismatch ϵ_0. The performance of the proposed estimators (8) is assessed after 500 Monte-Carlo simulations through the calculation of the mean square error (MSE) of the reconstructed

estimated target scene $\boldsymbol{F}(\hat{\boldsymbol{\varepsilon}}_{\text{MMSE}})\hat{\boldsymbol{x}}_{\text{MMSE}}$ i.e.

$$\text{MSE}(\boldsymbol{F}(\hat{\boldsymbol{\varepsilon}}_{\text{MMSE}})\hat{\boldsymbol{x}}_{\text{MMSE}}) = \frac{1}{N_{mc}}\sum_{n=1}^{N_{mc}} \parallel \boldsymbol{F}(\hat{\boldsymbol{\varepsilon}}^{(n)}_{\text{MMSE}})\hat{\boldsymbol{x}}^{(n)}_{\text{MMSE}} - \boldsymbol{F}(\boldsymbol{\varepsilon})\boldsymbol{x} \parallel_2^2 . \quad (17)$$

This metric seems more relevant than the MSE of $\hat{\boldsymbol{x}}_{\text{MMSE}}$ and $\hat{\boldsymbol{\varepsilon}}_{\text{MMSE}}$. Indeed, the sparse representation sometimes induces ambiguity about the position of the target: is it in the $\bar{m}$th frequency bin with a mismatch of .5 ($\varepsilon_{\bar{m}} = .5$) or in the $\bar{m}+1$th frequency bin with a mismatch of -.5 ($\varepsilon_{\bar{m}+1} = -.5$)? With such an ambiguity, two problems might arise: the target might be shifted or split. Let us consider the case of a target in the $\bar{m}$th frequency bin with a mismatch of .5 ($\varepsilon_{\bar{m}} = .5$). First, the analysis may result in a shifted target: the target is estimated in the next range bin with a mismatch of -.5. The analysis may also result in a split target: a target is estimated in the $\bar{m}$th frequency bin with a positive mismatch ($\varepsilon_{\bar{m}} \to .5$), and another in the $\bar{m}+1$th frequency bin with a negative mismatch ($\varepsilon_{\bar{m}+1} \to -.5$). In both cases, the MSE of $\hat{\boldsymbol{x}}_{\text{MMSE}}$ and $\hat{\boldsymbol{\varepsilon}}_{\text{MMSE}}$ will be high. On the other hand, the MSE of the reconstructed estimated target scene $\boldsymbol{F}(\hat{\boldsymbol{\varepsilon}}_{\text{MMSE}})\hat{\boldsymbol{x}}_{\text{MMSE}}$ will be lower, thanks to the estimation of $\hat{\boldsymbol{\varepsilon}}_{\text{MMSE}}$. This metric may be more representative of the estimation quality.

We represent on Fig.4 the MSE of $\boldsymbol{F}(\hat{\boldsymbol{\varepsilon}}_{\text{MMSE}})\hat{\boldsymbol{x}}_{\text{MMSE}}$ as a function of grid mismatch for a SNR of 10 and 20 dB. We can see from Fig.4 that for a SNR of 20 dB, the robust analysis outperforms the non-robust analysis as soon as $\epsilon_0 > .02$ ($\epsilon_0 > .75$ for a SNR of 10 dB) meaning that it is almost always profitable to estimate $\boldsymbol{\varepsilon}$. We can also see that the benefits of the robust analysis compared to the non-robust analysis increase with the SNR. The clairvoyant case with respect to (wrt) grid mismatch is also represented on Fig.4 and corresponds to the robust analysis when $\boldsymbol{\varepsilon}$ is known. It is a reference and it shows that when we know the true value of the grid mismatch, it is worth taking it into account for the estimation of the target scene. It also indicates how reliable the sampling of $\boldsymbol{\varepsilon}$ is in the proposed SSR method: it is highly efficient.

B. Experimental data

Let us now discuss the performance of the SSR algorithm in a practical case. To that end, data collected from the PARSAX radar [20] on November 2010 are considered. For this dataset the radar was pointing on a freeway during a heavy traffic time. The exact number of targets as well as their amplitude and location in the range-velocity map are unknown. Note that compared to the synthetic case, the target amplitudes are very high, so the hyperparameters (β_0, β_1) are set such that $(\text{m}_{\sigma_x^2}, \sqrt{\text{var}_{\sigma_x^2}}) = M_{dB} + (60, 3.5)$ dB. In Fig. 5 are depicted the elements of $\hat{\boldsymbol{x}}_{\text{MMSE}}$ whose magnitude is greater than the lowest value of the colorbar. For the sake of comparison, the amplitude estimated by a conventional Capon's method is represented as a transparent background. (Note however that the Capon's output gives only an estimate of the target scene and should not be taken has the ground truth.) We see again the benefit of estimating the grid mismatch to avoid target spreading and/or miss detection. It is worth noticing that the target at range bin #4 and velocity ≈ -17 m/s is split by the robust analysis. This problem has been discussed before, and we can see that the non-robust estimation of this target is even worse. Besides, this anomaly is not inherent to our model and has been identified in [15] where a perturbation matrix is used.

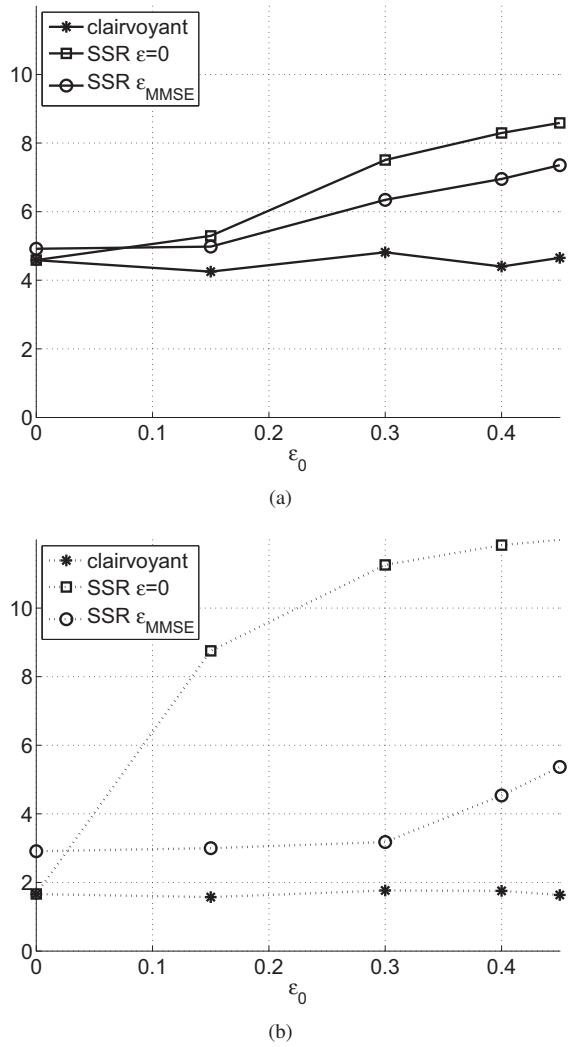
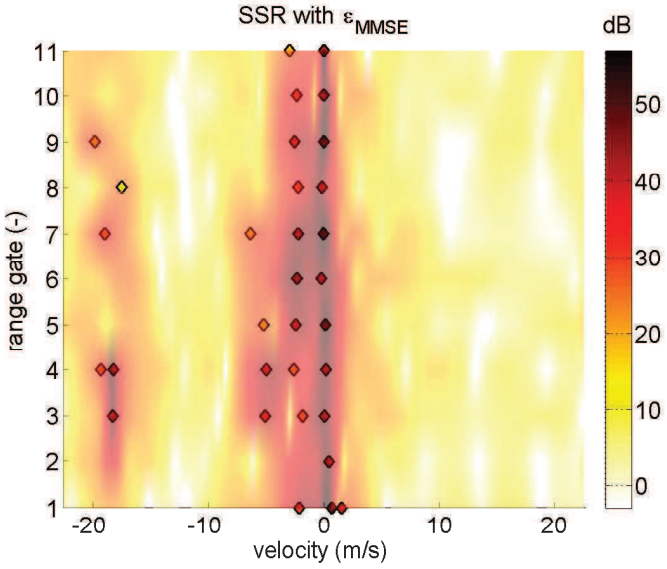


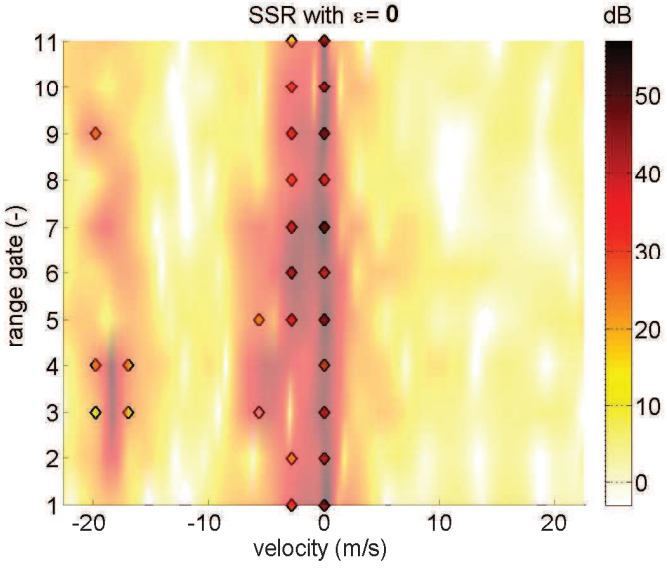
Fig. 4. Performance on synthetic target scene: $M = 32$, $N = 1$, $\sigma^2 = 1$, $\bar{M} = M$, $(\text{m}_{\sigma_x^2}, \sqrt{\text{var}_{\sigma_x^2}}) = M_{dB} + (0, 3.5)$ dB, $(\text{m}_{\sigma^2}, \sqrt{\text{var}_{\sigma^2}}) = (0, 2.4)$ dB. Comparison with the clairvoyant wrt ϵ and non-robust estimation. (a) SNR=10dB. (b) SNR=20dB

V. Conclusion

In this paper we have described a new Bayesian algorithm for the sparse representation of targets in Fourier basis. Particularly, an error vector is introduced to model the possible mismatch between the target frequency and the nearest frequency point in the Fourier basis. The method, though computationally intensive, allows the sparsity of the scene to be preserved in case of grid mismatch. However, a problem persists in the limiting case when $|\varepsilon_{\bar{m}}| \to .5$ but the proposed estimation method still remains more reliable than the non-robust estimation. A number of parameters need to be set by the radar operator, and they can change significantly the performance of the analysis (in particular the parameters of the target signal power). This point will be further investigated



(a)



(b)

Fig. 5. SSR of experimental target scene (PARSAX data): range resolution of 1.5 m, $M = 16$, N unknown, $\sigma^2 \approx 1$ (pre-normalization of the data), $\bar{M} = M$, $(\mathrm{m}_{\sigma_x^2}, \sqrt{\mathrm{var}_{\sigma_x^2}}) = M_{dB} + (60, 3.5)$ dB, $(\mathrm{m}_{\sigma^2}, \sqrt{\mathrm{var}_{\sigma^2}}) = (0, 2.4)$ dB. (a) Target amplitude vector $\hat{\boldsymbol{x}}_{\mathrm{MMSE}}$ (diamond marker). (b) Target amplitude vector $\hat{\boldsymbol{x}}_{\mathrm{MMSE}}$ when $\boldsymbol{\varepsilon} = \mathbf{0}$ (diamond marker).

in the near future. Furthermore, the proposed SSR technique will be integrated into more advanced hierarchical model.

ACKNOWLEDGMENT

The authors would like to thank O. Krasnov at TU Delft for kindly providing the PARSAX experimental data.

REFERENCES

[1] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, no. 8, pp. 1408–1418, Aug. 1969.

[2] J. Li and P. Stoica, "An adaptive filtering approach to spectral estimation and SAR imaging," *IEEE Trans. Signal Process.*, vol. 44, no. 6, pp. 1469–1484, Jun. 1996.

[3] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[4] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, jul 1989.

[5] M. Hayes, *Statistical digital signal processing and modeling*. New York: Wiley, Press, 1996.

[6] L. Anitori, A. Maleki, M. Otten, R. Baraniuk, and P. Hoogeboom, "Design and analysis of compressed sensing radar detectors," *IEEE Trans. Signal Process.*, vol. 61, no. 4, pp. 813–827, Feb. 2013.

[7] S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM review*, vol. 43, no. 1, pp. 129–159, 2001.

[8] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2346–2356, Jun. 2008.

[9] M. Herman and T. Strohmer, "General deviants: An analysis of perturbations in compressed sensing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 342–349, Apr. 2010.

[10] Y. Chi, L. Scharf, A. Pezeshki, and A. Calderbank, "Sensitivity to basis mismatch in compressed sensing," *IEEE Trans. Signal Process.*, no. 5, pp. 2182–2195, May 2011.

[11] D. Malioutov, M. Cetin, and A. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, Aug. 2005.

[12] H. Zhu, G. Leus, and G. Giannakis, "Sparsity-cognizant total least-squares for perturbed compressive sampling," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2002–2016, 2011.

[13] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse Bayesian inference," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 38–43, Jan. 2013.

[14] L. Hu, J. Zhou, Z. Shi, and Q. Fu, "A fast and accurate reconstruction algorithm for compressed sensing of complex sinusoids," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5744–5754, Nov. 2013.

[15] Z. Tan and A. Nehorai, "Sparse direction of arrival estimation using co-prime arrays with off-grid targets," *IEEE Signal Process. Lett.*, vol. 21, no. 1, January 2014.

[16] L. Hu, Z. Shi, J. Zhou, and Q. Fu, "Compressed sensing of complex sinusoids: An approach based on dictionary refinement," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3809–3822, Jul. 2012.

[17] S. Bidon, J.-Y. Tourneret, L. Savy, and F. Le Chevalier, "Bayesian sparse estimation of migrating targets for wideband radar," *IEEE Trans. Aerosp. Electron. Syst.*, to be published.

[18] C. Robert and G. Castella, *Monte Carlo Statistical Methods*. Springer, 2004.

[19] R. Gatto and S. Jammalamadaka, "The generalized von Mises distribution," *Statistical Methodology*, vol. 4, pp. 341–353, 2007.

[20] O. A. Krasnov, G. P. Babur, Z. Wang, L. P. Ligthart, and F. van der Zwan, "Basics and first experiments demonstrating isolation improvements in the agile polarimetric FM-CW radar – PARSAX," *International Journal of Microwave and Wireless Technologies*, vol. 2, pp. 419–428, 8 2010.