Ph.D. Dissertation

International Doctoral School in Information and Communication Technology

DISI - University of Trento

An Innovative Methodological Approach Based on Compressive Sensing for the Synthesis and Control of Antenna Arrays

Matteo Carlin

Tutor:
Andrea Massa, Professor
Università degli Studi di Trento
Advisor:
Paolo Rocca, Aggregate Professor
Università degli Studi di Trento
Co-Advisor:
Giacomo Oliveri, Aggregate Professor
Università degli Studi di Trento

December 2013

Abstract

In the framework of antenna array synthesis and control, this thesis focus on the development and analysis of techniques based on the Bayesian Compressive Sensing (BCS) for the design of sparse antenna arrays and for the estimation of the direction of arrival (DoA) of signals impinging on an antenna array. After formulating the sparse-array synthesis problem in a probabilistic fashion, the single-task BCS ($ST-BCS$) is applied to the synthesis of symmetrical antenna arrays with real weights. In order to deal with the synthesis of sparse arrays with complex weights, the multitask version of the BCS ($MT-BCS$) is employed to correlate the real and imaginary part of the resulting excitation distribution. Concerning the DoA estimation problem, starting from the observation that the signals impinging on the antenna array are sparse in the spatial domain, a single-snapshot $ST-BCS$-based technique is proposed. Moreover, the $MT-BCS$-based extension of this technique is introduced in order to enhance the quality of the estimations through the exploitation of the correlation among different snapshots. In the numerical validation, an exhaustive analysis has been performed to assess effectiveness, reliability, but also limitations of the proposed methodologies. Comparisons with state-of-the-art are reported and discussed, as well.

Keywords
Array synthesis, Bayesian Compressive Sensing, Direction-of-Arrival Estimation, Sparse Arrays

Published Conference Papers

[C1] G. Oliveri, M. Carlin, E.T. Bekele and A. Massa, "A BCS-Based Approach for the Synthesis of Conformal Arrays," 7th EUCAP, Gothenburg, Sweden, April 8-11, 2013.

[C2] G. Oliveri, M. Carlin, and A. Massa, "BCS-based formulations for antenna arrays synthesis," in 2012 6th European Conference on Antennas and Propagation (EUCAP), 2012, pp. 1500-1501.

[C3] M. Carlin and P. Rocca, "A Bayesian Compressive Sensing strategy for direction-of-arrival estimation," in 2012 6th European Conference on Antennas and Propagation (EUCAP), 2012, pp. 1508-1509.

[C4] G. Oliveri, M. Carlin, and A. Massa, "Synthesis of planar arrays through Bayesian Compressive Sensing," in Proceedings of the 2012 IEEE International Symposium on Antennas and Propagation, 2012.

[C5] M. Carlin, P. Rocca, G. Oliveri, and A. Massa, "Multi-task Bayesian compressive sensing for direction-of-arrival estimation," in 2012 IEEE International Conference on Wireless Information Technology and Systems (ICWITS), 2012.

[C6] A. Massa, P. Rocca, M. Carlin, and G. Oliveri, "Efficient synthesis of sparse arrays as the solution of an inversion problem within the bayesian compressive sensing framework," in 2012 IEEE International Conference on Wireless Information Technology and Systems (ICWITS), 2012.

[C7] G. Oliveri, F. Robol, M. Carlin and A. Massa, "Synthesis of Large Sparse Linear Arrays by Bayesian Compressive Sensing," 5th EUCAP, Rome, Italy, April 11-5, 2011.

[C8] M. Carlin, G. Oliveri, and A. Massa, "A CS-based strategy for the design of shaped-beam sparse arrays," in 2011 IEEE International Symposium on Antennas and Propagation (APSURSI), 2011.

[C9] G. Oliveri, P. Rocca, L. Poli, M. Carlin, and A. Massa, "Beyond the Nyquist Criterion - The Compressive Sensing Paradigm in Electromagnetics," in 2011 IEEE International Symposium on Antennas and Propagation (APSURSI), 2011.

[C10] A. Massa, P. Rocca, M. Carlin, and G. Oliveri, "Efficient synthesis of sparse arrays as the solution of an inversion problem within the bayesian compressive sensing framework," in 2012 IEEE International Conference on Wireless Information Technology and Systems (ICWITS), 2012.

[C11] G. Oliveri, A. Massa, L. Poli, P. Rocca and M. Carlin, "Compressive sensing strategies for active microwave imaging," PIERS, Marrakesh, Morocco, March 20-23, 2011.

Published Journals Papers

[R1] G. Oliveri, M. Carlin, and A. Massa, "Complex-Weight Sparse Linear Array Synthesis by Bayesian Compressive Sampling," IEEE Transactions on Antennas and Propagation, vol. 60, no. 5, pp. 2309-2326, May 2012.

[R2] M. Carlin, P. Rocca, G. Oliveri, F. Viani, and A. Massa, "Directions-of-Arrival Estimation Through Bayesian Compressive Sensing Strategies," IEEE Transactions on Antennas and Propagation, vol. 61, no. 7, pp. 3828-3838, Jul. 2013.

[R3] M. Carlin, P. Rocca, G. Oliveri, and A. Massa, "Bayesian compressive sensing as applied to directions-of-arrival estimation in planar arrays", Journal of Electrical and Computer Engineering, Special Issue on “Advances in Radar Technologies”, vol. 2013, Article ID 245867, 12 pages, 2013. doi:10.1155/2013/245867

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Thanks to its ability to allow the recovering of a signal starting from far fewer measurements than conventional techniques based on Shannon's theorem [85], the Compressive Sensing (CS) paradigm has attracted a lot of attention in several research areas of information theory, signal processing, computer science and electrical engineering [80, 81, 85], enabling the development of completely new approaches in these fields [80, 81, 85]. Traditional sampling approaches require that the sampling rate is higher than twice the maximum frequency value in the measured signal. However, in many applications such as imaging and radar, the signal to be acquired is often sparse with respect to a proper basis (i.e. it has a concise representation in that basis). As a consequence, a large amount of data samples can be represented with a small number of coefficients. For example, lossy image compression coders encode only the locations and the values of the most significant coefficients of an image, throwing away the majority of the coefficients with almost no perceptual loss. This process requires the acquisition of all the data in order to perform the compression, resulting in a waste of measuring resources. Differently from traditional compression schemes, instead of measuring the full signal, the CS acquires only the amount of data that is not discharged. This feature is useful in many applicative scenarios, where the acquisition of a large number of measurements is not practical for several reasons, like the high cost of the measurements, the limited number of sensors or the large time required for each measurement. The CS approach is based on finding an approximate solution $\mathbf{x}$ to an underdetermined linear problem $\mathbf{y} = \mathbf{A}\mathbf{x}$, minimizing at the same time the number of non-zero entries of $\mathbf{x}$ (see Equation 2.1). If suitable conditions are fitted, a high-dimensionality solution $\mathbf{x}$ can be retrieved from a small number of measurements $\mathbf{y}$. In addition to the advantages over classical sampling schemes, the popularity of the CS is related to (i) the flexibility and generality of its formulation, allowing its application to a wide range of problems, (ii) the effectiveness of the corresponding solution techniques and to (iii) the wide availability of software libraries implementing state-of-the-art CS algorithms [57, 58, 59, 60] for effectively dealing with complex engineering

problems.

Thanks to their efficiency, CS strategies have gained a lot of interest in the EM community. For example, in [84], the CS has been applied to radar remote sensing, a problem where the standard CS requirements (i.e. sparsity of the solution and linearity) are fitted in a natural way. However, by using suitable approximations or if some a-priori knowledge is at hand, several electromagnetic problems can be reformulated in order to fit the CS requirements. In this way, the CS has been recently extended to several fields of electromagnetics with very interesting results. These applications include array diagnosis [2], array synthesis [51, 52], direction-of-arrival estimation [76], inverse scattering and microwave imaging [53, 54].

On the other hand, in order to guarantee the practical exploitability of CS-based techniques, several issues like numerical stability and theoretical conditions to guarantee their optimality still need to be carefully addressed. Indeed, the validity of widely adopted assumptions concerning the features of the problems, such as the Restricted Isometry Property, cannot be always granted in EM problems of interest, whose properties are often constrained by the underlying physics [21, 86]. Accordingly, the use of several popular solvers relying on these assumptions, including those based on l1-norm minimization, may not be the optimal choice in EM synthesis and inverse problems [21, 86]. Recently, a set of effective techniques have been proposed in order to address the above issues and enable the effective application of the CS paradigm in EM problems [41, 42]. Such strategies are essentially based on the reformulation of the EM problems in suitable probabilistic scenarios, following a Bayesian paradigm comprising suitable sparsity priors [41, 42]. The arising "Bayesian CS" (BCS) solution strategies have been therefore adopted to properly address design/inversion problems arising in several different scenarios [3, 21, 86, 87, 95].

Differently from the approaches based on CS, the Bayesian Compressive Sensing (BCS) proposed in [41] searches for the most probable sparse solution fitting the measured data samples. Thanks to the probabilistic formulation, the kernel of the problem is not required to satisfy any specific theoretical feature, like the restricted isometry property (RIP) [85]. The verification of these features is often very difficult in practical applications. However, while the CS is able (under certain circumstances) to obtain the exact reconstruction [79][80], this is not the case of the BCS. Moreover, due to the real-valued nature of the BCS solver, its extension to the sampling/recovery of complex signals is not efficient [41]. In addition, the standard BCS approach is not able to correlate the information obtained from different measurement sets acquired in different time instants or by different acquisition systems to enhance the estimation performances. In order to avoid these problems, the Multi-Task Bayesian Compressive Sensing ($MT-BCS$) methodology has been introduced in [60]. Differently form the standard BCS implementation ($ST-BCS$), the $MT-BCS$ allows the probabilistic correlation [60] of different sets of measurements in order to improve the

accuracy of the reconstruction. Additionally, the problem of the estimation of complex signals can be handled in a similar way, by exploiting the $MT-BCS$ to correlate the real and imaginary components of complex measured data, enabling the methodology to recover complex signals. The basic formulation of the $ST-BCS$ and $MT-BCS$ approaches is resumed in the following Chapter.

Thesis outline

The thesis is organized as follows. Firstly, the BCS general formulation is introduced in Chapter 2. In Chapter 3, the problem of the synthesis of sparse linear arrays with real weights is addressed by means of a $ST-BCS$ strategy. Successively, the methodology is extended in Chapter 4 to the synthesis of antenna arrays with complex weights by means of a $MT-BCS$ approach. Chapter 5 presents the problem of the estimation of the DoAs of signals impinging on a linear antenna array from a BCS perspective, focusing on both the $ST-BCS$ and $MT-BCS$. This methodology is then extended to the planar array case in Chapter 6. Some general conclusions follow in Chapter 7.

Chapter 2

Mathematical Formulation

When the relation between the measured data and the unknowns is linear, the objective is to determine a K-dimensional vector $\mathbf{x} \in \mathbb{R}^{K\times 1}$ starting from a M-dimensional set of measured data $\mathbf{y} \in \mathbb{R}^{M\times 1}$, where $\mathbf{x}$ is related to the measured data $\mathbf{y}$ by the relation

$$\mathbf{y} = \mathbf{A} \quad \mathbf{x} \tag{2.1}$$

$\mathbf{A} \in \mathbb{R}^{M\times K}$ being a matrix modeling the linear relationship between the data and the unknowns. However, in many engineering and scientific problems, the number of measurements M is smaller than the number of unknowns K. In this case, the system of equations (2.1) results to be underdetermined with a non-unique solution. therefore, it is not possible to obtain an accurate reconstruction of the unknown $\mathbf{x}$ without adding some informations the problem.

In many circumstances, the unknown signal $\mathbf{x}$ can be represented by using a number of coefficients very small with respect to K. This means that only a small number of coefficients of the vector $\mathbf{x}$ is different from zero. In this case, the measured vector $\mathbf{y}$ is called compressible and the unknown vector $\mathbf{x}$ is called sparse. Under the sparsity hypothesis, the unknown signal $\mathbf{x}$ can be retrieved by solving the following compressive sensing (CS) problem[81, 85]

$$\min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_0} \quad \text{subject to} \quad \mathbf{y} = \mathbf{A}\mathbf{x} \tag{2.2}$$

where $\|\mathbf{x}\|_{\ell_0} = \sum_{k=1}^{K} |x_k|^0$. However, the problem (2.2) is non-convex, and its solution can be obtained only with an exhaustive combinatorial search. An alternative common approach is to consider the problem[81, 85]

$$\min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_1} \quad \text{subject to} \quad \mathbf{y} = \mathbf{A}\mathbf{x} \tag{2.3}$$

where $\|\mathbf{x}\|_{\ell_1} = \sum_{k=1}^{K} |x_k|$. This is a convex problem which can be recast as a linear programming problem and solved in an efficient way[81, 85].

2.1 Single Task BCS $(ST-BCS)$

Let us consider the system of linear equations (2.1). Under the $ST-BCS$ framework, the estimation problem is recast as: given $\mathbf{y} \in \mathbb{R}^{M\times 1}$ find the most sparse solution $\mathbf{x} \in \mathbb{R}^{K\times 1}$ which maximizes the a-posterior probability $\max_{\mathbf{x}} \wp(\mathbf{x}|\mathbf{y})$. In other words

$$\mathbf{x}_{ST} = \arg\left\{\max_{\mathbf{x}} \left[\wp(\mathbf{x}|\mathbf{y})\right]\right\} \quad \text{subject to} \quad \mathbf{x} \text{ is sparse} \tag{2.4}$$

In order to take into account the sparsity constraints imposed on the solution, the following sparseness prior is defined

$$\wp(\mathbf{x}) = \int \wp(\mathbf{x}|\mathbf{a})\,\wp(\mathbf{a})\,d\mathbf{a} \tag{2.5}$$

In (2.5) the sparseness of the signal vector $\mathbf{x}$ is controlled by the unknown hyper-parameter $\mathbf{a}$ [41]. By assuming a Gamma-distributed hyper-parameter vector, Equation 2.4 can be rewritten as follows[41]

$$\mathbf{x}_{ST} = \arg\left\{\max_{\mathbf{x}} \left[\wp\left(\mathbf{x}|\mathbf{y},\sigma^2,\mathbf{a}\right)\wp\left(\sigma^2,\mathbf{a}|\mathbf{y}\right)\right]\right\} \tag{2.6}$$

Since the term $\wp(\mathbf{x}|\mathbf{y},\sigma^2,\mathbf{a})$ in 2.6 is given by the signal model, if a Gaussian distribution is assumed (which is realistic if $AWGN$ is at hand [41]), it can be rewritten as [41]

$$\wp\left(\mathbf{x}|\mathbf{y},\sigma^2,\mathbf{a}\right) = \frac{1}{(2\pi)^{\frac{K+1}{2}}\sqrt{\det(\boldsymbol{\Xi})}} \exp\left\{-\frac{(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Xi}(\mathbf{x}-\boldsymbol{\mu})}{2}\right\} \tag{2.7}$$

where $\boldsymbol{\Xi} \in \mathbb{R}^{K\times K}$ and $\boldsymbol{\mu} \in \mathbb{R}^{K\times 1}$ are equal to

$$\boldsymbol{\Xi} = \left(\frac{1}{\sigma^2}\mathbf{A}^T\mathbf{A} + \operatorname{diag}(\mathbf{a})\right)^{-1} \tag{2.8}$$

$$\boldsymbol{\mu} = \frac{1}{\sigma^2}\boldsymbol{\Xi}\mathbf{A}^T\mathbf{y} \tag{2.9}$$

T denoting the transpose operation.

As it can be observed, the maximum value of (2.7) occurs when the posterior mean $\boldsymbol{\mu}$ is equal to $\mathbf{x}$. Hence, the problem of maximization of (2.6) is solved by finding the values of the parameters $\mathbf{a}$ and σ^2 that maximizes the term $\wp\left(\sigma^2, \mathbf{a}|\mathbf{y}\right)$. With the help of the Bayes Theory, it is possible to prove that the term $\wp\left(\sigma^2, \mathbf{a}|\mathbf{y}\right)$ in (2.6) complies with [41]

$$\wp\left(\sigma^2, \mathbf{a}|\mathbf{y}\right) \propto \wp\left(\mathbf{y}|\sigma^2, \mathbf{a}\right) \wp\left(\sigma^2\right) \wp\left(\mathbf{a}\right) \tag{2.10}$$

where, according to [41], the terms $\wp\left(\sigma^2\right)$ and $\wp\left(\mathbf{a}\right)$ can be assumed to be constant. Hence, the computation of the values of the parameters $\mathbf{a}$ and σ^2 that maximizes $\wp\left(\mathbf{y}|\sigma^2, \mathbf{a}\right)$ and hence the probability appearing in (2.6) can be performed by maximizing the logarithm of $\wp\left(\mathbf{y}|\sigma^2, \mathbf{a}\right)$.

Accordingly, the problem is solved by finding the parameters σ^2_{ST} and $\mathbf{a}_{ST}$ that maximize the following Maximum Likelihood function[41]

$$\begin{aligned} \mathcal{L}_{ST}\left(\sigma^2, \mathbf{a}\right) &= \log\left[\wp\left(\mathbf{y}|\sigma^2, \mathbf{a}\right)\right] = \\ & \frac{1}{2}\left\{N\log\left(2\pi\right) + \log\left[\det\left(\mathbf{C}_{ST}\right)\right] + \mathbf{y}^T\mathbf{C}_{ST}^{-1}\mathbf{y}\right\} \end{aligned} \tag{2.11}$$

where $\mathbf{C}_{ST} = \sigma^2\mathbf{I} + \mathbf{A}\mathrm{diag}\left(\mathbf{a}\right)^{-1}\mathbf{A}^T$, $\mathbf{C}_{ST} \in \mathbb{R}^{M\times M}$ and $\mathbf{I} \in \mathbb{R}^{M\times M}$ is the identity matrix. By following the guidelines in [41], the optimization of (2.11) is carried out by using a relevant vector machine (RVM), initialized with a user defined starting value for σ^2, $\sigma^2 = \sigma_0^2$. Finally, starting from the computed σ^2_{ST} and $\mathbf{a}_{ST}$ values, the estimated solution vector $\mathbf{x}_{ST}$ is obtained as

$$\mathbf{x}_{ST} = \frac{1}{\sigma^2_{ST}}\left(\frac{\mathbf{A}^T\mathbf{A}}{\sigma^2_{ST}} + \mathrm{diag}\left(\mathbf{a}_{ST}\right)\right)\mathbf{A}^T\mathbf{y} \tag{2.12}$$

It is worth noticing that this value correspond to the mean value of (2.7).

2.2 Multi Task BCS $(MT-BCS)$

As already discussed, the $ST-BCS$ methodology presented in the previous section is not efficient when dealing with multiple data sets (e.g. data sets acquired at multiple time instants or by different measurement systems) or when complex data are at hand. In these cases, the linear system (2.1) can be rewritten as

$$\mathbf{y}^{(w)} = \mathbf{A}\ \ \mathbf{x}^{(w)},\ w = 1, ..., W \tag{2.13}$$

where $\mathbf{x}^{(w)} \in \mathbb{R}^{K\times 1}$, $w = 1, ..., W$, can be, alternatively:

- the data measured by different sets of sensors or at different time instants ($w = 1, ..., W$).
- the real and imaginary part of a complex signal ($w = 1, 2$).

The only way to apply the $ST-BCS$ in this case is by solving W independent maximization problems (2.11), one for each set of data $\mathbf{x}^{(w)},\ w = 1, ..., W$, leading to independent solutions $\sigma_{ST}^{2,(w)}$ and $\mathbf{a}_{ST}^{(w)}$, $w = 1, ..., W$. This is a non-efficient way of using the data at hand, since the possible relation between different tasks (i.e. one of the W problems in (2.13)) is neglected. In order to address this limitation, the problem is formulated in a $MT-BCS$ framework as follows

$$x_{MT} = \frac{1}{W}\sum_{w=1}^{W} \arg\left\{\max_{\mathbf{x}^{(w)}}\left[\wp\left(\mathbf{x}^{(w)}, \mathbf{a}|\mathbf{y}^{(w)}\right)\right]\right\} \tag{2.14}$$

$\mathbf{a} \in \mathbb{R}^{K\times 1}$ being a shared hyper-parameter vector [42]. By following an approach similar to the BCS [42], the optimal value of the hyper-parameter vector $\mathbf{a}$ is computed by maximizing the marginal likelihood function

$$\mathcal{L}_{MT}(\mathbf{a}) = -\frac{1}{2}\sum_{w=1}^{W}\{\log\left[\det\left(\mathbf{C}_{MT}\right)\right] + \tag{2.15}$$

$$(N + 2\beta_1)\log\left[\left(\mathbf{y}^{(w)}\right)^T \mathbf{C}_{MT}\mathbf{y}^{(w)} + 2\beta_2\right]\Big\} \tag{2.16}$$

where $\mathbf{C}_{MT} = \mathbf{I} + \mathbf{A}\mathrm{diag}\left(\mathbf{a}\right)^{-1}\mathbf{A}^T$, $\mathbf{C}_{MT} \in \mathbb{R}^{M\times M}$, and β_1, β_2 are user-defined parameters. Finally, the $MT-BCS$ estimation of the signal $\mathbf{x}$ is computed as [42]

$$\mathbf{x}_{MT} = \frac{1}{W}\sum_{w=1}^{W}\left\{\left[\mathbf{A}^T\mathbf{A} + \mathrm{diag}\left(\mathbf{a}_{MT}\right)\right]^{-1}\mathbf{A}^T\mathbf{y}^{(w)}\right\} \tag{2.17}$$

Like in the $ST-BCS$ case, all the terms are unknown except the shared hyperparameter vector $\mathbf{a}_{MT}$, whose value is computed by applying a suitable RVM strategy applied to the multi-task case.

Chapter 3

Real-Weight Sparse Linear Array Synthesis by Bayesian Compressive Sensing

An innovative methodology for the synthesis of sparse linear arrays with prescribed pattern features is numerically analyzed when dealing with large aperture layouts. The technique is based on a probabilistic formulation of the synthesis problem which is solved through a Bayesian Compressive Sensing (BCS) technique. A set of numerical experiments are presented to assess the features and potentialities of the BCS design approach when layouts comprising several hundred elements are at hand.

3.1 Introduction and Motivation

The design of satellite communication systems, radars, and devices for biomedical imaging and remote sensing applications usually imposes severe constraints on the pattern features (in terms of peak sidelobe level, directivity, beam footprint and shape) of the array to be deployed [4]. The necessity of achieving these goals while obtaining inexpensive, light and simple architectures, especially when dealing with large antenna systems, has lead to the introduction of sparse arrays [4]. Despite their advantages, however, sparse layouts have the main limitation that they yield a reduced control of the beam shape [4, 11, 24, 9, 19, 29, 72, 18, 28]. In order to address this drawback, several different techniques have been proposed either for the minimization of the PSL in thinned arrays [11, 29, 72, 18, 28], or for the synthesis of maximally-sparse arrays with prescribed pattern characteristics [9, 19, 14]. While the first problem has been widely studied [4, 11, 29, 72, 18, 28], only few techniques have been introduced for the solution of the latter [21]. In this framework, numerically inexpensive approaches, such as the steepest descent method, the iterative least-square technique, the simplex search, and the linear programming, were among the first methodologies applied to sparse array design [14, 21]. However, these techniques exhibit some drawbacks in terms of flexibility, required a-priori information, and final obtained performances [21]. More recently, in order to overcome these limitations, the simulated annealing [19] and the Matrix Pencil Method [14] have been successfully applied to the design of sparse arrays with prescribed pattern features. Nevertheless, despite their excellent performances, these methodologies can lead either to high computational costs [19] or to sub-optimal performances when dealing with shaped beams [14]. An innovative approach for the synthesis of sparse arrays with prescribed pattern features has been recently proposed [21]. This methodology is based on the formulation of the sparse array synthesis problem as a "Compressive Sensing (CS) retrieval" one, in which the sparseness constraints are imposed on the final array layout [21]. The arising CS problem is then recast in a probabilistic framework exploiting the so-called Bayesian Compressive Sensing formulation [41], and then solved by means of an efficient Relevance Vector Machine (RVM) [45]. Thanks to this approach, *BCS* sparse array synthesis has proved to be effective in dealing with standard and reference sparse array synthesis problems [21]. However, an analysis of its performances (in terms of pattern matching accuracy and computational complexity) when dealing with large aperture arrays has never been presented. This Chapter is aimed at analyzing the performances, features and limitations of the *BCS*-based technique when dealing with the design of sparse arrays displaced over apertures of width up to several hundred wavelengths. Towards this end, a set of array synthesis problems dealing with different layouts and patterns features are presented to assess the potentialities and drawbacks of the considered technique.

3.2 Mathematical Formulation

The problem of finding the sparsest (real and symmetric [21]) linear array with desired radiating properties can be cast in terms of a pattern matching one as follows [21]:

Synthesis Problem: given a set of K samples of a reference pattern $\mathbf{F}_{REF} = [F_{REF}(u_1), ..., F_{REF}(u_k), ..., F_{REF}(u_K)]$, and a fidelity factor ε, find the sparsest set of array weights $\mathbf{w} = [w_1, ..., w_N]$ such that $\|\mathbf{F}_{REF} - \mathbf{F}_{BCS}\|^2 < \varepsilon$.

where $\mathbf{F}_{BCS} = [F_{BCS}(u_1), ..., F_{BCS}(u_k), ..., F_{BCS}(u_K)]$ is the vector of the samples of the sparse array radiation pattern, whose k-th component is

$$F_{BCS}(u_k) = \sum_{n=1}^{N} w_n \nu_n \cos\left(\frac{2\pi d_n u_k}{\lambda}\right), \; k = 1, ..., K \tag{3.1}$$

λ is the wavelenght, u_k $(k = 1, ..., K)$ are the matching angles, d_n $(n = 1, ..., N)$ are the allowed positions for the sparse array elements and ν_n is the Neumann's number [21, 11]. By modeling the radiation pattern as a Gaussian random variable [21], the above synthesis problem can be recast in the framework of BCS to obtain the following equivalent one [21]:

BCS Problem: given $\mathbf{F}_{REF}$, find $\mathbf{w}$, $\mathbf{a}$ and σ^2 which maximize the a-posteriori probability $p([\mathbf{w}, \mathbf{a}, \sigma^2] | \mathbf{F}_{REF})$.

where $\mathbf{a}$ and σ^2 are, respectively, the hyperparameter vector [45] and the estimated fidelity variance [21]. Following the RVM approach [41, 45], this BCS problem is then solved by the following procedure [21]:

1. *Input phase:* define the reference pattern samples $\mathbf{F}_{REF}$, the set of admissible element locations d_n $(n = 0, ..., N)$, and the initial estimate of the fidelity variance;

2. *Matrix Definition:* calculate the problem $\mathbf{\Phi} \in \mathbb{C}^{K \times N}$, with $\Phi(k, n) = \nu_n \cos\left(\frac{2\pi d_n u_k}{\lambda}\right)$;

3. *Hyperparameter Posterior Modes Estimation:* find a and σ^2 according to the RVM procedure [21];

4. *Array weights estimation:* find the optimal sparse weights by $\mathbf{w} = \mathbf{\Xi}\mathbf{\Phi}^H \mathbf{E}_{REF}/\sigma^2$, where $\mathbf{\Xi} = \left(\mathbf{A} + \frac{\mathbf{\Phi}\mathbf{\Phi}^H}{\sigma^2}\right)^{-1}$ and $\mathbf{A} = \mathrm{diag}(\mathbf{a})$.

3.3 Numerical Results

In order to assess the performances of the BCS design method when dealing with large aperture arrays, a set of experiments has been carried out considering either Dolph or Taylor reference patterns [4], and evaluating, for each obtained design, the pattern matching error

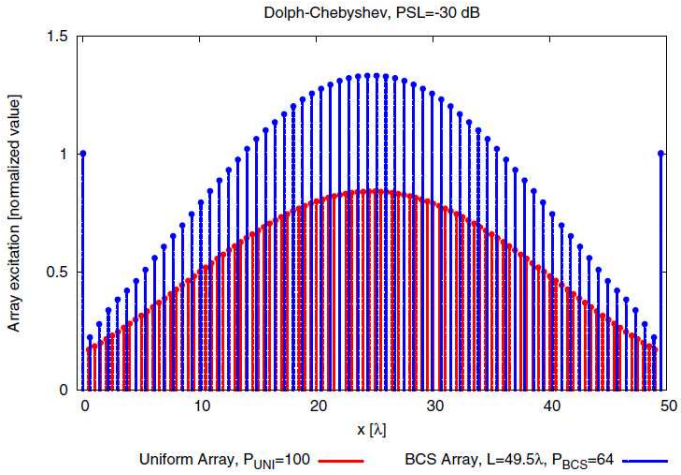
$$\xi \triangleq \frac{\int_{-1}^{1} |F_{REF}(u) - F_{BCS}(u)|^2 \, \mathrm{d}u}{\int_{-1}^{1} |F_{REF}(u)|^2 \, \mathrm{d}u}, \tag{3.2}$$

where $F_{BCS}(u)$ and $F_{REF}(u)$ are the sparse-array pattern and the reference pattern, respectively. Moreover, the number of elements of the sparse layout P_{BCS} and the total array size L_{BCS} have been compared to those obtained with a uniform layout, as well.

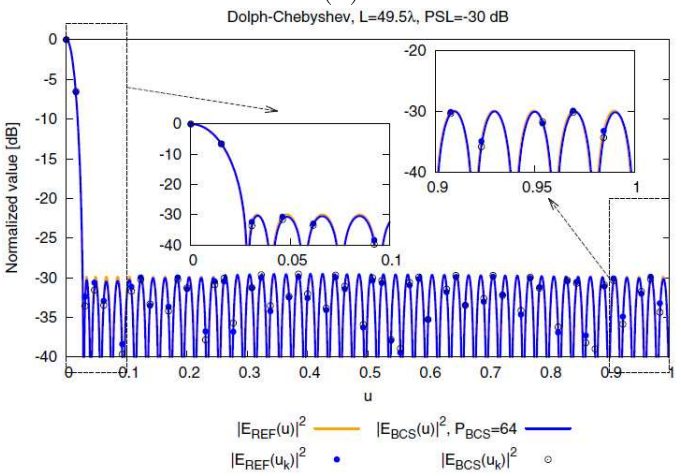
As a first numerical example, the synthesis of a sparse array exhibiting a Dolph pattern (uniform array aperture $L = 49.5\ \lambda$, $PSL = -30\ dB$) has been considered. The final obtained result (Fig. 3.1) indicates that the considered methodology is able to achieve a good accuracy [Fig. 3.1(b)], despite the exploitation of a reduced number of radiating elements [$P_{BCS} = 64$ - Fig. 3.1(a)]. This is actually confirmed by the achieved fidelity factor, which shows a matching error below 0.1% ($\xi = 5.3 \times 10^{-5}$ - Tab. 3.1), as well as by the obtained PSL value, which turns out very close to the reference one ($PSL_{BCS} = -29.5\ dB$ - Tab. 3.1). Moreover, the uniform and sparse weight arrangements indicate that a similar envelope is actually followed by both arrays [Fig. 3.1(a)], therefore suggesting that the BCS method actually samples in a nonuniform fashion the same Dolph distribution of the uniform layout.

It is also worth pointing out that a reduced synthesis time is observed in this case ($\Delta t = 0.23\ [s]$ - Tab. 3.1) notwithstanding the non-negligible problem size and the exploitation of a laptop for the synthesis (all the simulations have been performed on a single core PC running at 2.16 GHz). Similar conclusions can be drawn when dealing with a Taylor reference pattern for the same aperture ($PSL = -30\ dB$, 'transition index' $T = 6$). Indeed, the significant element reduction ($P_{BCS} = 66$ - Tab. 3.1), the numerical efficiency ($\Delta t = 0.25\ [s]$ - Tab. 3.1), and the good matching accuracy both in terms of fidelity factor ($\xi = 7.8 \times 10^{-5}$ - Tab. 3.1) as well as of PSL (which actually turns out improved - Tab. 3.1) are confirmed despite the presence of very low sidelobes at endfire [right inset of Fig. 3.2(b)]. Moreover, it is again noteworthy that the uniform and sparse layouts exhibit a similar weight envelope in the whole aperture, although the BCS yields a nonuniformly sampled architecture [Fig. 3.2(a)].

As a final numerical experiment, the synthesis of a $L = 499.5\ \lambda$, $PSL = -50\ dB$, $T = 6$ Taylor pattern has been considered to investigate the features of the considered methodology when dealing with significantly larger apertures and lower sidelobe levels. Also in this case, the plot of obtained sparse-array pattern [Fig. 3.3(b)] indicates that a good matching accuracy is obtained in the whole visible range ($\xi = 6.4 \times 10^{-5}$ - Tab. 3.1), despite the reduced number of radiating elements of the sparse layout ($P_{BCS} = 628$ - Tab. 3.1). Moreover, while the sparse

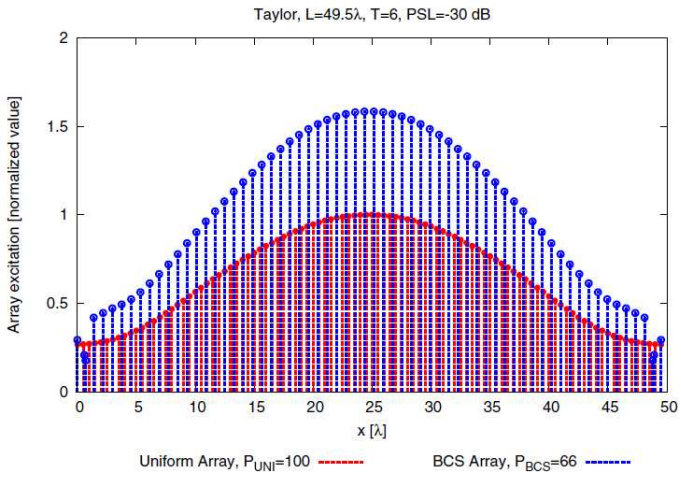


(a)

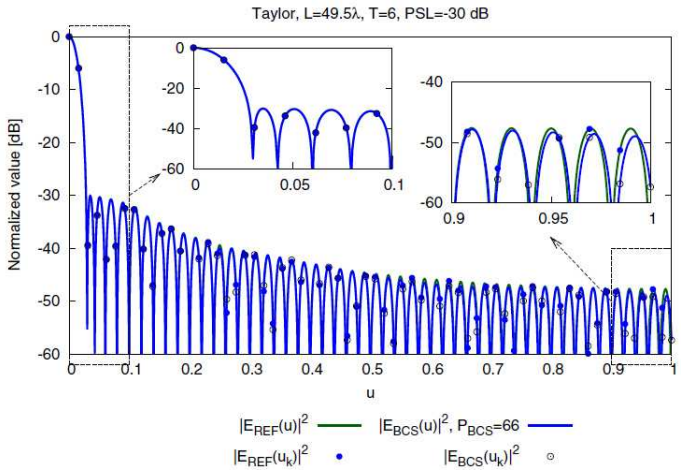


(b)

Figure 3.1: Dolph reference pattern ($L = 49.5\lambda$, $PSL = -30\ dB$) - Array layouts (a) and power pattern (b) of the reference and obtained array.

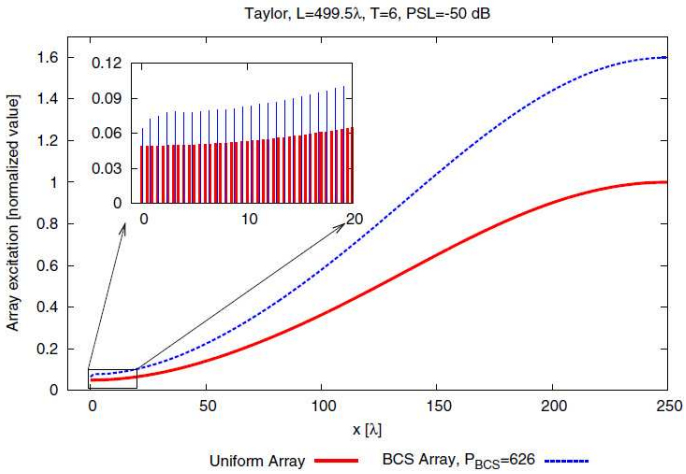


(a)

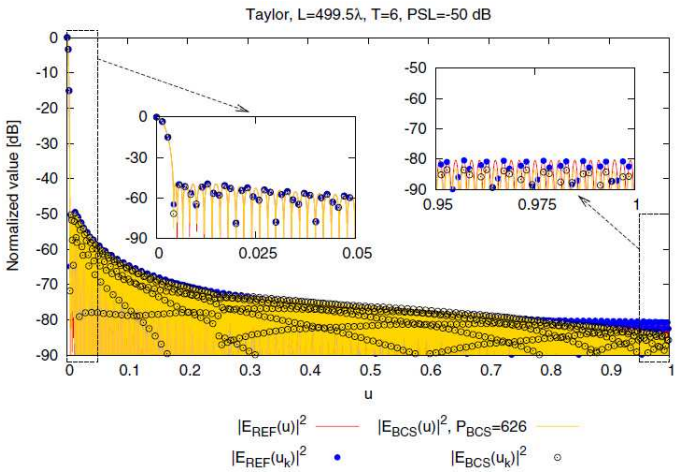


(b)

Figure 3.2: Taylor reference pattern ($L = 49.5\lambda$, $PSL = -30\ dB$, $T = 6$) - Array layouts (a) and power pattern (b) of the reference and obtained array.



(a)



(b)

Figure 3.3: Taylor reference pattern ($L = 499.5\lambda$, $PSL = -50\ dB$, $T = 6$) - Array layouts (a) and power pattern (b) of the reference and obtained array.

layout turns out slightly smaller than the reference one ($L_{BCS} = 499.3\ \lambda$ - Tab. 3.1) the above observations regarding the similarity of the envelopes shown by the BCS sparse and uniform layouts still hold true [Fig. 3.3(a)].

It is even more interesting to notice that, despite the wide aperture comprising several hundred elements, such synthesis was quite efficient also from the numerical viewpoint ($\Delta t = 2.24\ [s]$ - Tab. 3.1). These results further confirm the effectiveness and efficiency of the BCS synthesis approach in the design of large sparse layouts possibly comprising several hundreds elements.

Reference Pattern				*BCS*				
Test Case	L $[\lambda]$	PSL	P_{UNI}	L_{BCS} $[\lambda]$	PSL	$\frac{P_{BCS}}{P_{UNI}}$	ξ $[\times 10^{-5}]$	Δt
Fig. 3.1	49.5	-30	100	49.5	-29.5	0.64	5.3	0.23
Fig. 3.2	49.5	-30	100	49.5	-30.1	0.66	7.8	0.25
Fig. 3.3	499.5	-50	1000	499.3	-49.2	0.62	6.4	2.24

Table 3.1: Array Performance Indexes.

3.4 Discussions

The synthesis of large sparse linear arrays with prescribed pattern features has been carried out through an innovative methodology based on a Bayesian Compressive Sensing framework. The design approach, which formulates the synthesis problem in a probabilistic framework and then exploit a fast Relevance Vector Machine for its solution, has been numerically assessed when dealing with layouts possibly comprising several hundred elements. The presented analysis has shown that

- sparse layouts providing a good pattern fidelity ($\xi < 10^{-4}$) can be easily synthesized through the BCS methodology also when apertures of several hundred wavelengths are at hand (Tab. 3.1);
- the synthesis approach turns out efficient whatever the aperture size ($\Delta t < 3$ $[s]$ - Tab. 3.1);
- the arising sparse layouts usually exhibit an envelope close to that of their uniform counterparts, therefore indicating that the BCS method effectively tends to 'nonuniformly' sample the same current distribution [Figs. 3.1(a), 3.2(a), 3.3(a)].

Chapter 4

Complex-Weight Sparse Linear Array Synthesis by Multitask Bayesian Compressive Sensing

In this Chapter, an innovative method for the synthesis of maximally sparse linear arrays matching arbitrary reference patterns is proposed. In the framework of sparseness constrained optimization, the approach exploits the multi-task (MT) Bayesian Compressive Sensing (BCS) theory to enable the design of complex non-Hermitian layouts with arbitrary radiation and geometrical constraints. By casting the pattern matching problem into a probabilistic formulation, a Relevance-Vector-Machine (RVM) technique is used as solution tool. The numerical assessment points out the advances of the proposed implementation over the extension to complex patterns of [21] and it gives some indications about the reliability, flexibility, and numerical efficiency of the $MT-BCS$ approach also in comparison with state-of-the-art sparse-arrays synthesis methods.

4.1 Introduction and Motivation

Radar tracking, biomedical imaging, satellite and ground communications, and remote sensing applications require antenna patterns with suitable sidelobes, null positions, mainlobe size and shape, and directivity [4]. To synthesize shaped-beam arrays, several approaches especially concerned with uniformly-spaced arrangements [4][5][6][7][8] have been proposed over the last sixty years. Although successful in some applications, uniform arrays have the limitation to be expensive and heavy when wide apertures are at hand [4] since a huge amount of radiating elements spaced by $\frac{\lambda}{2}$ are needed to avoid grating lobes [9]. Therefore, non-uniform arrangements have been naturally proposed [9][10][11][12][13][14] because of their advantages over their regularly-spaced counterparts (e.g., resolution [15], sidelobe level control/reduction [16], and efficiency in dealing with physically constrained geometries [17]). State-of-the-art solutions usually consider *thinned* regular arrangements to yield a minimum peak sidelobe level (PSL) [4][13][18][19] or *sparse* layouts with the minimum number of radiating elements given a desired pattern [14][20][21]. Whether several techniques as random thinning [22][23], dynamic programming [24], genetic algorithms [25][26][27], analytical approaches [13][18][28], and hybrid methodologies [29][30][72][32] have been investigated for array thinning, few methods have been so far proposed for synthesizing sparse arrangements [14][20][21][33][34]. As for these latter, steepest descent [35], iterative least squares [36], simplex search [9], and linear programming [37] methodologies have been firstly developed because of their efficiency. Improved performances have been successively reached by using recursive inversion techniques [38][39], stochastic optimizers [20], generalized Gaussian quadrature approaches [40], and the matrix pencil method (MPM) [14][33][34]. More recently, a new approach based on the Bayesian Compressive Sensing (BCS) [41] has been proposed for the design of sparse layouts matching user-defined reference patterns [21]. The so-called "BCS technique" has been formulated starting from a probabilistic description of the array synthesis [21] then solved by exploiting an efficient fast relevance vector machine (RVM) [41]. Thanks to its efficiency, the BCS syntheses usually positively compares with state-of-the-art methodologies in terms of flexibility, synthesis time, and number of array elements, while guaranteeing an excellent pattern matching [21]. However, such a formulation deals with symmetric purely-real arrangements and its extension to complex syntheses is not efficient because of the real-valued nature of the BCS solver itself [41]. Consequently, this Chapter is aimed at proposing, still in the framework of the probabilistic sparseness constrained optimization, an innovative, flexible, and numerically efficient complements to state-of-the-art approaches for the synthesis of maximally sparse linear arrays matching a (possibly complex) reference pattern. Following the guidelines in [21] to recast the complex-valued synthesis in probabilistic terms and suitably reformulating the original pattern-matching problem in an equivalent 'fictitious' one (Eq. 4.11), a multi-task Bayesian Com-

pressive Sensing ($MT-BCS$) methodology [42] is applied. Unlike the BCS extension where the real and the imaginary components of the sparse excitation vector are dealt with as independent, a "shared-prior" [42] is exploited to enforce the synthesis of complex excitations rather than purely real and/or imaginary weights.
This Chapter is organized as follows. The sparse synthesis of complex-weight linear arrays is mathematically formulated in a probabilistic fashion and the $MT-BCS$ method is presented (Sect. 4.2). Representative results of an extensive set of numerical simulations are presented to validate the proposed approach, to assess its advances over the BCS extension to complex patterns, and to compare its performances with those of state-of-the-art techniques (Sect. 4.3). Finally, some conclusions are drawn (Sect. 4.4).

4.2 Mathematical Formulation

4.2.1 Array Synthesis Problem

The problem of synthesizing a (complex and non-symmetric) sparse linear array with a prescribed radiated pattern can be formulated as follows [21]

Array Synthesis Problem - Find the minimum P value and the corresponding sparse array descriptors $\mathbf{v} = \{v_p;\, p = 1, ..., P\}$ and $\mathbf{l} = \{l_p;\, p = 1, ..., P\}$ that satisfy the matching constraint

$$\sum_{k=1}^{K} \left| F_{REF}\left(u_k\right) - \sum_{p=1}^{P} v_p \exp\left(i2\pi l_p u_k\right) \right|^2 \leq \epsilon. \tag{4.1}$$

In (4.1), ϵ is the "fidelity factor", v_p and l_p are the complex ($v_p \in \mathbb{C}$) weight and the position in wavelengths ($l_p \in \mathbb{R}$) of the p-th array element, respectively, while $F_{REF}\left(u_k\right) \in \mathbb{C}$ is the k-th ($k = 1, ..., K$) sample of the reference pattern at the observation angle u_k within the angular range $[-1, 1]$. Similarly to [9][21], the P element positions are assumed to belong to a user-chosen set of N ($N \gg P$) arbitrary candidate locations $\mathbf{d} = \{d_n;\, n = 1, ..., N\}$ to straightforwardly integrate geometrical constraints in the synthesis process [21]. Equation (4.1) is then recast into the following sparse matrix form [43][21]

$$\mathbf{F}_{REF} - \Phi\mathbf{w} = \mathcal{D} \tag{4.2}$$

by introducing the sparse[1] weight vector $\mathbf{w} = \{w_n;\, n = 1, ..., N\}$

$$w_n = \begin{cases} v_p & if\ d_n = l_p \\ 0 & otherwise \end{cases}, \tag{4.3}$$

[1]It is worth remarking that $\mathbf{w}$ turns out a *sparse* vector since $N \gg P$.

Step 0. Get input values of N and w_n $(n = 1, ..., N)$;
Step 1. Set $p = 1$, $n = 1$;
Step 2. If $w_n \neq 0$, set $l_p = d_n$, $v_p = w_n$, and $p = p + 1$;
Step 3. If $n < N$, set $n = n + 1$ and goto **2.**; else goto **4.**
Step 4. Return output values of P, l_p and v_p $(p = 1, ..., P)$

Figure 4.1: Computation of the complex weights $v_p \in \mathbb{C}$ and element positions $l_p \in \mathbb{R}$ $(p = 1, .., P)$ starting from the sparse vector $\mathbf{w} \in \mathbb{C}^N$.

where $\mathbf{F}_{REF} = \{F_{REF}(u_k); k = 1, ..., K\}$, $\mathcal{D} = \{\Delta_k; k = 1, ..., K\}$ is a vector of zero-mean complex Gaussian entries with variance σ^2 proportional to ϵ [41][44][43], and

$$\Phi \triangleq \begin{bmatrix} \exp\left(\frac{i2\pi d_1 u_1}{\lambda}\right) & \cdots & \exp\left(\frac{i2\pi d_N u_1}{\lambda}\right) \\ \vdots & \ddots & \vdots \\ \exp\left(\frac{i2\pi d_1 u_K}{\lambda}\right) & \cdots & \exp\left(\frac{i2\pi d_N u_K}{\lambda}\right) \end{bmatrix} \tag{4.4}$$

is the "observation matrix" [41]. Thanks to this "sparse" description, the *Antenna Synthesis Problem* can be also formulated as follows

Sparse Vector Synthesis Problem - Find the minimum ℓ_0-norm weight vector $\mathbf{w}$ $(\mathbf{w} \in \mathbb{C}^N)$ that satisfies (4.2)

where

$$\|\mathbf{w}\|_{\ell_0} \triangleq \sum_{n=1}^{N} |w_n|^0 = \sum_{p=1}^{P} |v_n|^0 = P. \tag{4.5}$$

Once $\mathbf{w}$ is found, the unknowns $\mathbf{v}$ and $\mathbf{l}$ of the *Antenna Synthesis Problem* are computed as detailed in Fig. 4.1.

4.2.2 *BCS* Synthesis Method

The solution of the *Sparse Vector Synthesis Problem* cannot be yielded through the method described in [21], since the *BCS* approach addresses purely real-valued problems [21][41], while (4.2) generally includes complex-valued vectors and matrices. To directly extend the approach in [21] to the complex formulation at hand, Equation (4.2) is manipulated as follows

$$\widetilde{\mathbf{F}}_{REF} - \widetilde{\Phi}\widetilde{\mathbf{w}} = \widetilde{\mathcal{D}} \tag{4.6}$$

by defining $\widetilde{\mathbf{w}} = [\mathcal{R}\{\mathbf{w}\}, \mathcal{I}\{\mathbf{w}\}]$ $(\widetilde{\mathbf{w}} \in \mathbb{R}^{2N})$, $\widetilde{\mathbf{F}}_{REF} = [\mathcal{R}\{\mathbf{F}_{REF}\}, \mathcal{I}\{\mathbf{F}_{REF}\}]$ $(\widetilde{\mathbf{F}}_{REF} \in \mathbb{R}^{2K})$, $\widetilde{\mathcal{D}} = [\mathcal{R}\{\mathcal{D}\}, \mathcal{I}\{\mathcal{D}\}]$ $(\widetilde{\mathcal{D}} \in \mathbb{R}^{2K})$, and $\widetilde{\Phi} = \begin{bmatrix} \mathcal{R}\{\Phi\} & -\mathcal{I}\{\Phi\} \\ \mathcal{I}\{\Phi\} & \mathcal{R}\{\Phi\} \end{bmatrix}$ $(\widetilde{\Phi} \in \mathbb{R}^{2K \times 2N})$, where $\mathcal{R}\{\cdot\}$ and $\mathcal{I}\{\cdot\}$ stand for the real and the imaginary part, respectively. Accordingly, the following extended real-valued problem can be then formulated

BCS **'Deterministic' Synthesis Problem -** Find the minimum ℓ_0-norm "extended" weight vector $\widetilde{\mathbf{w}}$ ($\widetilde{\mathbf{w}} \in \mathbb{R}^{2N}$) that satisfies (4.6).,

and successively expressed in the *probabilistic* framework [21]

BCS **'Probabilistic' Synthesis Problem -** Find the minimum ℓ_0-norm "extended" weight vector $\widetilde{\mathbf{w}}$ ($\widetilde{\mathbf{w}} \in \mathbb{R}^{2N}$) subject to

$$\widetilde{\mathbf{w}}^{BCS} = \arg\left[\max_{\widetilde{\mathbf{w}}} \mathcal{P}\left(\widetilde{\mathbf{w}} | \, \widetilde{\mathbf{F}}_{REF}\right)\right] \tag{4.7}$$

whose (real-valued) solution is given by [21]

$$\widetilde{\mathbf{w}}^{BCS} = \frac{1}{\widetilde{\sigma}^2_{BCS}} \left(\frac{\widetilde{\Phi}^T \widetilde{\Phi}}{\widetilde{\sigma}^2_{BCS}} + \widetilde{\mathbf{a}}^{BCS}\right)^{-1} \widetilde{\Phi}^T \widetilde{\mathbf{F}}_{REF} \tag{4.8}$$

where $\widetilde{\sigma}^2_{BCS}$ is the estimated variance of Δ_k ($k = 1, ..., K$) and $\widetilde{\mathbf{a}}^{BCS}$ ($\widetilde{\mathbf{a}}^{BCS} \in \mathbb{R}^{2N}$) is the hyperparameter vector, whose n-th entry, $\widetilde{a}^{BCS}_n$, controls the strength of the sparseness prior over $\widetilde{w}^{BCS}_n$ [45]. These parameters are computed by maximizing the logarithm of the *BCS* "marginal likelihood", $\mathcal{L}^{BCS}(\widetilde{\mathbf{a}}, \sigma^2)$ [21]

$$\begin{aligned} \mathcal{L}^{BCS}(\widetilde{\mathbf{a}}, \sigma^2) = -\tfrac{1}{2} \Big[(2N) \log 2\pi + \log \left|\widetilde{C}\right| + \\ + \widetilde{\mathbf{F}}^T_{REF} \left(\widetilde{C}\right)^{-1} \widetilde{\mathbf{F}}_{REF} \Big] \end{aligned} \tag{4.9}$$

where $\widetilde{C} \triangleq \sigma^2 I + \widetilde{\Phi} \left[\widetilde{A}\right]^{-1} \widetilde{\Phi}^T$, and $\widetilde{A} = diag\,(\widetilde{\mathbf{a}})$.

Finally, the N entries of the weight vector $\mathbf{w}^{BCS}$ ($\mathbf{w}_{BCS} \in \mathbb{C}^N$) are found as

$$w^{BCS}_n = \widetilde{w}^{BCS}_n + i\widetilde{w}^{BCS}_{n+N}, \quad n = 1, ..., N. \tag{4.10}$$

Equation (4.8) provides a direct extension of the method in [21] to deal with complex and non-symmetric arrays. However, such a solution bears an intrinsic limitation. The real ($\mathcal{R}\left\{w^{BCS}_n\right\} = \widetilde{w}^{BCS}_n$, $n = 1, ..., N$) and imaginary ($\mathcal{I}\left\{w^{BCS}_n\right\} = \widetilde{w}^{BCS}_{n+N}$, $n = 1, ..., N$) parts of the weights are managed as independent quantities - see Eq. (4.6) - since each $\widetilde{w}^{BCS}_n \in \mathbb{R}$ ($n = 1, ..., 2N$) is treated as statistically independent. See Eqs. (4)-(6) in [21]. This in turns leads to sparse *BCS* layouts where the array weights v_p ($p = 1, ..., P$) are often either purely real or purely imaginary, neglecting that sparse *complex* layouts frequently exhibit non-negligible real and imaginary components at the same array locations. Such a drawback generally does not enable the approach to synthesize very sparse layouts with a good reference pattern matching, as it has been confirmed by the numerical analysis whose representative results will be presented in Section 4.3.

4.2.3 $MT-BCS$ Synthesis Method

To overcome the limitations of *BCS Synthesis Method* (Sect. 4.2.2), the $MT-BCS$ approach [42] is exploited and suitably customized for statistically modelling the relations between the real and imaginary parts of the array weights. Towards this end, Equation (4.2) is firstly rewritten in terms of the *fictitious* weights vectors $\mathbf{w}_R \triangleq \mathcal{R}\{\mathbf{w}\}$ and $\mathbf{w}_I \triangleq \mathcal{I}\{\mathbf{w}\}$ $(\mathbf{w}_I, \mathbf{w}_R \in \mathbb{R}^N)$

$$\begin{cases} \widehat{\mathbf{F}}_R - \widehat{\Phi}\mathbf{w}_R = \widehat{\mathcal{D}}_R \\ \widehat{\mathbf{F}}_I - \widehat{\Phi}\mathbf{w}_I = \widehat{\mathcal{D}}_I \end{cases} \tag{4.11}$$

where $\widehat{\mathcal{D}}_R \in \mathbb{R}^{2K}$ and $\widehat{\mathcal{D}}_I \in \mathbb{R}^{2K}$ are zero-mean complex Gaussian error vectors (with variance $\frac{\sigma^2}{2}$) such that $\widehat{\mathcal{D}}_R + \widehat{\mathcal{D}}_I = \widetilde{\mathcal{D}}$, $\widehat{\Phi} \triangleq [\mathcal{R}\{\Phi\}, \mathcal{I}\{\Phi\}]$ is the MT observation matrix, while $\widehat{\mathbf{F}}_R = [\mathcal{R}\{\mathbf{F}_R\}, \mathcal{I}\{\mathbf{F}_R\}]$ and $\widehat{\mathbf{F}}_I = [\mathcal{R}\{\mathbf{F}_I\}, \mathcal{I}\{\mathbf{F}_I\}]$ $(\widehat{\mathbf{F}}_R, \widehat{\mathbf{F}}_I \in \mathbb{R}^{2K})$. Moreover, $\mathbf{F}_R \in \mathbb{C}^K$ and $\mathbf{F}_I \in \mathbb{C}^K$ satisfy the following condition

$$\mathbf{F}_R + i\mathbf{F}_I = \mathbf{F}_{REF}. \tag{4.12}$$

Accordingly, the *multi-task* (real-valued) problem turns out to be

$MT-BCS$ **'Deterministic' Synthesis Problem -** Find the minimum ℓ_0-norm "fictitious" weight vectors $\mathbf{w}_R$ and $\mathbf{w}_I$ $(\mathbf{w}_I, \mathbf{w}_R \in \mathbb{R}^N)$ that satisfy (4.11)

and as follows into the *probabilistic* framework [42]

$MT-BCS$ **'Probabilistic' Synthesis Problem -** Find the minimum ℓ_0-norm "fictitious" weight vectors $\mathbf{w}_R$ and $\mathbf{w}_I$ $(\mathbf{w}_I, \mathbf{w}_R \in \mathbb{R}^N)$ subject to

$$\begin{cases} \mathbf{w}_R^{MT-BCS} = \arg\left[\max_{\mathbf{w}_R} \mathcal{P}\left(\mathbf{w}_R | \widehat{\mathbf{F}}_R\right)\right] \\ \mathbf{w}_I^{MT-BCS} = \arg\left[\max_{\mathbf{w}_I} \mathcal{P}\left(\mathbf{w}_I | \widehat{\mathbf{F}}_I\right)\right] \end{cases}. \tag{4.13}$$

whose (real-valued) solution are given by

$$\mathbf{w}_H^{MT-BCS} = \left(diag\left(\widehat{\mathbf{a}}^{MT-BCS}\right) + \widehat{\Phi}^T\widehat{\Phi}\right)^{-1}\widehat{\Phi}^T\widehat{\mathbf{F}}_H, \quad H \in \{R, I\}, \tag{4.14}$$

while the corresponding estimated weight vector turns out to be

$$\mathbf{w}^{MT-BCS} = \mathbf{w}_R^{MT-BCS} + i\mathbf{w}_I^{MT-BCS}. \tag{4.15}$$

See the Appendix.

4.2.4 $MT-BCS$ Algorithmic Implementations

The algorithmic implementation of the $MT-BCS$ technique consists of the following steps (Fig. 4.2(b)):

Input Phase - Set the reference pattern $F_{REF}(u)$, the grid of admissible locations ($\mathbf{d}$), the set of pattern sampling points ($\mathbf{u}$), the target variance σ^2 of the error term $\mathcal{D}$, and the user-defined *scale priors* β_1 and β_2 (Eq. (A.4)) [42];

Matrix Definition - Fill the entries of the vectors $\widehat{\mathbf{F}}_R$, $\widehat{\mathbf{F}}_I$, $\widehat{\Phi}$, $\widehat{\mathcal{D}}_R$, and $\widehat{\mathcal{D}}_I$;

Hyperparameter Posterior Modes Estimation - Find $\widehat{\mathbf{a}}^{MT-BCS}$ by maximizing (A.15) [42];

Array Weights Estimation - Find $\mathbf{w}^{MT-BCS}$ by (4.15);

Output Phase - Compute P_{MT-BCS}, $\mathbf{v}^{MT-BCS}$, and $\mathbf{l}^{MT-BCS}$ (Fig 4.1).

By comparing the algorithmic descriptions of the BCS (Sect. III of [21] and Fig. 4.2(a)) and $MT-BCS$ (Sect. 4.2.4 - Fig. 4.2(b)), it is observed that both approaches require $\mathbf{d}$, $\mathbf{u}$, and σ^2, while the $MT-BCS$ needs the definition of the scale priors β_1 and β_2 instead of the initial estimates σ_0^2 as for the BCS. Thanks to these differences and unlike the BCS approach, the $MT-BCS$

- enables the *explicit* model and control of the relationships between the real and imaginary parts of the array weights thanks to the specification of β_1 and β_2 in (A.4);
- requires neither some *a-priori* knowledge/information on the noise (e.g., σ_0^2) nor the estimation of the noise level (i.e., $\widetilde{\sigma}^2$) for determining the problem solution.

4.3 Numerical Results

The objectives of this section are two-fold: On the one hand, it provides guidelines for applying the $MT-BCS$ method to the synthesis of sparse complex layouts. On the other hand, it assesses the method's effectiveness in both reducing the number of array elements and accurately matching reference patterns, with the assessment made by comparing the $MT-BCS$ results with those of other reliable, state-of-the-art (regular and sparse) array synthesis methodologies. For the assessment, the following quantities are analyzed: the normalized matching error, ξ,

$$\xi \triangleq \frac{\int_{-1}^{1} \left| F_{REF}(u) - \sum_{p=1}^{P} v_p \exp\left(i2\pi l_p u\right) \right|^2 \mathrm{d}u}{\int_{-1}^{1} \left| F_{REF}(u) \right|^2 \mathrm{d}u}, \tag{4.16}$$

the aperture length, L, ($L \triangleq |l_P - l_1|$), the mean ($\Delta L \triangleq L/P - 1$), and the minimum ($\Delta L_{\min} \triangleq \min_{p=1,\ldots,P-1}\{|l_{p+1} - l_p|\}$) inter-element spacing.

4.3.1 Sensitivity Analysis

The first set of numerical experiments is concerned with the sensitivity of the $MT-BCS$ synthesis on its control parameters, while the reader is referred

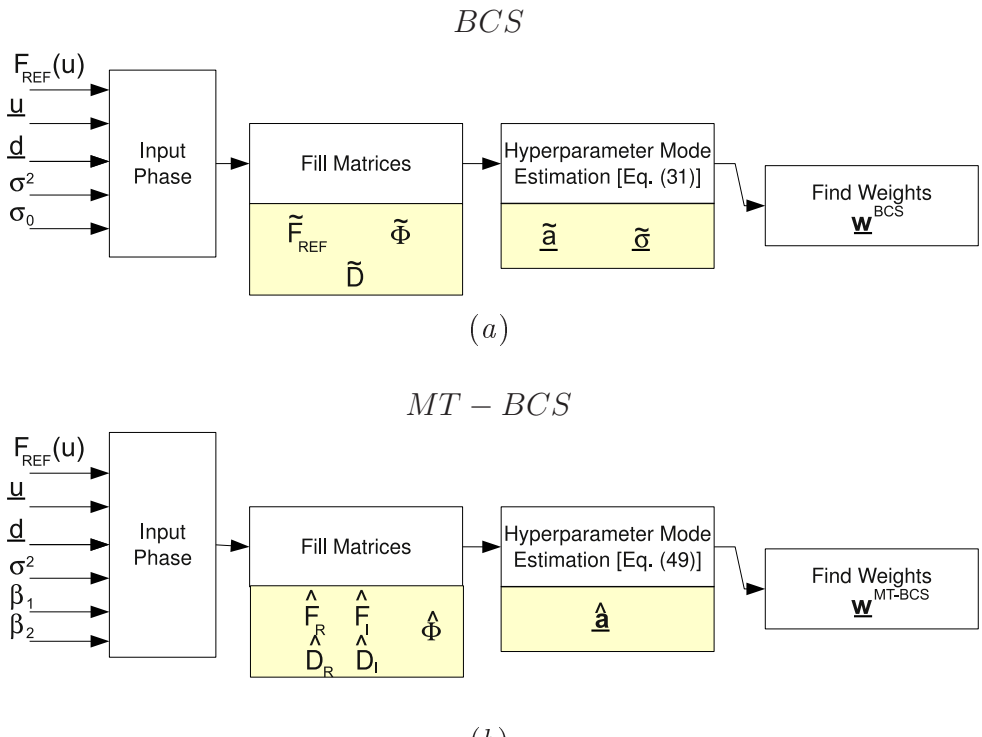


Figure 4.2: Sparse Synthesis Flowchart: (a) BCS method (σ_0 being the initial estimate of σ [21]) and (b) $MT-BCS$ method.

to [21] for the calibration of the BCS approach. Towards this purpose, the synthesis of a non-uniform array matching a complex-weight "cosecant" pattern with $L = 7.5\lambda$ and $PSL = -20$ dB is assumed as reference test case (Fig. 4.3(b)). Such a pattern can be synthesized by a uniform layout of $P_{UNI} = 16$ elements $\lambda/2$-spaced [5]. The $MT-BCS$ synthesis is carried out by assuming

$$u_k = -1 + \frac{2k}{K} \quad k = 1, ..., K \tag{4.17}$$

and setting the uniform grid of N candidate locations as follows $d_n = L\left(-\frac{1}{2} + \frac{n}{N}\right)$, $n = 1, ..., N$. Figure 4.3(a) shows the representative points of the synthesized $MT-BCS$ sparse layouts in the ξ-P_{MTBCS} plane, along with the associated Pareto front in such a plane, when varying the control parameters within the ranges: $N = \{25, ..., 800\}$, $K = \{10, ..., 30\}$, $\sigma^2 = \{10^{-5}, 5.0 \times 10^{-1}\}$, $\beta_1 = \{10^{-2}, 10^3\}$, and $\beta_2 = \{10^{-2}, 10^3\}$. These results show that the values of the pattern matching accuracy lie in the range $\xi \in [10^{-8}, 2]$ with a number of array elements ranging from a minimum of $P_{MTBCS} = 5$ up to a maximum of $P_{MTBCS} = 25$ (Fig. 4.3(a)). By analyzing the synthesized pattern for three Pareto solutions, namely $P_{MTBCS} = \{5, 13, 18\}$ [Fig. 4.3(b)], it turns out that the sparsest solution ($P_{MTBCS} = 5$) yields a poor approximation of the reference pattern as also confirmed by the corresponding matching error [$\xi = 2.86 \times 10^{-1}$ - Fig. 4.3(a)], while a good fitting is reached when $P_{MTBCS} = 13$ active elements are at hand [$\xi = 7.24 \times 10^{-5}$ - Fig. 4.3(a)]. A further reduction of the matching error [e.g., $\xi = 2.83 \times 10^{-7}$ - Fig. 4.3(a)] by using a larger number of elements ($P_{MTBCS} = 18$) does not provide significant improvements. Therefore, analogous to the guidelines deduced in [21], an accuracy index close to or below $\xi_{th} = 10^{-4}$ is identified as the optimal threshold for obtaining a suitable trade-off between pattern matching and reduction of the number of elements (i.e., $\frac{P_{MTBCS}}{P_{UNI}}$). As for the associated array structure, the optimal trade-off $MT-BCS$ layout (i.e., $P_{MTBCS} = 13$ - $\xi = 7.24 \times 10^{-5}$) exhibits a distribution of the array weights similar to that of the corresponding uniform architecture [5], although with a non-uniform, and larger, on the average, inter-element spacing [Figs. 4.3(c)-4.3(d)]. This suggests that the method performs an implicit non-uniform sampling of the ideal current distribution synthesizing $F_{REF}(u)$ [Fig. 4.3(c)-4.3(d)]. On the contrary, the non-optimal trade-off solutions differ quite significantly from the uniform distribution case [e.g., $\frac{L_{MTBCS}}{L_{UNI}} \approx 0.4$ when $P_{MTBCS} = 5$ - Fig. 4.3(c)].

Figure 4.4 completes the sensitivity analysis carried out for calibrating the $MT-BCS$. Each plot gives the values of ξ and P_{MTBCS} versus a control parameter (i.e., K, σ^2, β_1, β_2, and N) by setting the others to the optimal trade-off setup (i.e., $P_{MTBCS} = 13$ - $K = 33$, $N = 250$, $\sigma = 10^{-3}$, $\beta_1 = 10^3$, $\beta_2 = 10^2$).

By analyzing the behaviour of ξ as a function of K [Fig. 4.4(a)], it turns out that increasing the number of samples of the reference pattern up to the

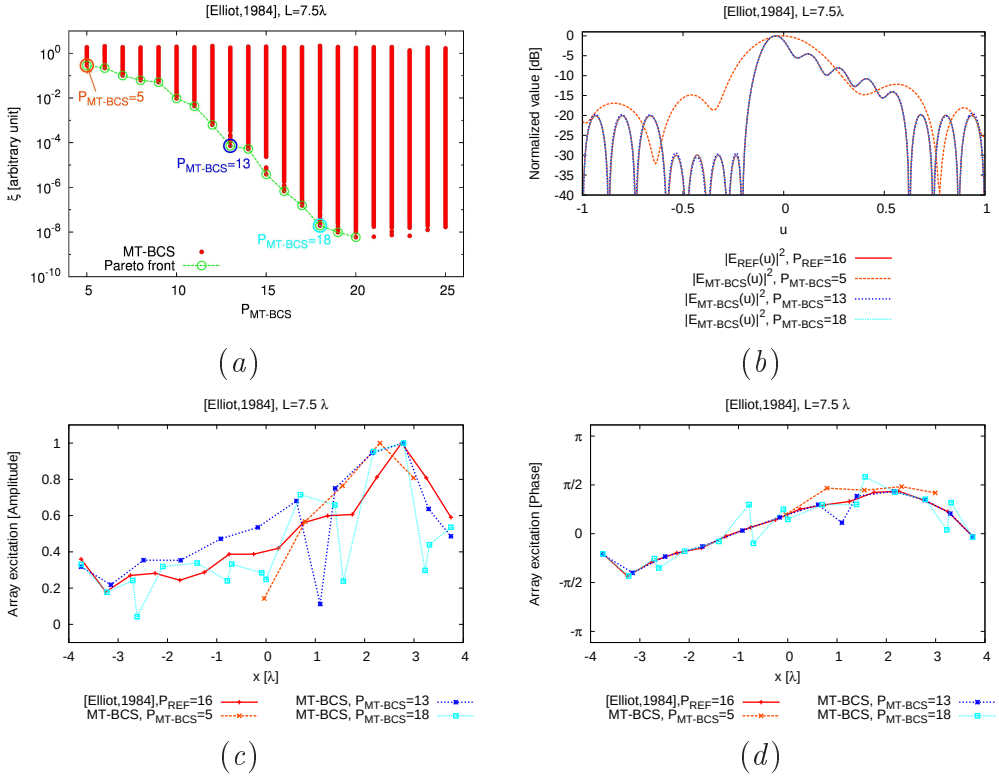


Figure 4.3: *MT−BCS Sensitivity Analysis* (*Shaped Pattern Synthesis:* $L = 7.5\lambda$ *[5]*) - Plot of the representative points of a set of $MT-BCS$ solutions in the (ξ, P_{MT-BCS}) plane (a). Power patterns (b), excitation amplitudes (c) and phases (d) of the reference and of the set of representative $MT-BCS$ arrays circled in (a).

Nyquist threshold ($K_{Nyquist} = 2 \times P_{UNI} - 1 = 31$ [14]) gives a non-negligible reduction of the matching error ξ, while further increments only slightly modify the matching accuracy or P_{MTBCS}. Accordingly, a sampling threshold within $K \in [K_{Nyquist}, 1.2K_{Nyquist}]$ has been assumed in the following analyses.

Concerning the dependence of ξ and P_{MTBCS} on σ^2, Figure 4.4(b) shows that the values of the two indexes are almost constant when $\sigma^2 \leq 3 \times 10^{-2}$, while they increase otherwise. Such a behavior is actually expected from the $MT-BCS$ theory. See Sect. 4.2.3 and the Appendix. Indeed, larger σ^2 values correspond to less accurate pattern approximations [see Eq. (4.2)] as well as less sparsely filled layouts. Consequently, good trade-offs between accuracy and sparseness are expected by choosing $\sigma^2 \in [10^{-4}, 10^{-2}]$.
With reference to the $MT-BCS$ sensitivity to the scale prior β_1, ξ reduces as the prior value is enlarged [Fig. 4.4(c)], even though such a matching improvement is obtained by increasing the number of radiating elements when $\beta_1 > 10^4$ [Fig. 4.4(c)]. Larger values of β_2 yield more sparsely filled layouts, while smaller priors provide higher accuracies [Fig. 4.4(d)]. Consequently, the ranges for the scale priors have been set to $\beta_1 \in [10^2, 10^4]$ and $\beta_2 \in [5 \times 10^1, 5 \times 10^2]$, respectively.
As far as the lattice grid is concerned, Figure 4.4(e) shows that the matching accuracy is quite stable if $N \gtrsim 2\frac{L_{UNI}}{\lambda}$, while larger/smaller N values result in a sharp increase of P_{MTBCS}/ξ. This is mainly caused by the increased numerical complexity of the problem at hand since its size grows with N. A trade-off value within $N \in \left[\frac{5L_{UNI}}{\lambda}, \frac{50L_{UNI}}{\lambda}\right]$ is then suggested.
The obtained tradeoff margins range from a $1:1.2$ ratio [for K - Fig. 4.4(a)] to a $1:100$ ratio [for β_1 and σ^2 - Figs. 4.4(b) and 3.4(c)]. Such a behaviour, caused by the different physical meaning of each parameter (see discussion above), does not actually represent a big issue for the proposed design methodology. In fact, quite wide ranges exist for which the method performances are almost constant. Furthermore, the MT-based BCS exhibits a "smoother" dependence on its control parameters than the single-task BCS approach. Indeed, unlike the BCS [21], ξ generally exhibits nearly monotone behaviour versus control parameters [e.g., Figs. 4.4(a)-4.4(e)] and P_{MTBCS} presents reduced oscillations given very large parameter variations [e.g., Fig. 4.4(c)]. Thus $MT-BCS$ provides better stability and robustness than BCS for any reference pattern or aperture.

4.3.2 $MT-BCS$ Assessment

For numerical assessment, we consider both unconstrained (Sect. 4.3.2.1) and constrained problems (Sect. 4.3.2.2), where forbidden regions are defined in the pattern region (Sect. 4.3.2.2.1) or on the array geometry (Sect. 4.3.2.2.2). Concerning the unconstrained syntheses, the analysis aims at performing a consistency check to assess the reliability of the $MT-BCS$ in dealing with problems also manageable by the original BCS approach [21] (Sect. 4.3.2.1.1) and successively detailing the $MT-BCS$ performance applied to the synthesis of arbitrary

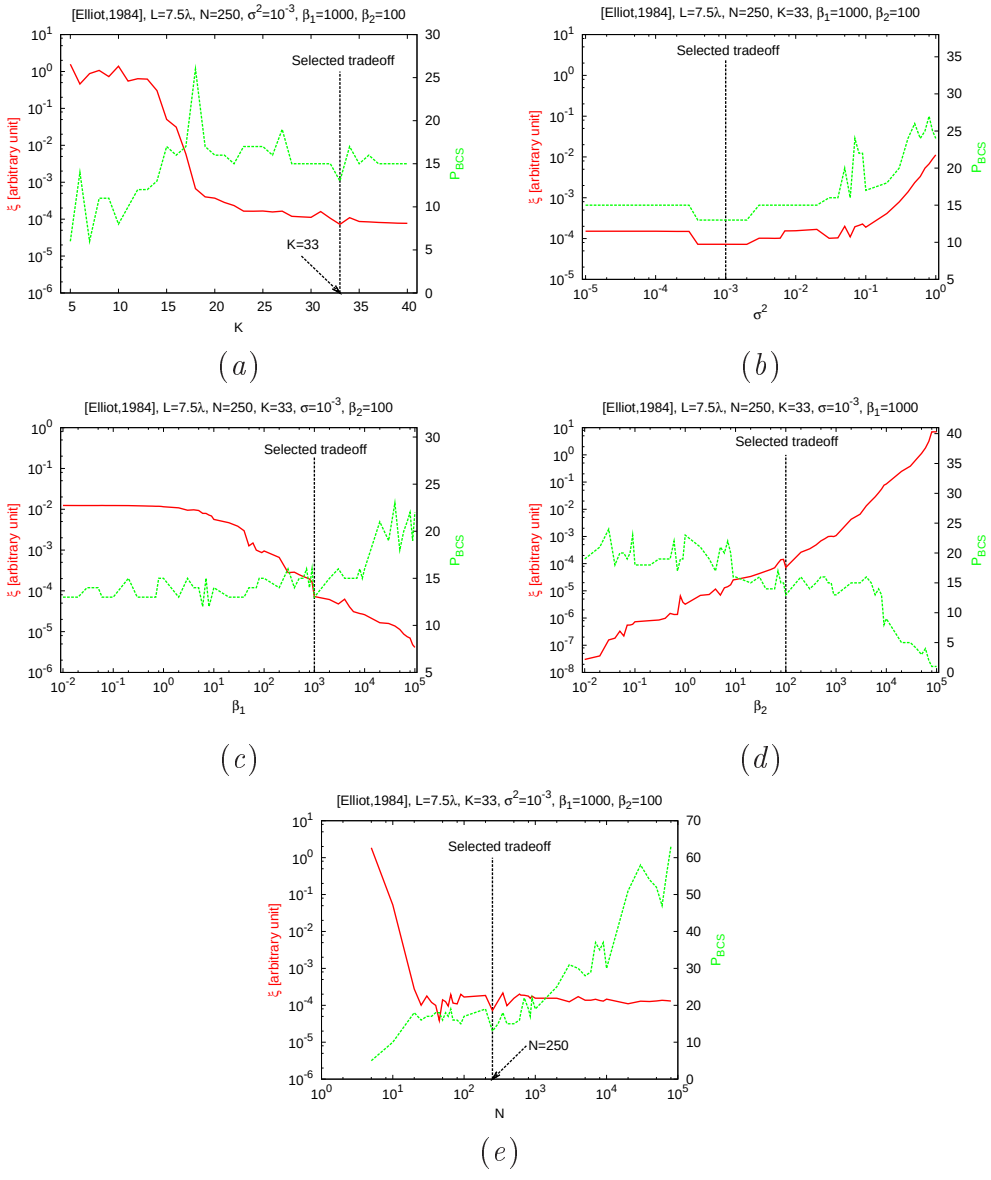


Figure 4.4: *MT−BCS Sensitivity Analysis* (*Shaped Pattern Synthesis:* $L = 7.5\lambda$ *[5]*) - Behaviours of ξ and P_{MT-BCS} versus (a) K, (b) σ^2, (c) β_1, (d) β_2, and (e) N.

Reference Pattern			BCS						$MT-BCS$					
$L\,[\lambda]$	P_{UNI}	*Type*	$\xi\,[\times 10^{-4}]$	$\frac{P_{BCS}}{P_{UNI}}$	$\frac{\Delta L_{min}}{\lambda/2}$	$\frac{\Delta L}{\lambda/2}$	$\frac{L_{BCS}}{L}$	Δt	$\xi\,[\times 10^{-4}]$	$\frac{P_{MTBCS}}{P_{UNI}}$	$\frac{\Delta L_{min}}{\lambda/2}$	$\frac{\Delta L}{\lambda/2}$	$\frac{L_{MTBCS}}{L}$	Δt
7.5	16	Tab. II [6]	1.33×10^{2}	1.18	0.12	0.81	0.97	0.17	0.59	0.81	0.2	1.20	0.96	0.21
18	37	Dolph	1.04	0.65	1.44	1.57	1.00	0.26	2.81	0.65	1.5	1.57	1.00	1.60
7.0	14	Tab. III [6]	0.52	1.47	0.018	0.65	0.98	0.52	0.22	0.73	0.93	1.38	0.98	0.45

Table 4.1: *Unconstrained Synthesis* - Array performance indexes.

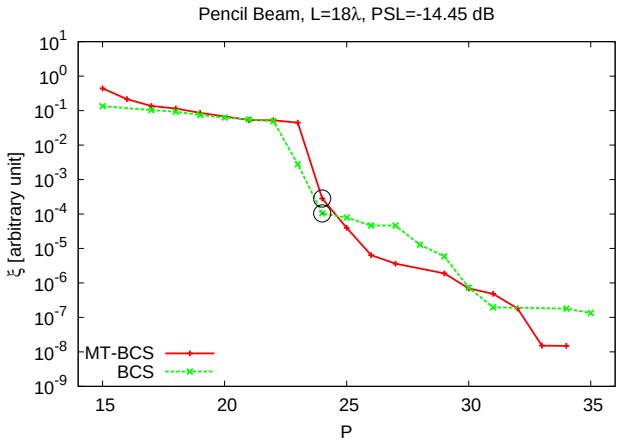
unconstrained patterns also in comparison with state-of-the-art methods (Sects. 4.3.2.1.2-4.3.2.1.3).

4.3.2.1 Unconstrained Synthesis

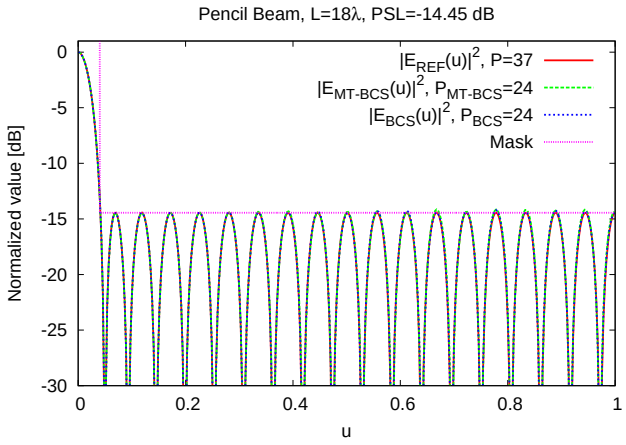
4.3.2.1.1 ***Consistency Check (Hermitian Patterns[2])*** In order to compare BCS and $MT-BCS$ approaches when dealing with Hermitian patterns, let us consider a $L = 18\lambda$ equi-ripple reference pattern ($PSL = -14.45$ dB) synthesized with the uniform array design method in [6] ($P_{UNI} = 37$). The plots of the Pareto fronts in the ξ-P plane indicate that, as expected, the two solutions' results are very close over a range of P [Fig. 4.5(a)]. The optimal trade-offs [$P_{MTBCS} = P_{BCS} = 24$, $\xi \approx \xi_{th}$ - Fig. 4.5(a)] turn out similar in both patterns [Fig. 4.5(b)] and weights [Fig. 4.5(c)] as also confirmed by the figures of merit in Table 4.1, notwithstanding the different synthesis processes. Both BCS and $MT-BCS$ behave similarly with Hermitian reference patterns, since a key difference between BCS and $MT-BCS$ is the numerical handling of the relation between the real and imaginary parts of the array weights, and $\mathcal{I}(v_p) = 0$ ($p = 1, ..., P$) when the reference pattern is Hermitian [Fig. 4.5(b)].

To further assess that such a behaviour is due to the symmetry properties of the pattern at hand, the next numerical experiment is concerned with a set of Hermitian patterns derived from [20]. The results of the synthesis of the three layouts with $L = \{19.5\,\lambda, 25\,\lambda, 50\,\lambda\}$ are presented in Table 4.2 and compared with the sparse arrangements generated by a stochastic methodology based on simulated-annealing (SA) [20]. As it can be observed, the BCS and $MT-BCS$ procedures achieve similar performances for each qualitative index (e.g., the matching accuracy and the array aperture) with an element saving equal or better than that of the stochastic approach (Table 4.2). This is also visually confirmed by the plots in Fig. 4.6 related to the representative example characterized by $L = 25\lambda$ and $PSL = -14.45$ dB [20]. With reference to the layout with $P_{BCS} = P_{MTBCS} = 20$ elements, it turns out that an acceptable fidelity [$\xi \leq 4.3\times 10^{-3}$ - Fig. 4.6(b)] is yielded by both BCS-based methods despite the reduction of the array elements with respect to the SA-optimized sparse solution ($P_{SA} = 24$). The similarities are not limited to the patterns, but as expected, are apparent also in the distribution of the real array coefficients [Fig. 4.6(c)].

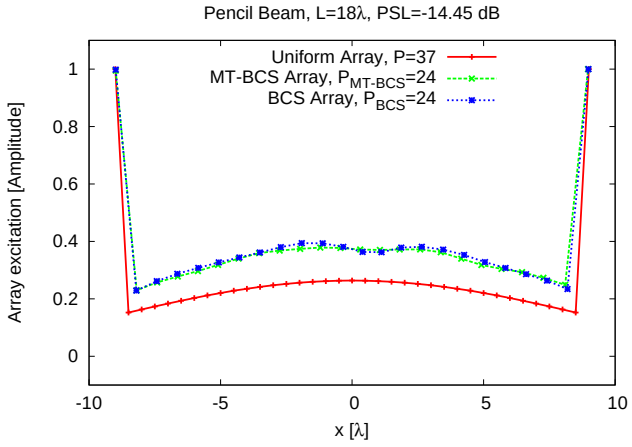
[2] *Hermitian Pattern* means symmetric pattern amplitude and anti-symmetric pattern phase that can also be generated by only real array weights.



(a)

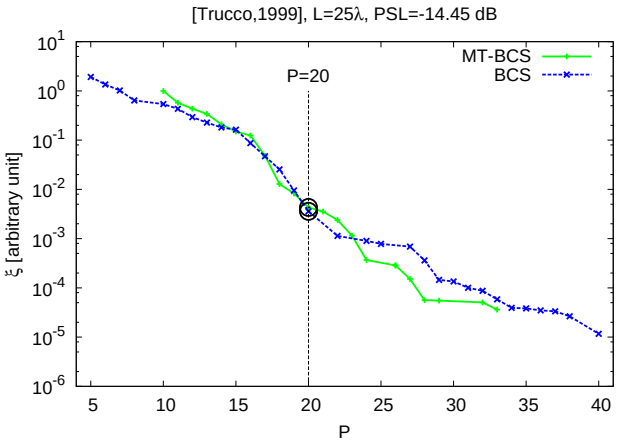


(b)

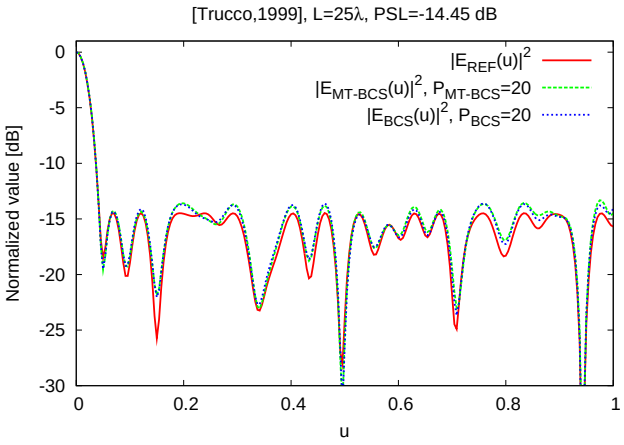


(c)

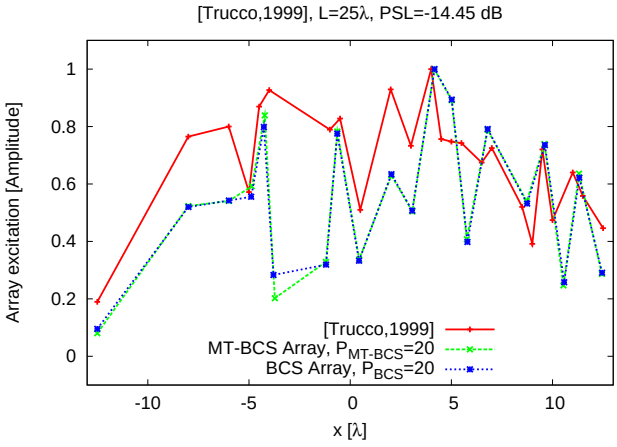
Figure 4.5: *Consistency Check* (*Hermitian Pattern Synthesis:* $L = 18\lambda$, $P_{REF} = P_{UNI} = 37$, $PSL = -14.45$ dB) - $MT-BCS$ vs. BCS: (a) Pareto fronts in the (ξ, P) plane, (b) optimal trade-off power patterns, and (c) the corresponding excitation amplitudes.



(a)



(b)



(c)

Figure 4.6: *Consistency Check* (*Hermitian Pattern Synthesis:* $L = 25\lambda$, $P_{REF} = 24$, $PSL = -14.45$ dB [20]) - $MT - BCS$ vs. BCS: (a) Pareto fronts in the (ξ, P) plane, (b) optimal trade-off power patterns, and (c) the corresponding excitations.

	L [λ]	PSL [dB]	P	ξ	$\frac{P}{P_{REF}}$	$\frac{\Delta L_{min}}{\lambda/2}$	$\frac{\Delta L}{\lambda/2}$	$\frac{L}{L_{REF}}$	Δt
[20]	19.50	−5.10	16	–	1.00	2.00	2.60	1.00	-
BCS	19.50	−5.10	16	2.34×10^{-7}	1.00	2.00	2.60	1.00	0.48
$MT-BCS$	19.50	−5.10	16	2.14×10^{-8}	1.00	2.00	2.60	1.00	0.30
[20]	25.00	−14.45	24	–	1.00	1.00	2.17	1.00	-
BCS	24.94	−13.63	20	3.58×10^{-3}	0.83	0.95	2.62	1.00	1.11
$MT-BCS$	24.95	−13.30	20	4.3×10^{-3}	0.83	1.00	2.62	1.00	2.23
[20]	50.00	−14.45	25	–	1.00	1.00	4.17	1.00	-
BCS	32.99	−11.70	22	2.06×10^{-2}	0.84	0.50	4.02	0.76	5.04
$MT-BCS$	32.99	−12.92	21	7.19×10^{-3}	0.84	1.00	3.30	0.76	4.52

Table 4.2: *Unconstrained Synthesis* (*Hermitian Pattern:* $P_{REF} = P_{UNI}$ [20]) - Array performance indexes.

4.3.2.1.2 ***Symmetric Power Patterns*** Unlike Hermitian patterns, BCS and $MT-BCS$ syntheses are expected to differ when only the reference power pattern is symmetric. The results from the synthesis of a non-Hermitian flat-top array ($P_{UNI} = 14$ - [6]) with symmetric power pattern [Fig. 4.7(a)] and asymmetric phase distribution [Fig. 4.7(b)] reveal the enhanced effectiveness of the MT procedure, which is due to its improved accuracy in modelling the statistical relations between the (non-negligible) real and imaginary parts of the array weights. As far as the optimal BCS-based trade-off solutions are concerned, it turns out that there is a halving of the array elements [$P_{BCS} = 22$ vs. $P_{MTBCS} = 11$ - Table 4.1] along with similar matching accuracies [$\xi_{BCS} = 0.52 \times 10^{-4}$ vs. $\xi_{MTBCS} = 0.22 \times 10^{-4}$ - Table 4.1]. This latter is mainly due to the intrinsic limitation of the BCS approach to deal with the two components of the array excitations as correlated unknowns [Eq. (4.6)]. Indeed, several BCS weights turn out either purely real or purely imaginary $[\angle\, v_p]_{BCS} \in \left\{0, \pm\frac{\pi}{2}, \pm\pi\right\}$ - Fig. 4.7(d)] unlike the $MT-BCS$ coefficients.

4.3.2.1.3 ***Asymmetric Power Patterns*** The improvements of the $MT-BCS$ approach are expected to be even more impressive when asymmetric patterns are at hand. In order to analyze such a case, the next example deals with the synthesis of a $L = 7.5\lambda$ cosecant pattern from [6] [$P_{UNI} = 16$, Fig. 4.8(b)]. The Pareto BCS solutions in the ξ-P plane [Fig. 4.8(a)] clearly indicate that the multi-task procedure is far more efficient than the single-task one. Indeed, the $MT-BCS$ yields sparser layouts for a fixed ξ threshold [e.g., $P_{MTBCS}/P_{BCS} = 0.68$ when $\xi \approx \xi_{th}$ - Fig. 4.8(a)], and a higher accuracy for a given P [e.g., $\xi_{MTBCS}/\xi_{BCS} \approx 8.0 \times 10^{-5}$ when $P = 15$ - Fig. 4.8(a)]. As an illustrative example, the patterns [Fig. 4.8(b)] and the array coefficients [Figs. 4.8(c)-4.8(d)] of the representative solutions circled in Fig. 4.8(a) [$P_{BCS} = 19$ vs. $P_{MTBCS} = 13$] are shown. As far as the array layouts are concerned, it is worth noticing that an element saving of $\approx 20\%$ ($P_{MTBCS}/P_{UNI} = 0.81$) and

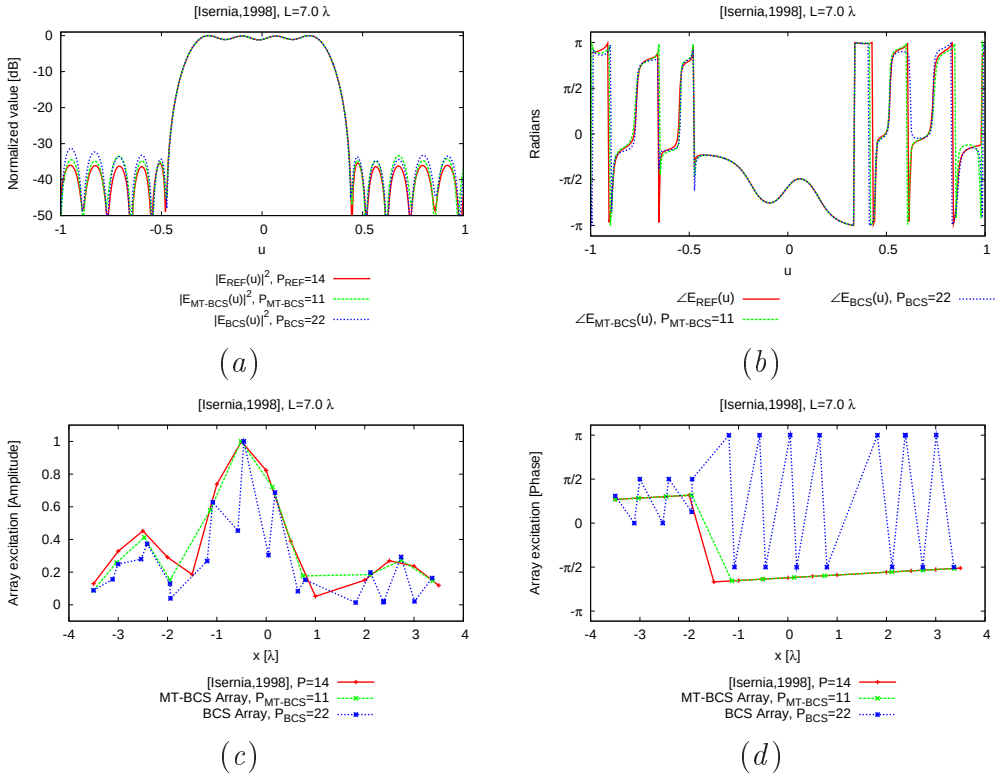


Figure 4.7: *Comparative Assessment* (*Symmetric Power Pattern Synthesis: 'Flat top'*, $L = 7\lambda$, $P_{REF} = P_{UNI} = 14$ [6]) - Pattern amplitudes (a), pattern phases (b), excitation amplitudes (c), excitation phases (d) of the uniform array [6] and of the optimal trade-off BCS and $MT - BCS$ layouts.

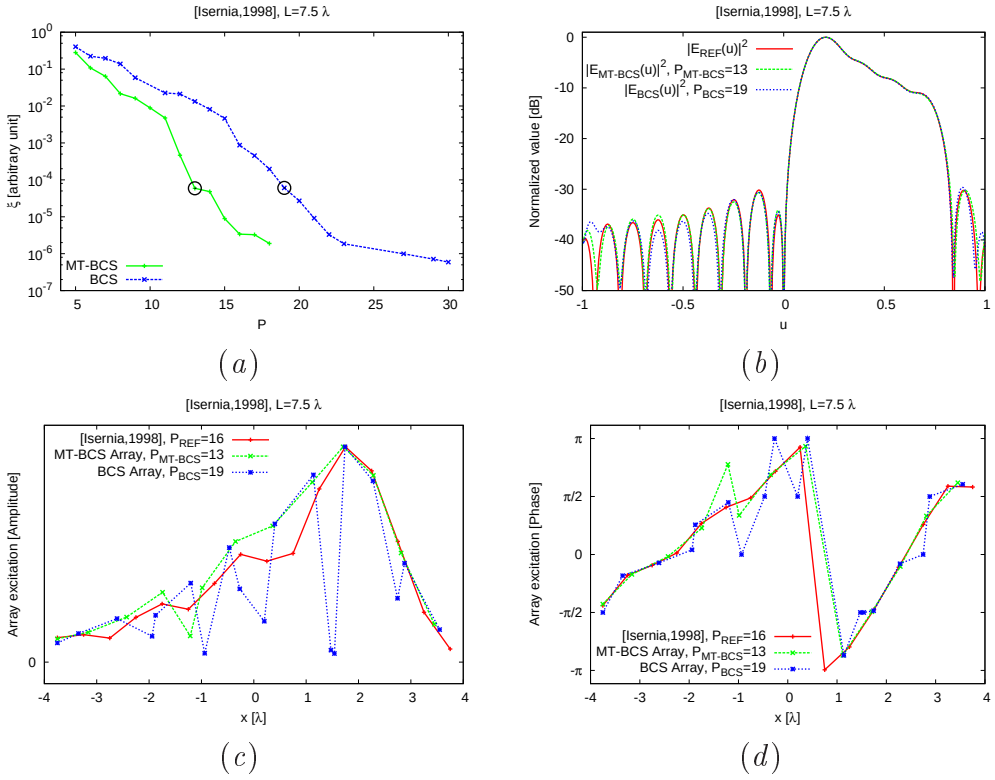


Figure 4.8: *Comparative Assessment* (*Asymmetric Power Pattern Synthesis: 'Cosecant'*, $L = 7.5\lambda$, $P_{REF} = P_{UNI} = 16$ [6]) - (a) $MT-BCS$ and BCS Pareto fronts in the (ξ, P) plane. Power patterns (b), excitation amplitudes (c), and excitation phases (d) of the uniform array [6] and of the optimal trade-off BCS and $MT-BCS$ layouts.

an aperture reduction of $\approx 4\%$ ($L_{MTBCS}/L_{UNI} = 0.96$) with respect to the uniform solution are obtained by the $MT-BCS$ without compromising the pattern matching accuracy (Table 4.1), while the BCS fails in reducing the array elements ($P_{MTBCS}/P_{UNI} = 1.18$). Moreover, the behaviour of the array excitations over the aperture confirms that the non-uniform $MT-BCS$ distribution follows the uniform one since the pattern matching refers to the complex reference pattern and not only to the power pattern, thus constraining both amplitudes and phases of the array coefficients.

To provide a more exhaustive comparison of the BCS methodologies, the results of an extensive analysis on asymmetric 'cosecant' reference patterns with constant sidelobes are presented. More specifically, the reference patterns have been chosen such that $L \in \{12\,\lambda, 19.5\,\lambda\}$ (i.e., $P_{UNI} \in \{25, 40\}$) and $PSL = \{-20\,dB, -30\,dB, -40\,dB\}$. The plots of P for the optimal (i.e., $\xi \sim \xi_{th}$) trade-off BCS and $MT-BCS$ layouts are shown in Fig. 4.9 as a function of P_{UNI}. By observing the case of the reference pattern with $PSL = -20$ dB, the MT

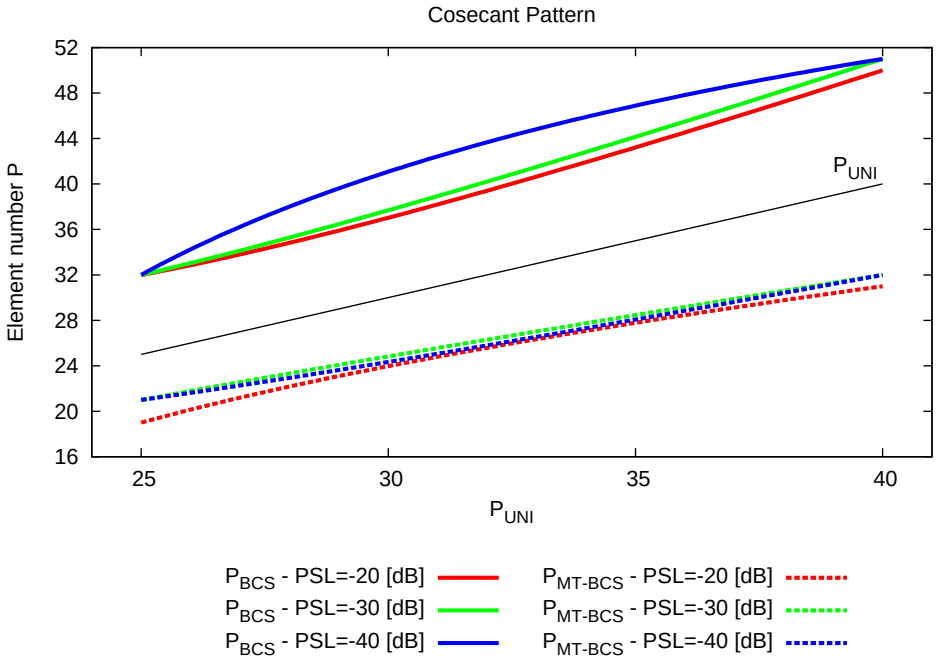


Figure 4.9: *Comparative Assessment* (*Asymmetric Power Pattern Synthesis: 'Cosecant'*, $P_{REF} = P_{UNI}$ [6]) - Plot of P vs. P_{UNI} for different values of the PSL [dB] of the reference pattern.

technique always outperforms the single-task method with significantly sparser solutions ($\frac{P_{MTBCS}}{P_{BCS}} \in [0.59, 0.66]$ - Table 4.3). This holds true also when lowering the sidelobe level (Fig. 4.9). On the other hand, although more array elements are necessary as P_{UNI} increases, the $MT-BCS$ always enables a reduction of the array elements with respect to the uniform architectures (Fig. 4.9 - $P_{MTBCS} < P_{UNI}$), while the condition $P_{BCS} > P_{UNI}$ is mandatory for the BCS to reach the accuracy threshold $\xi \sim 10^{-4}$ [$\frac{P_{MTBCS}}{P_{UNI}} \in [0.76, 0.84]$ vs. $\frac{P_{BCS}}{P_{UNI}} = 1.28$ - Table 4.3].

The effectiveness of the $MT-BCS$ to reduce the number of elements in the array arrangement is pictorially highlighted in the representative example analyzed in Fig. 4.10 ($PSL = -40$ dB). Whatever the matching accuracy, the $MT-BCS$ patterns exhibit a higher sparseness [Figs. 4.10(a)-4.10(b), 4.10(c)-4.10(d), 4.10(e)-4.10(f)] than the BCS. Furthermore, the pattern matching of the $MT-BCS$ solution is always better for a given value of P [Figs. 4.10(a), 4.10(c), 4.10(e)].

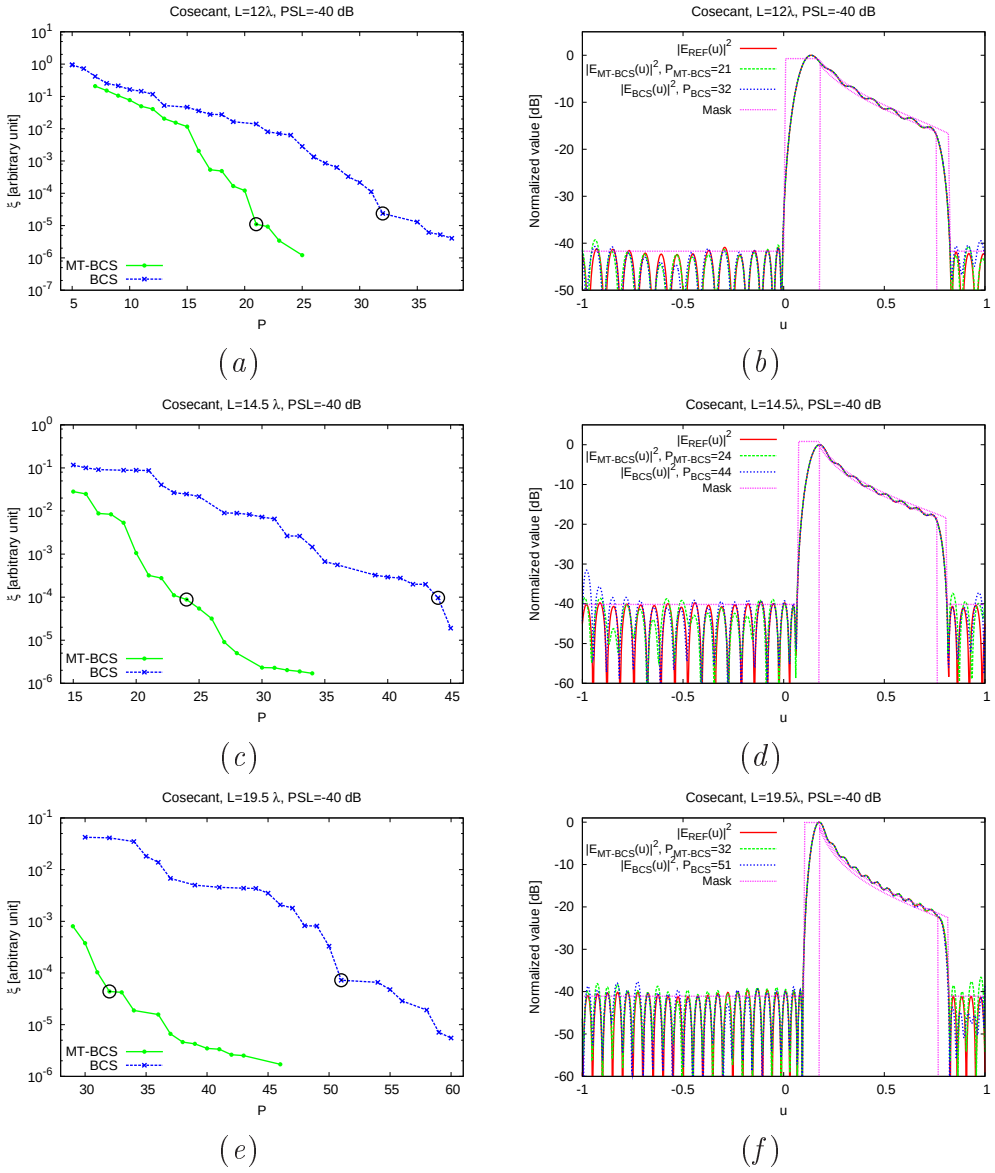


Figure 4.10: *Comparative Assessment* (*Asymmetric Power Pattern Synthesis:* '*Cosecant*', $P_{REF} = P_{UNI}$, $PSL = -40$ dB [6]) - BCS and $MT-BCS$ Pareto fronts in the (ξ, P) plane (*left column*), power patterns of the reference uniform array [6] and of the optimal trade-off BCS and $MT-BCS$ solutions (*right column*). $(a)(b)$ $L = 12\lambda$ ($P_{UNI} = 25$), $(c)(d)$ $L = 14.5\lambda$ ($P_{UNI} = 30$), and $(e)(f)$ $L = 19.5\lambda$ ($P_{UNI} = 40$).

Reference Pattern			*BCS*						*MT − BCS*					
L [λ]	PSL	P_{UNI}	ξ [$\times 10^{-4}$]	$\frac{P_{BCS}}{P_{UNI}}$	$\frac{\Delta L_{min}}{\lambda/2}$	$\frac{\Delta L}{\lambda/2}$	$\frac{L_{BCS}}{L_{UNI}}$	Δt	ξ [$\times 10^{-4}$]	$\frac{P_{MTBCS}}{P_{UNI}}$	$\frac{\Delta L_{min}}{\lambda/2}$	$\frac{\Delta L}{\lambda/2}$	$\frac{L_{MTBCS}}{L_{UNI}}$	Δt
12	−20	25	3.00	1.28	0.048	0.77	1.00	0.29	0.53	0.76	0.74	1.33	1.00	7.73
12	−30	25	2.86	1.28	0.048	0.77	1.00	0.38	0.38	0.84	0.51	1.2	1.00	1.39
12	−40	25	0.24	1.28	0.49	0.77	1.00	0.23	0.11	0.84	0.72	1.2	1.00	0.87
14.5	−20	30	0.48	1.2	0.59	0.83	1.00	0.23	0.46	0.80	0.49	1.26	1.00	0.56
14.5	−30	30	1.29	1.23	0.058	0.8	0.99	0.44	1.47	0.80	0.63	1.26	1.00	2.85
14.5	−40	30	0.96	1.47	0.33	0.67	1.00	0.25	0.81	0.77	0.40	1.28	0.98	3.82
19.5	−20	40	3.75	1.3	0.67	0.75	0.98	0.24	2.27	0.78	0.54	1.30	1.00	6.19
19.5	−30	40	1.29	1.43	0.31	0.70	1.00	1.30	0.80	0.78	0.19	1.30	1.00	6.99
19.5	−40	40	0.83	1.35	0.39	0.74	1.00	0.36	0.44	0.78	0.52	1.30	1.00	4.38

Table 4.3: *Unconstrained Synthesis* (*Asymmetric Pattern: 'Cosecant'*, $P_{REF} = P_{UNI}$ [6]) - Array performance indexes.

The BCS approach is usually faster[3] than the multi-task procedure, although both methods do not require heavy computations [$\Delta t \leq 8$ s - Table 4.3]. This is expected since neglecting the relationships between real and imaginary parts of the array excitations (see Sect. 4.2) simplifies the problem, but significantly degrades the synthesis performance with complex layouts.

As for the state-of-the-art comparisons, let us refer to recently introduced approaches based on the Matrix Pencil Method (MPM) [14][33][34][47]. Such a choice is mainly due to their effectiveness and numerical efficiency usually outperforming other sparse-synthesis methods in terms of convergence speed, reliability, and accuracy [14][33][34].

The first set of comparisons is concerned with the benchmark case in [46]. The synthesis results are reported in Figs. 4.11(a)-4.11(b) and quantitatively compared in Table 4.4. With reference to the (ξ, P)-plane [Fig. 4.11(a)], the stand-alone matrix pencil method [33] is, as expected, significantly less accurate ($P = 19$: $\xi_{MPM} = 1.43 \times 10^{-1}$ [Fig. 4.11(b)] vs. $\xi_{MTBCS} = 3.53 \times 10^{-3}$ - Table 4.4) than the sub-optimal (i.e., $\xi > \xi_{th}$) $MT-BCS$ because of the shaped-beam reference pattern [14], while the hybrid $TABU - MPM$ ($TMPM$) [47] reaches a comparable pattern matching ($P = 19$: $\xi_{MPM-TABU} = 3.21 \times 10^{-3}$ - Table 4.4) although requiring a non-negligible computational burden [47] because of the $TABU$-based stochastic optimization in the second step of the hybrid procedure.

Concerning the so-called forward-backward version of the matrix pencil method ($FBMPM$) [34], the results in Figs. 4.11(c)-4.11(f) derived from [48][49] (also discussed in [34]) point out that the $FBMPM$ exhibits performance close to that of the $MT - BCS$ when dealing with shaped-pattern problems [e.g., $P = 13$: $\xi_{FBMPM} = 8.09 \times 10^{-5}$ vs. $\xi_{MTBCS} = 5.32 \times 10^{-5}$ - Table 4.5 and Fig. 4.11(c); $P = 15$: $\xi_{FBMPM} = 4.94 \times 10^{-5}$ vs. $\xi_{MTBCS} = 1.68 \times 10^{-4}$ - Table 4.6 and Fig. 4.11(e)].

[3]In all cases, the synthesis time Δt refers to the execution of the Matlab code on a single core laptop running at 2.16 GHz.

				Optimal Tradeoff ($\xi \leq \xi_{th}$)		*Sub-Optimal Tradeoff* ($\xi > \xi_{th}$)	
	Uniform [46]	MPM [47]	$TMPM$ [47]	BCS	$MT-BCS$	BCS	$MT-BCS$
$L\,[\lambda]$	14.5	14.47	14.14	14.5	14.5	11.28	13.00
P	30	19	19	35	24	20	19
$\frac{P}{P_{UNI}}$	–	0.63	0.63	1.17	0.8	0.66	0.63
$\frac{\Delta L_{min}}{\Delta L_{UNI}}$	–	1.15	1.20	0.29	0.058	0.29	0.075
$\frac{\Delta L}{\Delta L_{UNI}}$	–	1.61	1.57	0.85	1.26	1.19	1.45
$\frac{L}{L_{UNI}}$	–	1	0.98	1.00	1.00	0.78	0.90
$t\,[s]$	–	–	–	0.24	0.53	0.22	0.97
ξ	–	1.43×10^{-1}	3.21×10^{-3}	9.85×10^{-5}	8.15×10^{-5}	3.71×10^{-2}	3.53×10^{-3}

Table 4.4: *Unconstrained Synthesis* (*Asymmetric Pattern: 'Cosecant'*, $L = 14.5\lambda$, $P_{REF} = P_{UNI} = 30$ [46]) - Array performance indexes.

	[48]	$FBMPM$ [34]	BCS	$MT-BCS$
$L\,[\lambda]$	7.5	7.51	7.50	7.46
P	16	13	14	13
$\frac{P}{P_{UNI}}$	–	0.81	0.88	0.81
$\frac{\Delta L_{min}}{\Delta L_{UNI}}$	–	1.06	0.042	0.74
$\frac{\Delta L}{\Delta L_{UNI}}$	–	1.25	1.15	1.24
$\frac{L}{L_{UNI}}$	–	1	1	1
$t\,[s]$	–	–	0.16	1.00
ξ	–	8.09×10^{-5}	1.89×10^{-2}	5.32×10^{-5}

Table 4.5: *Unconstrained Synthesis* (*Asymmetric Pattern: 'Cosecant'*, $L = 7.5\lambda$, $P_{REF} = P_{UNI} = 16$ [48]) - Array performance indexes.

	[49]	$FBMPM$ [34]	BCS	$MT-BCS$
$L\,[\lambda]$	9.5	9.375	9.5	9.34
P	20	15	15	15
$\frac{P}{P_{UNI}}$	–	0.75	0.75	0.75
$\frac{\Delta L_{min}}{\Delta L_{UNI}}$	–	1.23	0.39	0.97
$\frac{\Delta L}{\Delta L_{UNI}}$	–	1.34	1.36	1.35
$\frac{L}{L_{UNI}}$	–	0.99	1.00	0.98
$t\,[s]$	–	–	0.18	0.98
ξ	–	4.94×10^{-5}	4.62×10^{-2}	1.68×10^{-4}

Table 4.6: *Unconstrained Synthesis* (*Asymmetric Pattern: 'Cosecant'*, $L = 9.5\lambda$, $P_{REF} = P_{UNI} = 20$ [49]) - Array performance indexes.

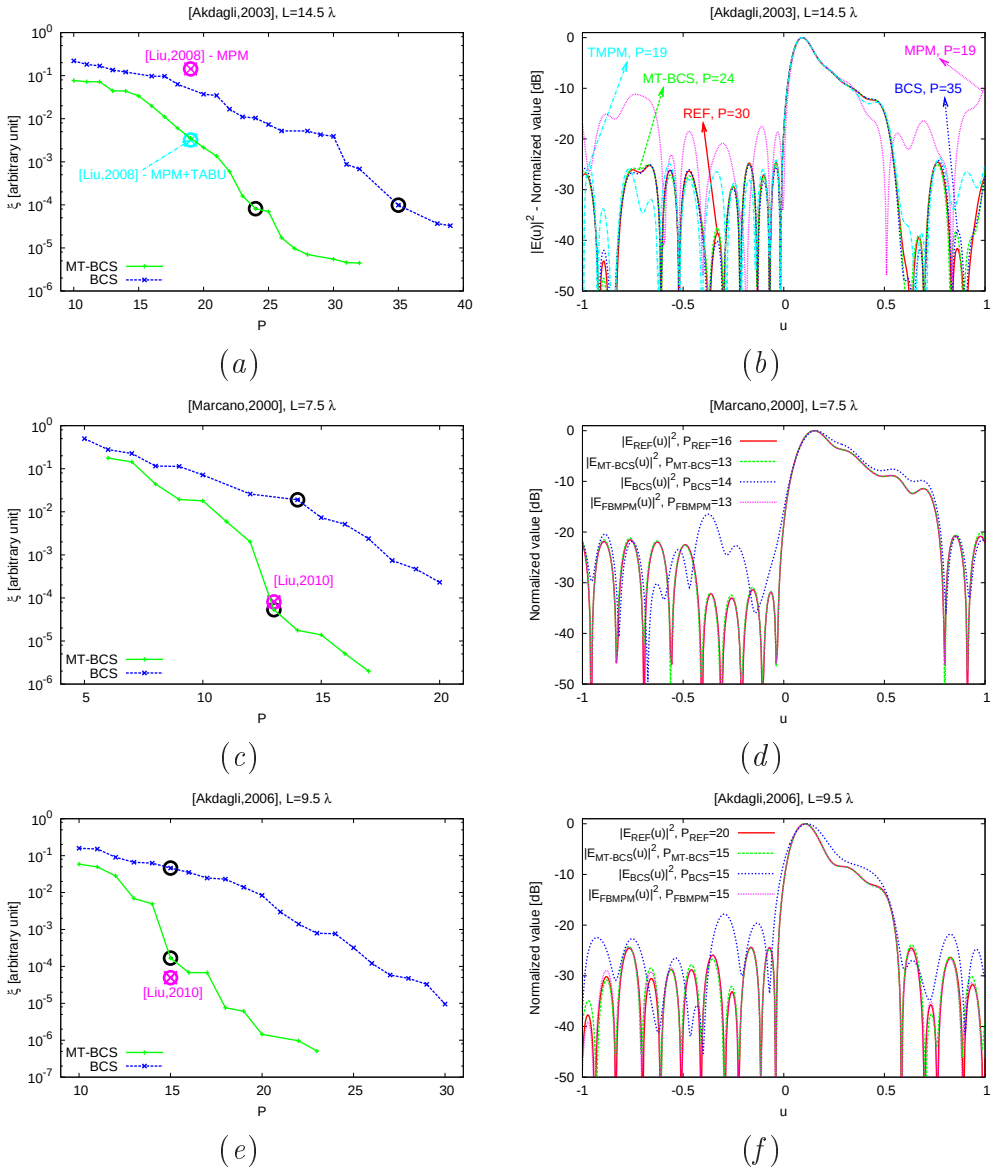


Figure 4.11: *Comparative Assessment* (*Asymmetric Power Pattern Synthesis: 'Cosecant', $P_{REF} = P_{UNI}$*) - Representative points in the (ξ, P) plane of the BCS and $MT-BCS$ Pareto fronts and of the MPM-based methods (*left column*), power patterns of the reference uniform array, the MPM-based methods, and the optimal trade-off BCS and $MT-BCS$ solutions (*right column*). $(a)(b)$ $L = 14.5\lambda$ ($P_{UNI} = 30$) [46], $(c)(d)$ $L = 7.5\lambda$ ($P_{UNI} = 16$) [48], and $(e)(f)$ $L = 9.5\lambda$ ($P_{UNI} = 20$) [49].

	[50] - $P_{UNI}=20$, $L_{UNI}=9.5\lambda$			$P_{UNI}=30$, $L_{UNI}=14.5\lambda$		
	$FBMPM$	BCS	$MT-BCS$	$FBMPM$	BCS	$MT-BCS$
$L\,[\lambda]$	9.5	9.46	9.5	14.5	14.5	14.5
P	16	29	16	24	44	24
$\frac{P}{P_{UNI}}$	0.8	1.45	0.8	0.8	1.47	0.8
$\frac{\Delta L_{min}}{\Delta L_{UNI}}$	0.97	0.019	0.31	0.00324	0.34	0.45
$\frac{\Delta L}{\Delta L_{UNI}}$	1.27	0.68	1.27	1.26	0.67	1.26
$\frac{L}{L_{UNI}}$	1.00	1.00	1.00	1.00	1.00	1.00
$t\,[s]$	7.83×10^{-1}	5.58×10^{-1}	6.70×10^{-1}	9.61×10^{-1}	2.5×10^{-1}	1.43
ξ	6.79×10^{-3}	7.14×10^{-5}	9.27×10^{-5}	3.98×10^{-3}	9.62×10^{-5}	7.93×10^{-5}

Table 4.7: *Unconstrained Synthesis* (*Asymmetric Pattern: 'Cosecant'*, $P_{REF}=P_{UNI}$) - Array performance indexes.

However, it cannot be neglected that the MPM (and, consequently, the $FBMPM$) can present some numerical instabilities (or no convergence) as it was pointed out in [14][18] and confirmed by the synthesis results of the 'cosecant' pattern with $PSL=-40$ dB generated by the uniform aperture $L=14.5\lambda$ (Fig. 4.13) as well as for the test case in [50] (Fig. 4.12). Unlike the BCS-based approaches, the fitting with the reference pattern of the $FBMPM$[4], ξ_{FBMPM}, does not monotonically improve as P grows [Fig. 4.12(a) and Fig. 4.13(a)]. For example [Fig. 4.12(a)], the $MT-BCS$ reaches the matching threshold $\xi\sim\xi_{th}$ (i.e., $\xi_{MTBCS}\rfloor_{P=16}=9.27\times10^{-5}$ - Table 4.7) just adding an element to the array with $P_{MTBCS}=15$, while the $FBMPM$ accuracy worsens when moving from $P_{FBMPM}=14$ to $P_{FBMPM}=16$ ($\xi_{FBMPM}\rfloor_{P=14}=8.50\times10^{-4}$ vs. $\xi_{FBMPM}\rfloor_{P=16}=6.79\times10^{-3}$). Therefore, the $MT-BCS$ faithfully reconstructs the reference pattern [Fig. 4.12(b)] reducing the uniform array elements of $\frac{1}{5}$ unlike the $FBMPM$ that does not provide the same accuracy (i.e., $\xi\leq10^{-4}$) unless using more radiators ($P_{FBMPM}=19\rightarrow\xi_{FBMPM}\rfloor_{P=19}=4.60\times10^{-6}$).

Similar outcomes can be drawn from the test case in Fig. 4.13 (Table 4.7) that allows us to point out also another interesting feature of the BCS-based approaches. By observing the $FBMPM$ arrangement in Fig. 4.13(c), it turns out that the minimum inter-element spacing is very small and equal to $\Delta L_{min}=1.62\times10^{-3}\,\lambda$ (Table 4.7). On the contrary, the BCS rationale with the choice of the candidate locations for the array elements, $\mathbf{d}$, gives the user the possibility to *a-priori* impose the lower bound for the distance between two adjacent elements. As for the CPU-time, the indexes in Table 4.7 indicate that the synthesis time for the $MT-BCS$ and the $FBMPM$ is generally of the same order in magnitude (e.g., $\Delta t_{FBMPM}=7.83\times10^{-1}$ [s] vs. $\Delta t_{MTBCS}=6.70\times10^{-1}$ [s]).

[4]A $MATLAB$ implementation of the $FBMPM$ (based on the *mpencil* function *http://www.mathworks.se/matlabcentral/index.html*) have been used assuming the parameters suggested in [34] for the following numerical tests.

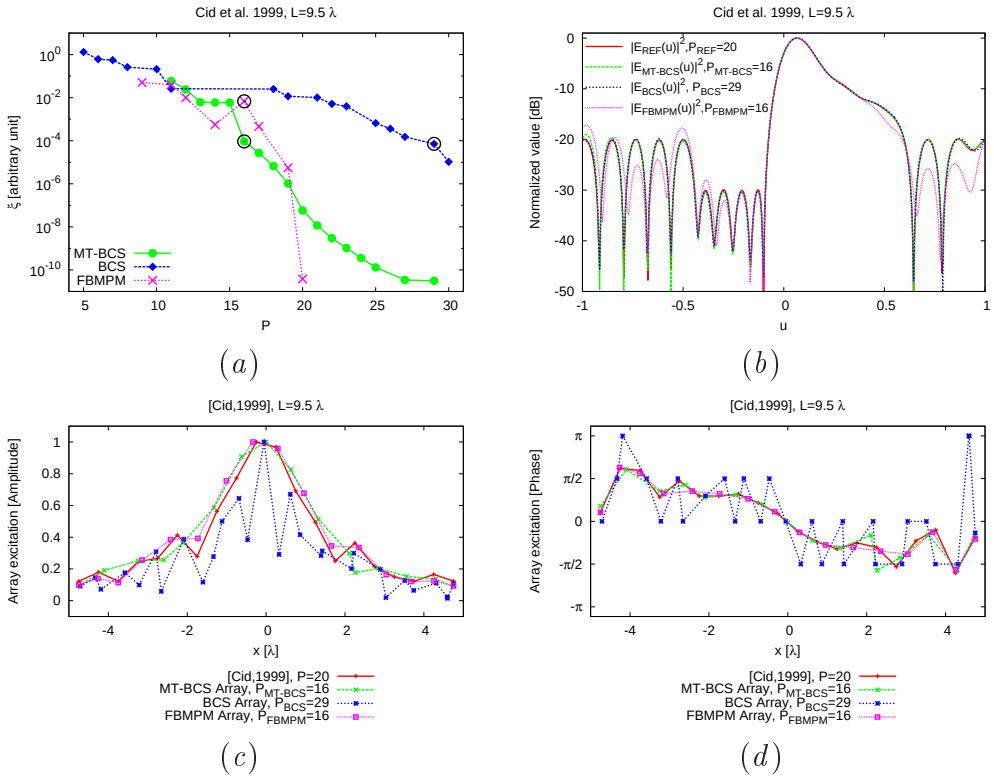


Figure 4.12: *Comparative Assessment* (*Asymmetric Power Pattern Synthesis: 'Shaped Cosecant'*, $P_{REF} = P_{UNI} = 20$, $L = 9.5\lambda$ [50]) - $FBMPM$, BCS, and $MT-BCS$ solutions: (a) Pareto fronts in the (ξ, P) plane, (b) power patterns, (c) excitation amplitudes, and excitation phases (d).

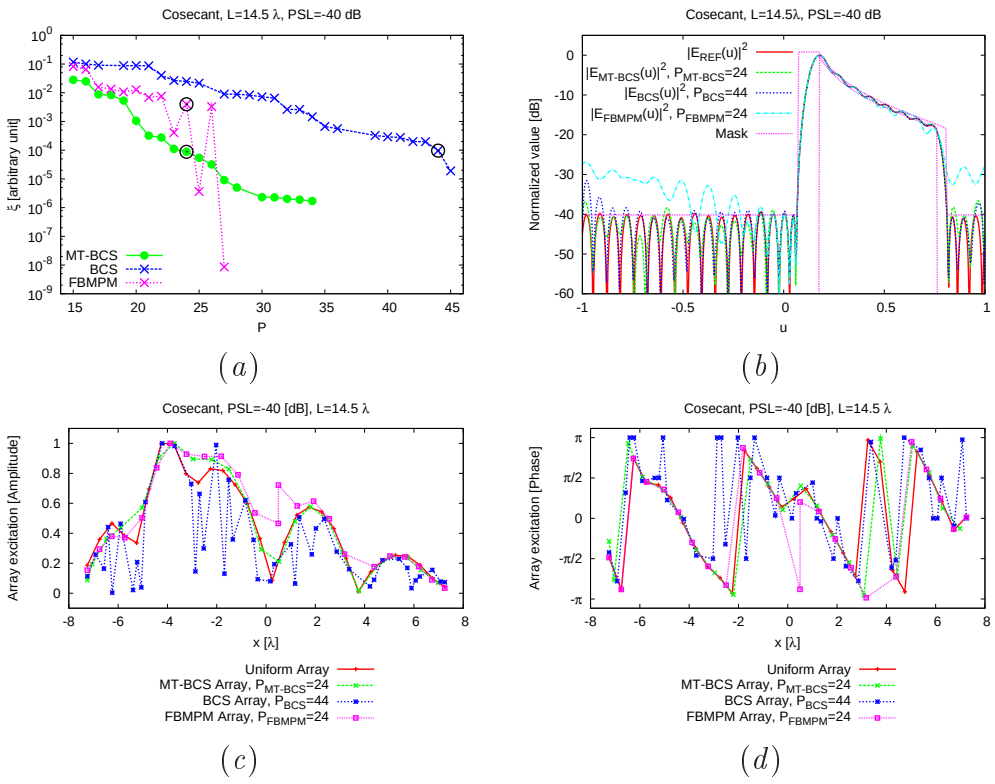


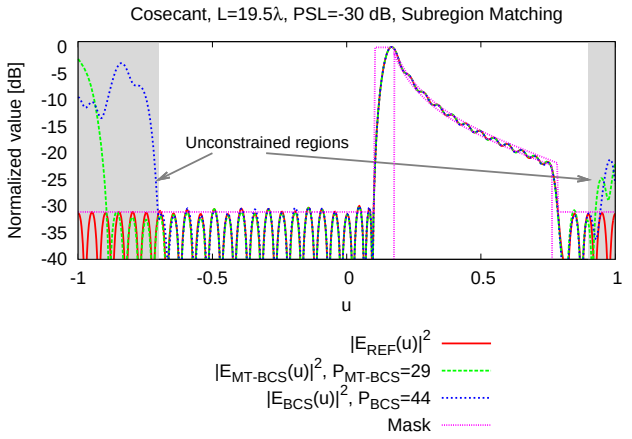
Figure 4.13: *Comparative Assessment* (*Asymmetric Power Pattern Synthesis:* '*Cosecant*', $L = 19.5\lambda$, $P_{REF} = P_{UNI} = 40$, $PSL = -40$ dB [6]) - $FBMPM$, BCS, and $MT-BCS$ solutions: (a) Pareto fronts in the (ξ, P) plane, (b) power patterns, (c) excitation amplitudes, and excitation phases (d).

4.3.2.2 Constrained Synthesis

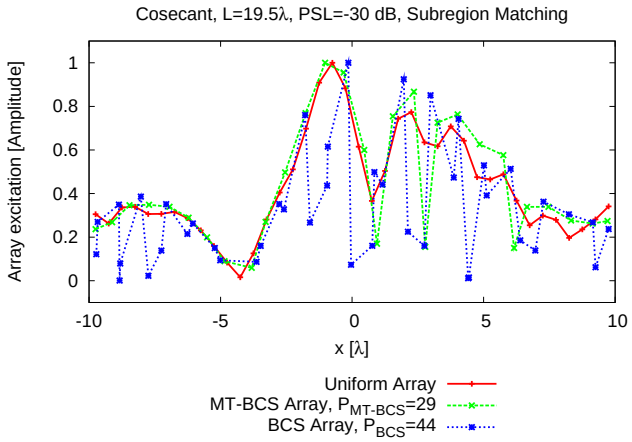
Previous discussions gave some insights about the efficiency of the $MT-BCS$ in dealing with shaped reference patterns as well as about its advances in terms of element sparseness, matching accuracy, final layout properties, and numerical efficiency over the standard BCS or in comparison with reference state-of-the-art methodologies. The final set of experiments, concerned with an equi-ripple ($PSL=-30$ dB) cosecant reference pattern generated by a uniform aperture of $L=19.5\lambda$, points out the flexibility of the BCS-based methodology to handle constrained sparse-array syntheses. Performing constrained sparse synthesis is not a trivial task for a wide range of state-of-the-art methods, except for optimization methods which, however, usually involve heavy computations when high-dimension solution spaces are at hand.

4.3.2.2.1 ***Pattern Constraints*** The first test case has been designed by limiting the reference pattern samples $F_{REF}(u_k)$ to the angular region $u\in(-0.7,0.9)$ [i.e., $u_k\notin\{[-1,-0.7]\cup[0.9,1]\}$, $k=1,...,K$]. As expected, the optimal trade-off $MT-BCS$ and BCS patterns faithfully match the reference pattern only within the constrained region [$\xi_{MTBCS}\rfloor_{P=29}=2.35\times10^{-5}$ vs. $\xi_{BCS}\rfloor_{P=44}=4.96\times10^{-5}$- Fig. 4.14($a$)] guaranteeing a reduction, more significant for the BCS even though still $P^{unc}_{BCS}=44>P_{UNI}=40$, of the element number with respect to the full-constrained case ($\frac{P^{con}}{P^{unco}}\rfloor_{MTBCS}=1.07$ and $\frac{P^{con}}{P^{unco}}\rfloor_{BCS}=1.29$).

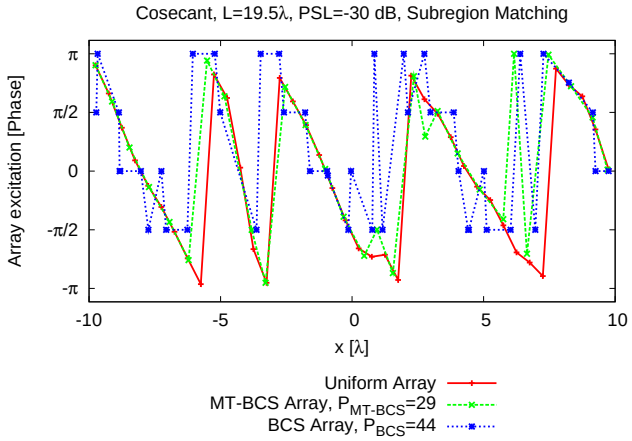
4.3.2.2.2 ***Geometry Constraints*** The last cases model aperture blockage constraints within the BCS syntheses by setting forbidden regions for the radiating elements [see Sect. 4.2]. More specifically, two different scenarios have been investigated either defining symmetric ($d_n\notin\{[-6\lambda,-5\lambda]\cup[5\lambda,6\lambda]\}$, $n=1,...,N$) or asymmetric ($d_n\notin\{[-7\lambda,-6\lambda]\cup[3\lambda,4\lambda]\}$, $n=1,...,N$) forbidden regions. The plots of the optimal trade-off layouts and associated patterns (Fig. 4.15) show that both compressive-sampling procedures succeed in carefully reproducing the reference pattern [$\xi_{MTBCS}=1.01\times10^{-5}$ vs. $\xi_{BCS}=2.32\times10^{-5}$ - Figs. 4.15(a); $\xi_{MTBCS}=6.08\times10^{-5}$ vs. $\xi_{BCS}=9.68\times10^{-5}$ - Figs. 4.15(b)] while also complying with the geometrical constraints [Figs. 4.15(c)-4.15(e) and Figs. 4.15(d)-4.15(f)] despite the non-negligible aperture blockage ($>10\%$ in both cases). Furthermore, the $MT-BCS$ technique confirms also in those scenarios its higher efficiency (than the BCS) in minimizing the array elements [$P_{BCS}=63$ vs. $P_{MTBCS}=37$ - Figs. 4.15(c)-4.15(e); $P_{BCS}=58$ vs. $P_{MTBCS}=34$ - Figs. 4.15(d)-4.15(f)] also with respect to the (unconstrained) uniform solution [$\frac{P_{MTBCS}}{P_{UNI}}=0.92$ - Figs. 4.15(c)-4.15(e); $\frac{P_{MTBCS}}{P_{UNI}}=0.85$ - Figs. 4.15(d)-4.15(f)]. Of course, the element saving turns out to be lower than that for the 'unconstrained' BCS-based syntheses because of the greater complexity of the synthesis at hand [i.e., $\frac{P^{con}}{P^{unco}}\rfloor_{MTBCS}=1.18$ and $\frac{P^{con}}{P^{unco}}\rfloor_{BCS}=1.09$ (*symmetric forbidden*



(a)



(b)



(c)

Figure 4.14: *Flexibility Check* (*Constrained Pattern Synthesis: 'Cosecant'*, $L = 19.5\lambda$, $PSL = -30$ dB, $P_{REF} = P_{UNI} = 40$, $u_k \notin \{[-1, -0.7] \cup [0.9, 1]\}$) - Power patterns (a) and array coefficients $(b)(c)$ of the optimal trade-off BCS and $MT-BCS$ layouts.

region) and $\frac{P^{con}}{P^{unco}}\big\rfloor_{MTBCS} = 1.09$ and $\frac{P^{con}}{P^{unco}}\big\rfloor_{BCS} = 1.02$ (*asymmetric forbidden region*)].

4.4 Discussions

An innovative, flexible, and efficient complement to the existing approaches for the synthesis of sparse layouts with arbitrary radiation features has been proposed. The proposed method extends the range of applicability of the technique in [21] by considering a MT Bayesian methodology. Towards this end, the original pattern matching problem has been formulated in a Bayesian fashion within the framework of the sparseness constrained optimization and afterwards it has been solved by a suitable RVM-derived methodology. Selected results from an extensive numerical validation have been presented to provide an evaluation of the sensitivity of the $MT-BCS$ method to its control parameters as well as on its accuracy, flexibility, and computational efficiency. Advantages and limitations of the proposed approach have been pointed out using comparisons with state-of-the-art approaches. In summary:

- the $MT-BCS$ technique is simpler to calibrate than the single-task BCS approach thanks to its smoother dependency on the control parameters (Sect. 4.3.1);
- the $MT-BCS$ methodology outperforms the single-task BCS procedure since, generally, the BCS extension to complex layouts often yields to (sub-optimal) arrangements mostly comprising purely-real and purely-imaginary excitations. As expected, BCS-based procedures provide very similar results when symmetric real layouts are at hand (Sub-Sect. 4.3.2.1.1);
- on average, the $MT-BCS$ guarantees an element saving with respect to ($\frac{\lambda}{2}$-spaced) uniform layouts of about $\frac{P_{MTBCS}}{P_{UNI}} \in [0.65, 0.81]$ when complex- or real-valued symmetric patterns are at hand still providing an excellent pattern matching [$\xi \lesssim 10^{-4}$];
- the $MT-BCS$ favorably compares with state-of-the-art sparse array design procedures in terms of pattern matching accuracy, element saving, numerical efficiency, and stability;
- additional constraints on the radiation pattern and/or the geometrical features of the sparse array can be easily and efficiently dealt with (Sect. 4.3.2.2).

In addition, other main and innovative contributions of this Chapter consist in the following methodological novelties:

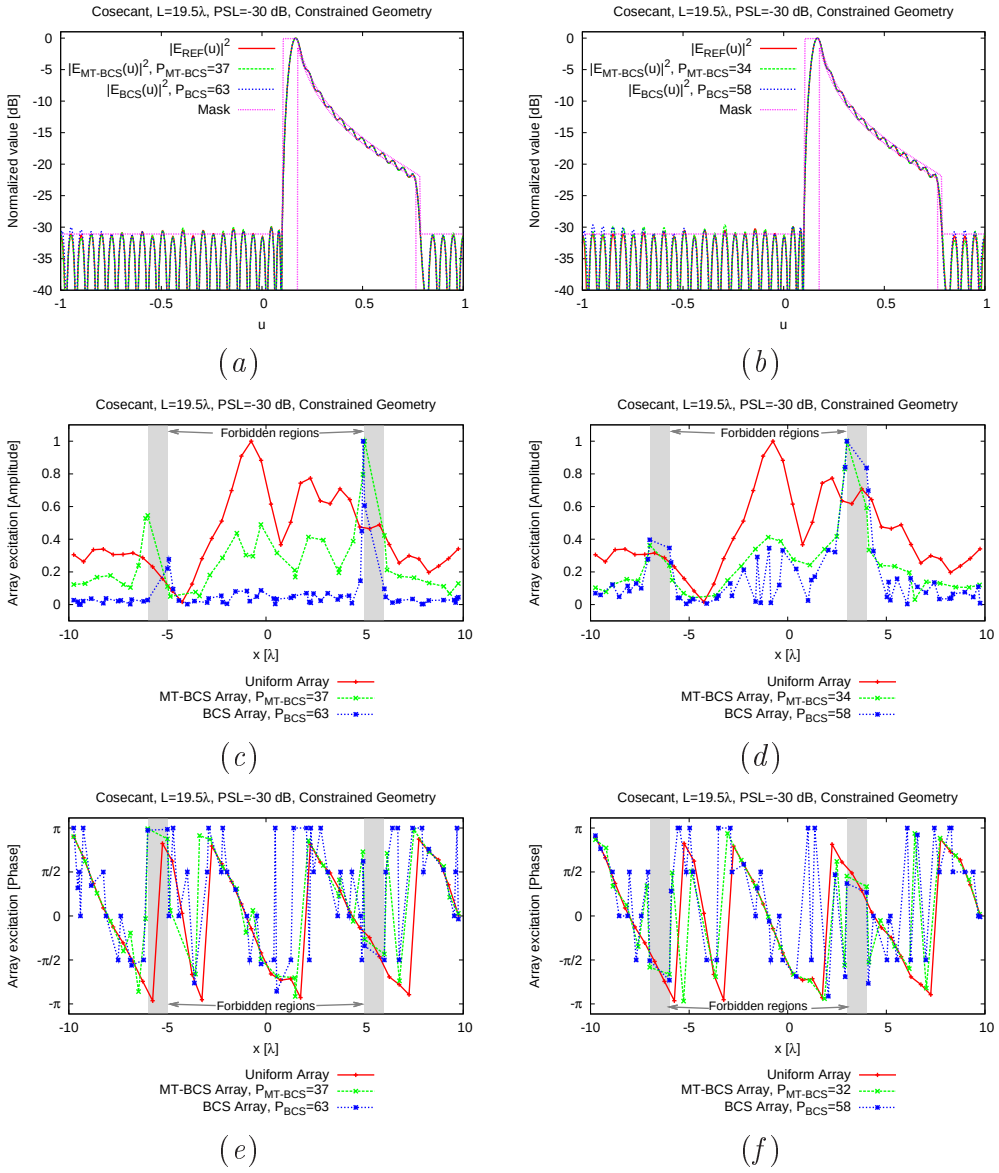


Figure 4.15: *Flexibility Check* (*Constrained-Geometry Pattern Synthesis: 'Cosecant'*, $L = 19.5\lambda$, $PSL = -30$ dB, $P_{REF} = P_{UNI} = 40$) - Power patterns $(a)(b)$, excitation amplitudes $(c)(d)$, and excitation phases $(e)(f)$ of the (unconstrained) uniform array and of the optimal trade-off constrained BCS and $MT-BCS$ layouts when $d_n \notin \{[-6\lambda, -5\lambda] \cup [5\lambda, 6\lambda]\}$ (*left column*) and $d_n \notin \{[-7\lambda, -6\lambda] \cup [3\lambda, 4\lambda]\}$ (*right column*).

an extension to the complex-valued synthesis problems of the BCS approach in [21];

an innovative and equivalent 'fictitious' formulation of the complex-weight pattern matching problem for enabling the application of the $MT-BCS$;

an innovative $MT-BCS$ method for dealing with complex-valued sparseness constrained optimization by statically correlating the real and the imaginary components of the sparse unknowns.

Future works, out-of-the-scope of the present Chapter, will be aimed at analyzing the mutual coupling effects between real elements in the sparse layouts as well as at taking into account in the synthesis process the presence of directive elements. Furthermore, the derivation of array processing algorithms (e.g., DOA-estimation [55] and adaptive beamforming [56] techniques) based on $MT-BCS$ geometries will be the subject of future analyses aimed at exploiting and integrating the features of such a sparse arrangements in an effective and customized way.

Chapter 5

Direction-of-Arrival Estimation in Linear Arrays Through Bayesian Compressive Sensing Strategies

In this Chapter, the estimation of the directions of arrival (DoAs) of narrow-band signals impinging on a linear antenna array is addressed within the Bayesian compressive sensing (BCS) framework. Unlike several state-of-the-art approaches, the voltages at the output of the receiving sensors are directly used to determine the DoAs of the signals thus avoiding the computation of the correlation matrix. Towards this end, the estimation problem is properly formulated to enforce the sparsity of the solution in the linear relationships between output voltages (i.e., the problem data) and the unknown DoAs. Customized implementations exploiting the measurements collected at a unique time instant (single-snapshot) and multiple time instants (multiple-snapshots) are presented and discussed. The effectiveness of the proposed approaches is assessed through an extensive numerical analysis addressing different scenarios, signal configurations, and noise conditions. Comparisons with state-of-the-art methods are reported, as well.

5.1 Introduction

Estimating the *DoA*s of signals is a topic of great interest in several research fields like electromagnetic, acoustic, and geophysical/seismic sensing [61][62][63][64]. As a matter of fact, the knowledge of the *DoA*s of the signals incoming on a receiver can be suitably exploited to localize the positions of the corresponding sources as well as to enable the adaptive beam forming of the receiving antenna pattern either to enhance the system sensitivity towards desired signal directions or to suppress unwanted interferences.

State-of-the-art literature gives to the interested reader several and effective approaches proposed in the last decades. The methods are most commonly used are: the multiple signal classification ($MUSIC$) [65][66], the signal estimation parameter via rotational invariance technique ($ESPRIT$) [67][68][69], and the maximum likelihood (ML) *DoA*s estimator [96][71]. A main drawback of these techniques is the need of an *a-priori* knowledge of the number of signals, which is rarely available especially nowadays with the huge proliferation of wireless devices/services and the presence of non-collaborative users. To avoid such a constraint, a learning-by-example (LBE) approach based on a support vector machine (SVM) has been proposed in [72] where the *DoA* estimation problem has been recast to a probabilistic framework looking for the identification of the smallest angular regions where the presence of incoming signals is most probable. While efficient for some applications, the rough estimation of the *DoA*s of the signals coming from the processing of the arising probability map is not adequate for high-resolution analyses since spatially-close signals cannot be satisfactorily detected. Therefore, the approach has been improved by implementing a multi-resolution strategy [72].

Despite the positive and attractive features of previous approaches, all of them share the same bottleneck. Indeed, they require the evaluation of the covariance matrix estimated from the measurements of each sensor at different time-instants (i.e., the *snapshots*). This implies an unavoidable increase of the receiver complexity and a delay in the *DoA*s recovery although LBE-based methods have proved to be promising solutions also for real-time localizations [73][74][75].

Starting from the key observation that the signals impinging on the antenna array are intrinsically *sparse* in the spatial domain, efficient strategies for *DoA*s estimation have been proposed [76][77][78] where the sparsity constraints have been imposed through a l_1-norm minimization. In this framework, approaches based on the compressive sensing (CS) theory [81] have recently been introduced because of the computationally efficiency, the accuracy, and the robustness to the noise. Thanks to these features, CS-based strategies have already been applied to a variety of applications in electromagnetic engineering [82][83][84]. However, the main issue to cope with when applying CS is the fact that the so-called 'sampling matrix' must satisfies the restricted isometry property (RIP) for guaranteeing reliable estimations. Unfortunately, such a condition cannot easily

verified since it results computationally demanding [85]. Alternatively, innovative approaches based on the Bayesian compressive sensing (BCS) [41] have been proposed. In such a case, the original deterministic problem is reformulated in its probabilistic counterpart then efficiently solved with the relevance vector machine (RVM) [45]. In this line of reasoning, preliminary attempts in the electromagnetic framework have been out to deal with microwave imaging [86][87] and array synthesis [21] (see also Chapter 4).
In this Chapter, the DoA estimation problem is formulated within the BCS framework thus avoiding constraints on the sampling (or observation) matrix, which directly links the measurements (i.e., voltages/currents) at the output of the array elements to the unknown signal directions. More specifically, two different strategies, extending and completing those preliminary introduced in [88] and [C3], are presented. The former is concerned with single time-instant measurements (i.e., single snapshot) to enable the real-time estimation, while the latter is aimed at giving high-resolution estimations, thanks to the processing over multiple snapshots, still avoiding any *a-priori* information on the number and the intensity of the unknown impinging signals.
The rest of the Chapter is organized as follows. The DoAs estimation problem is mathematically formulated in Sect. 5.2 where the single-snapshot and multiple-snapshots BCS-based approaches are described, as well. A set of representative numerical results is then reported and discussed in Sect. 5.3 where a comparative analysis with reference DoAs estimation methods is also performed. Finally, some conclusions are drawn (Sect. 5.4).

5.2 Mathematical Formulation

5.2.1 DoAs Estimation - Problem Formulation

Let us consider a set of L electromagnetic plane waves $\mathbf{E}_l^{inc}(\mathbf{r}) = E_l^{inc}\, e^{j\beta(x\,sin\theta_l + z\,cos\theta_l)}\hat{\mathbf{y}}$, $l = 1, ..., L$ arriving from unknown directions θ_l, $l = 1, ..., L$ on a linear array composed by M sensors placed along the x-axis with uniform inter-element spacing d (Fig. 5.1). The incident signals are supposed being narrow-band and characterized by the same frequency content. At the sensor locations, $x_m = \left(m - \frac{M+1}{2}\right) d$, $m = 1, ..., M$, the electromagnetic field can be assumed being the linear combination of the signals impinging on the antenna. Accordingly, the relationship between the (complex) open-circuit voltage induced on the receiving elements and the measured signal strengths and propagation delays across the array elements [89] turns out to be [72]

$$v_m = \sum_{l=1}^{L} E_l^{inc}\hat{\mathbf{y}} \cdot \mathbf{f} e^{j\beta x_m sin\theta_l} + n_m \ , \ m = 1, ..., M \tag{5.1}$$

where $\beta = \frac{2\pi}{\lambda}$, λ being the free space wavelength, $\mathbf{f}$ is the antenna effective length

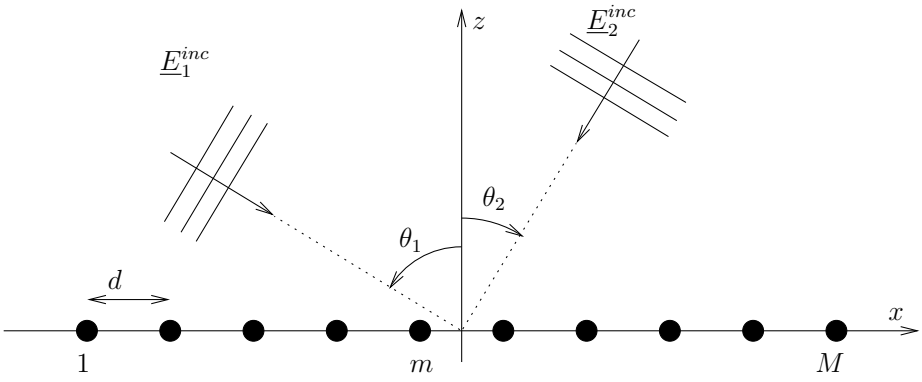


Figure 5.1: Sketch of the reference scenario: linear adaptive antenna array and impinging signals.

supposed identical for all elements[1], and n_m is the m-th ($m = 1, ..., M$) sample from a Gaussian distribution with zero mean that models the additive noise. In matrix form, (6.2) can be rewritten as follows

$$\mathbf{v} = A(\boldsymbol{\theta})\,\mathbf{s} + \mathbf{n} \tag{5.2}$$

where $\mathbf{v} = [v_1, v_2, ..., v_M]^T$ is a column vector of M complex entries ($\mathbf{v} \in \mathbb{C}^{M\times 1}$), T indicates the transpose, $\boldsymbol{\theta} = [\theta_1, .., \theta_L]$, $A(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), ..., \mathbf{a}(\theta_L)] \in \mathbb{C}^{M\times L}$ is the matrix of the steering vectors whose l-th column is given by $\mathbf{a}(\theta_l) = \left[e^{j\beta x_1 sin\theta_l}, e^{j\beta x_2 sin\theta_l}, ..., e^{j\beta x_M sin\theta_l}\right]^T \in \mathbb{C}^{M\times 1}$, $l = 1, ..., L$, $\mathbf{s} = [E_1^{inc}, E_2^{inc}, ..., E_L^{inc}]^T \in \mathbb{C}^{L\times 1}$, and $\mathbf{n} = [n_1, n_2, ..., n_M]^T \in \mathbb{C}^{M\times 1}$. It worth noticing that the problem at hand is non-linear with respect to the unknowns, θ_l, $l = 1, .., L$, which are present in the exponential terms of the elements of the matrix A.
To apply the BCS approach, the visible angular range is discretized with $K \gg L$ samples (Fig. 5.2) such that $A\left(\widetilde{\boldsymbol{\theta}}\right) \in \mathbb{C}^{M\times K}$ in (5.2) and the DoAs of the incoming signals are assumed to belong to the set of the K directions $\tilde{\theta}_k$, $k = 1, ..., K$. Now, the estimation problem turns out to be that of recovering the sparse signal vector $\tilde{\mathbf{s}} \in \mathbb{C}^{K\times 1}$ in correspondence with the user-defined K-sampling of the angular range, $\widetilde{\boldsymbol{\theta}} = \left[\widetilde{\theta}_1, .., \widetilde{\theta}_K\right]$. Since the problem is linear with respect to the unknown $\tilde{\mathbf{s}}$ and the solution is sparse in the spatial domain (i.e., few entries of $\tilde{\mathbf{s}}$ such that $\widetilde{\theta}_k = \theta_l$ are non-null), the BCS theory can be properly applied.

[1]Without loss of generality, isotropic elements are assumed (i.e., $\mathbf{f} = \mathbf{1}$). Extensions to directive or non-uniform arrangements is straightforward.

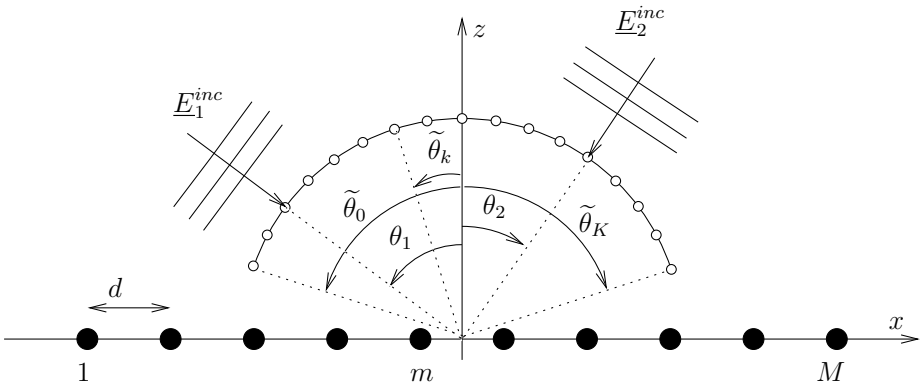


Figure 5.2: Angular region discretization.

5.2.2 Single-Snapshot *BCS*-Based Sparse Signal Estimation

By applying the guidelines of the *BCS* approach in [21] for dealing with complex data, (5.2) is firstly rewritten as

$$\begin{aligned}\begin{bmatrix} \Re\{\mathbf{v}\} \\ \Im\{\mathbf{v}\} \end{bmatrix} = \begin{bmatrix} \Re\left\{A\left(\widetilde{\boldsymbol{\theta}}\right)\right\} & -\Im\left\{A\left(\widetilde{\boldsymbol{\theta}}\right)\right\} \\ \Im\left\{A\left(\widetilde{\boldsymbol{\theta}}\right)\right\} & \Re\left\{A\left(\widetilde{\boldsymbol{\theta}}\right)\right\} \end{bmatrix} \begin{bmatrix} \Re\{\widetilde{\mathbf{s}}\} \\ \Im\{\widetilde{\mathbf{s}}\} \end{bmatrix} + \\ + \begin{bmatrix} \Re\{\mathbf{n}\} \\ \Im\{\mathbf{n}\} \end{bmatrix},\end{aligned} \tag{5.3}$$

$\Re\{\cdot\}$ and $\Im\{\cdot\}$ being the real and the imaginary part, respectively, to yield a real-valued problem suitable for *BCS*. As a matter of fact, although vectors and matrices in (5.3) have double dimensions compared to those in (5.2), all entries are now real. The sparse signal vector $\widehat{\mathbf{s}} = [\Re\{\widetilde{\mathbf{s}}\}, \Im\{\widetilde{\mathbf{s}}\}]^T \in \mathbb{R}^{2K\times 1}$ satisfying (5.3) and having minimum ℓ_0-norm is then obtained in a probabilistic way by solving the following [21]

$$\widehat{\mathbf{s}}_{BCS} = \arg\left[\max_{\widehat{\mathbf{s}}} \mathcal{P}r\left(\left[\widehat{\mathbf{s}},\, \sigma^2,\, \mathbf{p}\right] \middle|\, \mathbf{v}\right)\right] \tag{5.4}$$

where σ^2 is the (unknown) variance of the Gaussian noise and $\mathbf{p}$ is the hyper-parameter vector to be determined and controlling the sparseness of the signal vector $\widehat{\mathbf{s}}$ [45]. By virtue of the fact that

$$\mathcal{P}r\left(\left[\widehat{\mathbf{s}},\, \sigma^2,\, \mathbf{p}\right] \middle|\, \mathbf{v}\right) = \mathcal{P}r\left(\widehat{\mathbf{s}} \middle|\, \left[\mathbf{v},\, \sigma^2,\, \mathbf{p}\right]\right) \mathcal{P}r\left(\left[\sigma^2,\, \mathbf{p}\right] \middle|\, \mathbf{v}\right) \tag{5.5}$$

and the first term on the right of (5.5) is chosen, in the *BCS*-based approach, equal to the multivariate Gaussian distribution [21]

$$\begin{aligned}\mathcal{P}r\left(\widehat{\mathbf{s}} \middle|\, [\mathbf{v},\, \sigma^2,\, \mathbf{p}]\right) = \frac{1}{(2\pi)^{\frac{2K+1}{2}}\sqrt{\det(\Xi)}} \times \\ \exp\left\{-\frac{(\widehat{\mathbf{s}}-\boldsymbol{\mu})^H \Xi^{-1} (\widehat{\mathbf{s}}-\boldsymbol{\mu})}{2}\right\}\end{aligned} \tag{5.6}$$

whose maximum $\widehat{\mathbf{s}}_{BCS}$ coincides with the mean value, the maximization of the posterior probability (5.5) is obtained through the definition of the couple of parameters σ^2_{BCS} and $\mathbf{p}_{BCS}$ maximizing $\mathcal{P}r\left(\left[\sigma^2,\,\mathbf{p}\right]|\,\mathbf{v}\right)$. In (5.6), $\Xi = \left(\frac{1}{\sigma^2}\widehat{A}\left(\widetilde{\boldsymbol{\theta}}\right)^T\widehat{A}\left(\widetilde{\boldsymbol{\theta}}\right) + diag\left(\mathbf{p}\right)\right)^{-1}$ and $\boldsymbol{\mu} = \frac{1}{\sigma^2}\Xi\widehat{A}\left(\widetilde{\boldsymbol{\theta}}\right)^H\mathbf{v}$, where

$$\widehat{A}\left(\widetilde{\boldsymbol{\theta}}\right) = \begin{bmatrix} \Re\left\{A\left(\widetilde{\boldsymbol{\theta}}\right)\right\} & -\Im\left\{A\left(\widetilde{\boldsymbol{\theta}}\right)\right\} \\ \Im\left\{A\left(\widetilde{\boldsymbol{\theta}}\right)\right\} & \Re\left\{A\left(\widetilde{\boldsymbol{\theta}}\right)\right\} \end{bmatrix} \tag{5.7}$$

is the real-valued matrix of the steering vectors and H denotes the conjugate transpose operation. Since

$$\mathcal{P}r\left(\left[\sigma^2,\,\mathbf{p}\right]\big|\,\mathbf{v}\right) \propto \mathcal{P}r\left(\mathbf{v}|\left[\sigma^2,\,\mathbf{p}\right]\right)\mathcal{P}r\left(\sigma^2\right)\mathcal{P}r\left(\mathbf{p}\right) \tag{5.8}$$

and the two terms $\mathcal{P}r\left(\sigma^2\right)$ and $\mathcal{P}r\left(\mathbf{p}\right)$ are constant according to the guidelines of [45], the optimal parameters σ^2_{BCS} and $\mathbf{p}_{BCS}$ are computed through the relevance vector machine (RVM) by maximizing the logarithm of $\mathcal{P}r\left(\mathbf{v}|\left[\sigma^2,\,\mathbf{p}\right]\right)$ defined as [21]

$$\mathcal{L}^{BCS}\left(\sigma^2,\mathbf{p}\right) = -\frac{1}{2}\left[(2K)\log 2\pi + \log|C_{BCS}| + \mathbf{v}^T C^{-1}_{BCS}\mathbf{v}\right] \tag{5.9}$$

where an user-defined initial value for σ^2, $\sigma^2 = \sigma^2_0$, is chosen. Moreover in (5.9), $C_{BCS} \triangleq \sigma^2 I + \widehat{A}\left(\widetilde{\boldsymbol{\theta}}\right) diag\left(\mathbf{p}\right)^{-1}\widehat{A}\left(\widetilde{\boldsymbol{\theta}}\right)^T$. Once σ^2_{BCS} and $\mathbf{p}_{BCS}$ are determined, the estimated solution turns out to be

$$\begin{aligned} \widehat{\mathbf{s}}_{BCS} \;=\; & \frac{1}{\sigma^2_{BCS}}\left(\frac{\widehat{A}\left(\widetilde{\boldsymbol{\theta}}\right)^T\widehat{A}\left(\widetilde{\boldsymbol{\theta}}\right)}{\sigma^2_{BCS}} + diag\left(\mathbf{p}_{BCS}\right)\right)^{-1} \times \\ & \widehat{A}\left(\widetilde{\boldsymbol{\theta}}\right)^T\mathbf{v}\,. \end{aligned} \tag{5.10}$$

5.2.3 Multiple-Snapshot $MT-BCS$-Based Sparse Signal DoA Estimation

Unlike the $ST-BCS$, the $MT-BCS$ approach [60] correlates the DoAs estimation over multiple snapshots, thus avoiding the strong dependence of the estimation performance on the noise level of the collected measurements. With reference to the multiple-snapshots version of Eq. (5.2)

$$\mathbf{v}_w = A\left(\boldsymbol{\theta}\right)\mathbf{s}_w + \mathbf{n}_w,\;\; w = 1,...,W, \tag{5.11}$$

W being the number of snapshots, the sparse signal vector $\widehat{\mathbf{s}}$ is here determined as follows

$$\widehat{\mathbf{s}}_{MT-BCS} = \frac{1}{W}\sum_{w=1}^{W}\left\{\arg\left[\max_{\widehat{\mathbf{s}}_w}\mathcal{P}r\left(\left[\widehat{\mathbf{s}}_w,\,\mathbf{p}\right]|\,\mathbf{v}_w\right)\right]\right\} \tag{5.12}$$

where $\widehat{\mathbf{s}}_w$, $w = 1, ..., W$, are statistically-correlated through a proper definition of the "shared" hyperparameter vector correlating the different snapshots. The optimal value of $\mathbf{p}$, $\mathbf{p}_{MT-BCS}$, is computed as $\mathbf{p}_{MT-BCS} = \arg\max_{\mathbf{p}}\left\{\mathcal{L}^{MT-BCS}\left(\mathbf{p}\right)\right\}$ through the RVM according to the guidelines in Chapter 4, being

$$\begin{array}{c}\mathcal{L}^{MT-BCS}\left(\mathbf{p}\right) = -\frac{1}{2}\sum_{w=1}^{W}\left\{\log\left(\left|C_{MT-BCS}\right|\right) + \right. \\ \left.\left(K+2\psi_1\right)\log\left[\mathbf{v}_w^T\left(C_{MT-BCS}\right)\mathbf{v}_w + 2\psi_2\right]\right\}\end{array} \tag{5.13}$$

where $C_{MT-BCS} \triangleq I + \widehat{A}\left(\widetilde{\boldsymbol{\theta}}\right) diag\left(\mathbf{p}\right)^{-1}\widehat{A}\left(\widetilde{\boldsymbol{\theta}}\right)^T$ and ψ_1, ψ_2 are user-defined parameters [60]. Unlike the BCS approach, the knowledge/estimation of the variance σ^2 of the noise samples is not required in the $MT-BCS$ based method (see Chapter 4). Finally, the solution estimated by means of the $MT-BCS$ turns out equal to

$$\begin{array}{l}\widehat{\mathbf{s}}_{MT-BCS} = \\ = \frac{\sum_{w=1}^{W}\left\{\left[\widehat{A}\left(\widetilde{\boldsymbol{\theta}}\right)^T\widehat{A}\left(\widetilde{\boldsymbol{\theta}}\right)+diag\left(\mathbf{p}_{MT-BCS}\right)\right]^{-1}\widehat{A}\left(\widetilde{\boldsymbol{\theta}}\right)^T\mathbf{v}_w\right\}}{W}.\end{array} \tag{5.14}$$

5.2.4 *DoA* Estimation Procedure

In principle, the estimated number of impinging signals, $\widetilde{L}$, can be determined by simply counting the non-zero elements of the recovered signal vector $\widetilde{\mathbf{s}}$. However, many entries of $\widetilde{\mathbf{s}}$ can assume amplitudes close but not equal to zero that do not correspond to any actual signal due to the presence of the noise. Accordingly, the original L-sparse signal turns out being a compressible one where the strongest $\widetilde{L}$ signals have to be selected. Towards this aim, an energetic thresholding is applied to remove the lowest-energy components of $\widetilde{\mathbf{s}}$ in order to improve the reliability of the *DoA*s estimation. More specifically, the entries of the estimated sparse signal $\widetilde{\mathbf{s}}$ are sorted according to their energy content, $\left|\widetilde{s}_k\right|^2$, $k = 1, ..., K$, such that $\xi_1 \triangleq \max_k\left\{\left|\widetilde{s}_k\right|^2\right\}$ and $\xi_K \triangleq \min_k\left\{\left|\widetilde{s}_k\right|^2\right\}$. Successively, only the first $\widetilde{L}$ directions such that

$$\frac{1}{\left(\sum_{k=1}^{K}\xi_k\right)}\sum_{l=1}^{\widetilde{L}}\xi_l < \eta \tag{5.15}$$

are kept and assumed as those of the actual signals, η being a user-defined threshold (Fig. 5.3). Accordingly, the k-th thresholded element of $\widetilde{\mathbf{s}}$ turns out to be

$$\widetilde{s}_k\rfloor_\eta = \left\{\begin{array}{ll}\widetilde{s}_k & if\ \left|\widetilde{s}_k\right|^2 > \xi_{\widetilde{L}} \\ 0 & otherwise\end{array}\right. \tag{5.16}$$

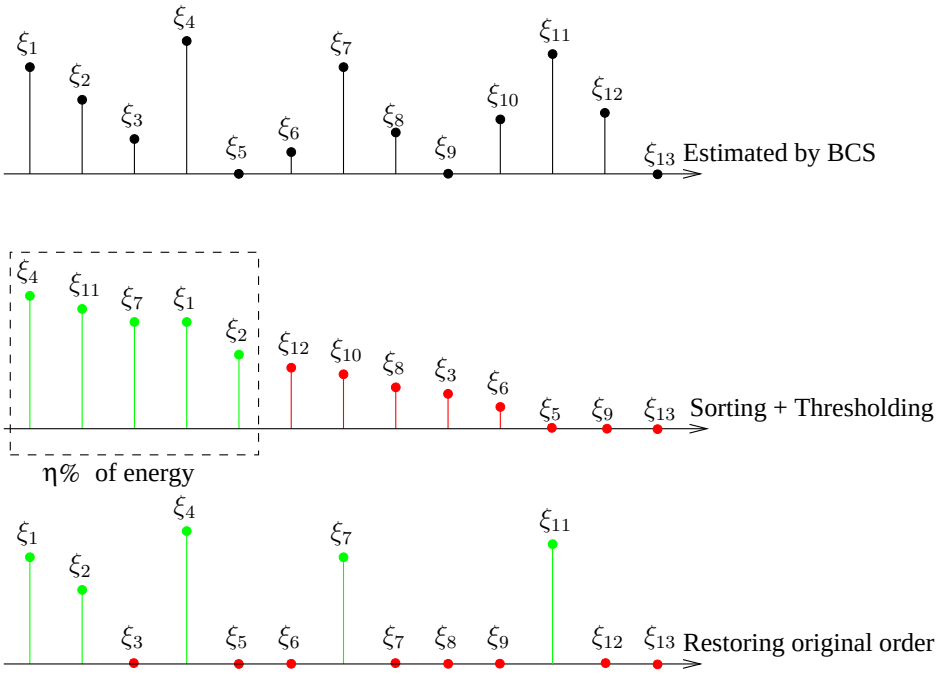


Figure 5.3: Sketch of the energy thresholding strategy for the estimation of the number of incident signals $\widetilde{L}$.

and the estimated DoAs, $\widetilde{\theta}_l$, $l = 1, ..., \widetilde{L}$, are determined as follows

1. Set $l = k = 1$;
2. If $\widetilde{s}_k\rfloor_\eta \neq 0$ then $\widetilde{\theta}_l = \widetilde{\theta}_k$, $l = l + 1$;
3. If $k < K$ then $k = k + 1$ and goto **2**; else stop.

5.3 Numerical Results

In the following, a set of numerical results is reported and discussed to show the behavior of the proposed approaches as well as to point out their advantages and drawbacks also in a comparison with state-of-the-art methods. Firstly, an analysis on the sensitivity on the calibration parameters (namely, the energy threshold η and the noise parameter σ_0^2) is carried out. Successively, the estimation capabilities of the BCS-based strategies are assessed dealing with single-snapshot and multiple-snapshots acquisitions. As for the $MT-BCS$, the parameters ψ_1, ψ_2 are set as in [95].

5.3.1 Sensitivity Analysis

In order to determine the optimal values of the control parameters σ_0^2 and η, the following benchmark test case has been considered: an antenna array of $M = 20$ isotropic sensors equally-spaced by $d = \frac{\lambda}{2}$ along the x-axis and a set of $L = \{2, 4, 6\}$ binary phase-shift keying ($BPSK$) signals ($E_l^{inc} = \pm 1$) impinging on the array from $\theta_l \in [-90^o, 90^o]$, $l = 1, ..., L$. The signals have been supposed arriving on the antenna with equal strength in order to perform an unbiased analysis of the accuracy of the method with respect to the angles of arrival. The minimum angular distance between the DoAs of two adjacent signals has been set to $\Delta\theta_{min} = 1^o$, while the angular range has been uniformly discretized into $K = 181$ samples. The measured data are characterized by signal-to-noise ratio equal to $SNR = \{2, 5, 10, 20\}$ dB, defined as

$$SNR = 10 \log \left[\frac{\sum_{m=1}^{M} \left| v_m^{Noiseless} \right|^2}{M\sigma^2} \right] \tag{5.17}$$

where σ^2 is the variance of the additive Gaussian noise and $v_m^{Noiseless}$, $m = 1, ..., M$ are the noise-free data. Since the actual DoAs are randomly chosen, $Q = 250$ different scenarios (i.e., $\theta_l^{(q)}$, $l = 1, ..., L$, $q = 1, ..., Q$) have been taken into account for each combination of L and SNR to give a consistent statistical validation. The BCS-based estimation has been applied varying the calibration parameters within the ranges $\eta \in [0.0, 1.0]$ and $\sigma_0^2 \in [10^{-6}, 1.0]$.
The optimal setup of the control parameters has been defined by choosing the values of η and σ_0^2 that minimize the modified root-mean-square error ($\overline{RMSE}$)

$$\left(\sigma_0^2, \eta\right)^{(opt)} = \arg \left\{ \min_{\left(\sigma_0^2, \eta\right)} \left[\overline{RMSE} \left(\sigma_0^2, \eta\right) \right] \right\} \tag{5.18}$$

where

$$\begin{aligned} &\overline{RMSE} \left(\sigma_0^2, \eta\right) = \\ &= \sum_L \int_{SNR} \frac{RMSE\left(\sigma_0^2, \eta | SNR, L\right)}{\max_{\left(\sigma_0^2, \eta\right)} \left\{ RMSE\left(\sigma_0^2, \eta | SNR, L\right) \right\}} dSNR \end{aligned} \tag{5.19}$$

and $RMSE = \frac{1}{Q} \sum_{q=1}^{Q} RMSE^{(q)}$, $RMSE^{(q)}$ being an indicator of the reliability of the method in predicting the q-th scenario. This latter takes into account both the errors in estimating the signal number $\widetilde{L}^{(q)}$ and the corresponding DoAs $\widetilde{\theta}_l^{(q)}$, $l = 1, ..., \widetilde{L}^{(q)}$. It is defined as follows

$$\begin{aligned} &RMSE^{(q)} = \\ &= \begin{cases} \sqrt{\frac{\left\{ \sum_{l=1}^{\widetilde{L}^{(q)}} \left| \theta_l - \widetilde{\theta}_l^{(q)} \right|^2 + \left| L - \widetilde{L}^{(q)} \right| (\Delta\theta_{max})^2 \right\}}{L}} \; if \;\; \widetilde{L}^{(q)} \leq L \\ \sqrt{\frac{\left\{ \sum_{l=1}^{L} \left| \theta_l - \widetilde{\theta}_l^{(q)} \right|^2 + \sum_{j=L+1}^{\widetilde{L}^{(q)}} \left| \widetilde{\theta}_j^{(q)} - \overline{\theta}_j^{(q)} \right|^2 \right\}}{L}} \; if \;\; \widetilde{L}^{(q)} > L \end{cases} \end{aligned} \tag{5.20}$$

where θ_l and $\widetilde{\theta}_l^{(q)}$ are the l-th actual and the closest (among the $\widetilde{L}^{(q)}$ estimates) retrieved DoA, respectively, while $\Delta\theta_{max}$ is a penalty term equal to the maximum admissible localization error (i.e., $\Delta\theta_{max} = 180^o$) and applied when the number of estimated signals is smaller than the actual one. Moreover,

$$\overline{\theta}_j^{(q)} = \arg\left\{ \min_{\theta_l,\, l\in[1,L]} \left| \theta_l - \widetilde{\theta}_j^{(q)} \right| \right\}. \tag{5.21}$$

It is worth pointing out that (5.20) coincides with the standard $RMSE$ definition of the literature when $\widetilde{L}^{(q)} = L$, while it penalizes the cases when $\widetilde{L}^{(q)} < L$ since it is assumed that, at the receiver, it is preferable to identify at least the signals which are really present in the environment, also admitting the prediction of non-existing signals, than missing the identification of one or more actual signals.

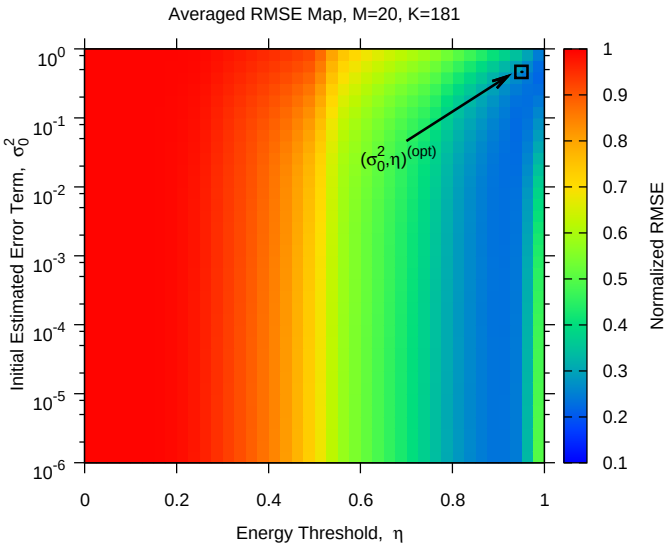


Figure 5.4: *BCS-Calibration* ($M = 20$; $d = 0.5\lambda$; $L = \{2, 4, 6\}$; $\theta_l \in [-90^o, 90^o]$; $SNR = \{2, 5, 10, 20\}$ dB; $Q = 250$; $\sigma_0^2 \in [10^{-6}, 1.0]$; $\eta \in [0, 1]$). Normalized average $RMSE$ (5.18) vs σ_0^2 and η.

	$RMSE\left(\sigma_0^2,\eta\right)^{(opt)}$			$P_L\left(\sigma_0^2,\eta\right)^{(opt)}$ [%]		
SNR $[dB]$	$L=2$	$L=4$	$L=6$	$L=2$	$L=4$	$L=6$
2	35.01	43.75	74.23	18.4	20.0	20.0
5	14.88	41.47	70.92	64.4	43.2	22.8
10	7.05	32.12	66.47	89.2	55.2	24.4
20	8.14	27.15	49.20	92.4	59.2	24.4

Table 5.1: *Single Snapshot* ($W=1$) *DoA Estimation* ($M=20$, $d=0.5\lambda$; $L=\{2,4,6\}$, $\theta_l\in[-90^o,90^o]$, $Q=250$; $SNR=\{2,5,10,20\}$ dB; $\sigma_0^2=0.46$, $\eta=0.95$). Average $RMSE$ and P_L values.

Figure 5.4 shows the plot of the normalized $\overline{RMSE}\left(\sigma_0^2,\eta\right)$, where the minimum value of $\overline{RMSE}$ occurs at $\left(\sigma_0^2,\eta\right)^{(opt)}=(0.46,0.95)$ which is assumed as the optimal setup hereinafter. As an example, Table 5.1 gives the $RMSE$ values for a set of representative combinations of L and SNR when setting $\left(\sigma_0^2,\eta\right)^{(opt)}$. As expected, the estimation accuracy improves for higher SNRs and decreasing $\frac{L}{M}$ values. For completeness, the percentage of faithfully detected scenarios (i.e., $\widetilde{L}^{(q)}=L$), P_L:

$$P_L\left(\sigma_0^2,\eta|SNR,L\right)=\frac{1}{Q}\sum_{q=1}^{Q}P_L^{(q)}\left(\sigma_0^2,\eta|SNR,L\right) \tag{5.22}$$

where

$$P_L^{(q)}\left(\sigma_0^2,\eta|SNR,L\right)=\begin{cases}1 & if\ \widetilde{L}^{(q)}=L\\ 0 & otherwise\end{cases},\ q=1,...,Q, \tag{5.23}$$

is reported, as well (Tab. 5.1). Similarly to the $RMSE$ behavior, the P_L improves when the noise level decreases and the number of impinging signals is smaller than the number of array sensors.

5.3.2 Performance Assessment (Single-Snapshot BCS-Based Estimation Approach)

With reference to the single-snapshot acquisition, let us consider the test case characterized by $L=4$ and $SNR=10\,dB$. To illustrate the behavior of the BCS-based estimation approach, the results in Fig. 5.5 refer to three representative situations: $\widetilde{L}^{(q)}=L$ [Figs. 5.5(a)-(b)], $\widetilde{L}^{(q)}>L$ [Figs. 5.5(c)-(d)], and $\widetilde{L}^{(q)}<L$ [Fig. 5.5(e)] corresponding to low [Figs. 5.5(a)-(c)] or high [Figs. 5.5(b)-(d)] $RMSE$ [when $\widetilde{L}^{(q)}<L$ the $RMSE$ value turns out being always high due to the presence of the penalty term in (5.20)]. For illustrative purposes, the symbols + and $\times$ indicate the actual DoAs and those estimated after thresholding, while the green dots are the BCS estimates before thresholding. Since

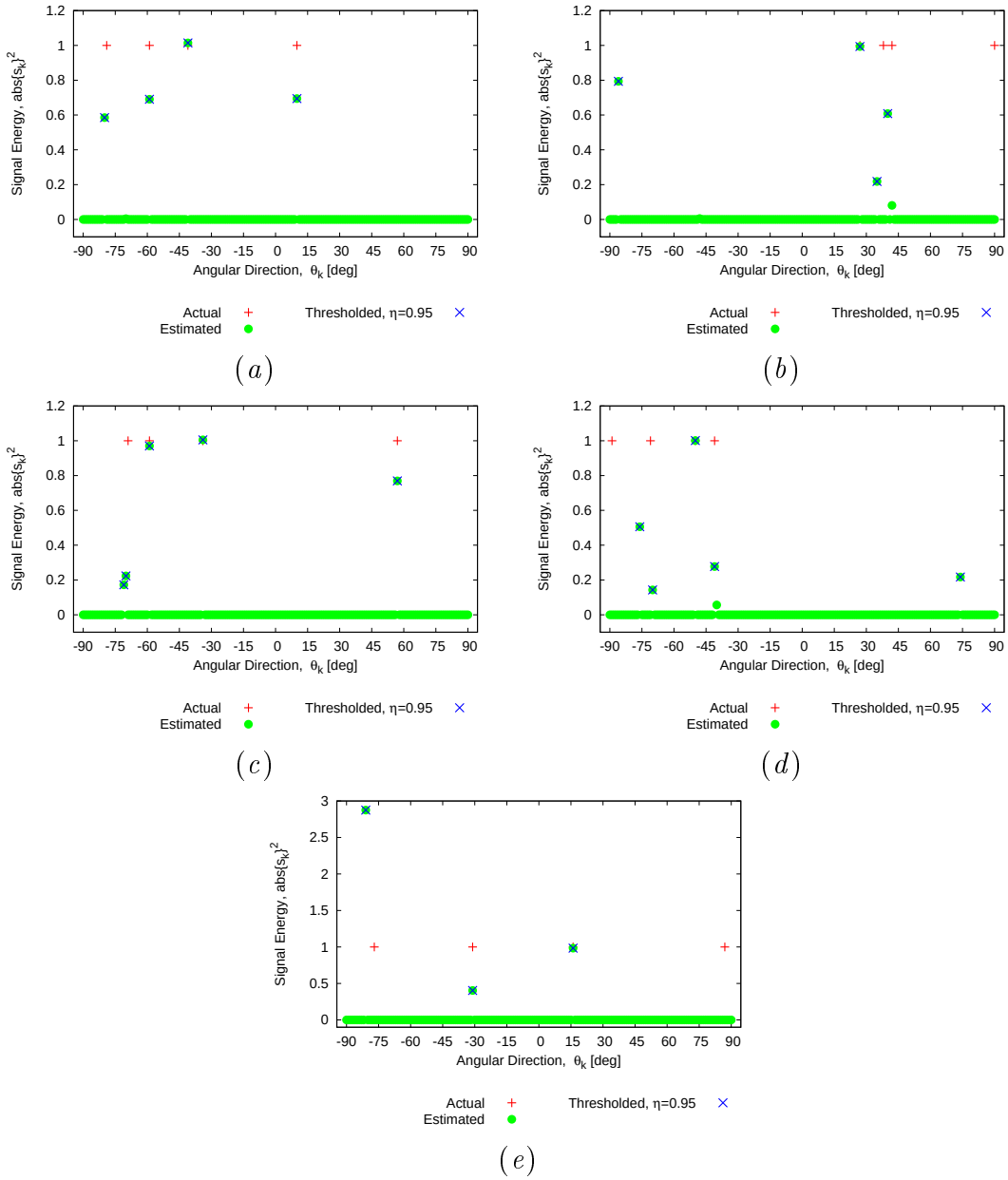


Figure 5.5: *Single Snapshot* ($W = 1$) *DoA Estimation* ($M = 20$, $d = 0.5\lambda$; $L = 4$; $SNR = 10\,dB$; $\sigma_0^2 = 0.46$, $\eta = 0.95$). Representative examples of actual and estimated *DoA*s when (a) $\widetilde{L} = L$ and low $RMSE$, (b) $\widetilde{L} = L$ and high $RMSE$, (c) $\widetilde{L} > L$ and low $RMSE$, (d) $\widetilde{L} > L$ and high $RMSE$, and (e) $\widetilde{L} < L$.

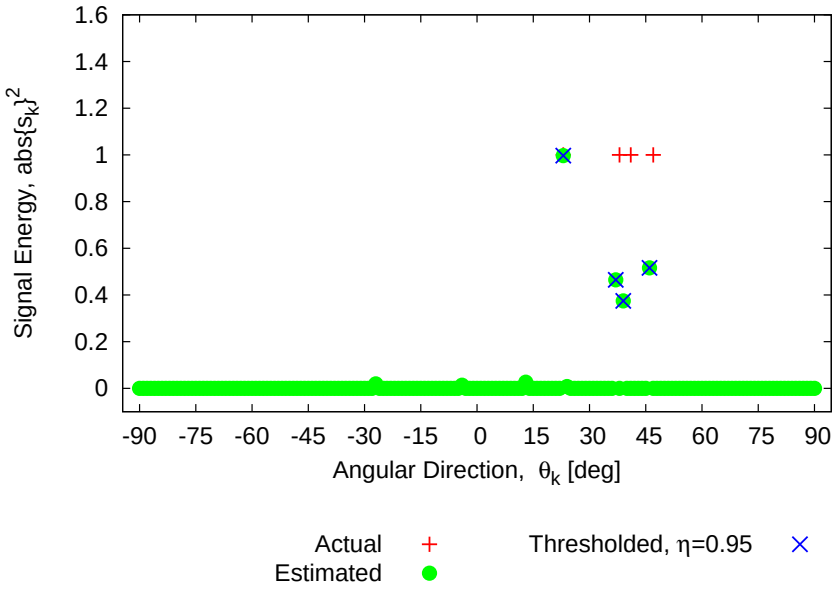
Figure	$\boldsymbol{\theta}$	$\widetilde{\boldsymbol{\theta}}$	$RMSE$
5.5(a)	$[-79, -59, -41, 10]$	$[-80, -59, -41, 10]$	0.50
5.5(b)	$[27, 38, 42, 90]$	$[-86, +27, +35, +40]$	62.13
5.5(c)	$[-69, -59, -34, 57]$	$[-71, -70, -59, -34, 57]$	1.12
5.5(d)	$[-89, -71, -50, -41]$	$[-76, -70, -50, -41, +74]$	58.87
5.5(e)	$[-77, -31, 16, 87]$	$[-81, -31, 16]$	90.02

Table 5.2: *Single Snapshot* ($W = 1$) *DoA Estimation* ($M = 20$, $d = 0.5\lambda$; $L = 4$; $SNR = 10\,dB$; $\sigma_0^2 = 0.46$, $\eta = 0.95$). Actual directions and estimated *DoA*s. $RMSE$ values.

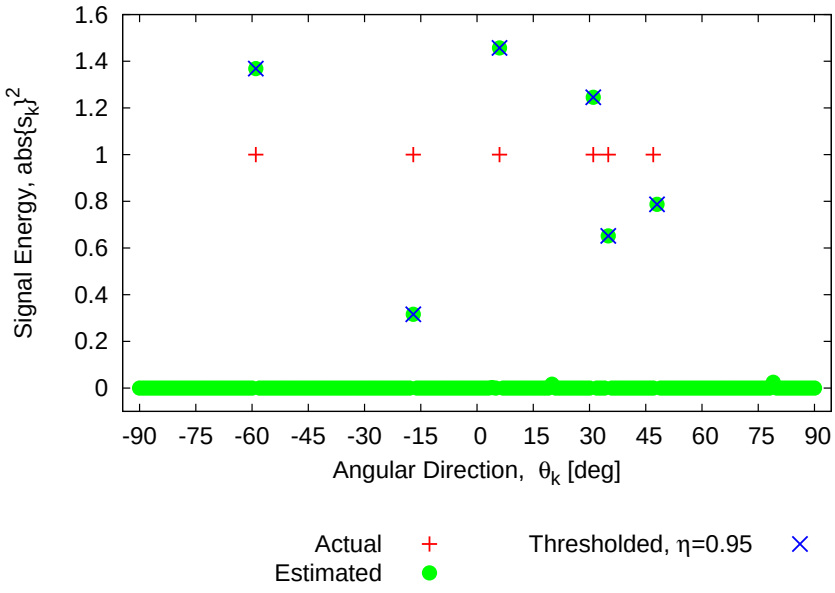
l	1	2	3	4	5	6
$L = \widetilde{L} = 4$						
$\boldsymbol{\theta}$	23	38	41	47	–	–
$\widetilde{\boldsymbol{\theta}}$	23	37	39	46	–	–
$L = \widetilde{L} = 6$						
$\boldsymbol{\theta}$	−59	−17	6	31	35	47
$\widetilde{\boldsymbol{\theta}}$	−59	−17	6	31	35	48

Table 5.3: *Single Snapshot* ($W = 1$) *DoA Estimation* ($M = 20$, $d = 0.5\lambda$; $SNR = 2\,dB$; $\sigma_0^2 = 0.46$, $\eta = 0.95$). Actual and estimated *DoA*s when $\widetilde{L} = L$: $L = 4$ and $L = 6$.

the method determines the signal vector $\widetilde{\mathbf{s}}$, the estimated signal strengths are reported, as well, to prove that several entries are null because of the sparsity constraint enforced through the BCS.
Because of the limited information from the single-snapshot acquisition and the unavoidable presence of the noise, the main concern is the detection of impinging sources located close to the end-fire angular direction [Fig. 5.5(b) - $\theta_4 = 90^o$, Fig. 5.5(d) - $\theta_1 = -89^o$, Fig. 5.5(e) - $\theta_2 = 87^o$] as pointed out by the $RMSE$ values in Tab. 5.2 where both actual and estimated *DoA*s are reported, as well. Otherwise, the *DoA*s are retrieved with a high precision [e.g., $RMSE = 0.50^o$ - Fig. 5.5(a) and $RMSE = 1.12^o$ - Fig. 5.5(c)] even in most severe noisy conditions (e.g., $SNR = 2\,dB$) for both more [Fig. 5.6(a)] and less [Fig. 5.6(b)] densely distributed signals. Quantitatively, the estimation errors amount to $RMSE_{L=4} = 1.22^o$ [Fig. 5.6(a)] and $RMSE_{L=6} = 0.41^o$ [Fig. 5.6(b)], respectively, and Tab. 5.3 reports the values of the actual and estimated *DoA*s.



(a)



(b)

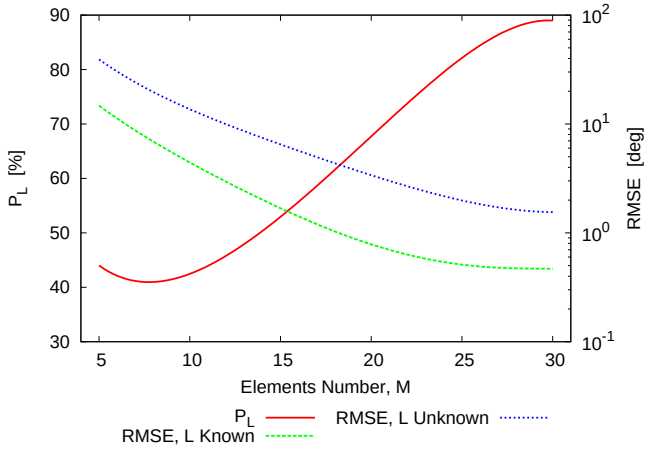
Figure 5.6: *Single Snapshot* ($W = 1$) *DoA Estimation* ($M = 20$, $d = 0.5\lambda$; $SNR = 2\,dB$; $\sigma_0^2 = 0.46$, $\eta = 0.95$). Actual and estimated *DoA*s when $\widetilde{L} = L$: (a) $L = 4$ and (b) $L = 6$.

To draw more general outcomes on the behavior of the single-snapshot BCS-based DoA estimator, further experiments have been carried out varying the number of receiving elements M, the minimum spacing between the signals $\Delta\theta$, and the signal-to-noise ratio. The values of $RMSE$ (5.20) and P_L (5.23) averaged over $Q = 100$ simulations for each scenario at hand have been assumed as reliability/accuracy indicators. More specifically, the $RMSE$ has been computed either without or with the *a-priori* knowledge on the number of incident signals, L. In this latter case, the energetic thresholding has not been applied and the first L signals having higher strength have been selected.
With reference to the representative test case with $L = 2$ and $SNR = 7\,dB$ (Fig. 5.7), Figure 5.7(a) shows that, as expected, increasing the number of data when using more sensors (i.e., using larger arrays since $d = \frac{\lambda}{2}$ has been set), drastically reduces the $RMSE$ of more than one order of magnitude when L is *a-priori* known ($RMSE_{M=5} = 14.75^o \rightarrow RMSE_{M=30} = 0.47^o$) or totally unknown ($RMSE_{M=5} = 39.11^o \rightarrow RMSE_{M=30} = 1.55^o$). Analogously, P_L improves from $P_L\rfloor_{M=5} \simeq 40\%$ up to $P_L\rfloor_{M=30}\, 90\%$.

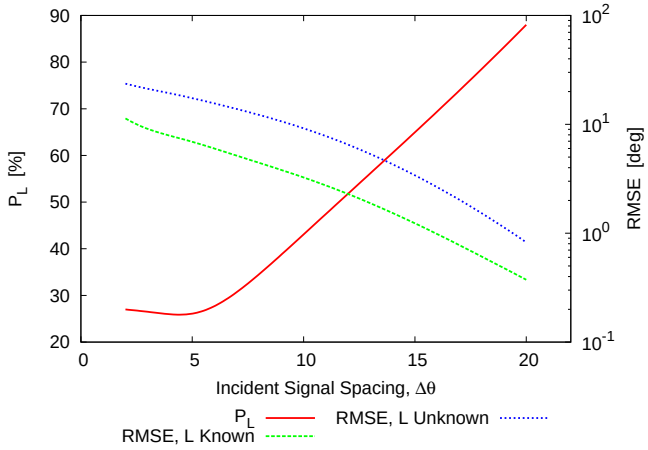
Similar conclusions hold true for the analyses whose results are summarized in Figs. 5.7(b)-5.7(c). Except for the behavior of P_L in Fig. 5.7(c), where an almost constant threshold is yielded from $SNR = 10\,dB$, both $RMSE$ and P_L behave as in Fig. 5.7(a). Of course, the knowledge of L gives reduced errors, but the gap between the two estimates is still close whatever the variable at hand (i.e., M, $\Delta\theta$, SNR): $\Delta RMSE = 0.64$ [Fig. 5.7(a)] ($\Delta RMSE \triangleq \frac{RMSE_{L\,known} - RMSE_{L\,unk}}{RMSE_{L\,unk}}$), $\Delta RMSE = 0.63$ [Fig. 5.7(b)], and $\Delta RMSE = 0.52$ [Fig. 5.7(c)]. Such a result further confirms a key-feature of the BCS estimation, that is, its high reliability also when no information on the scenario is available.

5.3.3 Performance Assessment ($MT-BCS$-Based Estimation Approach)

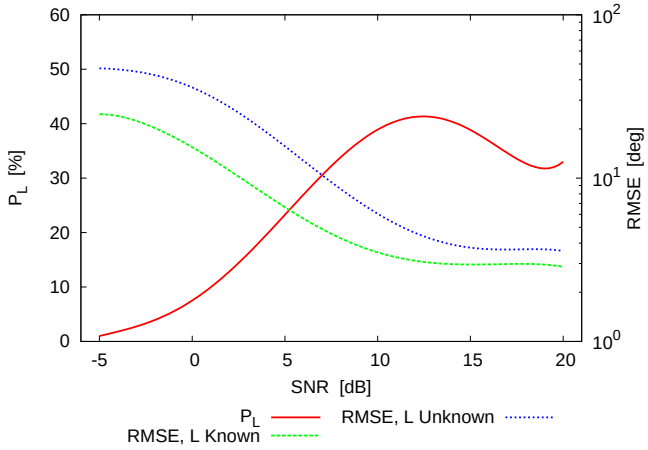
Dealing with multiple-snapshots, the MT implementation of the BCS-estimator ($MT-BCS$) has been used. Firstly, the same test cases of Fig. 5.5 have been considered to perform a comparison with the $ST-BCS$ performances. Towards this end, $W = 25$ consecutive time instants have been considered for modeling the multi-snapshots acquisition. Figure 5.8 shows the $MT-BCS$ estimates, while the corresponding $RMSE$ values are given in Tab. 5.4. As it can be observed, the $MT-BCS$ (Tab. 5.4) outperforms the single-snapshot ($W = 1$) $ST-BCS$ (Tab. 5.2) whatever the scenario at hand. As a matter of fact,



(a)



(b)



(c)

Figure 5.7: *Single Snapshot* ($W = 1$) *DoA Estimation* ($d = 0.5\lambda$; $L = 2$). Plots of P_L and $RMSE$: (a) $M \in [5, 30]$, (b) $\Delta\theta \in [2, 20]$, and (c) $SNR \in [-5, 20]$ dB.

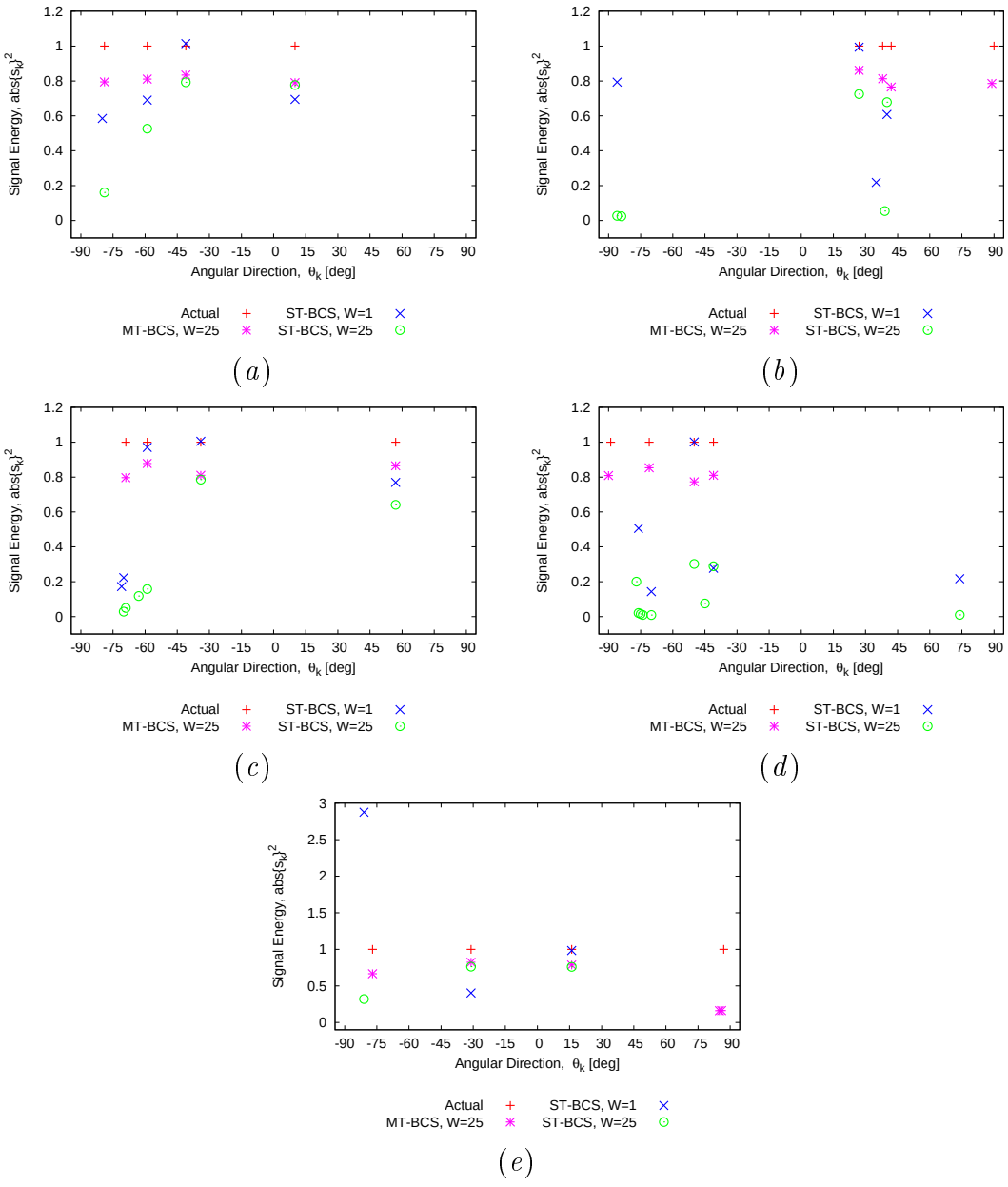


Figure 5.8: *Multiple Snapshots* ($W = 25$) *DoA Estimation* ($M = 20$, $d = 0.5\lambda$; $L = 4$; $SNR = 10\,dB$; $\sigma_0^2 = 0.46$, $\eta = 0.95$). Actual and estimated *DoA*s by means of the $MT - BCS$ and the multi-snapshots $ST - BCS$.

	$MT-BCS$, $W=25$		$ST-BCS$, $W=25$	
Figure	$\widetilde{\boldsymbol{\theta}}$	$RMSE$	$\widetilde{\boldsymbol{\theta}}$	$RMSE$
5.8(a)	$[-79,-59,$ $-41,10]$	0.00	$[-79,-59,$ $-41,10]$	0.00
5.8(b)	$[27,38,42,89]$	0.50	$[-86,27,39,40]$	83.25
5.8(c)	$[-69,-59,$ $-34,57]$	0.00	$[-63,-59,$ $-34,57]$	2.06
5.8(d)	$[-90,-71,$ $-50,-41]$	0.50	$[-77,-50,$ $-45,-41]$	57.96
5.8(e)	$[-77,-31,$ $16,86]$	0.50	$[-81,-31,16]$	90.02

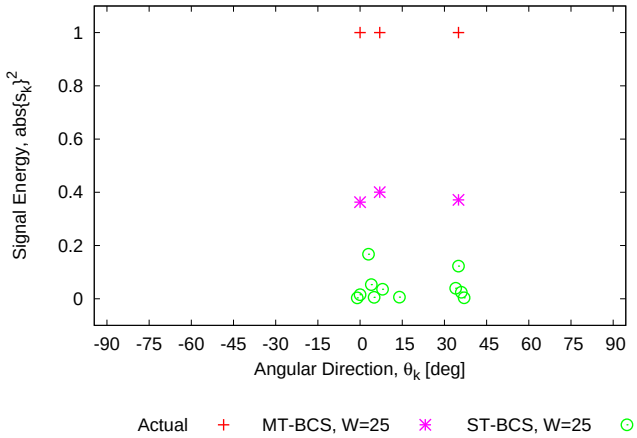
Table 5.4: *Multiple Snapshots* ($W=25$) *DoA Estimation* ($M=20$, $d=0.5\lambda$; $L=4$; $SNR=10\,dB$). *DoA*s estimated with the $MT-BCS$ and the multi-snapshots $ST-BCS$. $RMSE$ values.

although the intrinsic difficulty to correctly retrieve the *DoA*s of signals close to end-fire due to the fact that the antenna effective aperture tends to zero, better estimations than Fig. 5.5 have been obtained for such critical situations thanks to the $MT-BCS$ features [see Fig. 5.8(b), Fig. 5.8(d), and Fig. 5.8(e)]. On the other side, the *DoA*s of signals far from directions $\theta=\pm 90^o$ are instead precisely estimated (Tab. 5.4).

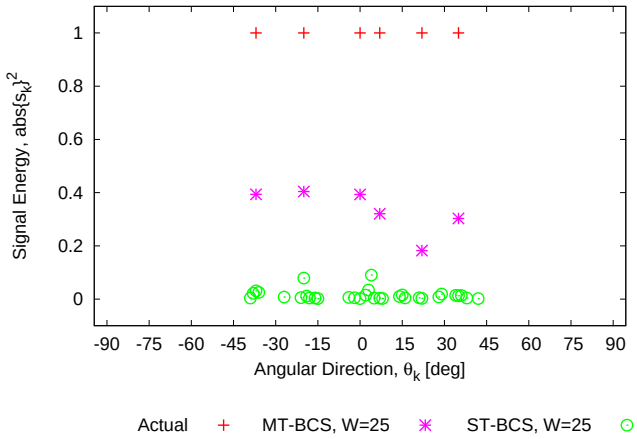
To investigate whether such an improvement is due to the MT implementation or only arises from the multi-snapshots acquisition, the multi-snapshot data ($W=25$) have been processed with the $ST-BCS$ as follows

$$\widetilde{\mathbf{s}}^{avg}_{ST-BCS} = \frac{\sum_{w=1}^{W}\left[\frac{1}{\sigma^2_{BCS}\rfloor_w}\left(\frac{\widehat{A}(\widetilde{\boldsymbol{\theta}})^T\widehat{A}(\widetilde{\boldsymbol{\theta}})}{\sigma^2_{BCS}\rfloor_w}+diag\left(\mathbf{p}_{BCS}\rfloor_w\right)\right)^{-1}\widehat{A}(\widetilde{\boldsymbol{\theta}})^T\mathbf{v}_w\right]}{W} \tag{5.24}$$

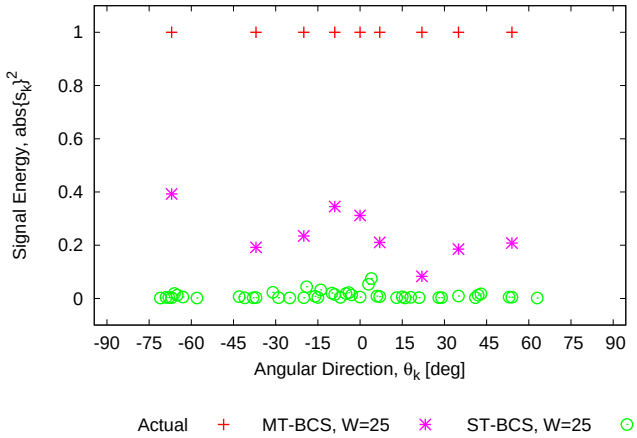
then applying the energetic filtering (5.15) on $\widetilde{\mathbf{s}}^{avg}_{ST-BCS}$. The results of such a processing are reported in Fig. 5.8 with the $RMSE$ values in Tab. 5.4. It is worth noting that the performance of the multi-snapshots $ST-BCS$ ($W=25$) does not significantly improve and the errors in estimating the *DoA*s turn out almost unaltered. This is caused by the impossibility for the $ST-BCS$ to correlate the estimates from different snapshots although related to the same scenario.



(a)

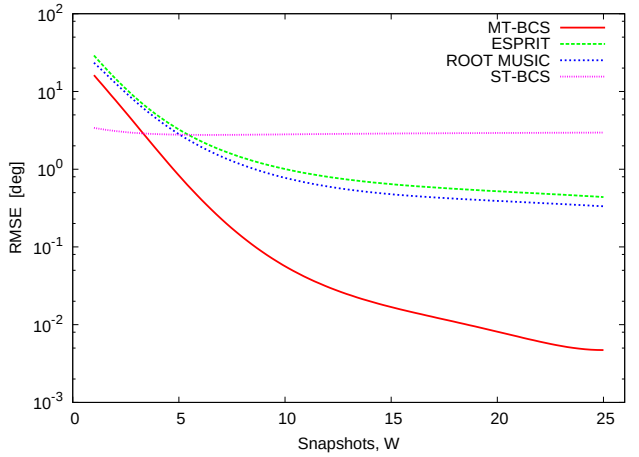


(b)

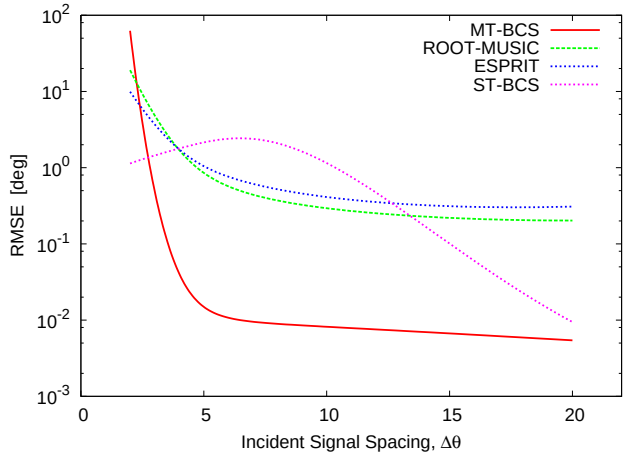


(c)

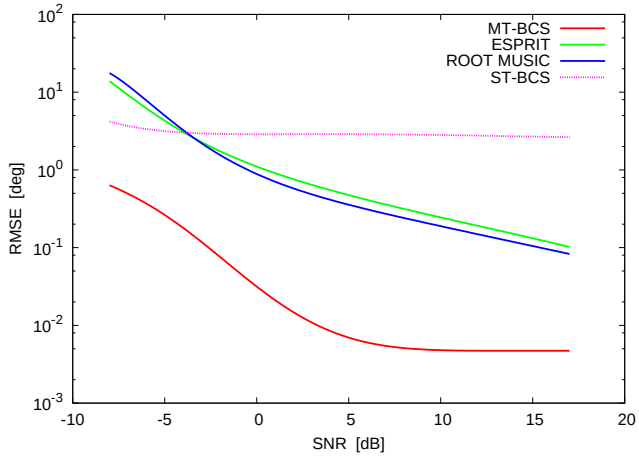
Figure 5.9: *Multiple Snapshots* ($W = 25$) *DoA Estimation* ($M = 10$, $d = 0.5\lambda$; $SNR = 7\,dB$). Actual and estimated *DoA*s by means of the proposed BCS and $MT - BCS$ methods when (a) $L = 3$, (b) $L = 6$, and (c) $L = 9$.



(a)



(b)



(c)

Figure 5.10: *Multiple Snapshots DoA Estimation* ($M = 10$, $d = 0.5\lambda$; $L = 2$) - Plot of the $RMSE$ when (a) $W \in [1, 25]$ - $\Delta\theta = 7^o$ - $SNR = 7\,dB$, (b) $\Delta\theta \in [2^o, 20^o]$ - $SNR = 7\,dB$ - $W = 20$, and (c) $SNR \in [-5, 20]\ dB$ - $\Delta\theta = 7^o$ - $W = 20$ by applying the $MT - BCS$, the multi-snapshots $ST - BCS$, the $ROOT - MUSIC$ [90], and the $ESPRIT$ [91] estimator.

		$MT-BCS$, $W=25$		$ST-BCS$, $W=25$	
Fig.	$\boldsymbol{\theta}$	$\widetilde{\boldsymbol{\theta}}$	$RMSE$	$\widetilde{\boldsymbol{\theta}}$	$RMSE$
5.9(a)	$[0, 7, 35]$	$[0, 7, 35]$	0.0	$[3, 4, 35]$	2.45
5.9(b)	$[-37, -20, 0, 7, 22, 35]$	$[-37, -20, 0, 7, 22, 35]$	0.0	$[-38, -37, -36, -20, 3, 4]$	24.62
5.9(c)	$[-67, -37, -20, -9, 0, 7, 22, 35, 54]$	$[-67, -37, -20, -9, 0, 7, 22, 35, 54]$	0.0	$[-66, -31, -19, -14, -10, -5, -4, 3, 4]$	22.38

Table 5.5: *Multiple Snapshots* ($W = 25$) *DoA Estimation* ($M = 10$, $d = 0.5\lambda$; $SNR = 7\,dB$; $L \in \{3, 6, 9\}$). *DoA*s estimated with the $MT-BCS$ and the multi-snapshots $ST-BCS$. $RMSE$ values.

These conclusions are further confirmed from the results in Fig. 5.9 and Tab. 5.5 concerned with an array of $M = 10$ elements spaced by $d = \frac{\lambda}{2}$ and different incident signals, $L = \{3, 6, 9\}$, in an environment characterized by $SNR = 7\,dB$. As expected, the $ST-BCS$, although in its multi-snapshots version, overestimates the unknown number of incident signals thus unavoidably increasing the $RMSE$, while the $MT-BCS$ correctly identifies the actual signal directions in all the examples [Figs. 5.9(a)-5.9(c)].

Still dealing with multiple sequential acquisitions, the final numerical analysis is concerned with a comparative assessment of the $MT-BCS$ and state-of-the-art approaches such as $ESPRIT$ [91] and $ROOT-MUSIC$ [90]. Figure 5.10 plots the $RMSE$ averaged over $Q = 100$ simulations for each scenario and yielded by the $MT-BCS$, the multi-snapshot $ST-BCS$, and the two reference methods as a function of W, the minimum spacing between two adjacent signals $\Delta\theta$, and the SNR. As shown in Fig. 5.10(a), the accuracy of the $MT-BCS$ improves with W and at the upper value ($W = 25$) the error is of some order in magnitude below that of the $ST-BCS$ [$RMSE_{MT-BCS} = (4.7 \times 10^{-3})^o$ vs $RMSE_{ST-BCS} = 2.90^o$]. Unlike the $ST-BCS$, the larger the number of snapshots, the better is the estimation of the actual *DoA*s for both the matrix-correlation approaches and the $MT-BCS$. Moreover, the $MT-BCS$ performs better than $ESPRIT$ and $ROOT-MUSIC$ with a non-negligible and increasing enhancement of the estimation accuracy as the acquisition time grows [Fig. 5.10(a)]. As a matter of fact, both $ESPRIT$ and $ROOT-MUSIC$ do not further improve their estimates after $W = 10$, while the precision of the $MT-BCS$ monotonically enhances [$RMSE_{MT-BCS}\rfloor_{W=25} < (10^{-2})^o$].

As for the results when varying $\Delta\theta$ and SNR, the arising outcomes still point out the effectiveness of the $MT-BCS$ and its enhanced accuracy if compared to state-of-the-art methods. As expected, the $ST-BCS$ turns out to be very reliable when the angular spacing is quite large [Fig. 5.10(b)].

5.4 Discussions

In this Chapter, innovative strategies for the estimation of the directions of arrival of signals impinging on linear arrays of electromagnetic sensors have been presented and assessed. Starting from a sparse representation of the problem solution, the DoA estimation problem has been addressed by means of two methodologies based on the BCS paradigm, the one devoted to the single-snapshot processing, the other exploiting multiple-snapshot acquisitions. Advantages and limitations of those implementations have been analyzed and highlighted also in comparison with well-assessed state-of-the-art DoA estimation strategies.
The proposed approaches have shown being able to:

- directly work on the voltages measured at the output of the array elements without requiring the computation of the covariance matrix;
- provide accurate and reliable DoAs estimation also without the *a-priori* knowledge on the number of incident signals;
- estimate the DoAs just processing a single snapshot, with more precision for signals closer to the boresight direction;
- provide robust and very accurate estimates when correlating the information from multiple snapshots.

Further advances, currently under investigation and out-of-the-scope of this Chapter, will consider potential improvements of the estimation accuracy thanks to a multi-resolution strategy, the possibility to estimate the DoAs of wideband signals by correlating the information available in the measurements at different frequencies thanks to the $MT-BCS$, and the definition of alternative sparse representations of the problem unknowns for straightforwardly exploiting similar formulations when dealing with different estimation problems still concerned with adaptive arrays. It is also important to point out that from a methodological viewpoint, the extension of the proposed strategies to deal with planar ($2D$) or conformal ($3D$) antenna configurations is straightforward. In this case, the number of array elements, usually larger than the linear array case, and the highest number of samples of the angular range, due to the fact that both θ and ϕ directions are present, will unavoidably increase the computational cost.

Chapter 6

Direction-of-Arrival Estimation in Planar Arrays by Bayesian Compressive Sensing

In this Chapter, the Bayesian Compressive Sensing (BCS) is applied to estimate the directions-of-arrival (DoAs) of narrow-band electromagnetic signals impinging on planar antenna arrangements. Starting from the measurement of the voltages induced at the output of the array elements, the performance of the BCS-based approach is evaluated when data are acquired at a single time instant and at consecutive time instants, respectively. Different signal configurations, planar array geometries, and noise conditions are taken into account, as well.

6.1 Introduction

In the last few years, we assisted to an extraordinary and still growing development and use of Compressive Sensing (CS)-based methods [81] in a wide number of applicative contexts such as communications [92], bio-medicine [93], radar [94], and microwave imaging [86][95]. CS has proven to be a very effective resolution tool when the relationship between the problem data and the unknowns is linear and these latter are sparse (or they can be sparsified) with respect to some representation bases.

In this Chapter, a probabilistic version of the CS, namely the Bayesian Compressive Sensing (BCS) [41], is used for estimating the directions of arrival (DoAs) of electromagnetic signals impinging on an array of sensors in a planar arrangement. Since the DoAs of the incoming signals are few with respect to the whole set of angular directions, they can be modeled as a sparse vector. Accordingly, the estimation problem at hand can be reformulated as the retrieval of such a sparse signal vector whose non-null entries are related to the unknown angular directions of the signals.

Compared to the state-of-the-art estimation methods (e.g., the multiple signal classification ($MUSIC$) [65], the signal parameters via rotational invariance technique ($ESPRIT$) [67], the maximum likelihood (ML) DoAs estimators [96], and the class of techniques based on learning-by-examples (LBE) strategies [73][74][72]), CS-based approaches have shown several interesting advantages. Likewise LBE-based methods, the computationally-expensive calculation of the covariance matrix is not necessary since the voltages measured at the output of the array elements can be directly processed. CS-based methods turn out to be fast and also work with single time-instant (*snapshot*) data acquisitions. Moreover, unlike $MUSIC$ and $ESPRIT$ that require the incoherence of the impinging signals and a set of measurements larger than the number of signals, careful DoA estimates can be yielded also when the number of arriving signals is greater than the array sensors as well as in the presence of highly-correlated sources.

Within the class of CS-based approaches, deterministic strategies recover the signal vector by enforcing the sparsity constraints through the l_1-norm, while the l_2-norm is adopted to quantify the mismatch between measured and estimated data as shown in [97] for the localization of narrowband sources when using a circular array. Hybrid l_1-norm and l_2-norm formulations have been considered [98][99], as well. Others CS-based methods have been proposed [76][100][101] also dealing with the DoAs estimation of correlated sources [102]. Unfortunately, common formulations of the CS (i.e., based on deterministic strategies) require a minimum number of measurements equal to twice the number of impinging signals to satisfy the necessary condition for the well-posedness of the problem (i.e., the restricted isometry property of the sapling matrix). To overcome such an issue, probabilistic CS-based approaches have been taken into account [103]

(see also Chapter 5) as the one considered in this Chapter.
The outline of the Chapter is as follows. The *DoA*s estimation problem, its sparse reformulation, and the *BCS*-based *DoA*s estimation approach are presented in Sect. 6.2. A selected set of representative numerical results is reported in Sect. 6.3 to discuss, in a comparative fashion, the performance of the single and multiple snapshot implementations of the two-dimensional extension of the *BCS* method presented in Chapter 5 for different array architectures. Eventually, some conclusions are drawn (Sect. 6.4).

6.2 Mathematical Formulation

Let us consider a planar antenna array made of N isotropic sensors located on the $x-y$ plane. An unknown set of I signals $\mathbf{s}_i(\mathbf{r},t) = \alpha_i(t)\, e^{j(2\pi f_0 t + \mathbf{k}_i \cdot \mathbf{r})}$, $i = 1, ..., I$ is supposed to impinge on the array from the unknown directions $\Psi_i = (\theta_i, \phi_i)$, $i = 1, ..., I$, being $0° \leq \theta_i \leq 90°$ and $0° \leq \phi_i \leq 360°$. Such signals are modeled as narrowband electromagnetic plane waves (i.e., $\alpha_i(t) \simeq \alpha_i$, $i = 1, ..., I$) at the carrier frequency f_0, $\mathbf{k}_i$ $(i = 1, ..., I)$ being the i-th wave vector having amplitude $k = |\mathbf{k}_i| = \frac{2\pi}{\lambda}$, $\forall i = 1, ..., I$, where λ is the free space wavelength.
By modelling the background noise as an additive Gaussian process with zero mean and variance σ^2, the phasor voltage measured at the n-th element is equal to

$$\upsilon_n(\tau) = \sum_{i=1}^{I} \upsilon_{i,n}(\tau) + \eta_n(\tau) \tag{6.1}$$

where τ is the measurement time-instant/snapshot and $\eta_n(\tau)$ is the noise sample at the same instant. Moreover,

$$\upsilon_{i,n}(\tau) = \alpha_i(\tau) e^{j\frac{2\pi}{\lambda}(x_n \sin\theta_i \cos\phi_i + y_n \sin\theta_i \sin\phi_i)} \tag{6.2}$$

is the open circuit voltage induced by the i-th impinging wave at the n-th planar array element located in the position $\mathbf{r}_n = (x_n, y_n)$.
The relationship between the measured data (i.e., $\upsilon_n(\tau)$, $n = 1, ..., N$, $\tau = 1, ..., T$) and the unknown *DoA*s [i.e., $\Psi_i = (\theta_i, \phi_i)$, $i = 1, ..., I$] can be then represented in a compact matrix form as follows

$$\boldsymbol{\upsilon}(\tau) \;=\; \mathbf{H}(\boldsymbol{\Psi})\, \mathbf{s}(\tau) + \boldsymbol{\eta}(\tau)\;,\; \tau = 1, ..., T \tag{6.3}$$

where $\boldsymbol{\upsilon}(\tau) = [\upsilon_1(\tau), \upsilon_2(\tau), ..., \upsilon_N(\tau)]^*$ is the complex measurement vector, $*$ denoting the transpose operation, and $\mathbf{H}(\boldsymbol{\Psi}) = [\mathbf{h}(\Psi_1), \mathbf{h}(\Psi_2), ..., \mathbf{h}(\Psi_I)]$ is the steering vector matrix where $\mathbf{h}(\Psi_i) = [h_{i,1}, h_{i,2}, ..., h_{i,N}]^*$ being $h_{i,n} = e^{j\frac{2\pi}{\lambda}(x_n \sin\theta_i \cos\phi_i + y_n \sin\theta_i \sin\phi_i)}$. Moreover, $\mathbf{s}(\tau) = [\alpha_1(\tau), \alpha_2(\tau), ..., \alpha_I(\tau)]^*$ is the signal vector and $\boldsymbol{\eta}(\tau) = [\eta_1(\tau), \eta_2(\tau), ..., \eta_N(\tau)]^*$ is the noise vector.
It is simple to observe that the solution of (6.3) is neither linear nor sparse with respect to the problem unknowns $\Psi_i = (\theta_i, \phi_i)$, $i = 1, ..., I$, while it is linear

versus $\mathbf{s}(\tau)$, $\forall\tau$. In order to apply the BCS to the DoAs estimation in planar arrays, the method in Chapter 5 for linear arrays has been exploited and here suitably customized to the dimensionality ($2D$) at hand.
To reformulate the original problem as a sparse one, the observation domain composed by all angular directions $0° \leq \theta \leq 90°$ and $0° \leq \phi \leq 360°$ is partitioned (Fig. 6.1) in a fine grid of K samples satisfying the condition $K \gg I$. Therefore, the terms $\mathbf{H}(\mathbf{\Psi})$ and $\mathbf{s}(\tau)$ in (6.3) turn out being equal to

$$\breve{\mathbf{H}}\left(\breve{\mathbf{\Psi}}\right) = \left[\breve{\mathbf{h}}\left(\breve{\Psi}_1\right), \breve{\mathbf{h}}\left(\breve{\Psi}_2\right), ..., \breve{\mathbf{h}}\left(\breve{\Psi}_k\right), ..., \breve{\mathbf{h}}\left(\breve{\Psi}_K\right)\right] \tag{6.4}$$

and

$$\breve{\mathbf{s}}(\tau) = [\breve{\alpha}_1(\tau), \breve{\alpha}_2(\tau), , ...\breve{\alpha}_k(\tau), ..., \breve{\alpha}_K(\tau)]^* . \tag{6.5}$$

By substituting (6.4) and (6.5) in (6.3), the problem is still linear with respect to also $\breve{\mathbf{s}}(\tau)$, but $\breve{\mathbf{s}}(\tau)$ [unlike $\mathbf{s}(\tau)$] is now sparse since $K \gg I$. Accordingly, only few coefficients $\breve{\alpha}_k(\tau)$, $k = 1, ..., K$ are expected to differ from zero and exactly in correspondence with the steering vectors $\breve{\mathbf{h}}\left(\breve{\Psi}_k\right)$ at the angular direction $\breve{\Psi}_k$ where the wave is estimated to impinge on the array. Accordingly, the original problem of determining the DoAs, $\Psi_i = (\theta_i, \phi_i)$, $i = 1, ..., I$, is reformulated as the estimation of the (sparse) signal vector $\hat{\mathbf{s}}(\tau)$. The signal DoAs are then retrieved as the directions $\hat{\Psi}_k = \left(\hat{\theta}_k, \hat{\phi}_k\right)$ whose corresponding signal amplitudes $\hat{\alpha}_k(\tau)$ are non null.

For single time-instant ($T = 1$) acquisitions, the Single-Task Bayesian Compressive Sensing ($ST - BCS$) is used and the sparsest vector $\hat{\mathbf{s}}(\tau)$ is retrieved by maximizing the posterior probability (see Chapter 5)

$$\mathcal{P}\left(\left[\hat{\mathbf{s}}(\tau), \hat{\sigma}^2, \mathbf{a}(\tau)\right] \middle| \boldsymbol{\upsilon}(\tau)\right) \tag{6.6}$$

where $\hat{\sigma}^2$ is the estimate of the noise power, supposed not varying in time, and $\mathbf{a}(\tau)$ is the hyper-parameter vector [45] enforcing the sparseness of the solution $\hat{\mathbf{s}}(\tau)$ at the τ-th snapshot. Accordingly, the analytic form of the solution turns out to be

$$\hat{\mathbf{s}}(\tau) = \frac{1}{\hat{\sigma}^2}\left(\frac{\hat{\mathbf{H}}\left(\breve{\mathbf{\Psi}}\right)^* \hat{\mathbf{H}}\left(\breve{\mathbf{\Psi}}\right)}{\hat{\sigma}^2} + diag\left(\mathbf{a}(\tau)\right)\right)^{-1} \hat{\mathbf{H}}\left(\breve{\mathbf{\Psi}}\right)^* \hat{\boldsymbol{\upsilon}}(\tau) \tag{6.7}$$

where all terms are real since the BCS works only with real numbers. The signal vector, $\hat{\mathbf{s}}(\tau) = [Re\{\hat{\mathbf{s}}(\tau)\}; Im\{\hat{\mathbf{s}}(\tau)\}]^*$, has dimension $2K \times 1$, $\hat{\boldsymbol{\upsilon}}(\tau) = [Re\{\hat{\boldsymbol{\upsilon}}(\tau)\}; Im\{\hat{\boldsymbol{\upsilon}}(\tau)\}]^*$ is a $2N \times 1$ vector, while

$$\hat{\mathbf{H}}\left(\breve{\mathbf{\Psi}}\right) = \begin{bmatrix} Re\left\{\breve{\mathbf{H}}\left(\breve{\mathbf{\Psi}}\right)\right\} & -Im\left\{\breve{\mathbf{H}}\left(\breve{\mathbf{\Psi}}\right)\right\} \\ Im\left\{\breve{\mathbf{H}}\left(\breve{\mathbf{\Psi}}\right)\right\} & Re\left\{\breve{\mathbf{H}}\left(\breve{\mathbf{\Psi}}\right)\right\} \end{bmatrix} \tag{6.8}$$

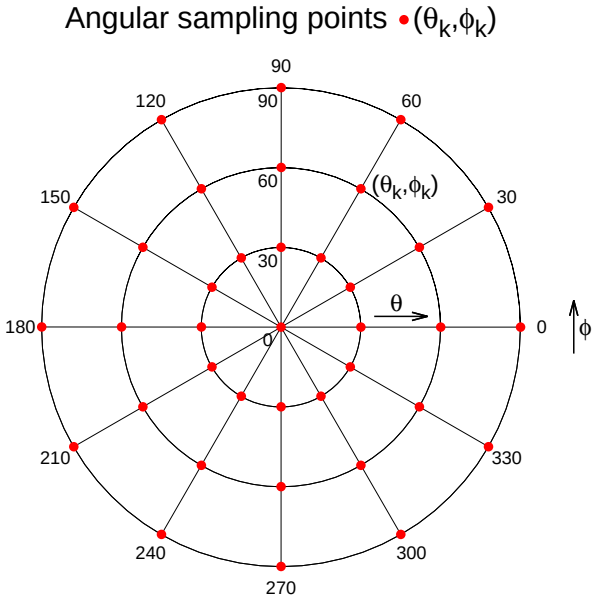


Figure 6.1: Sketch of the discretized observation domain for CS-based DoAs estimations.

is $2N \times 2K$ matrix, $Re\{\cdot\}$ and $Im\{\cdot\}$ being the real and imaginary part, respectively. The two control parameters in (6.7), $\mathbf{a}(\tau)$ and $\hat{\sigma}^2$, are obtained through the maximization of the function

$$\Pi_{ST}\left(\hat{\sigma}^2, \mathbf{a}(\tau)\right) = K \log\left(\frac{1}{2\pi}\right) - \frac{\log|\boldsymbol{\Omega}(\tau)| + (\hat{\boldsymbol{\upsilon}}(\tau))^* (\boldsymbol{\Omega}(\tau))^{-1} \hat{\boldsymbol{\upsilon}}(\tau)}{2} \tag{6.9}$$

by means of the relevance vector machine (RVM). In (6.7), $\boldsymbol{\Omega}(\tau) \triangleq \hat{\sigma}^2\mathbf{I} + \hat{\mathbf{H}}\left(\breve{\boldsymbol{\Psi}}\right) diag\left(\mathbf{a}(\tau)\right)^{-1} \hat{\mathbf{H}}\left(\breve{\boldsymbol{\Psi}}\right)^*$ where $\mathbf{I}$ is the identity matrix.
When a set of consecutive snapshots is available, the Multi-Task BCS ($MT-BCS$) implementation is used to statistically correlate the estimates derived for each snapshot by setting a common hyper-parameter vector: $\mathbf{a}(\tau) = \mathbf{a}$, $\forall \tau = 1, ..., T$. Hence, the final $MT-BCS$ solution is given by (see Chapter 4 and Chapter 5)

$$\hat{\mathbf{s}} = \frac{1}{T}\sum_{\tau=1}^{T}\left\{\left[\hat{\mathbf{H}}\left(\breve{\boldsymbol{\Psi}}\right)^* \hat{\mathbf{H}}\left(\breve{\boldsymbol{\Psi}}\right) + diag\left(\mathbf{a}\right)\right]^{-1} \hat{\mathbf{H}}\left(\breve{\boldsymbol{\Psi}}\right)^* \hat{\boldsymbol{\upsilon}}(\tau)\right\} \tag{6.10}$$

where $\mathbf{a}$ is computed through the RVM maximization of the following function

$$\Pi_{MT}\left(\mathbf{a}\right) = -\frac{1}{2}\sum_{\tau=1}^{T}\left\{\log\left(|\boldsymbol{\Omega}|\right) + (K + 2\beta_1)\log\left[(\hat{\boldsymbol{\upsilon}}(\tau))^* \boldsymbol{\Omega}\hat{\boldsymbol{\upsilon}}(\tau) + 2\beta_2\right]\right\} \tag{6.11}$$

where $\mathbf{\Omega} \triangleq \mathbf{I} + \hat{\mathbf{H}}\left(\breve{\mathbf{\Psi}}\right) diag\left(\mathbf{a}\right)^{-1} \hat{\mathbf{H}}\left(\breve{\mathbf{\Psi}}\right)^{*}$ and β_1 and β_2 are two user-defined parameters [60].
Although the condition $\hat{\alpha}_k(\tau) \simeq 0$ or $\hat{\alpha}_k \simeq 0$ usually holds true, the number of non-null coefficients in either $\hat{\mathbf{s}}(\tau)$ $(ST-BCS)$ or $\hat{\mathbf{s}}$ $(MT-BCS)$ could be larger because of the presence of the noise. Hence, the energy thresholding techniques described in Chapter 5 is exploited to firstly count the number of arriving signals, $\hat{I}$, and then to estimate the corresponding DoAs. More in detail, the coefficients $\hat{\alpha}_k(\tau)$ (or $\hat{\alpha}_k$) are firstly sorted according to their magnitude, then only the first $\hat{I}$ coefficients whose cumulative power content is lower than a percentage χ of the totally received signal power, namely $\|\hat{\mathbf{s}}(\tau)\| = \sum_{k=1}^{K}\left(\hat{\alpha}_k(\tau)\right)^2$ (or $\|\hat{\mathbf{s}}\| = \sum_{k=1}^{K}\left(\hat{\alpha}_k\right)^2$), are preserved. Hence, $\hat{I}$ is selected such that $\sum_{i=1}^{\hat{I}}\left(\hat{\alpha}_i(\tau)\right)^2 < \chi\,\|\hat{\mathbf{s}}(\tau)\|$ (or $\sum_{i=1}^{\hat{I}}\left(\hat{\alpha}_i\right)^2 < \chi\,\|\hat{\mathbf{s}}\|$).

6.3 Numerical Results

The planar array BCS-based estimation method is assessed by means of the following analysis devoted to evaluate (a) the performance of its different implementations in correspondence with single snapshot $(T=1)$ or multiple-snapshots $(T>1)$ acquisitions and (b) the impact of different array configurations. Throughout the numerical assessment, the array elements have been assumed uniformly-spaced of $d_x = \frac{\lambda}{2}$ and $d_x = \frac{\lambda}{2}$ along the x-axis and y-axis, respectively, and all signals have been characterized with the same amplitude $\alpha_i(\tau) = \alpha_{i+1}(\tau)$, $i = 1, ..., I-1$. The measurements have been blurred with an additive Gaussian noise of variance σ^2 such that the resulting signal-to-noise ratio turns out to be

$$SNR = 10 \times \log\left[\frac{\sum_{n=1}^{N}\left|\upsilon_n^{no-noise}\right|^2}{N\sigma^2}\right] \tag{6.12}$$

$\upsilon_n^{no-noise}$ $(n = 1, ..., N)$ being the voltage measured at the n-th array element in the noiseless case. The angular observation domain (Fig. 6.1) has been partitioned with a uniform grid characterized by a sampling step equal to $\Delta\theta = 1.25°$ and $\Delta\phi = 1.25°$ along the elevation and azimuthal direction, respectively. The energy threshold has been set to $\chi = 0.95$ according to the calibration results presented in Chapter 5.
In order to quantify the reliability and the effectiveness of the DoA estimation, the following indexes have been computed. For each i-th signal, the location index [72] is defined as

$$\xi_i \;\; = \xi\left(\Psi_i, \hat{\Psi}_i\right) \triangleq \;\; \frac{\Phi\left(\Psi_i, \hat{\Psi}_i\right)}{\Phi^{(max)}} \times 100 \tag{6.13}$$

I	$\{\Psi_i = (\theta_i,\, \phi_i)\,,\, i = 1, ..., I\}$
2	$\{(25,\, 60)\,;\, (60,\, 140)\}$
4	$\{(25,\, 60)\,;\, (60,\, 140)\,;\, (70,\, 210)\,;\, (60,\, 300)\}$
8	$\{(25,\, 60)\,;\, (60,\, 140)\,;\, (70,\, 210)\,;\, (60,\, 300)\,;\, (40,\, 210)\,;\, (80,\, 45)\,;\, (15,\, 5)\,;\, (30,\, 350)\}$

Table 6.1: *Fully Populated Array* - ($N = 25$; $d_x = d_y = 0.5\lambda$; $I \in [2:8]$; $SNR = 10\,dB$; $C = 50$) - Actual DoAs of the impinging signals.

where

$$\Phi\left(\Psi_i, \hat{\Psi}_i\right) = \sqrt{\left(\sin\theta_i\cos\phi_i - \sin\hat{\theta}_i\cos\hat{\phi}_i\right)^2 + \left(\sin\theta_i\sin\phi_i - \sin\hat{\theta}_i\sin\hat{\phi}_i\right)^2 + \left(\cos\theta_i - \cos\hat{\theta}_i\right)^2} \tag{6.14}$$

and $\Phi^{(max)} = \max_{\Psi_i,\hat{\Psi}_i}\left\{\Phi\left(\Psi_i, \hat{\Psi}_i\right)\right\} = 2$ is the maximum admissible error in the DoA retrieval. Since the number of arriving signals $\hat{I}$ is unknown and it is derived from the BCS processing, the global location index has been also evaluated as in Chapter 5

$$\xi = \begin{cases} \frac{1}{I}\left[\sum_{i=1}^{\hat{I}} \xi\left(\Psi_i, \hat{\Psi}_i\right) + \left(I - \hat{I}\right)\xi^{(penalty)}\right] & if \quad \hat{I} < I \\ \frac{1}{I}\left[\sum_{i=1}^{I} \xi\left(\Psi_i, \hat{\Psi}_i\right) + \sum_{i=I+1}^{\hat{I}} \xi\left(\overline{\Psi}_i, \hat{\Psi}_i\right)\right] & if \quad \hat{I} \geq I \end{cases} \tag{6.15}$$

where $\xi^{(penalty)} = \max_{\Psi_i,\hat{\Psi}_i}\{\xi_i\} = 100$ is the maximum of (6.13) and $\overline{\Psi}_i = \arg\left\{\min_{i=I+1}\left[\xi\left(\Psi_i, \hat{\Psi}_i\right)\right]\right\}$. Since it is preferred to detect all signals really present in the scenario, although overestimating their number then missing some of them, the penalty is considered only when $\hat{I} < I$.

A. *Single and Multiple Snapshot BCS-based DoAs Estimation Techniques*

Let us consider the fully populated array of Fig. 6.2 with $N = N_x \times N_y = 25$ elements, $N_x = N_y = 5$ being the number of elements along the x and y axes, collecting the data $\boldsymbol{\upsilon}(\tau)$. Several different electromagnetic scenarios have been considered in which $I = 2$, $I = 4$, and $I = 8$ signals are supposed to impinge on the planar array from the directions indicated in Tab. $6.1^{(1)}$ [1].

[1] (1) In the numerical results, the actual DoAs are chosen lying on the sampling grid of the observation domain. Whether this condition does not hold true, off-grid compensation methods [104][105], already proposed in the state-of-the-art literature, can be profitably used.

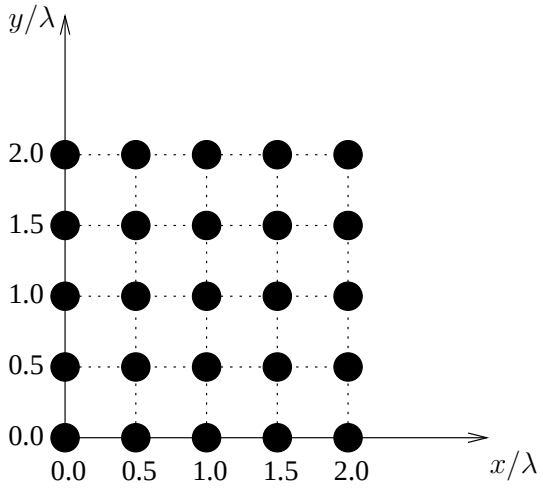


Figure 6.2: Geometry of the receiving fully populated array ($N = 25$).

The power of the background noise has been set to yield $SNR = 10dB$. In order to test the behavior of the $ST - BCS$ and the $MT - BCS$, the simulation for each signal configuration has been repeated $C = 50$ times, while varying the noise samples on the data. The *DoA*s estimation error has been therefore evaluated through the average location index defined as

$$\xi^{(avg)} = \frac{1}{C} \sum_{c=1}^{C} \xi^{(c)} \tag{6.16}$$

$\xi^{(c)}$ being computed as in (6.15).
As for the $ST - BCS$, a single snapshot has been processed each time ($T = 1$). Figure 6.3 shows the best (Fig. 6.3 - left column) and the worst (Fig. 6.3 - right column) solutions in terms of minimum ($\xi^{(min)} = \min_{c=1,\dots,C} \left\{\xi^{(c)}\right\}$) and maximum ($\xi^{(max)} = \max_{c=1,\dots,C} \left\{\xi^{(c)}\right\}$) location error, respectively, among the $C = 50$ *DoA*s estimations carried out when $I = 2$ [Figs. 6.3(a)-6.3(b)], $I = 4$ [Figs. 6.3(c)-6.3(d)], and $I = 8$ [Figs. 6.3(e)-6.3(f)]. In Fig. 6.3, the actual *DoA*s are denoted with a point at the center of a circle, while the color points indicate the estimated signal locations and amplitudes. For the sake of clarity, the retrieved *DoA*s are also reported in Tab. 6.2 where the number of estimated signals $\hat{I}$ is given, as well. As it can be observed, the strength of the estimated signals is different (Fig. 6.3), even though they impinge on the antenna with the same energy, because of the presence of the noise. On the other hand, the *DoA*s are predicted with a high degree of accuracy when $I = 2$ and $I = 4$ as confirmed by the values of the location error (Tab. 6.3). As a matter of fact, the error values are low also for the worst solutions among the C trials (i.e., $\xi^{(max)}\big|_{I=2} = 3.80\%$ and $\xi^{(max)}\big|_{I=4} = 3.89\%$). It is worth also noting that for $I = 2$ the location error is small even though the number of detected signals are greater than the actual ones ($\hat{I}^{(wst)}\Big|_{I=2} = 3$) because of two signals have very

I	$\hat{I}^{(bst)}$	$\left\{\hat{\Psi}_i^{(bst)} = \left(\hat{\theta}_i^{(bst)}, \hat{\phi}_i^{(bst)}\right), i = 1, ..., I\right\}$
2	2	{(25, 60) ; (60, 140)}
4	4	{(23.75, 65) ; (60, 140) ; (63.75, 300) ; (70, 210)}
8	7	{(23.75, 345) ; (32.5, 65) ; (67.5, 145) , (71.25, 300) , (72.5, 300) , (82.5, 40) , (90, 205)}
I	$\hat{I}^{(wst)}$	$\left\{\hat{\Psi}_i^{(wst)} = \left(\hat{\theta}_i^{(wst)}, \hat{\phi}_i^{(wst)}\right), i = 1, ..., I\right\}$
2	3	{(22.5, 60) ; (57.5, 135) ; (58.75, 137.5)}
4	4	{(23.75, 55) ; (63.75, 145) ; (61.25, 300) ; (77.5, 210)}
8	4	{(21.25, 345) ; (28.75, 70) ; (55, 210) ; (90, 45)}

Table 6.2: *Fully Populated Array* - ($N = 25$; $d_x = d_y = 0.5\lambda$; $I \in [2:8]$; $SNR = 10\,dB$; $T = 1$; $C = 50$) - Values of the DoAs for the best and worst estimation obtained by means of the $ST-BCS$ among the C different noisy scenarios.

I	$\xi^{(min)}$	$\xi^{(max)}$	$\xi^{(avg)}$	$\xi^{(var)}$	$t^{(avg)}$ $[sec]$
	$ST-BCS$				
2	0.00	3.80	1.36	1.24	4.48×10^{-1}
4	1.34	3.70	2.07	6.02×10^{-1}	1.37
8	3.02×10^{1}	8.23×10^{1}	6.06×10^{1}	2.96×10^{2}	1.77
	$MT-BCS$				
2	0.00	2.18	8.01×10^{-1}	4.06×10^{-1}	3.97
4	5.45×10^{-1}	1.91	1.37	1.19×10^{-1}	6.44
8	5.27	3.31×10^{1}	1.81×10^{1}	5.94×10^{1}	7.80

Table 6.3: *Fully Populated Array* - ($N = 25$; $d_x = d_y = 0.5\lambda$; $I \in [2:8]$; $SNR = 10\,dB$; $T = \{1, 2\}$; $C = 50$) - Statistics (minimum, maximum, average, and variance) of the location index ξ among C different noisy scenarios when using the $ST-BCS$ ($T = 1$) and the $MT-BCS$ ($T = 2$).

close DoAs (as compared to the sampling steps $\Delta\theta$ and $\Delta\phi$). However, if the $ST-BCS$ shows being robust and accurate in such scenarios ($I = 2$ and $I = 4$), it is not able to correctly locate the actual DoAs when the number of signals increases to $I = 8$ [Figs. 6.3(e)-6.3(f) - Tab. 6.2]. Indeed, the location error significantly increases as indicated by the indexes in Tab. 6.3.
As for the computational efficiency, the $ST-BCS$ is able to perform the DoAs estimation in a limited CPU time ($t^{(avg)} < 2.0\,[sec]$ - Tab. 6.3)[2(2)] also thanks to the single-snapshot processing.

In order to investigate the effects of the SNR on the DoAs estimation capabilities of the $ST-BCS$, the SNR has been varied from $-5dB$ up to $30dB$ with a step

[2(2)] The simulations have been run using a standard processing unit (i.e., $2.4GHz$ PC with $2GB$ of RAM) with a non-optimized code.

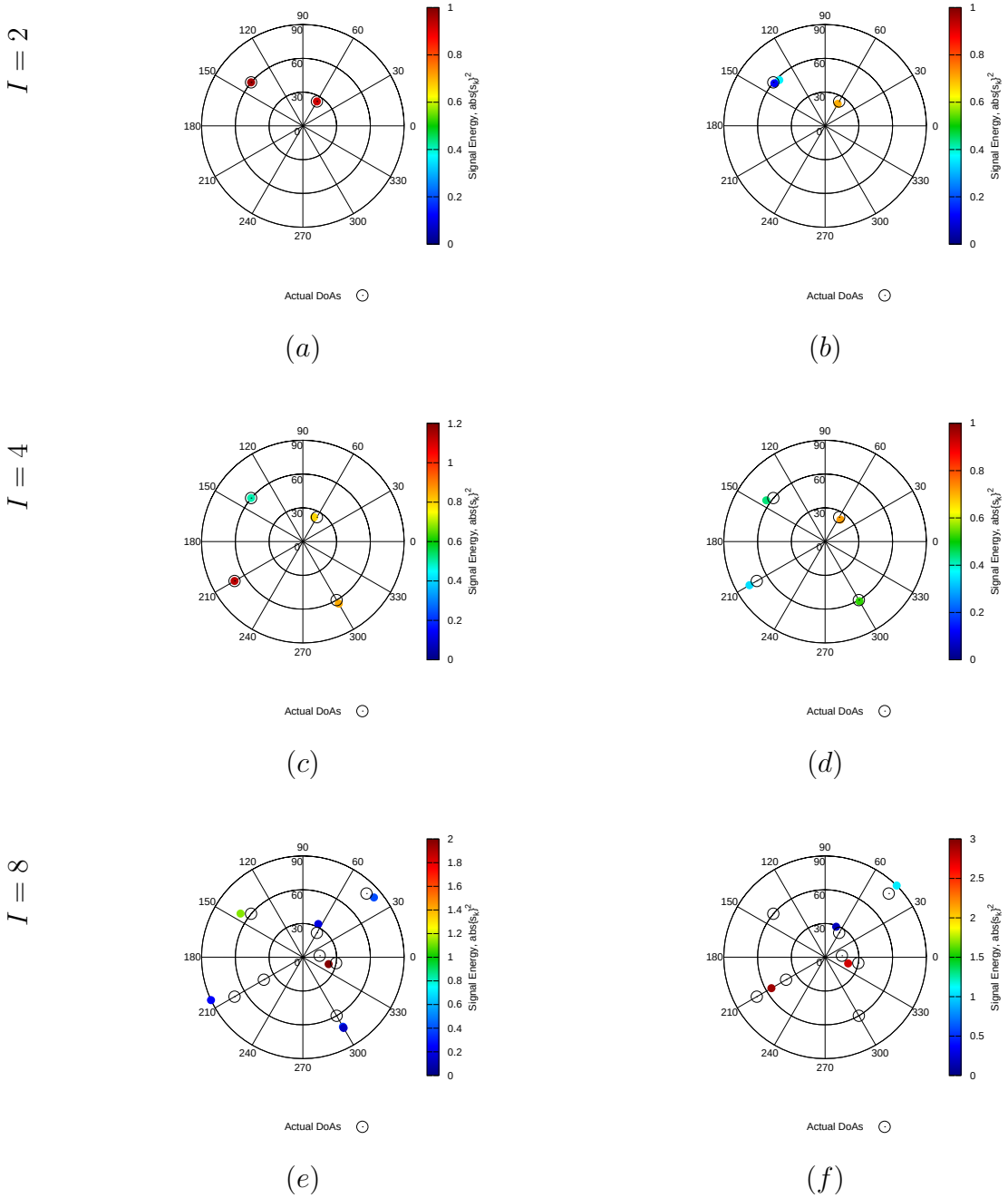


Figure 6.3: *Fully Populated Array* - ($N = 25$; $d_x = d_y = 0.5\lambda$; $I \in [2:8]$; $SNR = 10\,dB$; $T = 1$; $C = 50$) - Plot of the best (*left column*) and worst (*right column*) estimations obtained by means of the $ST-BCS$ among the C different noisy scenarios when $(a)(b)$ $I = 2$, $(c)(d)$ $I = 4$, and $(e)(f)$ $I = 8$.

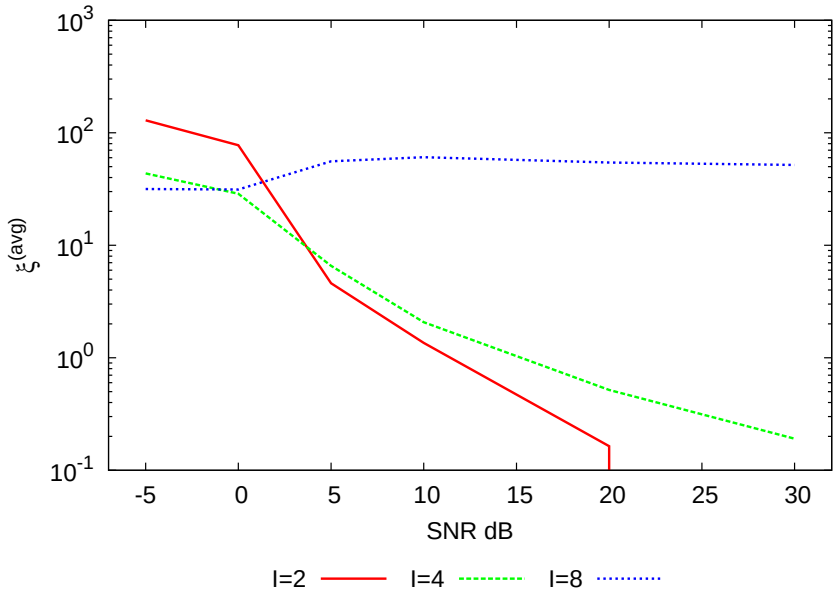


Figure 6.4: *Fully Populated Array* - ($N = 25$; $d_x = d_y = 0.5\lambda$; $I \in [2:8]$; $SNR \in [-5:30]\,dB$; $T = 1$; $C = 50$) - Behavior of the location index $\xi^{(ave)}$ averaged among C different noisy scenarios versus the SNR when using the $ST-BCS$.

of $5dB$, while keeping the same *DoA*s of Tab. 6.1. In Fig. 6.4, the values of the average location index are reported. As it can be noticed, the location index $\xi^{(avg)}$ for $I = 2$ and $I = 4$ monotonically decreases, as one should expect, with the increment of the SNR. However, the $ST-BCS$ estimates when $I = 8$ turn out to be still non-reliable also for higher SNR confirming the difficulty of dealing with such a complex scenario just processing one snapshot.

Let us now analyze the $MT-BCS$ behavior. Firstly, the same problems addressed by means of the $ST-BCS$ in Fig. 6.3 are considered by taking into account only $T = 2$ snapshots. The best and worst $MT-BCS$ results are reported in Fig. 6.5 and the corresponding *DoA*s are given in Tab. 6.4. Unlike the $ST-BCS$ (Tab. 6.2), the number of impinging signals is always correctly identify in the best case (Fig. 6.5 - left column), while in the worst case (Fig. 6.5 - right column), $\hat{I} = I$ only when $I = 2$ and $I = 4$ signals. As a matter of fact, the average location error when $I = 8$ is still high ($\xi^{(avg)}\big|_{I=8} = 18.1\%$). The use of only $T = 2$ snapshots does not guarantee reliable performance also with the $MT-BCS$, even though the advantages in terms of accuracy of the $MT-BCS$ over the $ST-BCS$ are non-negligible as pointed out by the values in Tab. 6.3. On the opposite, the computational cost of the $MT-BCS$ is higher than that of the $ST-BCS$ (Tab. 6.3).

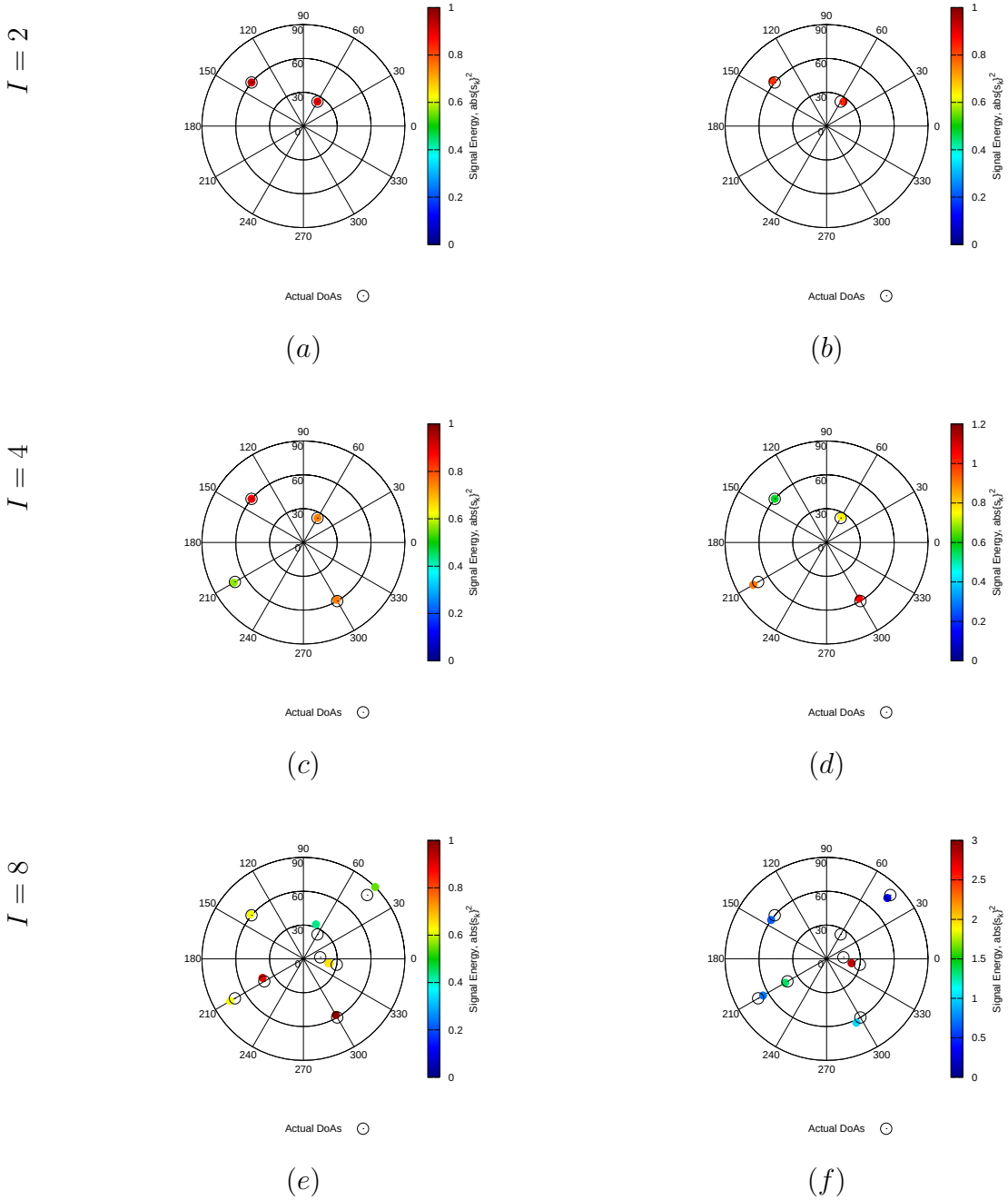


Figure 6.5: *Fully Populated Array* - ($N = 25$; $d_x = d_y = 0.5\lambda$; $I \in [2:8]$; $SNR = 10\,dB$; $T = 2$; $C = 50$) - Plot of the best (*left column*) and worst (*right column*) estimations obtained by means of the $MT-BCS$ among the C different noisy scenarios when $(a)(b)$ $I = 2$, $(c)(d)$ $I = 4$, and $(e)(f)$ $I = 8$.

I	$\hat{I}^{(bst)}$	$\left\{\hat{\Psi}_i^{(bst)} = \left(\hat{\theta}_i^{(bst)}, \hat{\phi}_i^{(bst)}\right), i = 1, ..., I\right\}$
2	2	{(25, 60) ; (60, 140)}
4	4	{(25, 60) ; (58.75, 300) ; (60, 140) ; (71.25, 210)}
8	8	{(22.5, 350) ; (23.75, 350) ; (32.5, 70) ; (40, 205) ; (57.5, 300) ; (61.25, 140) ; (75, 210) ; (90, 45)}
I	$\hat{I}^{(wst)}$	$\left\{\hat{\Psi}_i^{(wst)} = \left(\hat{\theta}_i^{(wst)}, \hat{\phi}_i^{(wst)}\right), i = 1, ..., I\right\}$
2	2	{(26.25, 55) ; (62.5, 140)}
4	4	{(26.25, 60) ; (57.5, 300) ; (60, 140) ; (75, 210)}
8	6	{(22.5, 350) ; (42.5, 210) ; (60, 145) ; (62.5, 295) ; (65, 210) ; (76.25, 45)}

Table 6.4: *Fully Populated Array* - ($N = 25$; $d_x = d_y = 0.5\lambda$; $I \in [2:8]$; $SNR = 10\,dB$; $T = 2$; $C = 50$) - Values of the *DoA*s for the best and worst estimation obtained by means of the $MT-BCS$ among the C different noisy scenarios.

More reliable $MT-BCS$ estimations can be yielded when processing a larger number of snapshots. Figure 6.6 shows that, also for complex electromagnetic scenarios (i.e., $I = 8$ - Tab. 6.1), the average location error gets lower when T increases. By considering $SNR = 10\,dB$ as a representative example, one can observe that $\xi^{(avg)}$ reduces of almost one order of magnitude from $\xi^{(avg)}\big|_{I=8} = 18.1\%$ ($T = 2$) to $\xi^{(avg)}\big|_{I=8} = 2.20\%$ ($T = 5$). As expected, more accurate estimations arise with even more data (i.e. $\xi^{(avg)}\big|_{I=8} = 1.23\%$ when $I = 10$ and $\xi^{(avg)}\big|_{I=8} = 0.95\%$ when $I = 25$ - Fig. 6.6). The benefits from the correlation of the information coming from different time instants thanks to the $MT-BCS$ are also highlighted by the behavior of the plots in Fig. 6.6: $\xi^{(avg)}$ more rapidly decreases for higher values of T when the quality of the data improves (i.e., higher SNR).

As long as the applications at hand do not require the fast or real-time identification of the *DoA*s and there is the possibility to collect the data at consecutive time instants, the robust estimation of a larger number of impinging signals is allowed. In this context, Figure 6.7 shows the results obtained with the $MT-BCS$ when $I = 12$ [Figs. 6.7(a)-(b)] and $I = 18$ [Figs. 6.7(c)-(d)] ($SNR = 10dB$). As for the case $I = 12$, the *DoA*s are estimated with a good degree of accuracy also in the worst case within the C experiments [Fig. 6.7(b) - $\xi^{(max)}\big|_{I=12} = 1.77\%$], while the average location error amounts to $\xi^{(avg)}\big|_{I=12} = 1.04\%$. Differently , the average error is $\xi^{(avg)}\big|_{I=18} = 4.70\%$ and in the worst case [Fig. 6.7(d)] is $\xi^{(max)}\big|_{I=18} = 7.85\%$ when $I = 18$. For the sake of completeness, the best solutions are reported in Figs. 6.7(a) and 6.7(c) when $I = 12$ and $I = 18$, respectively.

B. *DoAs Estimation Performance for Different Array Geometries*

In this section, the behavior of the BCS-based single-snapshot and multiple-snapshots *DoA*s estimators is analyzed for different array architectures. The

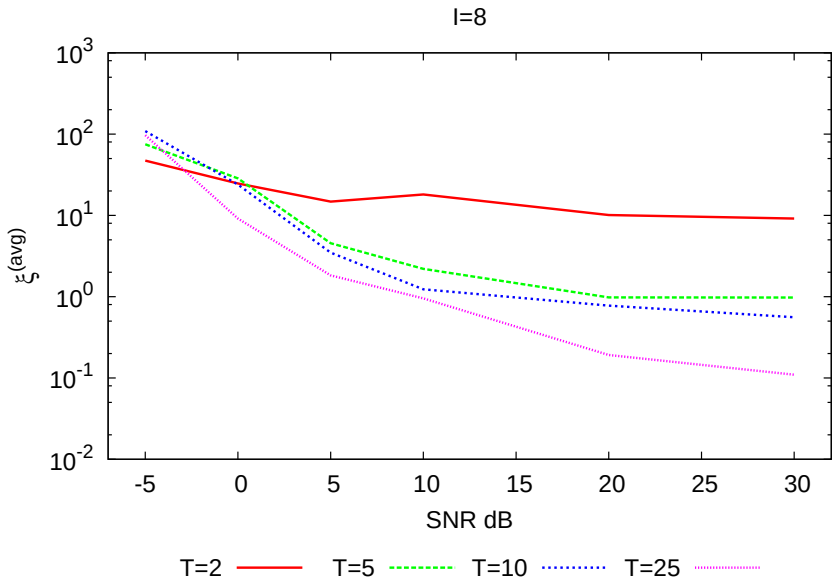
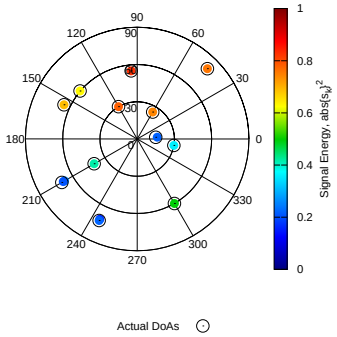


Figure 6.6: *Fully Populated Array* - ($N = 25$; $d_x = d_y = 0.5\lambda$; $I = 8$; $SNR \in [-5 : 30]\, dB$; $T \in [2 : 25]$; $C = 50$) - Behavior of the location index $\xi^{(ave)}$ averaged among C different noisy scenarios versus the SNR when using the $MT - BCS$ with different number of available snapshots T.

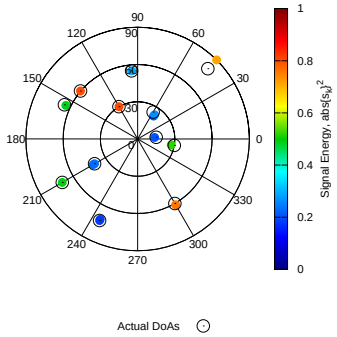
three array geometries in Fig. 6.8 are taken into account. As it can be noticed, the first array [Fig. 6.8(a)] has the same number of elements of the fully populated one but the sensors are randomly located on the antenna aperture. The other two arrays [Figs. 6.8(b) and 6.8(c)] have less elements (i.e., $N = 9$) but same aperture length of the fully populated array along the two coordinate axes.

In the first example, the performance of the $ST - BCS$ is assessed when changing the number of impinging signals from $I = 2$ up to $I = 8$, while keeping the noise level to $SNR = 10dB$. Figure 6.9 shows the average location error ($C = 50$) obtained in correspondence with the three arrays. Unlike the fully populated arrangement enabling good estimation features especially until $I = 4$ ($\xi^{(avg)}\big|_{I=2,3,4} < 2.00\%$), both the L-shaped array and the cross-shaped one do not allow reliable estimations also for the simplest scenario (i.e., $\xi^{(avg)}\big|_{I=2}^{L-Shaped} = 7.69\%$ and $\xi^{(avg)}\big|_{I=2}^{Cross-Shaped} = 10.87\%$). This is due, on the one hand, to the limited information collected from a single snapshot acquisition and, on the other hand, to the fact that the number of sensors is one third the elements of the fully-populated configuration (i.e., $N^{Fully-Populated}/N^{L/Cross-Shaped} = 2.78$). As for the random array, the achieved performance are almost equal to those of the fully populated solution thus confirming the higher reliability when having at disposal a larger number of sensors. When using the $MT - BCS$, no significant

$I = 12$

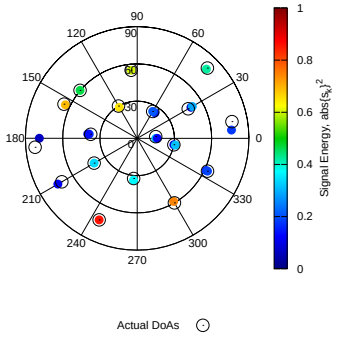


(a)

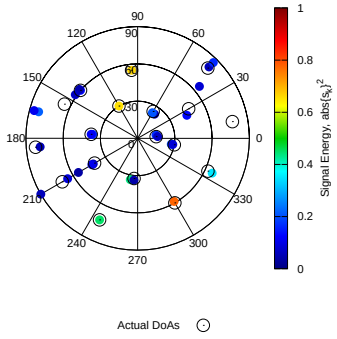


(b)

$I = 18$



(c)



(d)

Figure 6.7: *Fully Populated Array* - ($N = 25$; $d_x = d_y = 0.5\lambda$; $I = \{12, 18\}$; $SNR = 10\,dB$; $T = 25$; $C = 50$) - Plot of the best (*left column*) and worst (*right column*) estimations obtained by means of the $MT-BCS$ among the C different noisy scenarios when $(a)(b)$ $I = 12$ and $(c)(d)$ $I = 18$.

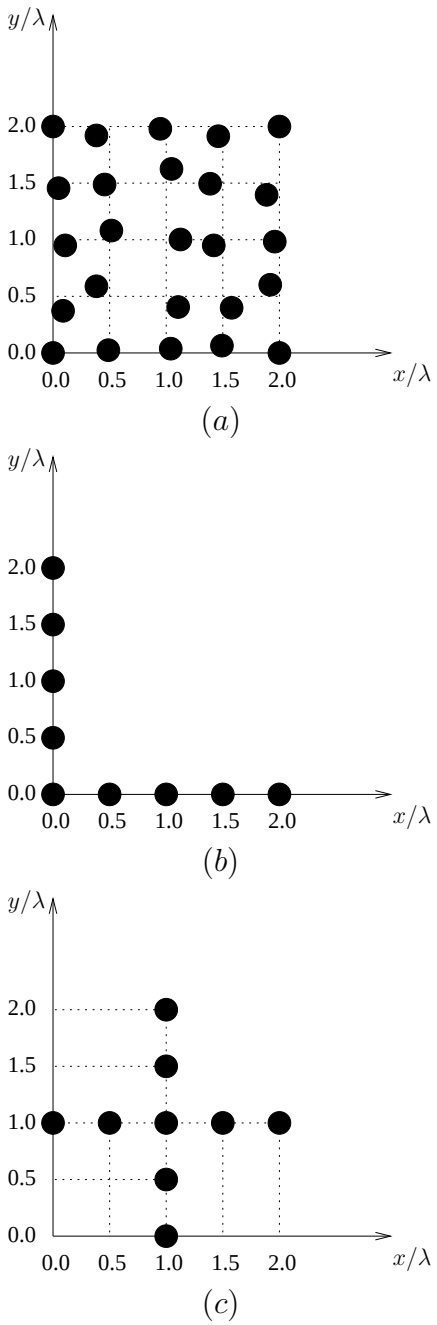


Figure 6.8: *Array Geometries Comparison* - ($N = \{9, 25\}$; $d_x = d_y = 0.5\lambda$; $I \in [2:8]$; $SNR = 10dB$; $T = 1$; $C = 50$) - Behavior of the location index $\xi^{(ave)}$ averaged among C different noisy scenarios versus the number of arriving signals I when using the $ST - BCS$.

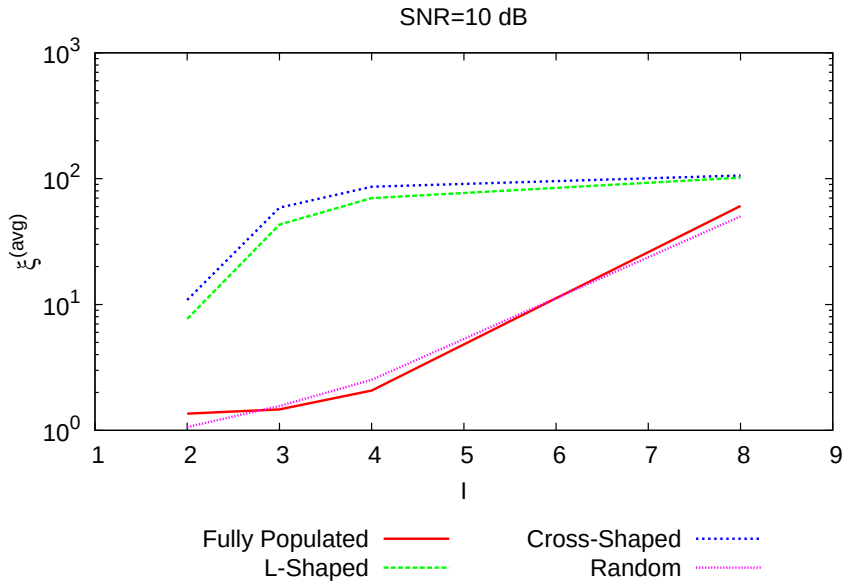


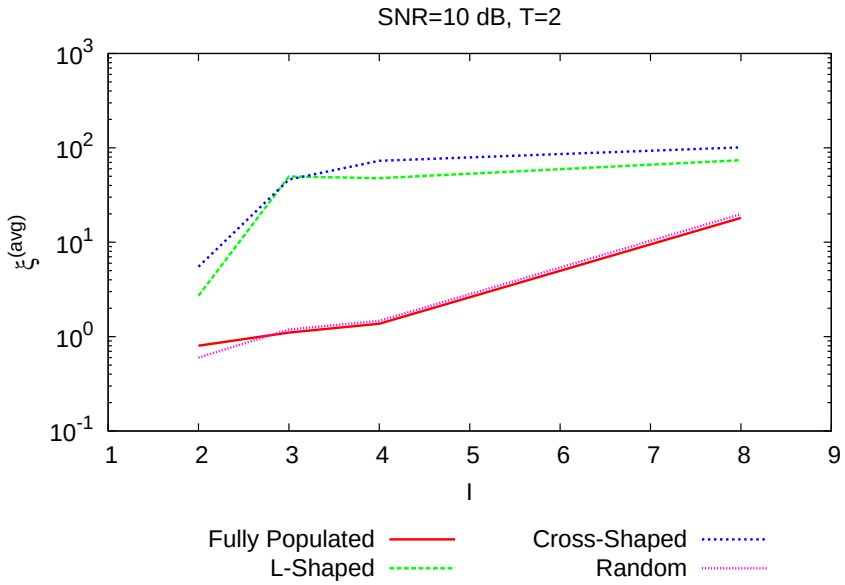
Figure 6.9: *Array Geometries Comparison* - ($N = \{9, 25\}$; $d_x = d_y = 0.5\lambda$; $I \in [2:8]$; $SNR = 10dB$; $T = 1$; $C = 50$) - Behavior of the location index $\xi^{(ave)}$ averaged among C different noisy scenarios versus the number of arriving signals I when using the $ST-BCS$.

improvements occur in comparison with the $ST-BCS$ when $T = 2$, since average errors higher than $\xi^{(avg)} = 2.00\%$ [Fig. 6.10(a)] are obtained with both the L-shaped or cross-shaped array. Whether $T = 25$ snapshots are at disposal [Fig. 6.10(b)], it turns out that the estimates from the L-shaped array present average location errors below $\xi^{(avg)} = 2.00\%$ until $I = 5$. Differently, always worse performance are achieved with the cross-shaped array [Fig. 6.10(b)].

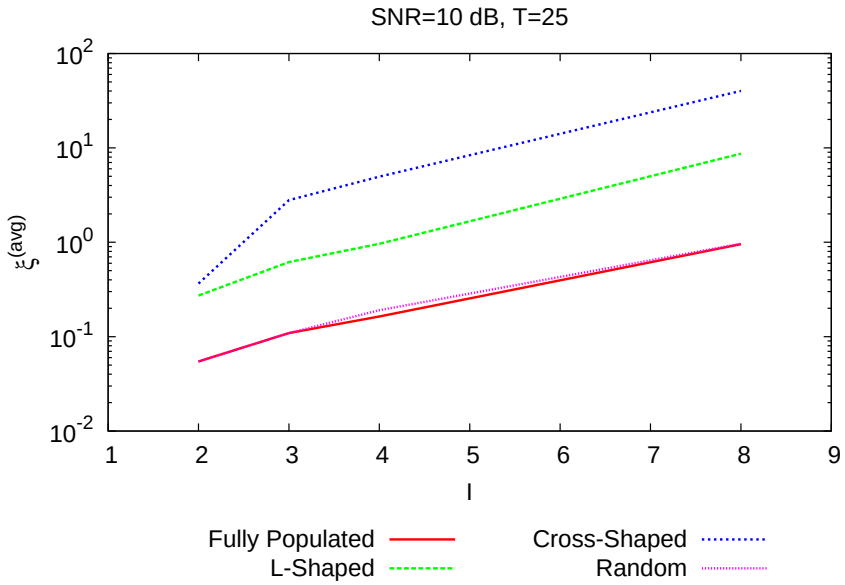
In order to give some insight on the effects of the SNR, let us consider the case $I = 2$ as a representative example. The results from the $ST-BCS$ and the $MT-BCS$ are reported in Fig. 6.11(a) and Fig. 6.11(b), respectively. The location error tends to reduce as the SNR increases for all array structures, even though the L-shaped array outperforms the cross-shaped one and the random array behavior is always very close to that of the fully populated configuration.

6.4 Discussions

The BCS method has been customized for the DoAs estimation of multiple signals impinging on planar arrays. Two different implementations, one requiring the data measured at a single snapshot and the other using the data collected

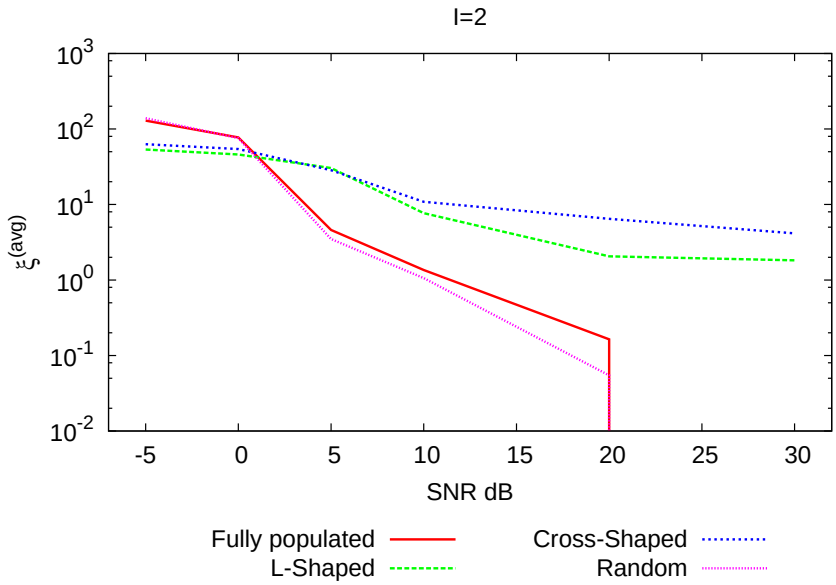


(a)

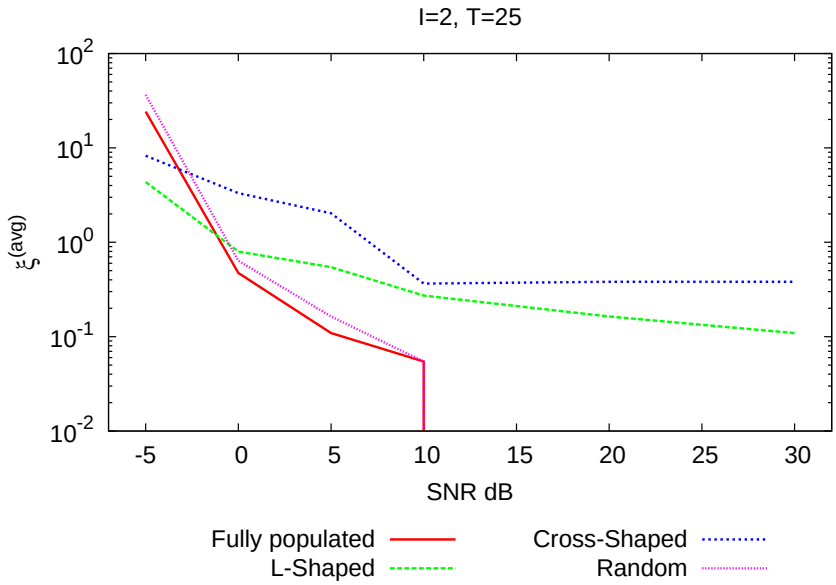


(b)

Figure 6.10: *Array Geometries Comparison* - ($N = \{9, 25\}$; $d_x = d_y = 0.5\lambda$; $I \in [2:8]$; $SNR = 10dB$; $T = \{2, 25\}$; $C = 50$) - Behavior of the location index $\xi^{(ave)}$ averaged among C different noisy scenarios versus the number of arriving signals I when using the $MT-BCS$ with (a) $T = 2$ snapshots and (b) $T = 25$ snapshots.



(a)



(b)

Figure 6.11: *Array Geometries Comparison* - ($N = \{9, 25\}$; $d_x = d_y = 0.5\lambda$; $I = 2$; $SNR = 10dB$; $T = \{1, 25\}$; $C = 50$) - Behavior of the location index $\xi^{(ave)}$ averaged among C different noisy scenarios versus the SNR when using (a) the $ST - BCS$ with $T = 1$ snapshot and (b) the $MT - BCS$ with $T = 25$ snapshots.

at multiple snapshots, have been tested on a wide number of different scenarios as well as using different array arrangements. Likewise the linear array case, the reported results have shown that:

- the two BCS-based implementations provide effective DoAs estimates although just processing the sensors output voltages and not the covariance matrix;
- the joint estimation of the signals number and DoAs is enabled;
- the correlation capability of the $MT - BCS$ allows one to yield better results than the $ST - BCS$ at the expenses of an increased computational burden.

As for the behavior of the two approaches versus the planar array geometry, it is possible to conclude that:

- the fully-populated and the random arrays give the best performance as compared to both the L-shaped and the cross-shaped array, but using a larger number of sensors;
- under the assumption of the same number of elements, the L-shaped configuration always outperforms the precision from the cross-shaped arrangement.

Chapter 7

Conclusions and future developments

In this chapter, some conclusions and ideas for future research are presented.

In this thesis, the problem of the synthesis and control of antenna arrays within the Bayesian Compressive Sensing (BCS) framework has been investigated. More in detail, the BCS method has been customized in order to deal with (i) the problem of the synthesis of linear antenna arrays and (ii) the problem of Direction-of-Arrival (DoA) estimation of signals impinging on an antenna array.

The array pattern synthesis problem has been reformulated in a Bayesian Compressive fashion as a pattern matching problem with sparseness constraints ad then it has been solved by using a suitable RVM approach. In addition, the $MT-BCS$ approach has been adopted to extend the BCS array synthesis method in order to deal with the synthesis of asymmetrical patterns (arrays with complex weights). A set of representative results have been presented in order to assess the performances of the proposed method. Comparisons with the state of the art have been shown and discussed, as well. The main features shown by the proposed technique are summarized in the following:

- the BCS methodology is able to approximate the pattern produced by a uniform array arrangement with a high degree of accuracy, providing at the same time a consistent reduction in the total element count.
- the $MT-BCS$ approach improves the $ST-BCS$ one, allowing the accurate and efficient synthesis of complex-weights arrays with non-symmetrical patterns.
- with the proposed BCS strategy is very easy to take into account of application specific constraints in the radiation pattern or in the array geometry.
- the BCS-based proposed methodology positively compares with recently introduced state-of-the-art approaches, such as the $FBMPM$.

The DoA estimation problem has been addressed by means of two methodologies based on the Bayesian Compressive Sensing paradigm, one exploiting single-snapshot measurements, the other one devoted to the processing of multiple-snapshots data. A set of representative examples concerning the DoA estimation in different scenarios have been presented and discussed. Some additional numerical results concerning the comparison with other state-of-the art methodologies have been presented, as well. The main outcomes of this work are:

- the computation of the covariance matrix is not required and the estimation can be performed by directly processing the measured voltages.
- the a-priori knowledge of the number of incoming signals is not required in order to obtain an accurate and reliable estimation.
- the method is able to provide accurate results with a limited number of snapshots. In some scenarios is is possible to obtain good estimations with only one snapshots.

- the $MT-BCS$ approach outperforms the $ST-BCS$ one thanks to the efficient correlation of multiple snapshots data.
- the approach has been extended to the azimuth-elevation estimation with planar arrays, achieving accurate results also in this case.

Concerning the array synthesis problem, future works will concern the analysis of the mutual coupling effect in the synthesized configuration and directive elements synthesis. In addition, the synthesis of reconfigurable arrays as well as arrays for wideband applications will be matter of future studies.
Regarding the DoA estimation problem, future study will deal with the synthesis of wideband signals by correlating the information available at multiple frequencies. Moreover, in order to reduce the computational burden of the algorithm, suitable multiresolution strategies will be implemented and assessed.

Bibliography

[1] O. M. Bucci, M. D. Migliore, G. Panariello, and G. Sgambato, "Accurate diagnosis of conformal arrays from near-field data using the matrix method," IEEE Trans. Antennas Propag., vol. 53, no. 3, pp. 1114-1120, Mar. 2005.

[2] M. D. Migliore, "A compressed sensing approach for array diagnosis from a small set of near-field measurements," IEEE Trans. Antennas Propag., vol. 59, no. 6, pp. 2127-2133, Jun. 2011.

[3] G. Oliveri, P. Rocca, and A. Massa "Reliable diagnosis of large linear arrays - a Bayesian Compressive Sensing approach", IEEE Trans. Antennas Propag., vol. 60, no. 10, pp. 4627-4636, Oct. 2012.

[4] R. J. Mailloux, *Phased Array Antenna Handbook*, 2nd ed. Norwood, MA: Artech House, 2005.

[5] R.S. Elliot and J. G. Stern, "Optimizing synthesis of shaped beam antenna patterns," *IEEE Trans. Antennas Propag.*, vol 32, no. 10, pp. 1129-1133, Oct. 1984.

[6] T. Isernia, O. M. Bucci, and N. Fiorentino, "Shaped beam antenna synthesis problems: Feasibility criteria and new strategies," *J. Electromagn. Waves Appl.*, vol. 12, pp. 103-137, 1998.

[7] T. Isernia and G. Panariello, "Optimal focusing of scalar fields subject to arbitrary upper bounds," *Electron. Lett*,, vol. 34, no. 2, pp. 162-164, Jan 1998.

[8] O. M. Bucci, L. Caccavale, and T. Isernia, "Optimal far-field focusing of uniformly spaced arrays subject to arbitrary upper bounds in nontarget directions," *IEEE Trans. Antennas Propag.*, vol. 50, no. 11, pp. 1539-1554, Nov. 2002.

[9] R. M. Leahy and B. D. Jeffs, "On the design of maximally sparse beamforming arrays", *IEEE Trans. Antennas Propagat.*, vol. 39, no. 8, pp. 1178-1187, Aug. 1991.

[10] R. F. Harrington, "Sidelobe reduction by nonuniform element spacing," *IEEE Trans. Antennas Propagat.*, vol. 9, p. 187, Mar. 1961.

[11] M. G. Andreasan, "Linear arrays with variable interelement spacings," *IEEE Trans. Antennas Propagat.*, vol. 10, pp. 137-143, Mar. 1962.

[12] A. Ishimaru, "Theory of unequally-spaced arrays," *IEEE Trans. Antennas Propagat.*, vol. 11, pp. 691-702, Nov. 1962.

[13] D. G. Leeper, "Isophoric arrays - massively thinned phased arrays with well-controlled sidelobes," *IEEE Trans. Antennas Propag.*, vol. 47, no. 12, pp. 1825-1835, Dec. 1999.

[14] Y. Liu, Z. Nie, and Q. H. Liu, "Reducing the number of antenna elements in a linear antenna array by the matrix pencil method," *IEEE Trans. Antennas Propagat.*, vol. 56, no. 9, pp. 2955-2962, Sep. 2008.

[15] D. King, R. Packard, and R. Thomas, "Unequally spaced, broad-band antenna arrays," *IRE Trans. Antennas Propagat.*, vol. 8, pp. 380-384, Jul. 1960.

[16] A. Maffett, "Array factors with nonuniform spacing arrays," *IRE Trans. Antennas Propagat.*, vol. 10, pp. 131-136, Mar. 1962.

[17] N. Balakrishan, P. Murthy, and S. Ramakrishna, "Synthesis of antenna arrays with spatial and excitation constraints," *IEEE Trans. Antennas Propagat.*, vol. 29, pp. 690-696, Sep. 1962.

[18] G. Oliveri, M. Donelli, and A. Massa, "Linear array thinning exploiting almost difference sets," *IEEE Trans. Antennas Propag.*, vol. 57, no. 12, pp. 3800-3812, Dec. 2009.

[19] V. Murino, A. Trucco, and C. S. Regazzoni, "Synthesis of unequally spaced arrays by simulated annealing," *IEEE Trans. Signal Processing*, vol. 44, no. 1, pp. 119-123, Jan. 1996.

[20] A. Trucco and V. Murino, "Stochastic optimization of linear sparse arrays," *IEEE J. Oceanic Engineering*, vol 24, no. 3, pp. 291-299, Jul. 1999.

[21] G. Oliveri and A. Massa, "Bayesian compressive sampling for pattern synthesis with maximally sparse non-uniform linear arrays," *IEEE Trans. Antennas Propag.*, vol. 59, no. 2, pp. 467-481, Feb. 2011.

[22] Y. T. Lo, "A mathematical theory of antenna arrays with randomly spaced elements," *IEEE Trans. Antennas Propag.*, vol. 12, no. 3, pp. 257-268, May 1964.

[23] B. Steinberg, "The peak sidelobe of the phased array having randomly located elements," *IEEE Trans. Antennas Propag.*, vol. 20, no. 2, pp. 129-136, Mar. 1972.

[24] M. I. Skolnik, G. Nemhauser, and J. W. Sherman, "Dynamic programming applied to unequally-space arrays", *IRE Trans. Antennas Propagat.*, vol. AP-12, pp. 35-43, Jan. 1964.

[25] R. L. Haupt, "Thinned arrays using genetic algorithms," *IEEE Trans. Antennas Propag.*, vol. 42, no. 7, pp. 993-999, Jul. 1994.

[26] T. G. Spence and D. H. Werner, "Thinning of aperiodic antenna arrays for low side-lobe levels and broadband operation using genetic algorithms," *Proc. IEEE Antennas and Propagation Society International Symposium*, pp. 2059-2062, 9-14 Jul. 2006.

[27] R. L. Haupt and D. H. Werner, *Genetic algorithms in electromagnetics.* Hoboken, NJ: Wiley, 2007.

[28] G. Oliveri, L. Manica, and A. Massa, "ADS-based guidelines for thinned planar arrays," *IEEE Trans. Antennas Propag.*, vol. 58, no. 6, pp. 1935-1948, Jun. 2010.

[29] S. Caorsi, A. Lommi, A. Massa, and M. Pastorino, "Peak sidelobe reduction with a hybrid approach based on GAs and difference sets," *IEEE Trans. Antennas Propag.*, vol. 52, no. 4, pp. 1116-1121, Apr. 2004.

[30] P. J. Bevelacqua and C. A. Balanis, "Minimum sidelobe levels for linear arrays," *IEEE Trans. Antennas Propag.*, vol. 55, pp. 2210-2217, Dec. 2007.

[31] M. Donelli, A. Martini, and A. Massa, "A hybrid approach based on PSO and Hadamard difference sets for the synthesis of square thinned arrays," *IEEE Trans. Antennas Propag.*, vol. 57, no. 8, 2491-2495, Aug. 2009.

[32] G. Oliveri and A. Massa, "Genetic algorithm (GA)-enhanced almost difference set (ADS)-based approach for array thinning," *IET Microw. Antennas Propag.*, vol. 5, no. 3, pp. 305-315, Feb. 2011.

[33] Y. Liu, Z. Nie, and Q. H. Liu, "A new method for the synthesis of non-uniform linear arrays with shaped power patterns," *Prog. Electromagn. Res.*, vol. 107, pp. 349-363, 2010.

[34] Y. Liu, Q. H. Liu, and Z. Nie, "Reducing the number of elements in the synthesis of shaped-beam pattern by the forward-backward matrix pencil method," *IEEE Trans. Antennas Propag.*, vol. 58, no. 2, pp. 604-608, Feb. 2010.

[35] J. Perini and M. Idselis, "Note on antenna pattern synthesis using numerical iterative methods," *IEEE Trans. Antennas Propag.*, vol. 19, no. 2, pp. 284-286, Mar. 1971.

[36] R. W. Redlich, "Iterative least-squares of nonuniformly spaced linear arrays," *IEEE Trans. Antennas Propag.*, vol. 21, no. 1, pp. 106-108, Jan. 1973.

[37] S. Holm, B. Elgetun, and G. Dahl, "Properties of the beampattern of weight- and layout-optimized sparse arrays," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 44, no. 5, pp. 983-991, Sep. 1997.

[38] B. P. Kumar and G. R. Branner, "Design of unequally spaced arrays for performance improvement," *IEEE Trans. Antennas Propag.*, vol. 47, pp. 511-523, Mar. 1999.

[39] B. P. Kumar and G. R. Branner, "Generalized analytical technique for the synthesis of unequally spaced arrays with linear, planar, cylindrical or spherical geometry," *IEEE Trans. Antennas Propag.*, vol. 53, pp. 621-633, Feb. 2005.

[40] F. B. T. Marchaud, G. D. de Villiers, and E. R. Pike, "Element positioning for linear arrays using generalized Gaussian quadrature," *IEEE Trans. Antennas Propag.*, vol. 51, no. 6, pp. 1357-1363, Jun. 2003.

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

[42] S. Ji, D. Dunson and L. Carin, "Multi-task compressive sampling," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 92-106, Jan. 2009.

[43] M. E. Tipping and A. C. Faul, "Fast marginal likelihood maximization for sparse Bayesian models," in Proc. 9th Int. Workshop Artificial Intelligence and Statistics, C.M. Bishop and B. J. Frey, Eds., 2003 [Online]. Available: *http://citeseer.ist.psu.edu/611465.html*

[44] A. C. Faul and M. E. Tipping, "Analysis of sparse Bayesian learning," in Advances in Neural Information Processing Systems (NIPS 14), T. G. Dietterich, S. Becker, and Z. Ghahramani, Eds., 2002, pp. 383-389 [Online]. Available: *http://citeseer.ist.psu.edu/faul01analysis.html*

[45] M. E. Tipping, "Sparse bayesian learning and the relevance vector machine", *J. Machine Learning Res.*, vol 1, pp. 211-244, 2001.

[46] A. Akdagli and K. Guney, "Shaped-beam pattern synthesis of equally and unequally spaced linear antenna arrays using a modified tabu search algorithm," *Microwave Opt. Technol. Lett.*, vol. 36, no. 1, pp. 16-20, 2003.

[47] S. Yang, Y. Liu, and Q. H. Liu, "Combined strategies based on matrix pencil method and tabu search algorithm to minimize elements of nonuniform antenna array", *Prog. Electromagn. Res. B*, vol. 18, pp. 259-277, 2009.

[48] D. Marcano and F. Duran, "Synthesis of antenna arrays using genetic algorithms," *IEEE Antennas Propag. Mag.*, vol. 42, no. 3, pp. 12-22, Jun. 2000.

[49] A. Akdagli and K. Guney, "Touring Ant Colony Optimization algorithm for shaped beam pattern synthesis of linear antenna arrays," *Electromagnetics*, vol. 25, no. 6, pp. 615-628, Aug. 2006.

[50] J. M. Cid, J. A. Rodriguez, and F. Ares, "Shaped power patterns produced by equispaced linear arrays: Optimized synthesis using orthogonal sin(Nx)/sin(x) beams," *J. Electromagn. Waves Appl.,* vol. 13, no. 7, pp. 985-992, 1999.

[51] B. Fuchs, "Synthesis of Sparse Arrays With Focused or Shaped Beampattern via Sequential Convex Optimizations," IEEE Transactions on Antennas and Propagation, vol. 60, no. 7, pp. 3499-3503, Jul. 2012.

[52] G. Prisco and M. D'Urso, "Maximally sparse arrays via sequential convex optimizations," IEEE Antenna Wireless Propag. Lett., vol. 11, pp. 192-195, 2012.

[53] A. Fannjiang, P. Yan, and T. Strohmer, "Compressed remote sensing of sparse objects," arXiv, 0904.3994v2, pp. 1-22, 2009.

[54] A. C. Fannjiang, "Compressive inverse scattering II. SISO measurements with Born scatterers," Inverse Problems, vol. 26, no. 3, pp. 1-17, Mar. 2010.

[55] S. A. Vorobyov, A. B. Gershman, and K. W. Wong, "Maximum likelihood direction-of-arrival estimation in unknown noise fields using sparse sensor arrays", *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 34-43, Jan. 2005.

[56] L. Lei, J. P. Lie, A. B. Gershman, and C. M. Samson See, "Robust adaptive beamforming in partly calibrated sensor arrays", *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1661-1667, Mar. 2005.

[57] E. Candes and J. Romberg, L1-Magic Code [online]. Available: http://users.ece.gatech.edu/~justin/l1magic/

[58] V. Stodden, L. Carin, D. Donoho, I. Drori, D. Dunson, M. Elad, S. Ji, J.-L. Starck, J. Tanner, V. Temlyakov, Y. Tsaig, and Y. Xue, SparseLab Code [online]. Available: http://sparselab.stanford.edu/

[59] K. Koh, S.-J. Kim, and S. Boyd, ell-1LS Code [online]. Available: http://www.stanford.edu/~boyd/l1_ls/

[60] S. Ji, Y. Xue, and L. Carin, "Single-task and multi-task Bayesian compressive sensing code", 2009 [Online]. Available: *http://people.ee.duke.edu/~lihan/cs/*

[61] Y. Nagata, T. Fujioka, and M. Abe, "Two-dimensional DOA estimation of sound sources based on weighted Wiener gain exploiting two-directional microphones," *IEEE Trans. Audio, Speech, Language Process.,* vol. 15, no. 2, pp. 416-429, Feb. 2007.

[62] P. Stoica, P. Babu, L. Jian, "SPICE: A sparse covariance-based estimation method for array processing" *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 629-638, Feb. 2011.

[63] L. Fulai, W. Jinkuan, S. Changyin, and D. Ruiyan, "Spatial differencing method for DOA estimation under the coexistence of both uncorrelated and coherent signals," *IEEE Trans. Antennas Propag.*, vol. 60, no. 4, pp. 2052-2062, Apr. 2012.

[64] C. H. Niow and H. T. Hui, "Improved noise modeling with mutual coupling in receiving antenna arrays for direction-of-arrival estimation," *IEEE Trans. Wireless Comm.*, vol. 11, no. 4, pp. 1616-1621, Apr. 2012.

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

[66] A. Swindlehurst and T. Kailath, "A performance analysis of subspace-based methods in the presence of model errors. I. The MUSIC algorithm," *IEEE Trans. Signal Process.*, vol. 40, no. 7, pp. 1578-1774, Jul. 1992.

[67] 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.

[68] M. D. Zoltowski, M. Haardt, and C. P. Mathews, "Closed-form 2-D angle estimation with rectangular arrays in element space or beamspace via unitary ESPRIT," *IEEE Trans. Signal Process.*, vol. 44, no. 2, pp. 316-328, Feb. 1996.

[69] F. Gao and B. Gershman, "A generalized ESPRIT approach to direction-of-arrival estimator," *IEEE Signal Process. Lett.*, vol. 12, no. 3, pp. 254-257, Mar. 2005.

[70] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 10, pp. 1553-1560, Oct. 1988.

[71] P. Stoica and A. B. Gershman, "Maximum-likelihood DOA estimation by data-supported grid search," *IEEE Signal Process. Lett.*, vol. 6, no. 10, pp. 273-275, Oct. 1999.

[72] M. Donelli, F. Viani, P. Rocca, and A. Massa, "An innovative multi-resolution approach for DOA estimation based on a support vector classification," *IEEE Trans. Antennas Propag.*, vol. 57, no. 8, pp. 2279-2292, Aug. 2009.

[73] A. H. El Zooghby, C. G. Christodoulou, and M. Georgiopulos, "A neural network-based smart antenna for multiple source tracking," *IEEE Trans. Antennas Propag.*, vol. 48, no. 5, pp. 768-776, May 2000.

[74] M. Pastorino and A. Randazzo, "The SVM-based smart antenna for estimation of the directions of arrival of electromagnetic waves," *IEEE Trans. Antennas Propag.*, Vol 55, No. 6, pp. 1918-1925, Dec. 2006.

[75] A. Randazzo, M. A. Abou-Khousa, M. Pastorino, and R. Zoughi, "Direction of arrival estimation based on support vector regression: Experimental validation and comparison with MUSIC," *IEEE Antennas Wireless Propag. Lett.*, vol. 6, pp. 379-382, 2007.

[76] D. Malioutov, M. Cetin, and A. S. 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.

[77] M. M. Hyder and K. Mahata, "A robust algorithm for joint-sparse recovery," *IEEE Signal Process. Lett.*, vol. 16, no. 12, pp. 1091-1094, Dec. 2009.

[78] Z.-M. Liu, z.-T. Huang, and Y.-Y. Zhou, "Direction-of-arrival estimation of wideband signals via covariance matrix sparse representation," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4256-4270, Sep. 2011.

[79] E. J. Candes, J. K. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," IEEE Transactions on Information Theory, vol. 52, n. 2, pp. 489-509, 2006.

[80] Donoho D., "Compressed sensing," IEEE Trans. Inf. Theory, vol. 52, no. 4, pp. 1289-1306, Apr. 2006

[81] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21-30, Mar. 2008.

[82] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing MRI," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 72-82, Mar. 2008.

[83] W. L. Chan, M. Moravec, R. Baraniuk, and D. Mittleman, "Terahertz imaging with compressed sensing and phase retrieval," *Optics Lett.*, vol. 33, pp. 974 - 976, 2008.

[84] L. C. Potter, E. Ertin, J. T. Parker, and M. Cetin, "Sparsity and compressed sensing in radar imaging," *Proc. IEEE*, vol. 98, no. 6, pp. 1006-1020, Jun. 2010.

[85] R. G. Baraniuk, "Compressive sampling," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 118-124, Jul. 2007.

[86] G. Oliveri, P. Rocca, and A. Massa, "A Bayesian compressive sampling-based inversion for imaging sparse scatterers," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3993-4006, Oct. 2011.

[87] G. Oliveri, L. Poli, P. Rocca, and A. Massa, "Bayesian compressive optical imaging within the Rytov approximation," *Optics Lett.*, vol. 37, no. 10, pp. 1760-1762, 2012.

[88] S. Lei and W. Huali, "Direction-of-arrival estimation based on modified Bayesian compressive sensing method," *Proc. 2011 Int. Conf. Wireless Comm Sig. Proc.* (*WCSP*), Nanjing,China, 9-11 Nov. 2011, pp. 1-4

[89] J. Foutz, A. Spanias, and M. K. Banavar, *Narrowband direction of arrival estimation for antenna arrays*. Morgan & Claypool, 2008.

[90] Bhaskar. D. Rao, and K. V. S. Hari, "Performance analysis of Root-Music," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 12, pp. 1939-1949, Dec. 1989.

[91] B. Otterstern, M. Viberg and T. Kailath, "Performances analysis of the total least squares ESPRIT algorithm," *IEEE Trans. Signal Process.*, vol. 39, no. 5, pp. 1122-1135, May 1991.

[92] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *IEEE Proc.*, vol. 98, no. 6, pp. 1058-1076, Jun. 2010.

[93] A. Majumdar, R. K. Ward, and T. Aboulnasr, "Compressed sensing based real-time dynamic MRI reconstruction," *IEEE Trans. Medical Imaging*, vol. 31, no. 12, pp. 2253-2266, Dec. 2012.

[94] J. Yang, J. Thompson, X. Huang, T. Jin, Z. Zhou, "Random-frequency SAR imaging based on compressed sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 2, pp. 983-994, Feb. 2013.

[95] L. Poli, G. Oliveri, P. Rocca, and A. Massa, "Bayesian compressive sensing approaches for the reconstruction of two-dimensional sparse scatterers under TE illuminations," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2920-2936, May 2013.

[96] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 10, pp. 1553-1560, Oct. 1988.

[97] J.-J. Fuchs, "On the application of the global matched filter to DOA estimation with uniform circular arrays," *IEEE Trans. Signal Process.*, vol. 49, no. 4, pp. 702-709, Apr. 2001.

[98] D. Model and M. Zibulevsky, "Signal reconstruction in sensor arrays using sparse representations," *Signal Process.*, vol. 86, pp. 624-638, 2006.

[99] A. C. Gubruz, V. Cevher, and J. H. McClellan, "Bearing estimation via spatial sparsity using compressive sensing," *IEEE Trans. Aerospace Electron. Syst.*, vol. 48, no. 2, pp. 1358-1369, Apr. 2012.

[100] M. M. Hyder and K. Mahata, "Direction-of-arrival estimation using a mixed l2,0 norm approximation," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4647-4655, Sep. 2010.

[101] H. Zhu, G. Leus, and G. B. Giannakis, "Sparsity-cognizant total least-squares for perturbed compressive sampling," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2202-2016, May 2011.

[102] I. Bilik, "Spatial compressive sensing for direction-of-arrival estimation of multiple sources using dynamic sensor arrays," *IEEE Trans. Aerospace Electron. Syst.*, vol. 47, no. 3, pp. 1754-1769, Jul. 2011.

[103] Z. Zhang and B. D. Rao, "Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning," *IEEE J. Select. Topic Signal Process.*, vol. 5, no. 5, pp. 912-926, Sep. 2011.

[104] E. T. Northardt, I. Bilik, and Y. I. Abramovich, "Spatial compressive sensing for direction-of-arrival estimation with bias mitigation via expected likelihood," *IEEE Trans. Signal Process.*, vol. 61, no. 5, pp. 1183-106, Mar. 2013.

[105] 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.

Appendix A

Derivation of (4.14)

To solve (4.13), the conditional probability $\mathcal{P}\left(\mathbf{w}_H|\,\widehat{\mathbf{F}}_H\right)$ is written according to the Bayes theorem, as

$$\mathcal{P}\left(\mathbf{w}_H|\,\widehat{\mathbf{F}}_H\right) \triangleq \frac{\mathcal{P}\left(\widehat{\mathbf{F}}_H\middle|\,\mathbf{w}_H\right)\mathcal{P}\left(\mathbf{w}_H\right)}{\mathcal{P}\left(\widehat{\mathbf{F}}_H\right)} \tag{A.1}$$

where $\mathcal{P}\left(\widehat{\mathbf{F}}_H\middle|\,\mathbf{w}_H\right)$ is the 'likelihood', whereas $\mathcal{P}\left(\widehat{\mathbf{F}}_H\right)$ and $\mathcal{P}\left(\mathbf{w}_H\right)$ are the priors of $\widehat{\mathbf{F}}_H$ and $\mathbf{w}_H$, respectively. Equation (A.1) is substituted in (4.13) to yield

$$\mathbf{w}_H^{MT-BCS} = \arg\left\{\max_{\mathbf{w}_H}\left[\frac{\mathcal{P}\left(\widehat{\mathbf{F}}_H\middle|\,\mathbf{w}_H\right)\mathcal{P}\left(\mathbf{w}_H\right)}{\mathcal{P}\left(\widehat{\mathbf{F}}_H\right)}\right]\right\} \tag{A.2}$$

Analogously to the BCS case, $\mathcal{P}\left(\mathbf{w}_H\right)$ in (A.2) is used to enforce the 'sparseness' of $\mathbf{w}_H$ (i.e., the minimization of $\|\mathbf{w}_H\|_{\ell_0}$) [21], but besides the BCS definition, the $MT-BCS$ prior also establishes the interrelationships between $\mathbf{w}_R$ and $\mathbf{w}_I$. Towards this end, a shared prior is placed across the two (i.e., $H=R$ and $H=I$) CS "tasks" in Eq. (A.2) [42]. Mathematically, it is assumed that [42]

$$\mathcal{P}\left(\mathbf{w}_H\right) = \int \mathcal{P}\left(\mathbf{w}_H|\,\widehat{\mathbf{a}},\widehat{\sigma}^2\right)\mathcal{P}\left(\widehat{\mathbf{a}}\right)\mathcal{P}\left(\widehat{\sigma}^2\right)d\widehat{\mathbf{a}}d\widehat{\sigma}^2 \tag{A.3}$$

where $\widehat{\mathbf{a}} = \{\widehat{a}_n;\, n=1,...,N\}$, $\widehat{\mathbf{a}} \in \mathbb{R}^N$, is the "shared" hyperparameters vector [42], whose associated hyperpriors still comply with the Gamma distribution [42]

$$\mathcal{P}\left(\widehat{\mathbf{a}}\right) = \prod_{n=1}^{N}\left[\frac{\beta_2^{\beta_1}\left(\widehat{a}_n\right)^{\beta_1-1}e^{-\beta_2\widehat{a}_n}}{\int_0^\infty t^{\beta_1-1}e^{-t}\mathrm{d}t}\right] \tag{A.4}$$

as for the BCS [see Eq. (5) - [21]]. Moreover, a "shared" Gamma hierarchical prior is enforced on $\widehat{\sigma}^2$ [42] with the same form as in the BCS (see Eq. (6) -

[21])

$$\mathcal{P}\left(\widehat{\sigma}^2\right) = \frac{\beta_4^{\beta_3}\left(\frac{1}{\widehat{\sigma}^2}\right)^{(\beta_3-1)} e^{-\frac{\beta_4}{\widehat{\sigma}^2}}}{\int_0^\infty t^{\beta_3-1}e^{-t}\mathrm{d}t} \tag{A.5}$$

where the user-defined coefficients β_1-β_4 are the so-called 'scale priors' [42].

Concerning $\mathcal{P}\left(\mathbf{w}_H|\,\widehat{\mathbf{a}},\widehat{\sigma}^2\right)$, the following hierarchical Gaussian model is assumed [42]

$$\mathcal{P}\left(\mathbf{w}_H|\,\widehat{\mathbf{a}},\widehat{\sigma}^2\right) = \left[(2\pi\widehat{\sigma})^{-N}\right]\prod_{n=1}^{N}\sqrt{\widehat{a}_n}\exp\left[-\frac{\widehat{a}_n\left(w_n^H\right)^2}{2\widehat{\sigma}^2}\right]. \tag{A.6}$$

Back substituting (A.3) in (A.2), it results that

$$\begin{aligned}&\mathbf{w}_H^{MT-BCS} = \\ &= \arg\left\{\max_{\mathbf{w}_H}\left[\int\frac{\mathcal{P}\left(\mathbf{w}_H|\widehat{\mathbf{a}},\widehat{\sigma}^2\right)\mathcal{P}\left(\widehat{\mathbf{F}}_H\middle|\mathbf{w}_H\right)\mathcal{P}(\widehat{\mathbf{a}})\mathcal{P}\left(\widehat{\sigma}^2\right)}{\mathcal{P}\left(\widehat{\mathbf{F}}_H\right)}d\widehat{\mathbf{a}}d\widehat{\sigma}^2\right]\right\}\end{aligned} \tag{A.7}$$

and, by integrating over $\widehat{\sigma}^2$ and performing simple mathematical manipulations, the relation (A.7) can be rewritten as

$$\mathbf{w}_H^{MT-BCS} = \arg\left\{\max_{\mathbf{w}_H}\left[\int\mathcal{P}\left(\mathbf{w}_H|\,\widehat{\mathbf{F}}_H,\widehat{\mathbf{a}}\right)\mathcal{P}\left(\widehat{\mathbf{a}}\middle|\widehat{\mathbf{F}}_H\right)d\widehat{\mathbf{a}}\right]\right\}. \tag{A.8}$$

As far as the first term in (A.8) is concerned, one can notice that [42]

$$\mathcal{P}\left(\mathbf{w}_H|\,\widehat{\mathbf{F}}_H,\widehat{\mathbf{a}}\right) = \int\mathcal{P}\left(\mathbf{w}_H|\,\widehat{\mathbf{F}}_H,\widehat{\mathbf{a}},\widehat{\sigma}^2\right)\mathcal{P}\left(\widehat{\sigma}^2\right)d\widehat{\sigma}^2 \tag{A.9}$$

whose integrand is given by

$$\mathcal{P}\left(\mathbf{w}_H|\,\widehat{\mathbf{F}}_H,\widehat{\mathbf{a}},\widehat{\sigma}^2\right)\mathcal{P}\left(\widehat{\sigma}^2\right) = \frac{\mathcal{P}\left(\widehat{\mathbf{F}}_H\middle|\mathbf{w}_H,\widehat{\sigma}^2\right)\mathcal{P}\left(\mathbf{w}_H|\widehat{\mathbf{a}},\widehat{\sigma}^2\right)\mathcal{P}\left(\widehat{\sigma}^2\right)}{\int\mathcal{P}\left(\widehat{\mathbf{F}}_H\middle|\mathbf{w}_H,\widehat{\sigma}^2\right)\mathcal{P}(\mathbf{w}_H|\widehat{\mathbf{a}},\widehat{\sigma}^2)d\mathbf{w}_H} \tag{A.10}$$

according to Bayes' theorem. By using (A.5) and (A.6), and observing that [see (4.11)]

$$\mathcal{P}\left(\widehat{\mathbf{F}}_H\middle|\,\mathbf{w}_H,\widehat{\sigma}^2\right) = \frac{1}{\left(2\pi\widehat{\sigma}^2\right)^{K/2}}\exp\left(-\frac{1}{2\widehat{\sigma}^2}\left\|\widehat{\mathbf{F}}_H-\widehat{\Phi}\mathbf{w}_H\right\|^2\right), \tag{A.11}$$

it results that

$$\begin{aligned}\mathcal{P}\left(\mathbf{w}_H|\,\widehat{\mathbf{F}}_H,\widehat{\mathbf{a}}\right) = &\left(\int_0^\infty t^{\beta_1+N/2-1}e^{-t}\mathrm{d}t\right)\times \\ &\times\frac{\left[1+\frac{1}{2\beta_2}(\mathbf{w}_H-\widehat{\mu}_H)^T\widehat{\Sigma}^{-1}(\mathbf{w}_H-\widehat{\mu}_H)\right]^{-(\beta_1+N/2)}}{\left(\int_0^\infty t^{\beta_1-1}e^{-t}\mathrm{d}t\right)(2\pi\beta_2)^{N/2}\sqrt{\left|\widehat{\Sigma}\right|}}\end{aligned} \tag{A.12}$$

where $\widehat{\mu}_H \triangleq \widehat{\Sigma}\widehat{\Phi}^T\widehat{\mathbf{F}}_H$ and $\widehat{\Sigma} \triangleq \left(\widehat{A}+\widehat{\Phi}^T\widehat{\Phi}\right)^{-1}$, being $\widehat{A} \triangleq diag\left(\widehat{\mathbf{a}}\right)$.

By analyzing the expression of $\mathcal{P}\left(\mathbf{w}_H|\,\widehat{\mathbf{F}}_H,\widehat{\mathbf{a}}\right)$, it is worth noticing that the posterior distribution over $\mathbf{w}_H$ is now a multivariate *Student-t* distribution (A.12) instead of the multivariate Gaussian distribution of the *BCS* (Eq. (9) - [21]). Moreover, the scale terms β_3 and β_4 do not have to be specified unlike β_1 and β_2 since the corresponding distributions are not explicitly required for the computations.

Concerning the remaining term in the integral of (A.8), a "delta-function" approximation is adopted analogously to the *BCS* case [42] since its closed-form computation is not feasible. Towards this end, let us firstly notice that

$$\mathcal{P}\left(\widehat{\mathbf{a}}\left|\widehat{\mathbf{F}}_H\right.\right)\propto\mathcal{P}\left(\widehat{\mathbf{F}}_H\left|\widehat{\mathbf{a}}\right.\right)\mathcal{P}\left(\widehat{\mathbf{a}}\right)$$

or in a different fashion

$$\begin{aligned}\mathcal{P}\left(\widehat{\mathbf{a}}\left|\widehat{\mathbf{F}}_H\right.\right)\propto\Big[\int\mathcal{P}\left(\widehat{\mathbf{F}}_H\left|\mathbf{w}_H,\widehat{\sigma}^2\right.\right)\mathcal{P}\left(\mathbf{w}_H|\,\widehat{\mathbf{a}},\widehat{\sigma}^2\right)\times\\ \times\,\mathcal{P}\left(\widehat{\sigma}^2\right)d\mathbf{w}_H d\widehat{\sigma}^2\Big]\mathcal{P}\left(\widehat{\mathbf{a}}\right)\end{aligned}\tag{A.13}$$

whose *mode* (over the *two* tasks $H\in\{R,\,I\}$) can be computed, by using (A.11), (A.5), and (A.6), as [42]

$$\widehat{\mathbf{a}}^{MT-BCS}=\arg\max_{\widehat{\mathbf{a}}}\left\{\mathcal{L}^{MT-BCS}\left(\widehat{\mathbf{a}}\right)\right\}\tag{A.14}$$

where $\mathcal{L}^{MT-BCS}\left(\widehat{\mathbf{a}}\right)$ is the logarithm of the $MT-BCS$ "marginal likelihood" given by

$$\begin{aligned}\mathcal{L}^{MT-BCS}\left(\widehat{\mathbf{a}}\right)=-\tfrac{1}{2}\textstyle\sum_H\left\{\log\left(\left|I+\widehat{\Phi}\left[\widehat{A}\right]^{-1}\widehat{\Phi}^T\right|\right)+\right.\\ \left.+\left(N+2\beta_1\right)\log\left[\widehat{\mathbf{F}}_H^T\left(I+\widehat{\Phi}\left[\widehat{A}\right]^{-1}\widehat{\Phi}^T\right)\widehat{\mathbf{F}}_H+2\beta_2\right]\right\}.\end{aligned}\tag{A.15}$$

By using (A.14), the *delta-function* approximation is then applied to obtain

$$\mathcal{P}\left(\widehat{\mathbf{a}}\left|\widehat{\mathbf{F}}_H\right.\right)\approx\delta\left(\widehat{\mathbf{a}}-\widehat{\mathbf{a}}^{MT-BCS}\right).\tag{A.16}$$

By substituting (A.12) and (A.16) in (A.8) and since the mode of a multi-variate *Student-t* distribution is equal to its average value (i.e., $\widehat{\mu}_H$), it turns out that

$$\begin{aligned}&\mathbf{w}_H\rfloor_{MTBCS}=\\&=\arg\left\{\max_{\mathbf{w}_H}\left[\int\mathcal{P}\left(\mathbf{w}_H|\,\widehat{\mathbf{F}}_H,\widehat{\mathbf{a}}\right)\delta\left(\widehat{\mathbf{a}}-\widehat{\mathbf{a}}^{MT-BCS}\right)d\widehat{\mathbf{a}}\right]\right\}=\\&=\arg\left\{\max_{\mathbf{w}_H}\left[\mathcal{P}\left(\mathbf{w}_H|\,\widehat{\mathbf{F}}_H,\widehat{\mathbf{a}}\right)\Big\rfloor_{\widehat{\mathbf{a}}=\widehat{\mathbf{a}}^{MT-BCS}}\right]\right\}\\&=\widehat{\mu}_H\rfloor_{\widehat{\mathbf{a}}=\widehat{\mathbf{a}}^{MT-BCS}}.\end{aligned}\tag{A.17}$$