UCLouvain

epl

École polytechnique de Louvain

Sparsity-Driven Moving Target Detection in Distributed Multistatic FMCW Radars

Author: **Gilles Monnoyer de Galland de Carnières**
Supervisors: **Laurent Jacques, Luc Vandendorpe**
Readers: **Jérôme Louveaux, Thomas Feuillen**
Academic year 2018–2019
Master [120] in Electrical Engineering

Abstract

We investigate target detection by employing joint sparse signal modeling and recovery techniques for widely distributed multistatic *Frequency Modulated Continuous Wave* (FMCW) radar systems using chirp modulation. The sparse domain considered is the space-velocity domain. While a vast literature on applying joint sparse modeling and recovery to distributed multistatic radars exists, the proposed models are often not scalable in terms of complexity and required processing power for high resolution in the space-velocity domain. Furthermore, while sparse modeling studies for multistatic FMCW radar only consider static targets, we introduce a complete and generic sparse modeling in the space-velocity domain to jointly represent baseband radar signals. From this model, simplifications are introduced and a trade-off is drawn between the gain in computational complexity and the impact on the accuracy due to added model mis-match. A low-complexity adaptation of the Matching Pursuit algorithm relying on this simplified FMCW model is designed, and an iterative algorithm to compensate for the errors caused by these model simplifications is introduced. The effectiveness of the proposed method is demonstrated through Monte-Carlo simulations. The method achieves a fast estimation of moving target's parameters on dense grids compared to previous sparsity-driven works on multistatic FMCW radars.

Acknowlegments

First of all, I want to express my thankfulness to my supervisors Prof. Laurent Jacques and Prof. Luc Vandendorpe for their constant support throughout the realization of my Master thesis. This work would never have been what it is without their help, rigour, and continuous desire to make me improve. I really appreciated their interest in this work, and their involvement in guiding me with their considerable knowledge, that made possible to produce this final rendering of my work.

I also want to thank Thomas Feuillen for his constant availability and manifest interest for my work, but also for all the advice and motivation he always gave to me. Many problems faced in my work would never have been solved without his help.

Finally, I would like to thank my family and friends who supported me all along, and give a particular mention to Mirian and to my godfather Vincent for their decisive encouragement and support in the most difficult times of this journey. I also want to give a particular thank to my friends Meidi and Dan for the help they provided to accomplish the measurements I needed.

Table of Contents

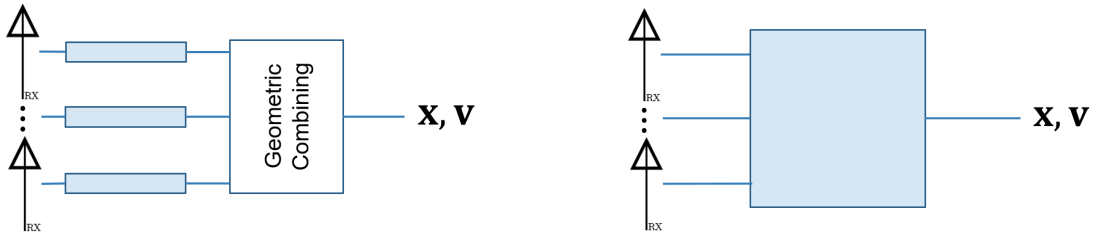
Introduction

In both military and civil domains, the need of localization and speed measurement systems has lead to the development of a large variety of sensors and signal processing techniques. In this framework, the widely-known *Radar* (RAdio Detection And Ranging) system has been developed several decades ago and has grown very popular to perform target detection and parameters (target location and velocity) estimation [1]. Depending on several signal parameters such as central frequency, bandwidth and modulation format, a Radar may provide information about one or several targets. While military and homeland security applications mainly require long-range devices (for instance, able to detect air planes from ground), the civil domain is more prone to require short-range devices (for instance, suited to observe the traffic in smart-cities applications). This last application is precisely the one targeted by this work.

In this context, the overarching purpose of this work is the derivation and improvement of algorithms for short-range Radar systems able to detect targets and estimate both their location and velocity vectors. To reach this goal, basic radar configurations containing a single transmitter (TX) and a single receiver (RX) are not sufficient because the lonely signal resulting from such systems may only provide information about a single speed component and a single position component for each target. We thus need Radar systems containing several TX antennas and/or several RX antennas, providing several distinct signals containing distinct information about the targets. Sensor configurations meeting this property are typically classified into two categories. The most conventional systems are composed of co-located Multiple-Input Multiple-Output (MIMO) antennas where all TX antennas and all RX antennas are close to one another. While more uncommon systems, yet growing in interest, can also be composed of distributed multi-static sensors, whose TX and RX antennas are all widely separated without any particular arrangement concerning their location.

Conventional modeling and processing treat each received signal independently to obtain estimates of distinct position and speed components, and subsequently combine them to localize targets in the 2-D plane (thanks to, e.g., trilateration [2]). Unfortunately, each independent process does not take advantage of information carried by other signals to make their decisions of detection. Motivated by this potential advantage, another approach, investigated in this work, consists in a joint modeling of received signals, hence implying joint estimation of the target parameters, as depicted by Fig. 1.

Last but not least, an important feature characterizing the aspects of this work is the assumption according to which the targets to detect are seen as point targets. Under



(a) Disjoint signals modeling

(b) Joint signals modeling

Figure 1: Schematic illustration comparing disjoint signals modeling and processing involving later combination of information extracted independently on each signal, and joint signals modeling and processing.

such an assumption, the crucial key element of driving the modeling and algorithmic developments is the *sparsity* prior on the received signals. More precisely, this assumption suggests that there exists some basis allowing a description of the signals as vectors containing only a very small amount of non-zero elements.

In this introduction, some general Radar and signal processing concepts are briefly described in order to highlight the challenges imposed by the problem addressed in this work. This analysis is followed by an overview of the existing solution proposed in the literature and a proper description of the contribution this document claims to bring.

Modulations

A great richness of radars is the large variety of modulations that can be used by transmitters. Each modulation has its own properties and advantages. There are two fundamental classes of radars regarding the type of signal they transmit. The *Pulse-Doppler Radars* (PDR), i.e., sending a sequence of very short signals separated by much longer breaks, and the *Frequency-Modulated Continuous Wave* (FMCW) Radars, i.e., sensing an uninterrupted modulated signal. While most contributions on multi-static Radars deal with PDRs, we are targeting FMCW Radars through this work. In order to both motivate this choice and highlight useful concepts to describe the literature on multi-static Radars, let us briefly compare these two types of Radars.

FMCW Radars

A *Frequency-Modulated Continuous Wave* Radar sends an uninterrupted signal $s_T(t)$ with an instantaneous carrier frequency $f_c(t)$ resulting from a modulation around a central carrier f_0. More precisely, the transmitted radio frequency (RF) signal writes

$$s_T(t) = e^{j2\pi \int_0^t f_c(t')\mathrm{d}t'}. \tag{1}$$

In the demodulation step of a FMCW Radar, the received signal is mixed with the currently transmitted signal. The most basic CW Radars send a constant carrier, and thus may provide unambiguous information only about the speeds of the targets thanks to Doppler frequencies appearing in the demodulated signal. To get unambiguous information on the targets location, modulated waves are required. In all generality, the resolution on

the distance estimate (denoted by R_d) is linked to the bandwidth B of the transmitted signal by

$$R_d = \frac{c}{2B}, \tag{2}$$

where c is the speed of light. More precisely, the distance resolution R_d corresponds to the minimal distance between two targets such that they are distinguishable. Therefore we desire a low value of R_d. Some popular frequency modulations are given in Fig. 2. In the step-frequency modulation, the echo resulting from the presence of a target is such that the sample of a given step has a phase shift with respect to the previous one, caused by the delay between transmission and reception, hence providing information on the distance of the target. In the chirp, or saw-tooth modulation, the delayed echo of the signal is demodulated with a slightly different frequency than the one it was modulated with. Therefore, the frequency of the obtained baseband signal gives information about the distance of the target. In both modulations, several stairs (or ramps) have to be acquired to estimate both distance and speed. Finally, a third example is given: the triangular modulation, whose advantage with respect to the chirp modulation is the fact that having two different slopes in the function $f_c(t)$ enables to easily distinguish the frequency resulting from the distance and the frequency resulting from the speed.

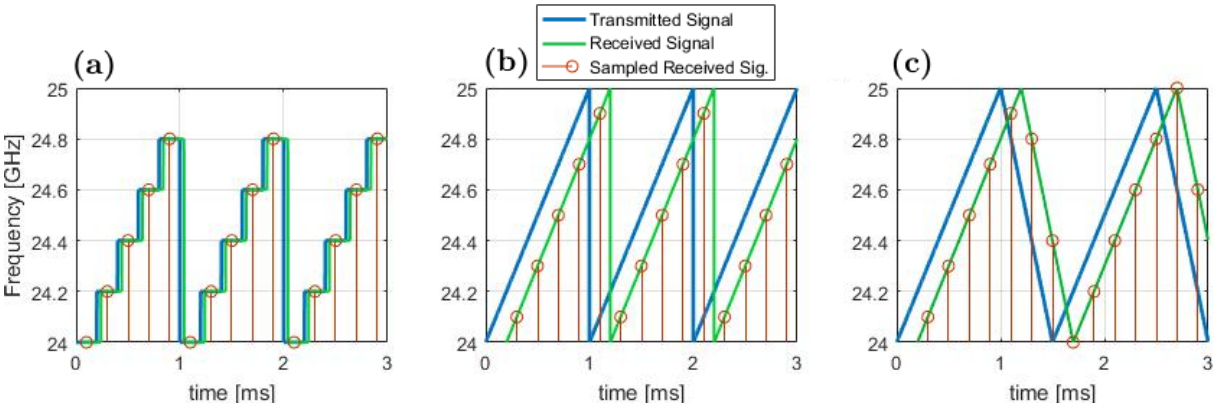


Figure 2: Example of popular FMCW modulations $f_c(t)$. Theoretical illustrations of transmitted signals, received signals (assuming one echo) and sampled received signals are displayed. **(a)** Step frequency, **(b)** Chirps (Saw-tooth function), **(c)** Triangular.

Pulse-Doppler Radars

Pulse Radars send short pulses, typically around 0.01 to 1 μs long, separated by breaks of typically 0.1 to 1ms long. The signal, $s_T(t)$, transmitted during a given pulse is obtained by the modulation of a waveform $w(t)$ around a single carrier of frequency f_c, such that

$$s_T(t) = w(t)e^{j2\pi f_c t}. \tag{3}$$

The information about the distances of the targets is contained in the delay of the received signal. Figure 3 shows a example of transmitted and received baseband signals of a PDR using Gaussian waveform (as used, for example in [3]), in the presence of a target. Note that we may acquire (sample) data only during some specific time periods if we

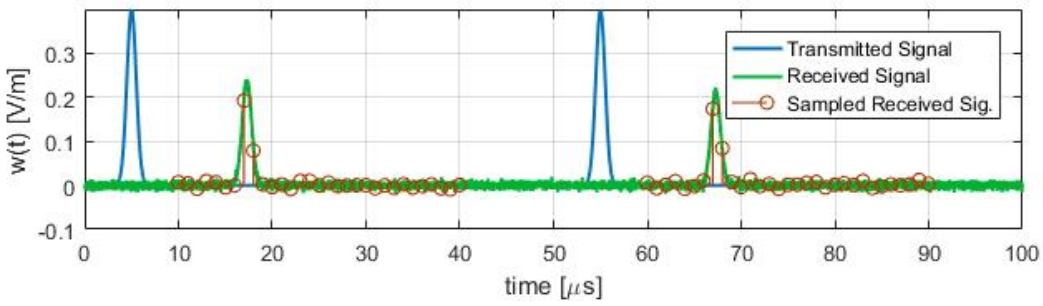


Figure 3: Example of the baseband representation of the transmitted and received signals we would obtain from a PDRs using a Gaussian waveform $w(t)$. The Gaussian represented has a variance of $0.5\mu s$, and thus a bandwidth of $0.63MHz$. The sampling frequency is $1MHz$.

desire to only consider a smaller domain of possible distances. Finding the delay, hence the distance of a target, is performed by applying a sliding correlation between the transmitted waveform and the received one. The speed is observed through the Doppler-effect, affecting the received waveform. More precisely, the Doppler effect, generally neglected within a pulse but significant between pulses, induces a phase shift between successive received pulses. The resolution on the distance estimation is implicitly linked to the duration of the pulse (denoted by T_p) via the relation

$$R_d = \frac{c}{2B}, \tag{4}$$

where $B \propto \frac{1}{T_p}$ is the bandwidth of the waveform $w(t)$. However, according to the Shannon-Nyquist theorem, the sampling frequency must be greater than the bandwidth. Therefore in PDRs, a better resolution requires a higher sampling rate. The maximal unambiguous range is related to the time between each pulse because the echo of the $n'th$ pulse is supposed to arise before the transmission of the $(n+1)'th$ pulse.

Application field of each molulation

From the descriptions given above, we can show that PDRs are more suited for long range applications while FMCW Radars are more suited to short range applications targeted by this work. There are three reasons supporting this:

- PDRs have the main advantage of distributing efficiently the transmitted power. Indeed, for a given average power, FMCW Radars spread this power almost uniformly across the time, while PDRs concentrate this power into very strong pulses. Therefore PDRs are more robust against strong path loss occurring in long range applications and thus, able to see much further away than their FMCW counterparts.

- In PDR, increasing B to improve the resolution requires to also increase the sampling frequency. Therefore, the distance resolution usually remains limited for PDRs (e.g., in the multi-static PDR developed by Derham et al. in [4, 5], $f_s = 100MHz$ and $B = 50MHz$ are chosen, leading to a resolution of 6m). Having such a limited resolution is not an issue for long range applications but is problematic for short-range applications. In FMCW, the sampling frequency is not linked to the choice of B, yet is somehow linked to the maximal unambiguous speed. Therefore, if the bandwidth

allows it, FMCW are able to reach much better resolutions while keeping a small sampling frequency (e.g., a resolution of 75cm with $B = 200MHz$ and $f_s = 50kHz$ in this work).

- Finally, PDRs offer the possibility to use the same antenna for both the transmission and reception. The only cost of doing so is the existence of a minimal detectable range because the echo must not arrive during the transmission, when we are not listening.

Radar Configurations

The most basic Radar configuration, consisting of a single TX antenna and the single RX antenna located at the same location, is the *monostatic* configuration. We shall call a location containing at least one antenna a *node*. Regardless of the modulation chosen, the single signal acquired may only provide information about the distances travelled by the wave and the radial speeds of the targets reflecting it. In other terms, we only know the radius of the circles on which the targets are localized and the radial components of their speeds. More generally, we can see the monostatic Radar as a particular case of *bistatic* Radar, i.e., with a single TX antenna located far from a single RX antenna, hence composed of two nodes. In this case, we only have information about the targets *bistatic range*: the total distance travelled by the wave, defining an ellipse on which the targets are localized. The speed component estimated is called the *bistatic speed*, which is the sum of the projections of the velocity vector on the paths travelled by the wave (i.e., the sum of the lengths of the green arrows in Fig. 4).

If we want to recover the location and velocity vectors, multiple distinct range and speed information are required. This means that at least 2 distinct paths are required for 2D vectors (3 distinct paths for 3D vectors). In other words, we have to add antennas. From the bistatic configuration, we have two possibilities. We can keep the nodes and increase the number of antennas per node to get a *co-located MIMO* configuration as shown in Fig. 5 b. More precisely, we replace the antennas by linear antenna array (typically spaced by a half-wavelength). Or, we can keep one antenna per node and add nodes far from one another to obtain a *multi-static SISO* configuration such as shown in Fig. 5 c. Finally, we may also set up a configuration composed of multiple nodes composed of antenna arrays, i.e., a *multi-static MIMO* system (Fig. 5 d).

From a signal model point of view, the concepts of MIMO and multi-static are defined by the hypothesis we use.

- All the signals resulting from one pair of nodes containing several antennas (i.e., one MIMO path), are assumed to be affected by the same *Radar Cross Section* (RCS)[1], path loss and scattering. Only the phase shift due to the slight difference of travelled distances is considered, hence providing some angular resolution.

[1]RCS measures how detectable an object is by a Radar. It reflects many parameters such as the material, shape and size of the target but also the angle of reflections and the polarization

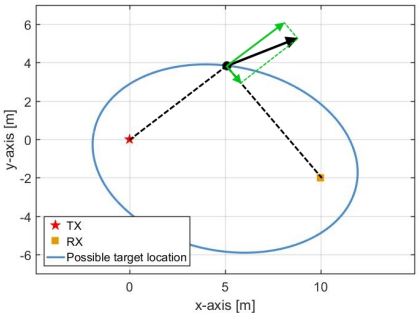


Figure 4: Schematic illustration of the target localization in bistatic Radar configuration. The *bistatic speed* is the sum of the green arrow.

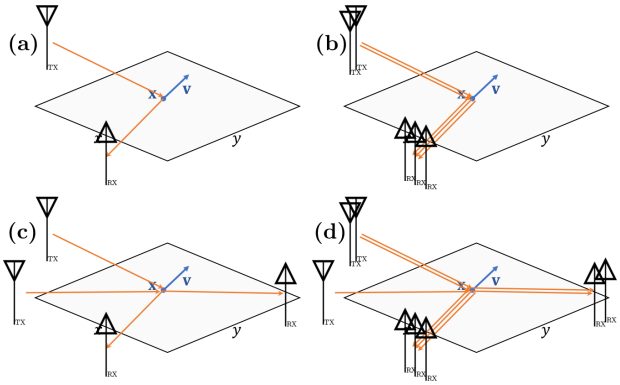


Figure 5: Illustration of the different configurations. **(a)** bistatic SISO, **(b)** bistatic MIMO, **(c)** multi-static SISO, **(d)** multi-static MIMO.

- All the signals resulting from distinct pairs of nodes (multi-static configuration), are assumed to be affected by i.i.d random (RCS) and different path losses and phase shifts, hence providing diversity by viewing the targets from different angles.

From these definitions, we can expect MIMO nodes to provide an array gain because we are possibly able to add signals coherently. Multi-static configurations are expected to provide a diversity gain and enhance the target recognition [6] and tracking [7].

Sparsity Principle

The sparsity principle relies on the fact that some vectors can be decomposed, in some basis, with only a few non-zero elements. For example, a sine with an appropriate frequency is a Dirac in the Discrete Fourier basis, and is thus expressed with only one non-zero element. More precisely, a signal $\boldsymbol{y}$ is said K-sparse if, in some basis $\boldsymbol{\Psi}$, we can write

$$\boldsymbol{y} = \boldsymbol{\Psi}\boldsymbol{s}, \tag{5}$$

where $\boldsymbol{s}$ has, at most, K non zero elements. The sparsity principle applies when we assume, as a prior knowledge, that a signal $\boldsymbol{y}$ is sparse in a known basis $\boldsymbol{\Psi}$. In the case of the Fourier basis, $\boldsymbol{\Psi}$ is orthogonal. In fact, $\boldsymbol{\Psi}$ does not have to be orthogonal, neither it has to even be a basis (i.e., be a square matrix of linearly independent rows/columns). It is common to interpret the sparsity principle by translating the fact that $\boldsymbol{y}$ can be obtained with a weighted addition of a few elementary signals, called *atoms*, selected from a large bank of atoms. This bank forms a matrix we call a *dictionary*. Unlike a basis, a dictionary is usually not a square matrix. More precisely, if $\boldsymbol{y} \in \mathbb{C}^M$, $\boldsymbol{s} \in \mathbb{C}^N$ and if we now consider that $\boldsymbol{\Psi}$ is a dictionary, we have $\boldsymbol{\Psi} \in \mathbb{C}^{M \times N}$ with possibly $N \gg M$ linking $\boldsymbol{y}$ and $\boldsymbol{s}$ the same way as (5).

When we have the prior knowledge of the sparsity of $\boldsymbol{y}$ in a dictionary , it is often desired to find its expression $\boldsymbol{s}$ in the sparse domain. Because there is an infinity of possible vectors $\boldsymbol{s}$ when $N \gg M$, specific algorithms dedicated to the reconstruction of sparse signals have been developed. Some of these algorithms are explained in Chap. 2. Such computations find many applications such as data compression, denoising and

deconvolution [8]. More recently, *Compressed Sensing* (CS) has emerged from the sparsity principle [9] as a generalization of the Shannon-Nyquist sampling theory. In a few words, CS suggests that the sparsity prior on a signal $\boldsymbol{y}$ would allow us to recover $\boldsymbol{s}$ with a number of acquired samples related to the intrinsic complexity of the signal (typically its sparsity level) and hence much lower than the ambient domain dimension, i.e., much lower than the number of samples M necessary according to the Shannon-Nyquist theory. While this document is only focused on the application of the sparsity principle to multi-static radars, the CS theory and its possible application to multi-static radars is discussed in Chap. 5.

Sparsity-Driven Contributions on Multi-static Radars

To sum up the concepts introduced so far, this work focuses on the joint modeling and algorithmic development for distributed multi-static Radar systems using the FMCW *sawtooth* modulation, suited for short-range applications, with the goal to estimate target states (location and velocity vectors) of one or multiple point targets. Moreover, the prior of the sparsity of the received signals (according to some known acquisition matrices) will appear to be both useful for the derivation of such algorithms and advantageous for further development of low-cost systems involving, for instance, *Compressed Sensing* (CS).

The sparse modeling of the signals acquired from Radar systems has been studied for more than a decade and has been mainly motivated by the use of CS. Some of these contributions are summarised in [10]. Regardless of the configuration, the sparse domain, defined to build such sparse representations, corresponds to a chosen set of possible target states to be tested. The best choice of this domain depends on the configuration and the desire or not to form a joint model. For monostatic or bistatic configurations, the target states defining the sparse domain are some set of bistatic ranges and bistatic speeds, leading to a *Delay-Doppler Shifted* dictionary for PDRs (e.g., in [11]), or possibly a Fourier based dictionary for FMCW Radars, if the set of states is well chosen (detailed in Chap. 1). If we have a bistatic MIMO Radar, the set of target states grows as a third feature is added: the angle(s) of arrival and/or of departure (see Fig. 6).

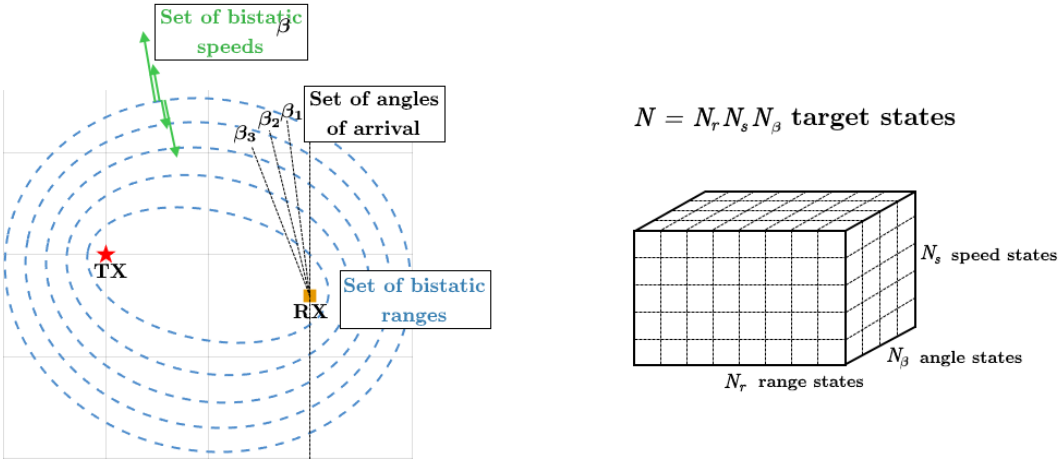


Figure 6: Schematic illustration of the (bistatic range, bistatic speed, angle) sampled domain.

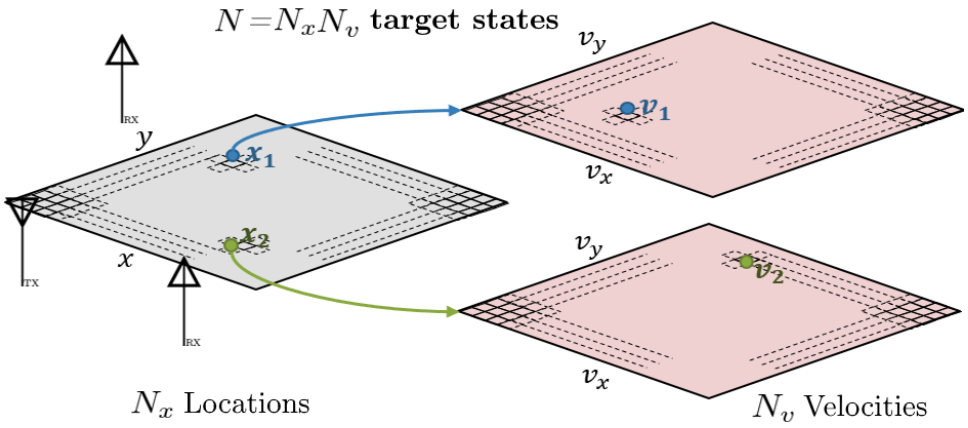


Figure 7: Schematic illustration of the location-velocity sampled domain according to a uniform set of target states to be tested

When dealing with multi-static Radars, we have two possibilities to build a sparse model. In the first method, we can consider each pair of nodes as a bistatic Radar. Therefore, for each pair, we define independently from the others a set of target states and the associated dictionary as explained above. Each problem can be solved independently and the solutions are combined using, for instance, trilateration. The second method consists in building a joint model, where a unique set of target states is defined for all pairs. Therefore the concept of bistatic range and bistatic speed no longer holds and the states are defined as vectors belonging to the location-velocity 2D×2D domain (Fig. 7). Such modeling was performed for PDRs in [12, 13], where the location-velocity domain was sampled into N possible states to be tested, hence building N basis functions for each TX-RX pair. While this case is still stated as multiple reconstruction problems, they are now coupled by a common sparsity pattern. Such model enables the use of an algorithm known as *Block Matching Pursuit* (BMP) described in Chap. 2, suited to solve the problem we just defined [14].

However, the number of states N can be tremendously large, even at the minimum sampling rate required to fully exploit the potential of the radars. To overcome this issue, a lower complexity strategy has been proposed for PDRs [3]. The idea consists in estimating separately the location and the velocity, which dramatically reduces the dimensions of the problem. To do so, the authors are taking advantage of the pulse structure of the transmitted signal. More precisely, they are first selecting arbitrarily one single pulse per received signal and build a joint model considering only the location states of the targets. This can easily be done in PDRs because the joint modeling still preserves the intrinsic location-velocity decoupling (coming from the delay-Doppler decoupling) of the signal structure. Similar solutions have not yet proposed to overcome the dimensional issue for FMCW Radars. For example, in [15], the authors are working with static targets, hence with a reduced number of target states to avoid facing this complication. Unlike multi-static PDRs signals, jointly modeled multi-static FMCW signals do not have simple properties directly exploitable to perform decoupled location-velocity estimation. In the present work, we perform a rigorous analysis of the atoms of the dictionaries resulting from the joint modeling of multi-static FMCW Radars signals. From this analysis, we

found a way to express this dictionary in a factored fashion, suggesting the possibility of decoupling the target states estimation. To harness this property, we developped an adaptation of MP and BMP we named the *Factored Matching Pursuit* (FMP) and *Factored Block Matching Pursuit* (FBMP).

The next chapters are organized according to the following methodological steps. In Chap. 1, the received signals description is provided according to several hypothesis that will be highlighted. These expressions enable the formulation of sparse modeling, possibly involving further approximations to enable simplifications. In Chap. 2, a relevant overview of sparse recovery algorithms is proposed. Adaptations and improvements of existing algorithms are developed to efficiently process signals whose properties correspond to model previously obtained, while meeting the desired low-complexity. Then, in Chap. 3, these algorithms are applied to the Radar systems of interest, in order to be evaluated and discussed. Thanks to simulations, the different sources of inaccuracies in the targets' parameters estimation are described and enable the derivation of iterative corrections to improve estimation performances of the low-complexity algorithms. The next step is the validation and performances analysis, performed in Chap. 4, first using extensive simulations and then completed with real measurements. Finally, the many perspectives offered by this work are discussed in Chap. 5, encouraging further works that will allow the development of efficient applications of CS algorithms, and super-resolution algorithms to multistatic FMCW radars.

Notations and Conventions

Notation	Meaning
j	Imaginary unit $\sqrt{-1}$
a, A	A scalar is denoted by a non-bold symbol
$\boldsymbol{a}$	A vector is denoted by a bold lowercase symbol
$\boldsymbol{A}$	A matrix is denoted by a bold uppercase symbol
$A_{i,:}$	i-th row of the matrix $\boldsymbol{A}$
$A_{:,i}$	i-th column of the matrix $\boldsymbol{A}$
$[\boldsymbol{A}]_i$	i-th column of the matrix $\boldsymbol{A}$ when $A_{:,i}$ would be a heavy or ambiguous notation
$\boldsymbol{A}^T$	Transpose of the matrix $\boldsymbol{A}$
$\boldsymbol{A}^*$	Conjugate, element by element, of the matrix $\boldsymbol{A}$
$\boldsymbol{A}^H$	Conjugate Transpose of the matrix $\boldsymbol{A}$
$\mathrm{Supp}\,(\boldsymbol{a})$	Support of the vector $\boldsymbol{a}$, i.e., $\{i : a_i \neq 0\}$
$\langle \boldsymbol{a}, \boldsymbol{b} \rangle$	Scalar product of $\boldsymbol{a}$ and $\boldsymbol{b}$ ($\boldsymbol{a}$ and $\boldsymbol{b}$ must be vectors of the same length)
$\|\boldsymbol{a}\|_0$	ℓ_0 norm of the vector $\boldsymbol{a}$, i.e., $\lvert \mathrm{Supp}\,(\boldsymbol{a}) \rvert$
$\|\boldsymbol{a}\|_1$	ℓ_1 norm of the vector $\boldsymbol{a}$, i.e., $\sum_i \lvert a_i \rvert$
$\|\boldsymbol{a}\|_2$	ℓ_2 norm of the vector $\boldsymbol{a}$, i.e., $\langle \boldsymbol{a}, \boldsymbol{a} \rangle$
$\mathrm{Vec}\,(\boldsymbol{A})$	Vector reshape of the matrix $\boldsymbol{A}$, by concatenating its rows
$\mathrm{Mat}_N\,(\boldsymbol{a})$	Matrix Reshape of the vector $\boldsymbol{a}$ by stacking chunks of N consecutive components of $\boldsymbol{a}$ in the rows of the output matrix.
$f(t)$	A continuous function of $t \in \mathbb{R}$
$y[n]$	A discrete signal of $n \in \mathbb{Z}$
$\Omega\,\{n\}$	the n-th element of the set Ω
$:=$	"Defined as"
$\propto$	"Proportional to"
$\overset{\propto}{\sim}$	"Approximately proportional to"
s.t.	"Such that"
$[N]$	Set of N first positive integers: $[N] := \{1, .., N\}$
$\overset{\text{i.i.d.}}{\sim}$	Independent and Identically Distributed
$\mathbb{C}\mathcal{N}\,(0, \sigma^2)$	Centered complex normal distribution of variance σ^2

Table 1: List of notations

Chapter 1

RADAR Signal Models and Acquisition

The first step to solve the problem stated in the introduction is the establishment of a proper joint acquisition, or sensing, model of the signals acquired by multi-static FMCW Radars. Therefore, the focus of this chapter is the formulation of this acquisition as a sparse sensing model, i.e., with dictionaries linking each received signal to a sparse vector.

The multi-static signals form a group of distinct, yet linked, bistatic signals. Therefore, in this chapter, we start with the most simple case of a bistatic SISO Radar and progressively build the models describing the complete multi-static MIMO Radar system. Regardless of the configuration, we need to provide a transmission model, describing every alteration occurring between a transmitted waveform and the corresponding received waveform. The analysis of a complete transmission model is a wide topic because such transmissions are affected by many specific system parameters such as the properties of the antennas, of the targets and of the surrounding environment. In the first section of this chapter, in order to define a very simple and general transmission model, we state many hypothesis on the transmission scene and derive an expression of the received waveform. Next, the acquisition step is considered and requires several additional assumptions that enable the derivation of a complete expression of the sampled received signal. This expression allows us to define a dictionary linking it to a sparse representation. Also, simplifications are proposed to reformulated this sparse model in a factored fashion.

Finally, with the help of additional assumptions, we extend this bistatic Radar acquisition sparse model to both multi-static and MIMO systems, taking into account their respective properties stated in the introduction. Moreover, the possibility to either build a joint acquisition model or separate bistatic models is discussed.

1 Transmission Model

1 .1 Problem statement and hypothesis

Let us address the problem of the transmission of a modulated wave propagating in the 2D-plane from a transmitter located in $\boldsymbol{X_t} \in \mathbb{R}^2$, reflected by a single point target located in $\boldsymbol{x} \in \mathbb{R}^2$ and arriving to a receiver located in $\boldsymbol{X_r} \in \mathbb{R}^2$. The target is moving with a constant velocity $\boldsymbol{v} \in \mathbb{R}^2$. The transmission scene is described in Fig. 1.1, it is assumed that there is no direct transmission between the transmitter and the receiver. Also, there are no ground reflection nor any other path taken by the wave than the one shown in Fig. 1.1. As in Fig. 1.2, we assume a single target located in the far field of both the transmitter and the receiver, and such that the wave reaching it is seen as a plane wave propagating along the direction of $\boldsymbol{x} - \boldsymbol{X_t}$. Let us define two unitary vectors, $\hat{\boldsymbol{u}}$ and $\hat{\boldsymbol{u}}'$, respectively pointing from the transmitter to the target and from the target to the receiver,

$$\begin{aligned} \hat{\boldsymbol{u}} &= \frac{\boldsymbol{x} - \boldsymbol{X_t}}{\|\boldsymbol{x} - \boldsymbol{X_t}\|_2}, \\ \hat{\boldsymbol{u}}' &= \frac{\boldsymbol{X_r} - \boldsymbol{x}}{\|\boldsymbol{X_r} - \boldsymbol{x}\|_2}. \end{aligned} \tag{1.1}$$

We are going establish the expression of the received waveform in $\boldsymbol{X_r}$ as a function of time t for $t \in [0, T_{acq}]$, where T_{acq} is the data acquisition duration. Although the target is moving, we assume negligible the impact its displacement has on both $\hat{\boldsymbol{u}}$ and $\hat{\boldsymbol{u}}'$ within an acquisition, i.e., $\hat{\boldsymbol{u}}$ and $\hat{\boldsymbol{u}}'$ are constant. We also define two referentials, $\mathcal{Z}$ and $\mathcal{Z}'$, respectively inertial and attached to the target. If $\boldsymbol{z}$ and $\boldsymbol{z}'$ describe the same location in respectively, $\mathcal{Z}$ and $\mathcal{Z}'$, they are linked by the following Galilean transformation:

$$\boldsymbol{z}' = \boldsymbol{z} - \boldsymbol{x} - \boldsymbol{v}t. \tag{1.2}$$

Because the Maxwell's equations are not invariant with respect to Galilean transformations, the model we describe only makes sense if $\|\boldsymbol{v}\| \ll c$.

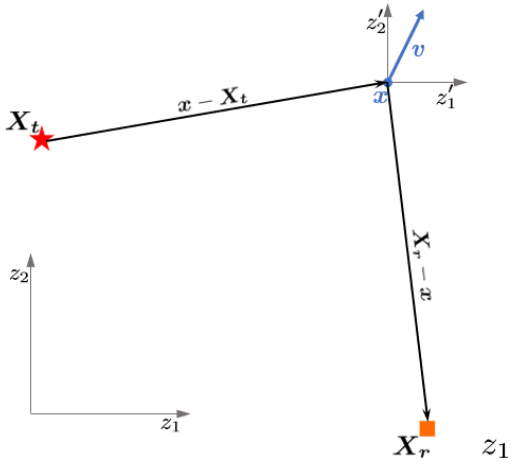


Figure 1.1: Schematic illustration of the propagation path of the wave between the transmitter located in $\boldsymbol{X}_t$ and the received located in $\boldsymbol{X}_r$.

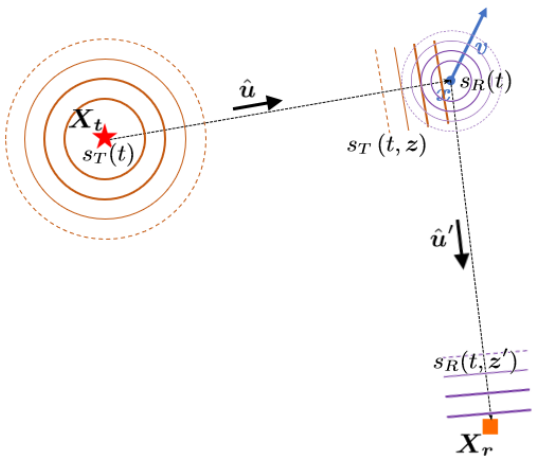


Figure 1.2: Schematic illustration of the propagation of a wave assumed as a plane wave when arriving to the target's location and to the receiver.

1 .2 Trasmission model description

At the transmitter side, we consider a wave with a frequency modulated by a general function $f_c(t)$, such that the transmitted signal is

$$s(t) = \sqrt{P} e^{j2\pi \int_0^t f_c(t')\mathrm{d}t'} := \sqrt{P} e^{j\phi_T(t)}, \tag{1.3}$$

where P is the transmitted power and we define $\phi_T(t)$ as the instantaneous phase of the transmitted signal at instant t. For the rest of this whole document, we normalize the transmitted power to $P = 1$. Therefore, the transmitted propagating wave, denoted $s_T(t, \boldsymbol{z})$ is such that $s_T(t, \boldsymbol{z} = \boldsymbol{X_t}) = s(t) = e^{j\phi_T(t)}$, hence the presumably plane wave in the surroundings of the target is

$$s_T(t, \boldsymbol{z}) = e^{j\phi_T\left(t - \frac{1}{c}\hat{\boldsymbol{u}}\cdot(\boldsymbol{z}-\boldsymbol{X_t})\right)}. \tag{1.4}$$

In the target's referential $\mathcal{Z}'$, this function is equivalent to

$$s_T(t, \boldsymbol{z}') = e^{j\phi_T\left(t - \frac{1}{c}\hat{\boldsymbol{u}}\cdot(\boldsymbol{z}'+\boldsymbol{x}+\boldsymbol{v}t-\boldsymbol{X_t})\right)}. \tag{1.5}$$

At the target's location at instant t , i.e., in $\boldsymbol{z}' = \boldsymbol{0}$, the target reflects the incident wave $s_T(t, \boldsymbol{z}' = \boldsymbol{0})$, and hence emits a reflected wave $s_R(t, \boldsymbol{z}')$. Let us define, as a approximated model, a complex *scattering coefficient* α modeling the effects of all scattering occurring in this process (which, in reality, depends on the target's shape and material, the wave polarization, and the incident and reflected angles). Therefore in $\boldsymbol{z}' = \boldsymbol{0}$, the incident and reflected waveforms are such that

$$s_T(t, \boldsymbol{z}' = \boldsymbol{0}) = e^{j\phi_T\left(t - \frac{1}{c}\hat{\boldsymbol{u}}\cdot(\boldsymbol{x}+\boldsymbol{v}t-\boldsymbol{X_t})\right)}, \tag{1.6}$$

$$s_R(t, \boldsymbol{z}' = \boldsymbol{0}) = \alpha e^{j\phi_T\left(t - \frac{1}{c}\hat{\boldsymbol{u}}\cdot(\boldsymbol{x}+\boldsymbol{v}t-\boldsymbol{X_t})\right)} := \alpha e^{j\phi_R(t)}, \tag{1.7}$$

where we define $\phi_R(t)$ as the instantaneous phase of the reflected wave emitted by the target at instant t. As shown in Fig. 1.2, in the surroundings of the receiver, the reflected wave is assumed to be a plane wave propagating along $\hat{\boldsymbol{u}}'$ and fulfilling the equation in (1.7). Its expression in the surroundings of the receiver is

$$s_R(t, \boldsymbol{z}') = \alpha e^{\phi_R\left(t - \frac{1}{c}\hat{\boldsymbol{u}}'\cdot\boldsymbol{z}'\right)}. \tag{1.8}$$

At the receiver's location, i.e., in $\boldsymbol{z}' = \boldsymbol{X_r} - \boldsymbol{x}$, the wave is expressed as

$$s_R(t, \boldsymbol{z}' = \boldsymbol{X_r} - \boldsymbol{x}) = \alpha e^{\phi_R\left(t - \frac{1}{c}\hat{\boldsymbol{u}}'\cdot(\boldsymbol{X_r}-\boldsymbol{x})\right)}. \tag{1.9}$$

$$= \alpha e^{\phi_T\left(t - \frac{1}{c}\hat{\boldsymbol{u}}\cdot(\boldsymbol{x}-\boldsymbol{X_t}) - \frac{1}{c}\hat{\boldsymbol{u}}'\cdot(\boldsymbol{X_r}-\boldsymbol{x}) - \frac{1}{c}\hat{\boldsymbol{u}}\cdot\boldsymbol{v}t + \frac{1}{c}\hat{\boldsymbol{u}}'\cdot\boldsymbol{v}t + \frac{1}{c^2}(\hat{\boldsymbol{u}}\cdot\boldsymbol{v})\left(\hat{\boldsymbol{u}}'\cdot(\boldsymbol{X_r}-\boldsymbol{x}-\boldsymbol{v}t)\right)\right)}. \tag{1.10}$$

Because our computation may only make sense for speeds much smaller than the speed of light, the last term in (1.10) –which is an artefact of the inconsistency of the Galilean transformation in Maxwell's equations– is negligible with respect to all the others. Therefore, the bistatic transmission model is

$$r(t) = \alpha s\left(t - \tau(t)\right), \tag{1.11}$$

where $\tau(t)$ is a *Delay-Doppler* term. Also, note that we inject into α the effect of path loss and antennas gains. In our model, α is a random variable following a Complex Normal distribution $\mathbb{CN}(0,1)$. Using (1.1), $\tau(t)$ is defined by

$$\tau(t) = \frac{1}{c}\left\|\boldsymbol{X_t} - \boldsymbol{x}\right\|_2 + \frac{1}{c}\left\|\boldsymbol{X_r} - \boldsymbol{x}\right\|_2 + \frac{1}{c}\left\langle\left(\frac{\boldsymbol{X_t} - \boldsymbol{x}}{\left\|\boldsymbol{X_t} - \boldsymbol{x}\right\|_2} + \frac{\boldsymbol{X_r} - \boldsymbol{x}}{\left\|\boldsymbol{X_r} - \boldsymbol{x}\right\|_2}\right), \boldsymbol{v}\right\rangle t \tag{1.12}$$

$$\tau(t) = \frac{1}{c}\left(d_b(\boldsymbol{x}) + v_b(\boldsymbol{x}, \boldsymbol{v})t\right), \tag{1.13}$$

where $d_b(\boldsymbol{x})$ is the bistatic range and $v_b(\boldsymbol{x}, \boldsymbol{v})$ is the bistatic speed, respectively described by

$$d_b(\boldsymbol{x}) = \left\|\boldsymbol{X_t} - \boldsymbol{x}\right\|_2 + \left\|\boldsymbol{X_r} - \boldsymbol{x}\right\|_2, \tag{1.14}$$

$$v_b(\boldsymbol{x}, \boldsymbol{v}) = \left\langle\left(\frac{\boldsymbol{X_t} - \boldsymbol{x}}{\left\|\boldsymbol{X_t} - \boldsymbol{x}\right\|_2} + \frac{\boldsymbol{X_r} - \boldsymbol{x}}{\left\|\boldsymbol{X_r} - \boldsymbol{x}\right\|_2}\right), \boldsymbol{v}\right\rangle. \tag{1.15}$$

2 The Mono- and Bistatic Radar Configuration

The expression of the sampled received signal of a bistatic FMCW Radar is derived here for the chirp modulation depicted in Fig. 1.3. We can already note that the resulting expression, at the end of this section, is adaptable to others FMCW modulations such as the triangular introduced before. Notably, it is similar, up to a negligible factor, to the corresponding expression for step-frequency modulation. The transmitted signal has a frequency increasing linearly over time between f_0 and $f_0 + B$ within a duration $M_s T_s$, where B is its bandwidth, T_s is the sampling period of the acquisition, and $M_s \in \mathbb{N}_+$ is the number of samples acquired per *ramp*. There is possibly a break between consecutive ramps, such that the duration between the beginning instants of two consecutive ramps is $T \geq MT_s$.

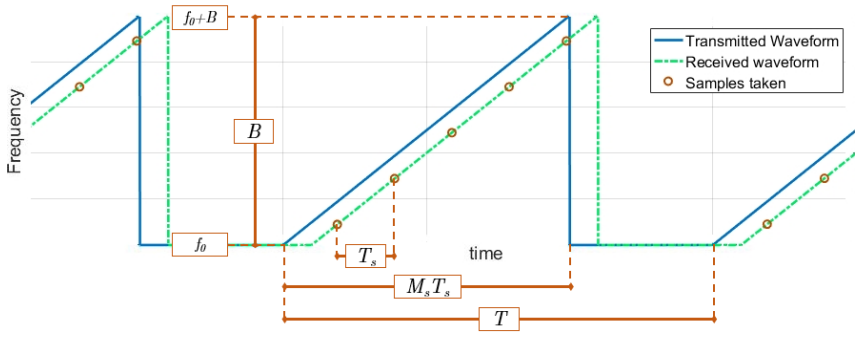


Figure 1.3: Illustration of the *saw-tooth* shaped modulation with timing parameters

Let us define M_r, the number of ramps sent and received during one acquisition process and $\mathcal{T}_{m_r} := [m_r T,\ \ m_r T + M_s T_s]$ with $m_r \in [M_r]$[1], the period of time over which the m_r-th

[1]Note that, this way $t \in [T, M_r T + M_s T_s]$ instead of $t \in [0, (M_r - 1)T + M_s T_s]$, this is an arbitrary choice made for the simplicity of the notations, avoiding the need to write $(m_r - 1)$ in all equations while keeping the consistency of the notation $[M_r] := \{1, .., M_r\}$

ramp is sent. The expression of $f_c(t)$ to insert in the expression (1.3) of the transmitted FMCW signal is

$$f_c(t) = \begin{cases} f_0 + \frac{B}{M_s T_s}(t - m_r T) & \text{if } t \in \mathcal{T}_{m_r} \quad \forall\, m_r \in [M_r]\,, \\ f_0 & \text{else.} \end{cases} \tag{1.16}$$

2 .1 Hypothesis on the System

In addition to the assumptions already stated in Sec. 1 , we need to define hypothesis about the acquisition system.

- **Hypothesis 1 (H1):** The transmitted signal is *narrowband*, i.e., $B \ll f_0$.
- **Hypothesis 2 (H2):** The demodulator is perfect. To put it simply, the actual physical received signal is $\Re(r(t))$. In FMCW Radars, the demodulation is performed by mixing it with the signal currently being transmitted to obtain the baseband equivalent of the received signal, denoted by $e_r(t)$. More precisely, we compute the rice components as $I = \mathrm{LP}\left(\Re(r(t))\Re(s(t))\right)$ and $Q = \mathrm{LP}\left(\Re\{r(t)\}\Im(s(t))\right)$, where LP() denotes a *Low-Pass* operation. Finally, $e_r(t) = I + jQ$. If this operation is performed perfectly (i.e., with no phase shift, no frequency offset and an ideal LP filter), we can write

$$e_r(t) = s^*(t)r(t). \tag{1.17}$$

- **Hypothesis 3 (H3):** All the noise and remaining non-idealities of the whole acquisition process are modeled as an Additive White Gaussian Noise (AWGN) on the sampled signal. If there are multiple receiving antennas, each one is affected by its own AWGN independent from the others.
- **Hypothesis 4 (H4):** The delay caused by the transmission is small with respect to T_s. This way, at the instants the samples are taken, the received ramp is the echo of the same ramp currently transmitted by its transmitter, hence used by the demodulator.
- **Hypothesis 5 (H5):** The received signal resulting from the presence of several targets is the sum of the signals that would be received in the presence of each target if it was alone. This implies that our model does not take target occlusion into account.

2 .2 Received signal model of a bistatic Radar

The samples are only taken a time instants within a chirp. Therefore, Let us compute the received signal's expression when $t \in \mathcal{T}_{m_r}$ for an arbitrary value of $m_r \in [M_r]$. From the modulated frequency expression is given in (1.16), we compute

$$\int_0^t f_c(t')\mathrm{d}t' = f_0 t + \sum_{i=0}^{m_r-1} \int_{iT}^{iT+M_sT_s} \frac{B}{MT_s}(t' - m_rT)\mathrm{d}t' + \int_{m_rT}^{t} \frac{B}{M_sT_s}(t' - m_rT)\mathrm{d}t' \quad (1.18)$$

$$= f_0 t + \frac{BM_sT_s}{2} m_r + \frac{B}{M_sT_s}(t - m_rT)^2. \quad (1.19)$$

Therefore, the transmitted signal is

$$s(t) = \exp\left(j2\pi\left(f_0 t + \frac{BM_sT_s}{2} m_r + \frac{B}{M_sT_s}(t - m_rT)^2\right)\right). \quad (1.20)$$

Because (1.20) is only describing the signal for $t \in \mathcal{T}_{m_r}$, we need to also have $t - \tau(t) \in \mathcal{T}_{m_r}$ to apply the transmission model described in (1.11). Thanks to Hyp. 4, we have $\tau(t) < T_s$, hence for $t \in [m_rT + T_s, m_rT + MT_s]$, we can to write the following:

$$\begin{aligned} r(t) &= \alpha \exp\left(j2\pi\left(f_0(t - \tau(t)) + \frac{BM_sT_s}{2} m_r + \frac{B}{M_sT_s}(t - \tau(t) - m_rT)^2\right)\right) \\ &= \alpha\, s(t)\, \exp\left(-j2\pi\left(f_0\tau(t) + \frac{B}{M_sT_s}\tau(t)(t - m_rT) - \frac{B}{2M_sT_s}\tau^2(t)\right)\right). \end{aligned} \quad (1.21)$$

The perfect coherent demodulation, assumed in Hyp. 2 leads to the baseband signal

$$e_r(t) = \alpha\, \exp\left(-j2\pi\left(f_0\tau(t) + \frac{B}{M_sT_s}\tau(t)(t - m_rT) - \frac{B}{2M_sT_s}\tau^2(t)\right)\right). \quad (1.22)$$

We may replace $\tau(t)$ by its expression (1.13). Finally, the baseband signal is uniformly sampled. The m_s-th sample of the m_r-th ramp is taken at the time instant $t = m_rT + m_sT_s$, with $m_s \in [M_s]$ and $m_r \in [M_r]$, for a total of $M = M_rM_s$ samples. With the addition of an AWGN according to Hyp. 3, the sampled signal in the presence of a single target located in $\boldsymbol{x}$, moving with a constant velocity $\boldsymbol{v}$ and characterized by a scattering coefficient α is

$$\begin{aligned} y[m_s, m_r] = \alpha &\exp\left(-j2\pi\tfrac{f_0}{c} d_b(\boldsymbol{x})\right) \exp\left(-j2\pi\tfrac{1}{c}\left(\tfrac{B}{M_s} d_b(\boldsymbol{x}) + f_0T_s v_b(\boldsymbol{x},\boldsymbol{v})\right) m_s\right) \\ &\exp\left(-j2\pi\tfrac{f_0}{c} T v_b(\boldsymbol{x},\boldsymbol{v}) m_r\right) \exp\left(-j2\pi\tfrac{1}{c}\tfrac{B}{M_s} v_b(\boldsymbol{x},\boldsymbol{v})(m_rT + m_sTs) m_s\right) \\ &\exp\left(j2\pi\tfrac{B}{2M_sT_s}\tfrac{1}{c^2}\left(d_b(\boldsymbol{x}) + v_b(\boldsymbol{x},\boldsymbol{v})(m_rT + m_sT_s)\right)^2\right) \\ &+ e[m_s, m_r], \end{aligned} \quad (1.23)$$

where $e[m_s, m_r]$ stands for the noise. The expression above is composed of three distinguishable parts we identify in the following manner:

- The exponential factor involving only a linear expression of m_s, that we shall call the *inner signal* because it correspond to the inside of a ramp,

$$\begin{aligned} \psi_{\boldsymbol{x},\boldsymbol{v}}[m_s] := &\exp\left(-j2\pi\tfrac{f_0}{c} d_b(\boldsymbol{x})\right) \exp\left(j2\pi\tfrac{B}{2M_sT_s}\tfrac{1}{c^2}(d_b(\boldsymbol{x}))^2\right) \\ &\exp\left(-j2\pi\tfrac{1}{c}\left(\tfrac{B}{M_s} d_b(\boldsymbol{x}) + f_0T_s v_b(\boldsymbol{x},\boldsymbol{v})\right) m_s\right). \end{aligned} \quad (1.24)$$

  Notice that we arbitrarily chose to gather in the inner signal the factors that independent of both m_s and m_r.

- The exponential factor involving only a linear expression of m_r, that we shall call the *outer signal*,

$$\phi_{\boldsymbol{x},\boldsymbol{v}}[m_r] \;:=\; \exp\Big(-j2\pi\tfrac{f_0}{c}Tv_b(\boldsymbol{x},\boldsymbol{v})m_r\Big). \tag{1.25}$$

- The exponential factor involving non-linear expressions of both m_s and m_r, that we shall call the *coupling signal*.

$$\begin{aligned}\theta_{\boldsymbol{x},\boldsymbol{v}}[m_s,m_r] \;:=\; & \exp\Big(-j2\pi\tfrac{1}{c}\tfrac{B}{M_s}v_b(\boldsymbol{x},\boldsymbol{v})(m_rT+m_sTs)m_s\Big)\\ & \exp\Big(j2\pi\tfrac{B}{M_sT_s}\tfrac{1}{c^2}(d_b(\boldsymbol{x})v_b(\boldsymbol{x},\boldsymbol{v})(m_rT+m_sT_s))\Big).\\ & \exp\Big(j2\pi\tfrac{B}{2M_sT_s}\tfrac{1}{c^2}(v_b(\boldsymbol{x},\boldsymbol{v})(m_rT+m_sT_s))^2\Big).\end{aligned} \tag{1.26}$$

In the presence of K distinct targets, such that the k-th target is located in $\boldsymbol{x}_k$, moving with a velocity $\boldsymbol{v}_k$ and characterized by the scattering coefficient α_k, for $k\in[K]$, we build the model by adding the received signals provided by the presence of each target, according to Hyp. 5. Therefore, the complete signal model of a bistatic Radar is

$$y[m_s,m_r]=\sum_{k\in[K]}\alpha_k\psi_{\boldsymbol{x}_k,\boldsymbol{v}_k}[m_s]\phi_{\boldsymbol{x}_k,\boldsymbol{v}_k}[m_r]\theta_{\boldsymbol{x}_k,\boldsymbol{v}_k}[m_s,m_r]+e[m_s,m_r], \tag{1.27}$$

$$y[m_s,m_r]=\sum_{k\in[K]}\alpha_k d_{\boldsymbol{x}_k,\boldsymbol{v}_k}[m_s,m_r]+e[m_s,m_r], \tag{1.28}$$

where we define

$$d_{\boldsymbol{x},\boldsymbol{v}}[m_s,m_r]=\psi_{\boldsymbol{x},\boldsymbol{v}}[m_s]\phi_{\boldsymbol{x},\boldsymbol{v}}[m_r]\theta_{\boldsymbol{x},\boldsymbol{v}}[m_s,m_r], \tag{1.29}$$

for $m_r\in[M_r]$ and $m_s\in[M_s]$. Also, note that α_k are i.i.d for $k\in[K]$.

In appendix B , we demonstrate that, as a very good approximation, the coupling signal $\theta_{\boldsymbol{x},\boldsymbol{v}}[m_s,m_r]=1$ if $\ll f_0$, which is asserted by Hyp. 1. Moreover, we will see in Chap. 3 that because of the properties of the outer signal $\phi_{\boldsymbol{x},\boldsymbol{v}}[m_r]$, we generally work with targets such that $|v_b(\boldsymbol{x}_k,\boldsymbol{v}_k)|\leq\frac{c}{2f_0T}$. As also demonstrated in appendix B , this constraint on the value of the velocity allows us to neglect the effect of the velocity in the inner signal, i.e.,

$$\begin{aligned}\psi_{\boldsymbol{x},\boldsymbol{v}}[m_s]\simeq\psi_{\boldsymbol{x}}[m_s] := & \exp\Big(-j2\pi\tfrac{f_0}{c}d_b(\boldsymbol{x})\Big)\exp\Big(j2\pi\tfrac{B}{2M_sT_s}\tfrac{1}{c^2}(d_b(\boldsymbol{x}))^2\Big)\\ & \exp\Big(-j2\pi\tfrac{1}{c}\tfrac{B}{M_s}d_b(\boldsymbol{x})m_s\Big).\end{aligned} \tag{1.30}$$

In practice this approximation corresponds to neglect the Doppler effect inside the ramps, and only consider its effects between the ramps, i.e., in the outer signal. The simplified signal model of a bistatic Radar is

$$y[m_s,m_r]\simeq\sum_{k\in[K]}\alpha_k\psi_{\boldsymbol{x}_k}[m_s]\phi_{\boldsymbol{x}_k,\boldsymbol{v}_k}[m_r]+e[m_s,m_r], \tag{1.31}$$

2 .3 Sparse modeling

The formulation of a sparse model in the Radar context is typically performed by determining a set of target states, i.e., a set of N location-velocity couples $(\boldsymbol{x}_n, \boldsymbol{v}_n) \in \mathbb{R}^4$, $n \in [N]$, the targets may be defined with. The definition of the target states leads to the definition of a discretized subset of $\mathbb{R}^4$ we call a *grid* and denote by Ω,

$$\Omega := \{(\boldsymbol{x}_1, \boldsymbol{v}_1), .., (\boldsymbol{x}_N, \boldsymbol{v}_N)\}. \tag{1.32}$$

Each target exactly *On-Grid*, i.e., such that $(\boldsymbol{x}_k, \boldsymbol{v}_k) \in \Omega$, for all $k \in [K]$. Moreover, from a physical point of view, each bin of Ω is only assigned by, at most, one target. The n-th target state of the grid Ω is now referred to as $\Omega\{n\}$.

Let us define the measurement vector $\boldsymbol{y} := (y[1,1], .., y[1, M_r], y[2,1].., y[M_s, M_r])^T \in \mathbb{C}^M$, with $M = M_r M_s$ and where $y[m_s, m_r]$ is as in (1.28). Because we assume only *On-Grid* targets, we can define a sparse vector $\boldsymbol{s} := (s[1], .., s[N]))^T \in \mathbb{C}^N$ such that

$$s[n] = \begin{cases} \alpha_k & \text{if the } k\text{-th target is such that } (\boldsymbol{x}_k, \boldsymbol{v}_k) = \Omega\{n\}, \\ 0 & \text{else,} \end{cases} \tag{1.33}$$

and a dictionary $\boldsymbol{D} \in \mathbb{C}^{M \times N}$ whose n-th atom is $\boldsymbol{d}_n := (d_{\Omega\{n\}}[1,1], .., d_{\Omega\{n\}}[1, M_r], d_{\Omega\{n\}}[2,1].., d_{\Omega\{n\}}[M_s, M_r])^T \in \mathbb{C}^M$ where $d_{\boldsymbol{x},\boldsymbol{v}}[m_s, m_r]$ is defined as in (1.29), with the subscript n now corresponding to the index of the n-th target state in Ω_d. The sparse model therefore reads

$$\boldsymbol{y} = \boldsymbol{D}\boldsymbol{s} + \boldsymbol{e}, \tag{1.34}$$

where $\boldsymbol{e} := (e[1,1], .., e[1, M_r], .., e[M_s, M_r]$ is the AWGN vector.

2 .4 Factored sparse modeling for bistatic Radars

The sparse model in (1.34) is a naive and general formulation that does not take account of the properties of the atoms in the dictionary $\boldsymbol{D}$. Analyzing these properties may enable the definition of an efficient grid Ω, possibly leading to a factorable dictionary. To reach this purpose, let us start from the simplified signal model proposed in (1.31). When we inspect the the definitions of the both the outer and the simplified inner signals given in (1.30) and (1.25) respectively, we notice that the target's location only affects the received signal via the *bistatic range*. This matches the statement established in the introduction, i.e., that bistatic Radars may only determine the ellipses on which the targets are located. It goes the same for the velocity, only affecting the received signal via the *bistatic speed*.

Therefore, an efficient definition for a set of target states is the following. We define N_1 possible values of bistatic ranges d_{b,n_1} and N_2 possible values of bistatic speed v_{b,n_2}, leading to the definitions of a bistatic range grid Ω_{db} and a bistatic speed grid Ω_{vb},

$$\Omega_{db} := \{d_{b,1}, .., d_{b,N_1}\}, \tag{1.35}$$

$$\Omega_{vb} := \{v_{b,1}, .., v_{b,N_2}\}. \tag{1.36}$$

With this definition of target states, the target parameters of interest are no longer the 2D location vector $\boldsymbol{x}$ and the 2D velocity vector $\boldsymbol{v}$, but the scalar bistatic range and bistatic speed. Therefore, we notice that the simplified model is fully decoupled in the sense that the simplified inner signal is exclusively linking the sample index m_s to the bistatic range, and the outer signal is exclusively linking the ramp index m_r to the bistatic speed. Let us define a measurement matrix $\boldsymbol{Y} := \mathrm{Mat}_{M_r}(\boldsymbol{y}) \in \mathbb{C}^{M_s \times M_r}$, i.e., such that the element in its m_s-th row and m_r-th column is $y[m_s, m_r]$. Let us also define a sparse matrix $\boldsymbol{S} \in \mathbb{C}^{N_1 \times N_2}$ such that the element in its n_1 row and n_2-th column is α_k if $d_b(\boldsymbol{x}_k) = \Omega_{db}\{n_1\}$ and $v_b(\boldsymbol{x}_k, \boldsymbol{v}_k) = \Omega_{vb}\{n_2\}$.

The decoupling mentioned above suggests the definition of two dictionaries: the inner dictionary $\boldsymbol{\Psi} \in \mathbb{C}^{M_s \times N_1}$ whose n_1-th atom is $\psi_{\Omega_{db}\{n_1\}}[m_s]$, and the outer dictionary $\boldsymbol{\Phi} \in \mathbb{C}^{M_r \times N_2}$ whose n_2-th atom is $\phi_{\Omega_{vb}\{n_2\}}[m_r]$. A factored sparse representation is directly emerging from these definitions,

$$\boldsymbol{Y} \simeq \boldsymbol{\Psi}\boldsymbol{S}\boldsymbol{\Phi}^T + \boldsymbol{E}, \tag{1.37}$$

where $\boldsymbol{E} := \mathrm{Mat}_{M_r}(\boldsymbol{e}) \in \mathbb{C}^{M_s \times M_r}$ is the noise vector $\boldsymbol{e}$ of (1.34) reshaped in a matrix form. Finally, it is important to note that choosing an appriopriate uniformly spaced set of bistatic ranges and an appriopriate uniformly spaced set of bistatic speed yield to a 2D Discrete Fourrier Transform (DFT2) dictionary,

$$\boldsymbol{Y} \simeq \mathrm{DFT2}(\boldsymbol{S}) + \boldsymbol{E}. \tag{1.38}$$

For instance, this formulation appears when we define the following grid:

$$\begin{aligned} \Omega_{db}\{n_1\} &= n_1 \frac{c}{B} && n_1 \in [M_s] \\ \Omega_{vb}\{n_2\} &= \left(n_2 - \frac{M_r}{2}\right) \frac{c}{f_0 T M_r} && n_2 \in [M_r] \end{aligned} \tag{1.39}$$

The DFT2 relation between the target states, somehow contained in $\boldsymbol{S}$ and the measurement $\boldsymbol{Y}$ is a famous result in the field of monostatic and bistatic FMCW Radars. Still, the detailed step-by-step approach we proposed to build this representation is instrumental for a proper extension to the definition of a joint acquisition modeling for multistatic Radars. Indeed, we saw that in order to reach the simple definition, there are several assumptions and choices (i.e., of specific grid) that are not all directly translatable to the multi-static Radar systems.

3 The distributed Multi-static Radar Configuration

The multistatic Radar system was depicted in fig 5.c, where multiple widely separated nodes, each composed of a single antenna, communicate with one another. There are Q_t transmitting nodes and Q_r receiving nodes. Multistatic Radar systems provide multiple signals observing the same scene. As detailed in the hypothesis description proposed below, we receive Q different signals where Q is not necessarily equal to $Q_t Q_r$.

3 .1 Additional hypothesis on the system

If several transmitters send exactly the same frequency at the same time, a receiver listening to this frequency will acquire the sum of the signals it would acquire if each transmitter was alone. This is asserted by an extension of Hyp. 5 to the presence of several transmitters. In opposition to this, we assume that each RX node is able to determine which signal is provided by which TX node. In practice, the TXs need to send different signals. Fig. 1.4 and Fig 1.5 show two examples of possible transmitted signals, each having its limitations. A RX may be programmed to listen to either all TXs or only some of them, which is why we can have $Q < Q_t Q_r$. Although fundamental, the topic of the physical implementation of a multistatic Radar system is very wide and goes beyond the scope of this document focused on the signal processing. Therefore, as a strong abstraction regarding this implementation, we assume the receiving of $Q \leq Q_t Q_r$ perfectly synchronized signals, each resulting from a connected TX-RX pair of nodes that we call a *bistatic pair*,

$$y^q[m] \quad \forall\, q \in [Q]\,. \tag{1.40}$$

In other words, the signal obtained from each bistatic pair is not influenced by the presence of the others bistatic pairs. For the simplicity of notations, we assume in that each receiver samples the signal the same way, i.e., with the same number of samples M taken a same time instants.

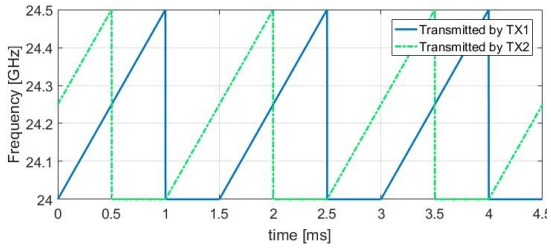


Figure 1.4: Illustration of the signals transmitted by two transmitters. The transmitted waveforms are shifted in order to make them distinguishable by a receiver. Requires an excellent synchronization between the receivers.

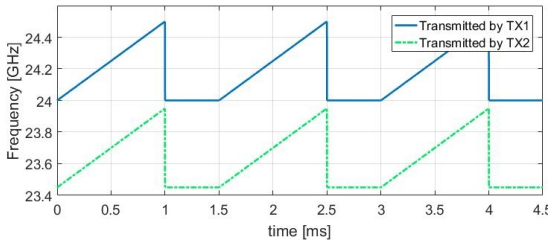


Figure 1.5: Illustration of the signals transmitted by two transmitters in a Frequency-Division Multiple Access fashion in order to make them distinguishable by a receiver. Requires a large Bandwidth to split.

3 .2 Received signals model of a multi-Static Radar

Let us assume K targets, such that the k-th target is located in $\boldsymbol{x}_k$, moving with a velocity $\boldsymbol{v}_k$ and is characterized by the scattering coefficients α_k^q, $q \in [Q]$, for all $k \in [K]$. Thanks to the above hypothesis, we can extend the complete received signal model of a bistatic Radar, given in (1.28), to obtain the complete received signals model of a multi-static Radar,

$$y^q[m_s, m_r] = \sum_{k \in [K]} \alpha_k^q d^q_{\boldsymbol{x}_k, \boldsymbol{v}_k}[m_s, m_r] + e^q[m_s, m_r], \tag{1.41}$$

where,

$$d^q_{\boldsymbol{x},\boldsymbol{v}}[m_s, m_r] = \psi^q_{\boldsymbol{x},\boldsymbol{v}}[m_s]\phi^q_{\boldsymbol{x},\boldsymbol{v}}[m_r]\theta^q_{\boldsymbol{x},\boldsymbol{v}}[m_s, m_r], \tag{1.42}$$

and $\psi^q_{\boldsymbol{x},\boldsymbol{v}}[m_s]$, $\phi^q_{\boldsymbol{x},\boldsymbol{v}}[m_r]$, $\theta^q_{\boldsymbol{x},\boldsymbol{v}}[m_s, m_r]$ are respectively the inner, outer and coupling signals, of the q-th pair. That is, defined as in (1.24), (1.25) and (1.26) where the bistatic range and bistatic speed are now defined with respect to the bistatic pair they are referring to, i.e.,

$$d^q_b(\boldsymbol{x}) = \|\boldsymbol{X_t}^q - \boldsymbol{x}\|_2 + \|\boldsymbol{X_r}^q - \boldsymbol{x}\|_2 \,, \tag{1.43}$$

$$v^q_b(\boldsymbol{x}, \boldsymbol{v}) = \left\langle \left(\frac{\boldsymbol{X_t}^q - \boldsymbol{x}}{\|\boldsymbol{X_t}^q - \boldsymbol{x}\|_2} + \frac{\boldsymbol{X_r}^q - \boldsymbol{x}}{\|\boldsymbol{X_r}^q - \boldsymbol{x}\|_2} \right), \boldsymbol{v} \right\rangle, \tag{1.44}$$

where $\boldsymbol{X_t}^q$ and $\boldsymbol{X_r}^q$ respectively stand for the location of the TX node and the RX node of the q-th bistatic pair. The simplified received signal model of a bistatic Radar, given in (1.31), is also extended to get the simplified received signals model of a multi-static Radar,

$$y^q[m_s, m_r] \simeq \sum_{k \in [K]} \alpha_k \psi^q_{\boldsymbol{x}_k}[m_s] \phi^q_{\boldsymbol{x}_k, \boldsymbol{v}_k}[m_r] + e^q[m_s, m_r], \tag{1.45}$$

and $\psi^q_{\boldsymbol{x}}[m_s]$ is the simplified inner signal of the q-th pair. That is, defined as in (1.30) where the bistatic range and bistatic speed are now defined with respect to the bistatic pair they are referring to, i.e., see (1.43) and (1.44).

3 .3 Joint sparse modeling

With the above definition of the received signals, we could consider each signal as an independent bistatic measurement. Therefore, the definition of the sparse model described in Sec. 2 .3 and the simplifications of Sec. 2 .4 hold individually for each signal. This way, we can model each signal as the DFT2 of a sparse vector, as we did in 1.38. However, in such a model, the sparse vectors are defined in sparsity domains that are not referring to the same definition of target states, hence not consistent with one another. This totally excludes the possibility of a joint acquisition model, further enabling a joint target states estimation.

To establish a joint acquisition model, we need to define a common sparsity domain. In this purpose, let us select N_x possible locations, named $\boldsymbol{x}_{n_x}$ with $n_x \in [N_x]$, and N_v possible velocities named $\boldsymbol{v}_{n_v}$ with $n_v \in [N_v]$, hence defining $N = N_x N_v$ target states. In other terms, we define a location grid denoted by Ω_X and a velocity grid denoted by Ω_V,

$$\Omega_X := \{\boldsymbol{x}_1, .., \boldsymbol{x}_{N_x}\} \,, \tag{1.46}$$

$$\Omega_V := \{\boldsymbol{v}_1, .., \boldsymbol{v}_{N_v}\} \,. \tag{1.47}$$

As we did before, we assume that the targets are always *On-Grid*, i.e., located in $\boldsymbol{x} \in \Omega_X$ and moving with a velocity $\boldsymbol{v} \in \Omega_V$. A simple example of such grids is provided in Fig. 1.6. For the received signal of each bistatic pair, we can define a dictionary, similarly to what we did in Sec. 2 .3. Therefore, we have Q dictionaries, $\boldsymbol{D}^q \in \mathbb{C}^{M \times N}$, for all $q \in [Q]$. The q-th dictionnary's atom corresponding to the n_x-th location and n_v-th velocity is $\boldsymbol{d}^q_{n_x n_v}$, defined as

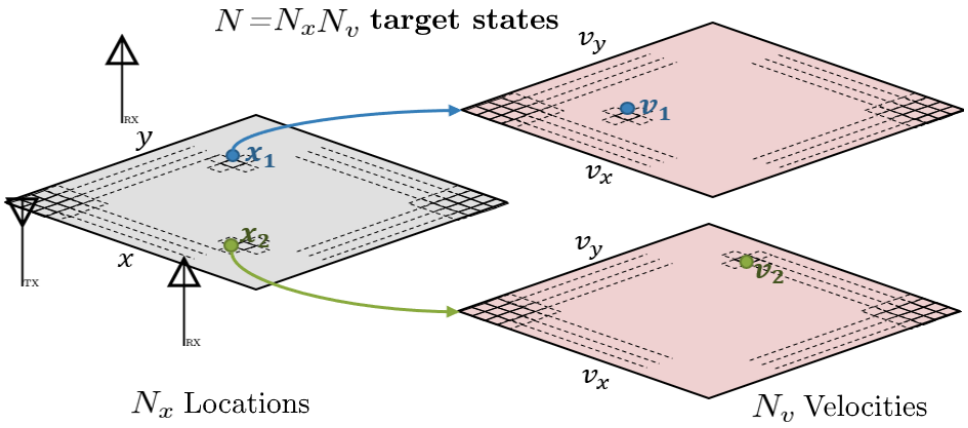


Figure 1.6: Schematic illustration of the definition of N_x possible locations and N_v possible velocities. In this simple example, rectangular subset of the 2D planes of, respectively, location and velocities are sampled uniformly.

$$\begin{aligned}\boldsymbol{d}^q_{n_x n_v} := \Big(&d^q_{\Omega_X\{n_x\},\Omega_V\{n_v\}}[1,1],..,d^q_{\Omega_X\{n_x\},\Omega_V\{n_v\}}[1,M_r],d^q_{\Omega_X\{n_x\},\Omega_V\{n_v\}}[2,1],..,\\ &..,d^q_{\Omega_X\{n_x\},\Omega_V\{n_v\}}[M_s,M_r]\Big)^T,\end{aligned} \tag{1.48}$$

These dictionaries enable the description of a joint sparse representation of the measurement vectors $\boldsymbol{y}^q \in \mathbb{C}^M$, for all $q \in [Q]$. We define the sparse vectors $\boldsymbol{s}^q := (s^q[1,1],..,s^q[1,N_v],\ \ s^q[2,1],..,s^q[N_x,N_v])^T$ for all $q \in [Q]$ such that

$$s^q[n_x,n_v] = \begin{cases} \alpha^q_k & \text{if the } k\text{-th target is located in } \boldsymbol{x}_k = \Omega_X\{n_x\} \\ & \text{and is moving with a velocity } \boldsymbol{v}_k = \Omega_V\{n_v\}. \\ 0 & \text{else.}\end{cases} \tag{1.49}$$

These sparse vectors have a common sparsity pattern as shown in Fig. 1.7, i.e., $\text{Supp}\,(\boldsymbol{s}^q) = \text{Supp}\left(\boldsymbol{s}^{q'}\right)$ for all $q,q' \in [Q]$ but the value of their corresponding non-zero elements are different because $\alpha^q_k \overset{\text{i.i.d.}}{\sim} \mathbb{C}\mathcal{N}(0,1)$ for all $k \in [K]$ and all $q \in [Q]$ as stated in the introduction. The joint sparse model reads,

$$\boldsymbol{y}^q = \boldsymbol{D}^q \boldsymbol{s}^q + \boldsymbol{e}^q, \tag{1.50}$$

where $\boldsymbol{e}^q := (e^q[1,1],..,e^q[1,M_r],e^q[2,1],..,e^q[M_s,M_r])^T$ is the AWGN vector associated to the q-th signal.

If now, we observe the simplified received signals model given in (1.45), we notice that the use of shared grids hinders the decoupling we had for bistatic Radars. More precisely, while the simplified inner signals exclusively link the target location to the sample index m_s, the outer signals link both the target location and the velocity to the ramp index m_r. However, this kind of half decoupling already enables the derivation of what we will call in Chap. 2, a *half-factored* sparse model.

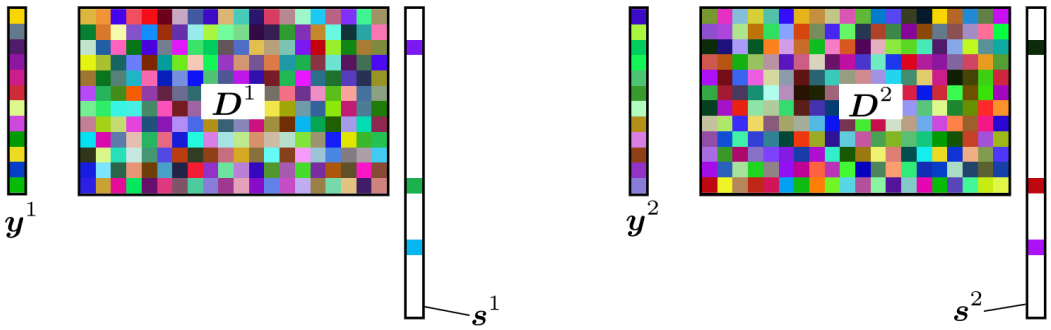


Figure 1.7: Description of two received signals having a joint sparse representation, i.e., such that $\boldsymbol{s}^1$ and $\boldsymbol{s}^2$ share a common sparsity pattern.

Similarly to what we did in Sec. 2 .4, let us define, for all $q \in [Q]$, the q-th measurement matrix $\boldsymbol{Y}^q \in \mathbb{C}^{M_s \times M_r} := \mathrm{Mat}_{M_r}(\boldsymbol{y}^q)$ and the q-th sparse matrix $\boldsymbol{S}^q \in \mathbb{C}^{N_x \times N_v}$ such that its element in its n_x-th row and n_v-th column $s^q[n_x, n_v]$. Also, for all $q \in [Q]$, let us define the following dictionaries: the q-th inner dictionary $\boldsymbol{\Psi}^q \in \mathbb{C}^{M_s \times N_x}$ whose n_x-th atom is $\psi_{\Omega_X\{n_x\}}[m_s]$, and for all $n_x \in [N_x]$, the q-th the outer dictionary for a fixed value of n_x, $\boldsymbol{\Phi}^q_{n_x} \in \mathbb{C}^{M_r \times N_v}$ whose n_v-th atom is $\phi_{\Omega_X\{n_x\},\Omega_V\{n_v\}}[m_r]$. Therefore, the factored sparse representation for multi-static Radars is

$$\begin{aligned} \boldsymbol{Y}^q &\simeq \boldsymbol{\Psi}^q \boldsymbol{P}^q + \boldsymbol{E}^q, \\ (\boldsymbol{P}^q)^T &= \left[\boldsymbol{\Phi}^q_1 \left[(\boldsymbol{S}^q)^T \right]_1 \cdots \boldsymbol{\Phi}^q_{n_x} \left[(\boldsymbol{S}^q)^T \right]_{n_x} \cdots \boldsymbol{\Phi}^q_{N_x} \left[(\boldsymbol{S}^q)^T \right]_{N_x} \right], \end{aligned} \tag{1.51}$$

where $\left[(\boldsymbol{S}^q)^T \right]_{n_x}$ is the n_x-th column of $(\boldsymbol{S}^q)^T$.

3 .4 Multi-static MIMO sparse modeling

The multi-static MIMO system was depicted in Fig. 5.d, where we assume that there are Q bistatic pairs. In the q-th pair, the transmitter is an antenna array composed of N^q_T antennas and the receiver is an antenna array composed of N^q_R antennas. Therefore the q-th bistatic pair provides N^q_R signals, each being the sum of N^q_T signals. If $\boldsymbol{y}^{q,r}$ denotes the measurement vector provided by the r-th RX antenna of the q-th bistatic pair and $\boldsymbol{y}^{q,t,r}$ is the measurement vector that the r-th RX antenna would provide if only the t-th TX antenna was emitting,

$$\boldsymbol{y}^{q,r} = \sum_{t=1}^{N^q_T} \boldsymbol{y}^{q,t,r} \quad \forall\, r \in [N^q_R]\,. \tag{1.52}$$

If we considered the q-th pair of nodes as a local multi-static system, (1.52) would not match the multi-static assumption according to which a receiver is able to separate the contribution of each transmitter. However, in accordance with the definition of a MIMO pair provided in the introduction, all the signals provided by the same pair of nodes are affected by the same scattering coefficient. Therefore, when expressed in the shared sparsity domain specified in Sec. 3 .3, they are all associated to a unique sparse vector $\boldsymbol{s}^q \in \mathbb{C}^N$. This implies the possibility to gather their dictionaries into a unique

sparse representation. More precisely, if we define $\boldsymbol{D}^{q,t,r}$, the dictionary linking $\boldsymbol{y}^{q,t,r}$ to $\boldsymbol{s}^q$, we have,

$$\boldsymbol{y}^{q,r} = \sum_{t=1}^{N_T^q} \boldsymbol{y}^{q,t,r} = \sum_{t=1}^{N_T^q} \boldsymbol{D}^{q,t,r} \boldsymbol{s}^q + \boldsymbol{e}^{q,r} \quad \forall\, r \in [N_R^q]\,, \quad \forall\, q \in [Q]\,, \tag{1.53}$$

where $\boldsymbol{e}^{q,r}$ is the additive white Gaussian noise originating from the r-th RX antenna of the q-th bistatic pair. The model in (1.53) can be reformulated as

$$\boldsymbol{y}^q = \boldsymbol{D}^q \boldsymbol{s}^q + \boldsymbol{e}^q \quad \forall\, q \in [Q]\,, \tag{1.54}$$

where,

$$\begin{aligned} \boldsymbol{y}^q &= \begin{bmatrix} \boldsymbol{y}^{q,1} \\ \vdots \\ \boldsymbol{y}^{q,N_R^q} \end{bmatrix} \in \mathbb{C}^{N_R^q M}, \\ \boldsymbol{D}^q &= \begin{bmatrix} \sum_{t=1}^{N_T^q} \boldsymbol{D}^{q,t,1} \\ \vdots \\ \sum_{t=1}^{N_T^q} \boldsymbol{D}^{q,t,N_R^q} \end{bmatrix} \in \mathbb{C}^{N_R^q M \times N}, \\ \boldsymbol{e}^q &= \begin{bmatrix} \boldsymbol{e}^{q,1} \\ \vdots \\ \boldsymbol{e}^{q,N_R^q} \end{bmatrix} \in \mathbb{C}^{N_R^q M}. \end{aligned} \tag{1.55}$$

Finally, in a similar manner, we can adapt the factored model given in (1.51). Let $\boldsymbol{\Psi}^{q,t,r}$ and $\boldsymbol{\Phi}_{n_x}^{q,t,r}$ be, respectively, the inner dictionary and the outer dictionary (for a fixed location indexed as $n_x \in [N_x]$) of the signal corresponding to the r-th receiving antenna and the t-th transmitting antenna of the q-th bistatic pair, $r \in [N_R^q]\,, t \in [N_T^q]\,, q \in [Q]$. We assume that $\boldsymbol{\Phi}_{n_x}^{q,t,r} = \boldsymbol{\Phi}_{n_x}^q$ for all $r \in [N_R^q]\,, t \in [N_T^q]$. Therefore, the simplified sparse modeling of the multi-static MIMO signals is

$$\begin{aligned} \boldsymbol{Y}^q &= \boldsymbol{\Psi}^q \boldsymbol{P}^q + \boldsymbol{E}^q, \\ (\boldsymbol{P}^q)^T &= \left[\boldsymbol{\Phi}_1^q \left[(\boldsymbol{S}^q)^T \right]_1 \cdots \boldsymbol{\Phi}_{n_x}^q \left[(\boldsymbol{S}^q)^T \right]_{n_x} \cdots \boldsymbol{\Phi}_{N_x}^q \left[(\boldsymbol{S}^q)^T \right]_{N_x} \right], \end{aligned} \tag{1.56}$$

where

$$\begin{aligned} \boldsymbol{Y}^q &= \begin{bmatrix} \boldsymbol{Y}^{q,1} \\ \vdots \\ \boldsymbol{Y}^{q,N_R^q} \end{bmatrix} \in \mathbb{C}^{N_R^q M_s \times M_r}, \\ \boldsymbol{\Psi}^q &= \begin{bmatrix} \sum_{t=1}^{N_T^q} \boldsymbol{\Psi}^{q,t,1} \\ \vdots \\ \sum_{t=1}^{N_T^q} \boldsymbol{\Psi}^{q,t,N_R^q} \end{bmatrix} \in \mathbb{C}^{N_R^q M_s \times N_x}, \\ \boldsymbol{E}^q &= \begin{bmatrix} \boldsymbol{E}^{q,1} \\ \vdots \\ \boldsymbol{E}^{q,N_R^q} \end{bmatrix} \in \mathbb{C}^{N_R^q M_s \times M_r}. \end{aligned} \tag{1.57}$$

Chapter 2

Multi-Signal and Multi-Dimensional Sparse Recovery Algorithms

In this chapter, the reconstruction problem of a sparse vector from observed data is first formulated, in a generic way, for the simple case of a single measurement vector. A well-known issue of this formal reconstruction problem is the ill-posed problem it describes, hence requiring NP-Hard algorithms [16]. Alternative approximated formulations overcome this challenge by describing convex optimization problems. Some of these formulations and associated algorithms are first briefly overviewed in this chapter, with emphasis given to greedy algorithms (i.e., iterative algorithms making the locally optimal choice at each step). Greedy algorithms have the significant advantage of being both easy to implement and often faster than the other sparse recovery algorithms. However, the prior knowledge of the sparsity level (i.e., the number of non-zero components in the sparse vector to be reconstructed) denoted by K, is required. The most widely known greedy algorithm for the reconstruction of sparse vectors is the Matching Pursuit (MP) [17], which has lead to the development of a large variety of enhanced versions of itself, such as the Orthogonal Matching Pursuit (OMP).

Next, we study the case of multiple measurements vectors corresponding to multiple sparse vectors sharing joint sparsity pattern. A suited existing extension of the MP algorithm for this context is the Block Matching Pursuit (BMP) [14], which will be described as well.

Although these greedy algorithms are known to be fast, both MP and BMP, as well as their *orthogonal* extensions (OMP and BOMP) may still require high computational complexity if their dictionaries grows extremely large. Such a situation happens, for instance in radar applications, if the sparsity domain is the 2D×2D targets' location-velocity domain uniformly sampled. In this case, each times the grid density is doubled, the size of the dictionary is multiplied by 16. Some dictionaries offer the possibility to express the sparse representation in a factored way. This is the case, for instance, for the multidimensional Fourier transform. As it will be detailed further in this chapter, the simple application of MP and BMP algorithms to problems modeled in a factored fashion already has a much lower algorithmic complexity than their application to non-factored representations. However, the main purpose of this chapter is to introduce a further

improvement of these algorithms allowing to further reduce the computational complexity. This new algorithm is named *Factored Matching Pursuit* (FMP). FMP takes advantage of the factored modeling by recovering each dimension in a recursive manner, hence sparing a lot a computation for highly redundant dictionaries. Still, as discussed at the end of this chapter, this improvement is also shown to be at the price of a slightly lower robustness against noise.

1 Formulation of the Sparse Recovery Problem

Let us consider a vector $\boldsymbol{y} \in \mathbb{C}^M$ resulting from the observation of a signal assumed to have a sparse representation $\boldsymbol{s} \in \mathbb{C}^N$ with a sparsity level equal to K, i.e., with K non-zero components such that $K \ll M$, The sparse vector $\boldsymbol{s}$ is linked to $\boldsymbol{y}$ by the dictionary $\boldsymbol{\Psi} \in \mathbb{C}^{M \times N}$, i.e., with possibly $N \gg M$,

$$\boldsymbol{y} = \boldsymbol{\Psi}\boldsymbol{s}. \tag{2.1}$$

Note that, to ensure the convergence of the algorithms described further, the columns of $\boldsymbol{\Psi}$, i.e., its atoms should be normalized. In order to ease further highlight of properties resulting from the application of this chapter to Radars, we chose the following normalization throughout this whole report,

$$\|\boldsymbol{\psi}_n\|_2^2 = M, \tag{2.2}$$

where $\boldsymbol{\psi}_n \in \mathbb{C}^M$ is the n-th atom (column) of $\boldsymbol{\Psi}$. This measurement is possibly affected by an additive noise, leading to

$$\boldsymbol{y} = \boldsymbol{\Psi}\boldsymbol{s} + \boldsymbol{e}, \tag{2.3}$$

where $\boldsymbol{e} \in \mathbb{C}^M$ is an AWGN vector whose variance of each component is σ_e^2. The global problem is formulated as the recovery of the vector $\boldsymbol{s}$ from the knowledge of the measurement $\boldsymbol{y}$. If the dictionary is orthogonal (for which $N = M$ is a necessary condition), a simple way to proceed is, strictly for (2.1), or in the least squared error sense for (2.3), by inverting the matrix $\boldsymbol{\Psi}$, which corresponds to make correlations with the dictionary's atoms. More precisely, $\hat{\boldsymbol{s}} = \boldsymbol{\Psi}^H\boldsymbol{y}$. However, when the dictionary is not orthogonal (typically, when $N \gg M$), correlations no longer work and algorithms for sparse signals recovery are required. The main underlying idea of these algorithms is the search for the *sparsest* of all the possible solutions. Indeed, if $N > M$, there are an infinity of possible vectors $\boldsymbol{s}$ such that $\boldsymbol{y} = \boldsymbol{\Psi}\boldsymbol{s}$. This problem is stated as

$$\hat{\boldsymbol{s}} = \underset{\boldsymbol{s}}{\operatorname{argmin}} \ \|\boldsymbol{s}\|_0 \ \text{ s.t. } \ \|\boldsymbol{y} - \boldsymbol{\Psi}\boldsymbol{s}\|_2^2 < \sigma_e^2. \tag{2.4}$$

As already explained, this formulation defines an non-convex optimization problem that is NP-Hard [16]. A classical reformulation consists of relaxing the ℓ_0 norm, and use an ℓ_1 norm instead [18],

$$\hat{\boldsymbol{s}} = \underset{\boldsymbol{s}}{\operatorname{argmin}} \ \|\boldsymbol{s}\|_1 \ \text{ s.t. } \ \|\boldsymbol{y} - \boldsymbol{\Psi}\boldsymbol{s}\|_2^2 < \sigma_e^2. \tag{2.5}$$

The algorithm solving this convex optimization problem is known as the Basis Pursuit (BP). With respect to Matching Pursuit introduced earlier, this algorithm does not require the prior knowledge of the sparsity level K of $\boldsymbol{s}$. However, MP-like algorithms are often faster. These greedy algorithms are based on the following alternative formulation.

$$\hat{\boldsymbol{s}} = \underset{\boldsymbol{s}}{\operatorname{argmin}} \ \|\boldsymbol{y} - \boldsymbol{\Psi}\boldsymbol{s}\|_2^2 \ \text{ s.t. } \ \|\boldsymbol{s}\|_0 \leq K. \tag{2.6}$$

In AWGN conditions (i.e., $e[m] \overset{\text{i.i.d.}}{\sim} \mathbb{C}\mathcal{N}(0, \sigma_e^2)$ for $m \in [M]$), the optimization problem described in (2.6) can be obtained from the derivation of the Maximum Likelihood Estimator (MLE) as follow. Let us define the space of K-sparse vectors with length N: $\Sigma_K^N := \left\{\boldsymbol{s} \in \mathbb{C}^N : \|\boldsymbol{s}\|_0 \leq K\right\}$, then,

$$\begin{aligned} \hat{\boldsymbol{s}} &= \underset{\boldsymbol{s}\in\boldsymbol{\Sigma}_{\boldsymbol{K}}^{\boldsymbol{N}}}{\operatorname{argmax}} \ \mathrm{P}(\boldsymbol{y}|\boldsymbol{s}) \\ &= \underset{\boldsymbol{s}\in\boldsymbol{\Sigma}_{\boldsymbol{K}}^{\boldsymbol{N}}}{\operatorname{argmax}} \ \prod_m \frac{1}{\sqrt{2\pi\sigma_e^2}} \exp\left(-\frac{|y[m] - \sum_n \psi_n[m]s[n]|^2}{2\sigma_e^2}\right), \end{aligned} \tag{2.7}$$

where $\mathrm{P}(\boldsymbol{y}|\boldsymbol{s})$ stands for the conditional probability of measuring $\boldsymbol{y}$ if its sparse representation is $\boldsymbol{s}$. Computing the Log-likelihood,

$$\begin{aligned} \hat{\boldsymbol{s}} &= \underset{\boldsymbol{s}\in\boldsymbol{\Sigma}_{\boldsymbol{K}}^{\boldsymbol{N}}}{\operatorname{argmin}} \ \sum_{m\in[M]} \left| y[m] - \sum_n \psi_n[m]s[n] \right|^2 \\ &= \underset{\boldsymbol{s}\in\boldsymbol{\Sigma}_{\boldsymbol{K}}^{\boldsymbol{N}}}{\operatorname{argmin}} \ \|\boldsymbol{y} - \boldsymbol{\Psi}\boldsymbol{s}\|_2^2 \\ &= \underset{\boldsymbol{s}}{\operatorname{argmin}} \ \|\boldsymbol{y} - \boldsymbol{\Psi}\boldsymbol{s}\|_2^2 \ \text{ s.t. } \ \|\boldsymbol{s}\|_0 \leq K. \end{aligned} \tag{2.8}$$

2 Two Useful Existing Greedy Algorithms

In this section, the MP algorithm is described, as well as its extension to block-sparse signals (BMP), which is notably useful to recover joint sparse signals observed through multiple measurement vectors. These descriptions enable the investigation of MP-like algorithms adapted to factored sparse representations, studied in the next section.

2 .1 Matching Pursuit

The MP algorithm takes its inspiration from (2.6), and its purpose is therefore to find the K atoms of the dictionary (and their associated values in $\boldsymbol{s}$) such that the error between the reconstructed signal $\boldsymbol{\Psi}\boldsymbol{s}$ and the actual signal $\boldsymbol{y}$ is minimized. The underlying idea to approximately, but quickly, solve this optimization problem is to iteratively find the atom that has the highest matching with $\boldsymbol{y}$, and subsequently remove the contribution of this atom to the signal. In other terms, at each iteration, the algorithm solves the sub-problem of finding the 1-sparse signal minimizing the approximation error. Let us denote by $\boldsymbol{r}^{(k)} \in \mathbb{C}^M$ the residue signal at the beginning of the k-th iteration and such that $\boldsymbol{r}^{(0)} = \boldsymbol{y}$. The vector $\hat{\boldsymbol{s}}^{(k)} \in \mathbb{C}^N$ is the current estimate of the sparse vector $\boldsymbol{s}$ and is filled up throughout the iterations, starting at $\hat{\boldsymbol{s}}^{(0)} = \boldsymbol{0}$. We describe MP in Alg. 1.

Algorithm 1: Matching Pursuit

Input : sparsity level K, measurement vector $\boldsymbol{y}$.
Output : sparse vector estimate $\hat{\boldsymbol{s}}$

begin

Initialization: $\boldsymbol{r}^{(0)} = \boldsymbol{y}$; $\hat{\boldsymbol{s}}^{(0)} = \boldsymbol{0}$; $k = 1$

While $k < K$ **:**

1. Find the index of the best matching atom:

$$n^* = \underset{n \in [N]}{\operatorname{argmax}} \left| \left\langle \boldsymbol{\psi}_n, \boldsymbol{r}^{(k)} \right\rangle \right|^2 .$$

2. Update the sparse vector estimate:

$$\hat{s}^{(k)}[n^*] = \frac{1}{M} \left\langle \boldsymbol{\psi}_{n^*}, \boldsymbol{r}^{(k)} \right\rangle .$$

3. Remove the contribution of the found atom:

$$\boldsymbol{r}^{(k+1)} = \boldsymbol{r}^{(k)} - \hat{s}^{(k)}[n^*] \boldsymbol{\psi}_{n^*} .$$

4. $k \leftarrow k + 1$

end

end

It is important to highlight that, even in the absence of noise, the algorithm has no guarantee to provide a perfect reconstruction for non-orthogonal dictionary (i.e., such that $\boldsymbol{y} = \boldsymbol{\Psi s}$). Indeed, if the dictionary is not orthogonal, the correlation between at least some of its atoms is non-zero. Therefore, the index corresponding to the maximal matching atom (step 1) might result from the existence of multiple *other* active atoms correlated with it. This problem is strongly linked to the concept of resolution in the framework of radar applications. In Chap 5, we discuss the ability of Orthogonal MP to overcome this issue under some specific conditions, and therefore the possibility to extend the other algorithms of this section to Orthogonal versions.

2 .2 Block Matching Pursuit

The BMP algorithm is an extension of MP adapted to signals exhibiting a block-sparse representation and has been proposed by [14]. A block sparse signal of order Q is a sparse signal whose N indices are grouped into L blocks of Q indices, such that $N = LQ$. A block is *inactive* if all the elements it contains are zeros, and *active* if at least one of its elements is non-zero. Block sparse signals are particular cases of sparse signals, hence suited for recovery based on known algorithms such as MP. However the block property is an additional prior on the signal that is not harnessed when using MP. BMP intends to adapt MP to take this prior into account. Let us define the block-K sparse representation

of a signal $\boldsymbol{y}$ using (2.1) where $\boldsymbol{s} \in \mathbb{C}^N$ is now composed of L blocks of Q elements, with at most $K \ll L$ *active* blocks. For the following developments, the index n is replaced by the double index (l, q) such that $s[l, q]$ denotes the q-th element of the l-th block of $\boldsymbol{s}$ and $\boldsymbol{\psi}_{l,q}$ denotes the atom of $\boldsymbol{\Psi}$ associated to $s[l, q]$.

The BMP algorithm

Similar to MP it is based on, BMP iteratively looks for indexes of atoms constituting the signal, and subsequently remove their contribution. Therefore, the initialization of BMP is identical to the one of MP. Finding a contributing atom is replaced by jointly finding a contributing block of atoms. Alg. 2 describes this algorithm.

Algorithm 2: Block Matching Pursuit

Input : block-sparsity level K, measurement vectors $\boldsymbol{y}$.
Output : sparse vector estimate $\hat{\boldsymbol{s}}$

begin

Initialization: $\boldsymbol{r}^{(0)} = \boldsymbol{y}$; $\hat{\boldsymbol{s}}^{(0)} = \boldsymbol{0}$; $k = 1$

While $k < K$ **:**

1. Find the index of the block of atoms matching with highest energy:

$$l^* = \underset{l \in [L]}{\operatorname{argmax}} \sum_{q \in [Q]} \left| \left\langle \boldsymbol{\psi}_{l,q}, \boldsymbol{r}^{(k)} \right\rangle \right|^2 .$$

2. Update the sparse vector estimate:

$$\hat{s}^{(k)}[l^*, q] = \frac{1}{M} \left\langle \boldsymbol{\psi}_{l^*,q}, \boldsymbol{r}^{(k)} \right\rangle \qquad \forall\, q \in [Q] \,.$$

3. Remove the contributions of the selected atoms:

$$\boldsymbol{r}^{(k+1)} = \boldsymbol{r}^{(k)} - \sum_{q \in [Q]} \hat{s}^{(k)}[l^*, q] \boldsymbol{\psi}_{l^*,q}.$$

4. $k \leftarrow k + 1$

end

end

Application to Multiple Measurements Vectors

In the framework of multiple measurements vectors, where a common event is observed through multiple channels, the joint acquisition model is equivalent to a particular case of block sparse signals. That is, let us assume the measurement of Q distinct signals resulting from multiple observations of a common event, hence sharing a common sparsity pattern. More precisely, the q-th measurement is $\boldsymbol{y}^q \in \mathbb{C}^{M^q}$ and the sparse vectors $\boldsymbol{s}^q, \boldsymbol{s}^{q'} \in \mathbb{C}^L$ are such that $\operatorname{Supp}(\boldsymbol{s}^q) = \operatorname{Supp}\left(\boldsymbol{s}^{q'}\right)$ for all $q' \in [Q]$,

$$\boldsymbol{y}^q = \boldsymbol{\Psi}^q \boldsymbol{s}^q \qquad \forall\, q \in [Q]\,. \tag{2.9}$$

We can combine the Q signals $\boldsymbol{y}^q$ into a single vector, as well as the Q sparse vectors $\boldsymbol{s}^q$ such that,

$$\begin{aligned} \boldsymbol{y} &:= \left(\left(\boldsymbol{y}^0\right)^T, ..., \left(\boldsymbol{y}^Q\right)^T \right)^T, \\ \boldsymbol{s} &:= \left(\left(\boldsymbol{s}^0\right)^T, ..., \left(\boldsymbol{s}^Q\right)^T \right)^T. \end{aligned} \tag{2.10}$$

These vectors are now linked by a common dictionary corresponding to a block diagonal matrix, such that the sensing model reads the the following,

$$\begin{bmatrix} \boldsymbol{y}^0 \\ \boldsymbol{y}^1 \\ \boldsymbol{y}^2 \\ \vdots \\ \boldsymbol{y}^{Q-1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Psi}^0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Psi}^1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\Psi}^2 & \cdots & \mathbf{0} \\ \vdots & & & \ddots & \vdots \\ \mathbf{0} & \cdots & & \mathbf{0} & \boldsymbol{\Psi}^{Q-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{s}^0 \\ \boldsymbol{s}^1 \\ \boldsymbol{s}^2 \\ \vdots \\ \boldsymbol{s}^{Q-1} \end{bmatrix}, \tag{2.11}$$

where the indices are grouped in blocks of length Q such that $s^q[l] = s[(q-1)L + l]$ is the q-th element of the l-th block of $\boldsymbol{s}$. $\boldsymbol{\psi}_l^q$ denotes the atom in $\boldsymbol{\Psi}^q$ associated to $s^q[l]$.

The BMP algorithm has been shown to lead to best improvements with respect to MP when the atoms of each block are orthogonal to one another [14]. This condition is obviously fulfilled by the representation in (2.11). In addition, let us notice that if we have the prior knowledge of, for instance, $\boldsymbol{s}^0 = \alpha \boldsymbol{s}^1$, where $\alpha \in \mathbb{C}$ then (2.11) can be re-written as

$$\begin{bmatrix} \boldsymbol{y}^0 \\ \boldsymbol{y}^1 \\ \boldsymbol{y}^2 \\ \vdots \\ \boldsymbol{y}^{Q-1} \end{bmatrix} = \begin{bmatrix} \alpha\boldsymbol{\Psi}^0 & \mathbf{0} & \cdots & \mathbf{0} \\ \boldsymbol{\Psi}^1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Psi}^2 & \cdots & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{\Psi}^{Q-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{s}^0 \\ \boldsymbol{s}^2 \\ \vdots \\ \boldsymbol{s}^{Q-1} \end{bmatrix}. \tag{2.12}$$

In radar applications, this example occurs if the signals $\boldsymbol{y}^0$ and $\boldsymbol{y}^1$ result from co-located receiving antennas. Also, note that according to requirement of MP, performing this operation requires to normalize the newly constructed atoms.

Finally, the general formulation of BMP described earlier can be adapted to the particular case of multiple measurements vectors described in (2.9). This adapted formulation is given in Alg. 3. Note that we added a weighting factors w^q. This generalization is motivated by the fact that the different measurements can suffer from different hazards. For instance in AWGN condition, suffer from different noise powers, or the same noise power but different measurement lengths presumably leading to different robustness against noise. These weights were not needed in the classical BMP formulation above because the authors of [14] implicitly assumed a single received signal, hence affected uniformly by the same hazards.

Algorithm 3: Block Matching Pursuit applied to Multiple Measurement Vectors

Input : shared sparsity level K, measurement vectors $\boldsymbol{y}^q$ for $q \in [Q]$.
Output : sparse vectors estimates $\hat{\boldsymbol{s}}^q$, for $q \in [Q]$.

begin

Initialization: $\boldsymbol{r}^{q,(0)} = \boldsymbol{y}^q, \quad \hat{\boldsymbol{s}}^{q,(0)} = \boldsymbol{0} \quad \forall\, q \in [Q]; \quad k = 1;$

While $k < K$ **:**

1. Find the common index of atoms matching with highest energy:

$$l^* = \underset{l \in [L]}{\operatorname{argmax}} \sum_{q \in [Q]} w^q \left| \left\langle \boldsymbol{\psi}_l^q, \boldsymbol{r}^{q,(k)} \right\rangle \right|^2 .$$

2. Update the sparse vectors estimates:

$$\hat{s}^{q,(k)}[l^*] = \frac{1}{M^q} \left\langle \boldsymbol{\psi}_{l^*}^q, \boldsymbol{r}^{q,(k)} \right\rangle \qquad \forall\, q \in [Q] \,.$$

3. Remove the contributions of the selected atoms:

$$\boldsymbol{r}^{q,(k+1)} = \boldsymbol{r}^{q,(k)} - \hat{s}^{q,(k)}[l^*] \boldsymbol{\psi}_{l^*}^q \qquad \forall\, q \in [Q] \,.$$

4. $k \leftarrow k + 1$

end

end

Finally, in AWGN conditions, a reasoning involving ML estimator similar to the one proposed in (2.7),(2.8) is useful to determine values to give to the weights w^q. Consider the signal $\boldsymbol{y}^q = \boldsymbol{\Psi}^q \boldsymbol{s}^q + \boldsymbol{e}^q$ where $e^q[m] \overset{\text{i.i.d.}}{\sim} \mathcal{N}\left(0, \sigma_q^2\right)$ for $m \in [M]$. If the dictionary $\boldsymbol{\Psi}^q$ is normalized to have atoms of energy $\|\boldsymbol{\psi}_l^q\|_2^2 = M^q$ for $l \in [L]$, then, according to the ML approach proposed in appendix A , the weight associated to this signal should fulfil the following,

$$w^q \propto \frac{1}{\sigma_q^2 M^q}. \tag{2.13}$$

The factor $1/M^q$ forces us to normalize the dictionaries. This might be against some intuition because we could think that longer measurements would be more trustworthy due to their higher robustness against noise at the output of the atom matching operation. The factor $1/\sigma_q^2$ gives more importance to signals affected by lower noise power. As a matter of convention, we decide to always normalize the weights such that $\sum_{q \in [Q]} w^q = 1$, therefore,

$$w^q = \frac{1}{\sum_{q \in [Q]} \frac{1}{\sigma_q^2 M^q}} \frac{1}{\sigma_q^2 M^q}. \tag{2.14}$$

3 Factored (Block) Matching Pursuit

3 .1 Multi-dimensional sparse representation and complexity

In both MP and BMP, the dominating operation in terms of computational complexity is the projection on the sparsity domain occurring in the first step of the loop. Let us first consider MP, in practice, such an operation requires the matrix product involving the conjugate transposed of the dictionary, in order to obtain the *projected* signal, that we denote by $\tilde{\boldsymbol{s}}$,

$$\tilde{\boldsymbol{s}} = \boldsymbol{\Psi}^H \boldsymbol{y}. \tag{2.15}$$

Because $\boldsymbol{\Psi} \in \mathbb{C}^{M \times N}$, this operation has a complexity in $\mathcal{O}(MN)$. When discussing this complexity, we are concerned about the best way of processing the *given* signal $\boldsymbol{y}$. While M is an intrinsic property of the received signal, N is a property of the sparse representation we attributed to $\boldsymbol{y}$, and is therefore the main parameter we have hands on concerning complexity. Typically, the asset of an increased value of N (i.e., of a larger dictionary) is the possibility to obtain a lower sparsity level K. A specific situation involving a high value of N is the case of sparse representations that can be expressed in a multi-dimensional fashion. This feature may result, for instance, from dictionaries built from the scaling and the shifting of a standard atom. Such methodical construction of dictionary is fairly common and is used, for example, to build the famous Discrete Wavelet Transform (DWT). The atoms of dictionaries built this way could be enumerated as $\boldsymbol{\psi}_{n_1,n_2}$, where $n_1 \in [N_1]$ indexes the shift and $n_2 \in [N_2]$ indexes the scale. More generally, we can express a sparse representation with dictionaries' atoms described with D features, or degrees of freedom, as

$$y[m] = \sum_{n_1 \in [N_1]} \cdots \sum_{n_D \in [N_D]} \psi_{n_1,..,n_D}[m] s[n_1,..,n_D]. \tag{2.16}$$

To use MP, this representation has to be reformulated as in (2.1), hence merging the D dimensions into one such that $\boldsymbol{\Psi} \in \mathbb{C}^{M \times N}$ where $N = \prod_{i=1}^{D} N_i$. Therefore, if we desire to refine the dictionary equally upon each feature (e.g., double both the number of possible shifts and the number of possible scales), the size of $\boldsymbol{\Psi}$ is tremendously affected (e.g., by a factor 2^D). On another flavour, if the measured signal's structure permits, the measurement signal $\boldsymbol{y}$ can be expressed using multiple indexes,

$$y[m_1,..,m_d] = \sum_{n_1 \in [N_1]} \cdots \sum_{n_D \in [N_D]} \psi_{n_1,..,n_D}[m_1,..,m_d] s[n_1,..,n_D], \tag{2.17}$$

where $m_i \in [M_i]$ for $i \in [d]$. Without additional assumption or property on the signal, these reformulations remain curiosities for the beauty of notations and interpretations. The computational complexity remains $\mathcal{O}(MN) = \mathcal{O}(M_1..M_d N_1..N_D)$.

3 .2 Fully factored sparse representation

From now, we assume $d = D$.[1] Starting from (2.17), we call a sparse model *order-D fully factored* if all of the D dimensions are fully factored with respect to the others, i.e., if the i-th dimension of $\boldsymbol{y}$ (indexed as m_i) is associated exclusively to the i-th dimension of $\boldsymbol{s}$ (indexed as n_i), for all $i \in [D]$. Of course, while any sparse representation (2.1) can be expressed like (2.17), only some of them can be expressed with a fully factored model of order $D > 1$ depending on the inner properties of the dictionary $\boldsymbol{\Psi}$. An order-D fully factored model is expressed as,

$$y[m_1, .., m_D] = \sum_{n_1 \in [N_1]} \psi^{(1)}_{n_1}[m_1] \sum_{n_2 \in [N_2]} \psi^{(2)}_{n_2}[m_2] \cdots \sum_{n_D \in [N_D]} \psi^{(D)}_{n_D}[m_D] s[n_1, .., n_D], \tag{2.18}$$

which also has an equivalent recursive description:

$$p^{(i)}[n_1, .., n_i; m_{i+1}, .., m_D] = \sum_{n_{i+1} \in [N_{i+1}]} \psi^{(i+1)}_{n_{i+1}}[m_{i+1}] p^{(i+1)}[n_1, .., n_{i+1}; m_{i+2}, .., m_D], \tag{2.19}$$

$$p^{(0)}[m_1, .., m_D] := y[m_1, .., m_D], \tag{2.20}$$

$$p^{(D)}[n_1, .., n_D] := s[n_1, .., n_D]. \tag{2.21}$$

In (2.19), $\boldsymbol{p}^{(i)} \in \mathbb{C}^{(N_1 \times .. \times N_i) \times (M_{i+1} \times .. \times M_D)}$ is an hybrid signal that has i dimensions in the sparsity domain and $D - i$ dimensions in the signal domain. This recursive model directly shows how a factored expression lowers the computational complexity involved in the projection to the sparsity domain. Indeed, from $\boldsymbol{y} \in \mathbb{C}^{M_1 \times .. \times M_D}$, we can compute the projected signal on the hybrid domain of $\boldsymbol{p}^{(1)}$, this projection is noted as $\tilde{\boldsymbol{p}}^{(1)}$. From $\tilde{\boldsymbol{p}}^{(i)}$ we can compute $\tilde{\boldsymbol{p}}^{(i+1)}$ recursively up to $\tilde{\boldsymbol{p}}^{(D)}$. This process has a complexity $\mathcal{O}\left(\sum_{i=1}^{D} N_1..N_i M_{i+1} M_D\right)$, which is much lower than $\mathcal{O}(MN) = \mathcal{O}(M_1..M_d N_1..N_D)$.

When $D = 2$, the representation and the projection gets simpler to write as we can use matrix notations. If $\boldsymbol{Y} \in \mathbb{C}^{M_1 \times M_2}$, $\boldsymbol{\Psi} \in \mathbb{C}^{M_1 \times N_1}$, $\boldsymbol{\Phi} \in \mathbb{C}^{M_2 \times N_2}$, $\boldsymbol{S} \in \mathbb{C}^{N_1 \times N_2}$, $\boldsymbol{P} \in \mathbb{C}^{N_1 \times M_2}$, we define $\boldsymbol{P} := \boldsymbol{S}\boldsymbol{\Phi}^T$ and the order-2 factored model is

$$\boldsymbol{Y} = \boldsymbol{\Psi}\boldsymbol{S}\boldsymbol{\Phi}^T = \boldsymbol{\Psi}\boldsymbol{P}, \tag{2.22}$$

such that the projection on the sparsity domain is easily obtained by

$$\tilde{\boldsymbol{S}} = \boldsymbol{\Psi}^H \boldsymbol{Y} \boldsymbol{\Phi}^* = \tilde{\boldsymbol{P}} \boldsymbol{\Phi}^*. \tag{2.23}$$

Let us insist on the fact that the method described in (2.23), which is the particular case of the recursive method proposed in the previous paragraph, is **not** yet the *Factored Matching Pursuit* we developed. It is simply the classical MP where we take advantage of a factored modeling. Indeed, alike MP, we are still computing the whole projection

[1] This situation may typically happen with $D = 2$ in bistatic Radars when we have the bistatic range and the bistatic speed as sparsity domain's feature on the one hand, and the ramp index and sample index structuring the measurement vector on the other hand. It can happen with $D = 3$ for bistatic-SIMO Radars if we add the direction of arrival as sparsity domain feature, and the index of the RX antenna in the measurement.

$\tilde{\boldsymbol{s}} = \text{Vec}\left(\tilde{\boldsymbol{S}}\right)$ (yet efficiently for factored dictionary) to find the index n of a non-zero element (where $n := \{n_1, .., n_D\}$ for factored dictionary). In opposition to this, the algorithm developed in this work and named *Factored Matching Pursuit* further spares a lot more complexity by determining successively $n_1, n_2, .., n_D$.

3 .3 Factored Matching Pursuit (FMP)

For simplicity of notations, further developments assume $D = 2$, yet the results are easily translated to $D > 2$ in a recursive fashion inspired by (2.19). In order for FMP to work, we need to assume a an additional structure on the sparse matrix $\boldsymbol{S}$. Not only $\boldsymbol{S}$ is K-sparse but each of its columns is sparse and each of its rows is sparse, formally,

$$\begin{aligned} |\{i : \|S_{i,:}\| \neq 0\}| &\ll N_1, \\ \|S_{i,:}\|_0 &\ll N_2 \quad \forall\, i \in [N_1]\,. \end{aligned} \tag{2.24}$$

The underlying idea of FMP resides in the property of the matrix $\boldsymbol{P} \in \mathbb{C}^{N_1 \times M_1}$. Because $\boldsymbol{P} = \left(\boldsymbol{\Phi}\boldsymbol{S}^T\right)^T$ and $\boldsymbol{S}$ is a sparse matrix, each row of $\boldsymbol{P}$ is either a zero-vector (if no non-zero element in the corresponding row of $\boldsymbol{S}$) or contains a non-zero signal (if at least one non-zero element in the corresponding row of $\boldsymbol{S}$). In other words, its columns are all sparse and share a common sparsity pattern. $\boldsymbol{P}$ is said to be *row-sparse*. This means that the matrix $\boldsymbol{Y}$ can be seen as a multiple measurements vectors, hence allowing the application of BMP. There are two versions of FMP, each corresponding to a different philosophy depicted, respectively, in Fig. 2.1 and 2.2.

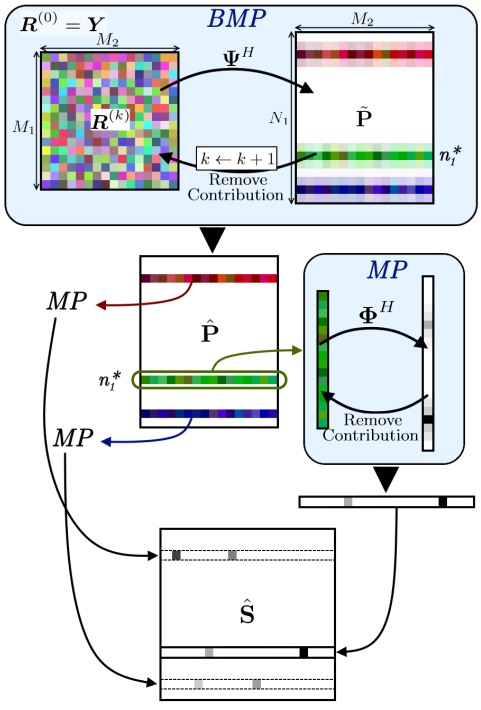


Figure 2.1: Schematic illustration of the first variant of FMP.

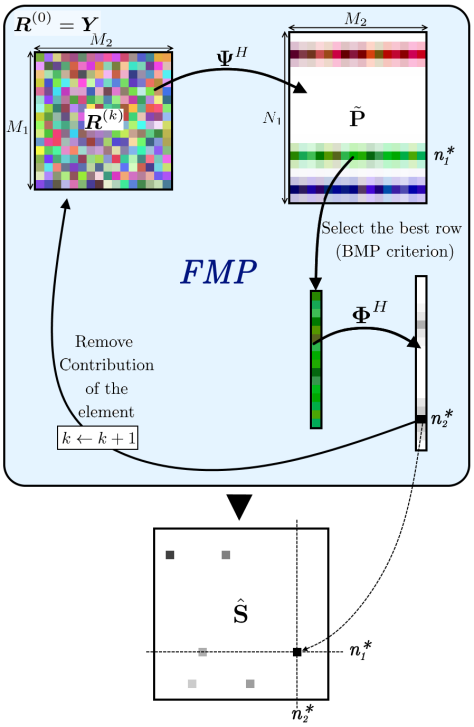


Figure 2.2: Schematic illustration of the second variant of FMP.

In the first version, we first perform BMP to recover the row-sparse matrix $\hat{\boldsymbol{P}}$ from $\boldsymbol{Y}$. Afterwards, we apply MP to each non-zero row of $\hat{\boldsymbol{P}}$ to recover all the elements of the associated rows of $\boldsymbol{S}$. In other words, this version first looks for all indices n_1 on the first dimension containing at least one non-zero element along the second dimension. Then for each such n_1, goes in the second dimension in order to find the associated indices n_2 of non-zero elements. On the other hand, in the second version, each time an index n_1 on the first dimension is found by the first step of one iteration of BMP, it directly goes to the second dimension to find one associated index n_2, then records the value in $\hat{s}[n_1^*, n_2^*]$. An algorithmic description of the second version is provided in Alg. 4, where $\boldsymbol{r}^{(k)}[m_2]$ denotes the m_2-th column of the residue matrix $\boldsymbol{R}^{(k)}$.

Algorithm 4: order-2 Factored Matching Pursuit

Input : sparsity level K, measurement matrix $\boldsymbol{Y}$.
Output : sparse matrix estimate $\hat{\boldsymbol{S}}$.
begin

Initialization: $\boldsymbol{R}^{(0)} = \boldsymbol{Y}, \quad \hat{\boldsymbol{S}}^{(0)} = \boldsymbol{0}; \qquad k = 1;$

While $k < K$ **:**

1. Find the first dimension index of atom matching with highest energy:

$$n_1^* = \underset{n_1 \in [N_1]}{\operatorname{argmax}} \sum_{m_2 \in [M_2]} \left| \left\langle \boldsymbol{\psi}_{n_1}, \boldsymbol{r}^{(k)}[m_2] \right\rangle \right|^2 .$$

2. Select the n_1^*-th row of $\tilde{\boldsymbol{P}}$:

$$\tilde{\boldsymbol{p}}_{n_1^*} := \left(\left\langle \boldsymbol{\psi}_{n_1^*}, \boldsymbol{r}^{(k)}[1] \right\rangle, .., \left\langle \boldsymbol{\psi}_{n_1^*}, \boldsymbol{r}^{(k)}[M_2] \right\rangle \right)^T .$$

3. Find the index of the best matching element of the selected row of $\tilde{\boldsymbol{P}}$:

$$n_2^* = \underset{n_2 \in [M2]}{\operatorname{argmax}} \left| \left\langle \boldsymbol{\phi}_{n_2}, \tilde{\boldsymbol{p}}_{n_1^*} \right\rangle \right|^2 .$$

4. Update the sparse matrix estimate $\hat{\boldsymbol{S}}$.

$$\hat{s}^{(k)}[n_1^*, n_2^*] = \frac{1}{M_1 M_2} \left\langle \boldsymbol{\phi}_{n_2^*}, \tilde{\boldsymbol{p}}_{n_1^*} \right\rangle$$

5. Remove the contribution of the selected atom:

$$\boldsymbol{R}^{(k+1)} = \boldsymbol{R}^{(k)} - \hat{s}^{(k)}[n_1^*, n_2^*] \boldsymbol{\psi}_{n_1^*} \boldsymbol{\phi}_{n_2^*}^T .$$

6. $k \leftarrow k + 1$

end
end

Comparatively, the particularity of the first version is the fact that it requires a strongest prior knowledge on the structure of sparsity level of $\boldsymbol{S}$. Indeed, so far, we assume the knowledge of K: the number of non-zero elements in $\boldsymbol{S}$. In order to use the first method, we should know the distribution of these elements across the dimensions of $\boldsymbol{S}$, i.e, $K = K_1K_2$ such that $\boldsymbol{S}$ contains K_1 non-zero row, each containing K_2 non-zero elements. If such a structure is known, both versions exhibit same reconstruction quality, yet the first version is faster up to a factor K_2 because this version needs a smaller total number of iterations to compute the estimate of $\boldsymbol{S}$. However, without such knowledge, we can only use the second version. In Chap 3, we shall see how FMP reacts to model mismatches and see that the second version is more prone to be adapted to counteract these mismatches. Still, some further improvements proposed in Chap. 5 are inspired from a feature of the first version.

Let us consider the second version as the reference in term of complexity, each FMP iteration has a complexity of $\mathcal{O}\left(M_2M_1N_1 + M_2N_2\right)$. Note that, because $M_2M_1N_1 + M_2N_2 \neq M_1M_2N_2 + M_1N_2$, the order chosen to process the dimensions matters. We shall see later that choosing an order favouring a low complexity increases the loss in robustness against noise. When $D > 2$, the complexity is $\mathcal{O}\left(\sum_{i=1}^{D} M_i..M_DN_i\right)$, which is smaller than the former complexity of MP for factored model of $\mathcal{O}\left(\sum_{i=1}^{D} M_1..M_iN_{i+1}..N_D\right)$. The difference in complexity becomes huge when $N \gg M$, which is therefore typically the type of models targeted by FMP.

3 .4 The robustness-complexity trade-off

Applying FMP allows to spare computational complexity but lowers the robustness against noise. To study this effect, let us consider a very simple case where $D = 2$, all dictionaries are orthogonals, and $\boldsymbol{S}$ is 1-sparse with $s[n_1^*, n_2^*] \neq 0$. We will compare the performances of two algorithms: MP and FMP. For a fair comparison, the MP we consider is the already faster version of MP we suggested in (2.23), which takes advantage of the factored model but still computes the whole projection $\tilde{\boldsymbol{S}}$. The FMP we consider is the second version. Under AWGN conditions, the model is thus formulated as

$$\boldsymbol{Y} = \boldsymbol{\Psi}\boldsymbol{P} + \boldsymbol{E} \text{ where } \boldsymbol{P} = \boldsymbol{S}\boldsymbol{\Phi}^T. \tag{2.25}$$

where $\boldsymbol{E}$ a white noise matrix such that each of its elements are follows $\mathbb{C}\mathcal{N}\left(0, \sigma_e^2\right)$. The first step of both algorithms is the computation of $\tilde{\boldsymbol{P}} = \boldsymbol{\Psi}^*\boldsymbol{Y}$. As we assumed orthogonal dictionaries, we can start the analysis from

$$\tilde{p}[n_1; m_2] = M \sum_{n_2 \in [N_2]} \phi_{n_2}[m_2]s[n_1, n_2] + u[n_1; m_1]. \tag{2.26}$$

where $u[n_1; m_1] \sim \mathbb{C}\mathcal{N}\left(0, M_1\sigma_e^2\right)$. From this point, MP for factored model of (2.23) performs the operation described in (2.27) while FMP performs the operation described in (2.28).

$$
\begin{aligned}
\tilde{s}[n_1, n_2] &:= \sum_{m_2 \in [M_2]} \phi_{n_2}[m_2] \tilde{p}[n_1; m_2], \\
\{n_1^*, n_2^*\} &= \underset{n_1, n_2}{\operatorname{argmax}} \ \left| \tilde{s}[n_1, n_2] \right|^2,
\end{aligned}
\tag{2.27}
$$

$$
\begin{aligned}
n_1^* &= \underset{n_1}{\operatorname{argmax}} \ \sum_{m_2 \in [M_2]} \left| \tilde{p}[n_1; m_2] \right|^2, \\
n_2^* &= \underset{n_2}{\operatorname{argmax}} \ \left| \sum_{m_2 \in [M_2]} \phi_{n_2}^*[m_2] \tilde{p}[n_1^*; m_2] \right|^2,
\end{aligned}
\tag{2.28}
$$

The loss in robustness occurs because the coherent summation performed in MP is replaced by an energy detector in FMP. More precisely, let us compare $\left|\tilde{s}[n_1, n_2]\right|^2$ and $\sum_{m_2 \in [M_2]} \left|\tilde{p}[n_1; m_2]\right|^2$, which are the decision variables for the determination of n_1^*. By injecting (2.26) in these expression, we obtain

$$
\left|\tilde{s}[n_1, n_2]\right|^2 = \begin{cases} (M_1 M_2)^2 + M_1 M_2 \sigma_e^2 \mu + w & \{n_1, n_2\} = \{n_1^*, n_2^*\} \\ M_1 M_2 \sigma_e^2 \mu' & \{n_1, n_2\} \neq \{n_1^*, n_2^*\} \end{cases}, \tag{2.29}
$$

$$
M_2 \sum_{m_2 \in [M_2]} \left|\tilde{p}[n_1^{\cdot} m_2]\right|^2 = \begin{cases} (M_1 M_2)^2 + M_1 M_2 \sigma_e^2 \nu + w' & \{n_1, n_2\} = \{n_1^*, n_2^*\} \\ M_1 M_2 \sigma_e^2 \nu' & \{n_1, n_2\} \neq \{n_1^*, n_2^*\} \end{cases}, \tag{2.30}
$$

where $w, w' \sim \mathbb{C}\mathcal{N}(0, 4M_1M_2)$, $\mu, \mu' \sim \chi_2^2$ and $\nu, \nu' \sim \chi_{2M_2}^2$. The only difference between the (2.29) and (2.30) is the random variable $\mu \sim \chi_2^2$ replaced by $\nu \sim \chi_{2M_2}^2$ when applying FMP instead of MP. This shows that the effective noise affecting the decision variable is stronger when applying FMP and that this gap is firmly affected by M_2, i.e., by the number of elements in the second dimension along which we perform a sum either coherently for MP, or incoherently for FMP. This matches the intuition. A more detailed analysis of the effect of these random variables would be tricky, notably because μ (resp. ν) is not independent from w (resp. w'). Moreover, it is important to remind that these equations relates a excessively simple situation. A lot of further work would be useful to rigorously characterize the impact of FMP on the robustness against noise for any dictionary (not orthogonal as assume here) and any sparsity level (we assumed here $K = 1$). However, the above analysis already succeeds to highlight the main effect and its relation with model parameters. We will therefore be satisfied with it in the framework of this document, and complete the analysis with some simulations.

Simulations using DFT-2 dictionary

A very simple example matching the assumptions of the above analysis is the DFT-2 dictionary. To measure the robustness of the reconstruction against noise, two metrics are considered. The first one is the average reconstruction error: $\|\boldsymbol{y} - \boldsymbol{\Psi}\boldsymbol{s}\|_2^2$, the second metric is the miss rate of the estimation of n_1^* and n_2^*. The first metric is more relevant in the sense that it is the metric that, in the first place, both MP and FMP try to minimize. The second metric is more related to the capability of support recovery, but is relevant in

many applications of sparse recovery algorithms. The simulations show the evolution of these metrics with respect to the noise power. In all the following figures, the x-axis is the Signal to Noise Ratio (SNR) define as

$$SNR := \frac{\mathbb{E}\left\{\left\|\boldsymbol{\Psi}\boldsymbol{S}\boldsymbol{\Phi}^T\right\|_F^2\right\}}{\mathbb{E}\left\{\|\boldsymbol{E}\|_F^2\right\}}. \tag{2.31}$$

As a first step, let us assume 1-sparse signals $s[n_1, n_2]$ randomly selected such that

$$s[n_1, n_2] = \begin{cases} \alpha & \{n_1, n_2\} = \{n_1^*, n_2^*\} \\ 0 & else \end{cases}, \tag{2.32}$$

with $\alpha \sim \mathbb{CN}(0, 1)$ and n_1^*, n_2^* uniformly selected among their possible values. In the first test, both MP and FMP where performed over 20.000 simulations per noise power for different values of M_1, M_2, selected such that $M_1 M_2 = 1024$ remain constant. Modifying M_1 and M_2 while keeping their product constant should have no influence on performances of MP because of the symmetry in (2.29). However, FMP should get worse when M_2 increases, according to (2.30). The results are shown in figure 2.3. Interestingly, regarding the reconstruction metric, there is no significant difference between the two algorithms, indicating similar reconstruction quality. However, regarding the support recovery metrics, we can clearly observe a small SNR-gap between the MP-curves and FMP-curves. As expected, The MP-curves are not affected by the modifications of the values of M_1 and M_2 when $M_1 M_2$ is constant, while the SNR-gap slightly increases when M_2 increases, as foreshadowed in the previous section.

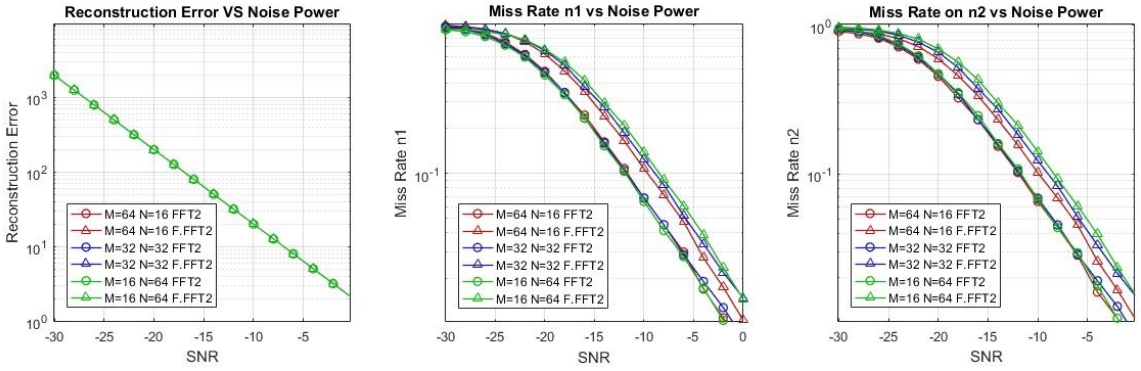


Figure 2.3: Comparison of the robustness against noise between MP and FMP for different sizes of Dictionaries, using DFT-2 Dictionaries. FFT2 is the 2D Fast Fourier Transform. F.FFT2 is the FMP applied to FFT2, i.e., a 1D FFT followed by the selection of the strongest row of the result, then another 1D FFT **only** on this row.

$M_1 = N_1$	$M_2 = N_2$	Complexity Reduction	SNR Gap [dB]
64	16	0.606	2.12
32	32	0.515	2.9
16	64	0.438	3.45

Table 2.1: Comparative results between MP and FMP showing the evolution of Complexity Reduction and SNR Gap with orthogonal dictionaries when the dimensions vary

Concerning the computational complexity, using the FFT-2 algorithm, MP has a complexity of $\mathcal{O}(M_1 M_2(\log_2 M_1 + \log_2 M_2))$ while FMP has a complexity of $\mathcal{O}(M_1 M_2 \log_2 M_1 + M_2 \log_2 M_2)$. Table 2.1 displays some examples comparing the *Complexity Redution* ($\mathcal{O}(FMP)/\mathcal{O}(MP)$) and the SNR Gap. The complexity reduction is only around 0.5, which is not a great result. Yet, it was said before that FMP is much more advantageous when $N \gg M$, which is not the case for the orthogonal dictionary considered. Figure 2.4 and Table 2.2 show the results of similar simulations performed with $M_1 = M_2 = 32$ fixed, but varying N_1 and N_2. The purpose is to observe the effect FMP has on highly redundant dictionaries[2] ($N \gg M$).

From the figure 2.4 and the table 2.2, we deduce that increasing the redundancy of the dictionary tends to decrease the SNR-Gap while strongly improving the complexity reduction. This firmly supports the conclusion of greater effectiveness of FMP for very redundant dictionaries. All the tests performed so far remain in the specific condition of 1-sparse signals. Figure 2.5 shows simulations result comparing different values of sparsity level. These results suggest that the previous conclusions hold for sparsity level K greater that 1. Finally, Fig. 2.6 sums up the conclusion about the comparisons of this section.

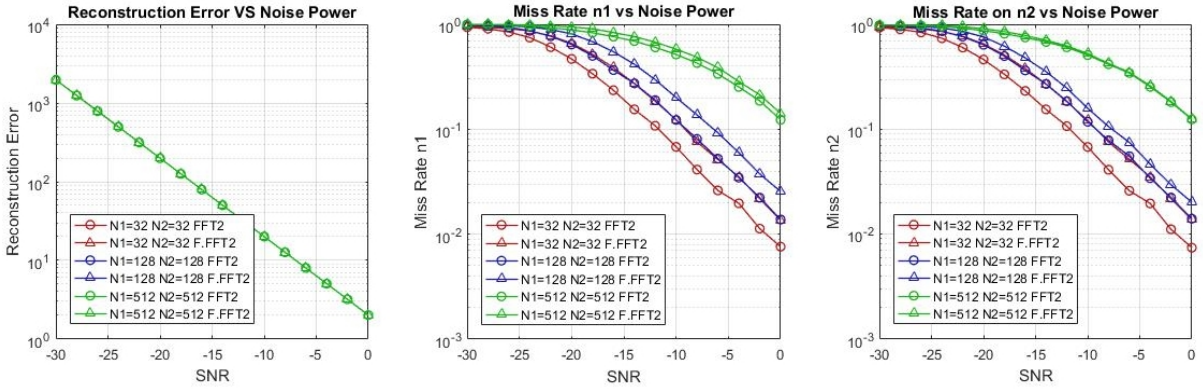


Figure 2.4: Comparison of the robustness against noise between MP and FMP for different sizes of Dictionaries, using oversampled DFT-2 Dictionaries

M_1	M_2	N_1	N_2	Complexity Reduction	SNR Gap [dB]
32	32	32	32	0.515	2.9
32	32	128	128	0.206	2.6
32	32	512	512	0.061	0.83

Table 2.2: Comparative results between MP and FMP showing the evolution of Complexity Reduction and SNR Gap when the dictionaries get more redundant

[2]A redundant DFT2 dictionary is the 2D extension of the following redundant 1D dictionary: $\boldsymbol{W}$ where its n-th atom is $\boldsymbol{w}_n := \left(e^{j2\pi\frac{n}{N}}, e^{j2\pi\frac{2n}{N}}..., e^{j2\pi\frac{(M-1)n}{N}}\right)^T$, with $n \in [N], N \gg M$.

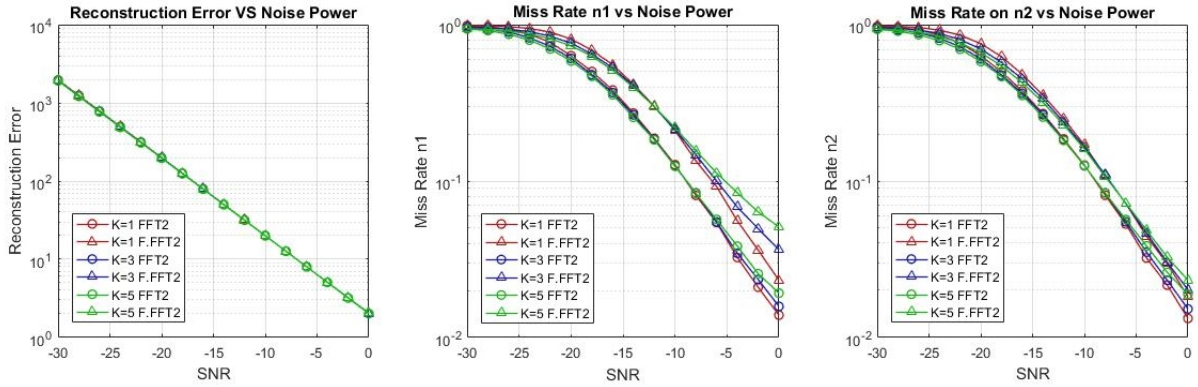


Figure 2.5: Comparison of the robustness against noise between MP and FMP for different sparsity level K. Fixed parameters: $M_1 = M_2 = 32$, $N_1 = N_2 = 128$

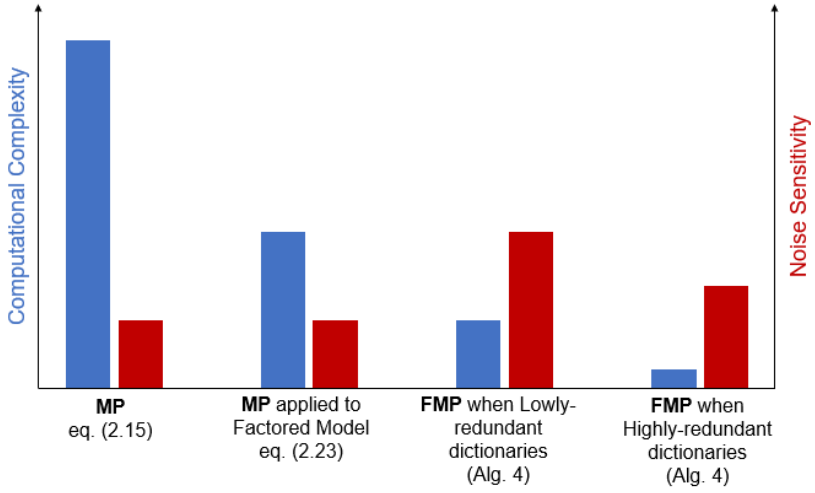


Figure 2.6: Schematic illustration of comparison the complexity of MP and FMP under different conditions. Both the computational complexity and Noise sensitivity are (symbolic) criterion desired to be low.

3 .5 FMP for half-factored sparse representation

So far, FMP addressed situations involving convenient sparse representations as in (2.18). Very interestingly, all the factored processes described above still hold for less convenient models we will call *half-factored*. The definition of an order-D half-factored representation is

$$y[m_1,..,m_D] = \sum_{n_1\in[N_1]} \psi^{(1)}_{n_1}[m_1] \sum_{n_2\in[N_2]} \psi^{(2)}_{n_1n_2}[m_2] \cdots \sum_{n_D\in[N_D]} \psi^{(D)}_{n_1,..,n_D}[m_D] s[n_1,..,n_D]. \tag{2.33}$$

It is important to note that the factored model described in (2.18) is simply a particular case of half-factored model as in (2.33). Similarly to the case of factored models, a recursive expression is provided,

$$p^{(i)}[n_1,..,n_i;m_{i+1},..,m_D] = \sum_{n_{i+1}\in[N_{i+1}]} \psi^{(i+1)}_{n_1,..,n_i,n_{i+1}}[m_{i+1}]p^{(i+1)}[n_1,..,n_{i+1};m_{i+2},..,m_D], \tag{2.34}$$

$$p^{(0)}[m_1,..,m_D] := y[m_1,..,m_D], \tag{2.35}$$

$$p^{(D)}[n_1,..,n_D] := s[n_1,..,n_D]. \tag{2.36}$$

This formulation proves the ability to apply previous reasoning to half-factored sparse representations. A constraining difference with respect to factored representations is the impossibility to chose in which order dimensions apply. As a final remark about half-factored representations, let us highlight that applying FMP instead of MP for factored model not only spares computational complexity but also offers the possibility to save memory. Indeed, consider $D = 2$ with the following half-factored matrix formulation,

$$\begin{aligned} \boldsymbol{Y} &= \boldsymbol{\Psi}\boldsymbol{P}, \\ \boldsymbol{P}^T &= \left[\boldsymbol{\Phi}_0\left[\boldsymbol{S}^T\right]_0, \boldsymbol{\Phi}_1\left[\boldsymbol{S}^T\right]_1, .., \boldsymbol{\Phi}_{N_1-1}\left[\boldsymbol{S}^T\right]_{N_1-1}\right], \end{aligned} \tag{2.37}$$

where $\boldsymbol{\Phi}_{n_1} \in \mathbb{C}^{M_2\times N_2}$ and $\left[\boldsymbol{S}^T\right]_{n_1}$ is the n_1-th column of $\boldsymbol{S}^T$. The $\boldsymbol{\Phi}_{n_1}$ dictionary contains N_2 atoms is such that its n_2-th atom is $(\psi_{n_1,n_2}[1],..,\psi_{n_1,n_2}[M_2])^T$. There are N_1 distinct $\boldsymbol{\Phi}_{n_1}$ dictionaries. FMP offers the possibility of only computing the required dictionaries, i.e., those such that n_1 has been found to contain at least one non-zero element. Fig. 2.7 shows how the second version of the order-2 FMP is adapted to the half-factored models.

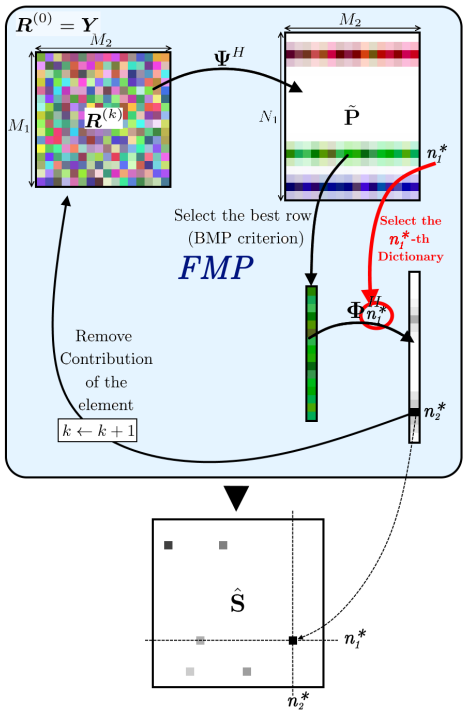


Figure 2.7: Second version of the order-2 FMP adapted to half-factored models

3 .6 Factored Block Matching Pursuit (FBMP)

Let us now address the question of extending FMP to multiple measurement vectors, i.e., create a Factored Block Matching Pursuit algorithm. Let us assume Q distinct signals order-D factored as in (2.18) (the reasoning still holds for signals order-D half-factored as in (2.33)). The dictionaries describing these signals are different, but their sparse representations share a common sparsity pattern. The joint sensing model holds,

$$y^q[m_1,..,m_D] = \sum_{n_1\in[N_1]} \psi^{(1)}_{n_1,q}[m_1] \sum_{n_2\in[N_2]} \psi^{(2)}_{n_2,q}[m_2] \cdots \sum_{n_D\in[N_D]} \psi^{(D)}_{n_D,q}[m_D] s^q[n_1,..,n_D]. \quad (2.38)$$

Although this model may be treated by FMP as an order-$(D+1)$ half-factored model, this interpretation does not take advantage of the common sparsity pattern along the new dimension. In this section, we will therefore start from BMP and adapt it to factored models. For simplicity, let us again consider $D = 2$, such that the received signals are written as

$$\boldsymbol{Y}^q = \boldsymbol{\Psi}^q \boldsymbol{S}^q \left(\boldsymbol{\Phi}^q\right)^T = \boldsymbol{\Psi}^q \boldsymbol{P}^q. \quad (2.39)$$

Starting from the BMP algorithm (Alg. 3), and injecting the FMP algorithm (Alg. 4), we obtain the definition of FBMP algorithm described for $D = 2$ in Alg. 5.

Finally, we can show that BMP, FMP and FBMP all are particular cases of a wider definition of FMP, i.e., of the application of FMP to even more generic forms of *partially*-factored model than the half-factored proposed in (2.33). However, such analysis goes beyond the scope of this work but may be the topic of a future document.

Algorithm 5: order-2 Factored Block Matching Pursuit

Input : block-sparsity level K, measurement matrices $\boldsymbol{Y}^q$ for $q \in [Q]$.

Output : sparse matrices estimates $\hat{\boldsymbol{S}}^q$ for $q \in [Q]$.

begin

Initialization: $\boldsymbol{R}^{q,(0)} = \boldsymbol{Y}^q$, $\hat{\boldsymbol{S}}^{q,(0)} = \boldsymbol{0} \quad \forall\, q \in [Q]; \quad k = 1;$

While $k < K$ **:**

1. Find the common first dimension index of atoms matching with highest energy:

$$n_1^* = \underset{n_1 \in [N_1]}{\operatorname{argmax}} \sum_{q \in [Q]} w^q \sum_{m_2 \in \left[M_2^q\right]} \left| \left\langle \boldsymbol{\psi}_{n_1}^q, \boldsymbol{r}^{q,(k)}[m_2] \right\rangle \right|^2 .$$

2. Select the n_1^*-th row of each $\tilde{\boldsymbol{P}}^q$ for $q \in [Q]$:

$$\tilde{\boldsymbol{p}}_{n_1^*}^q := \left(\left\langle \boldsymbol{\psi}_{n_1^*}^q, \boldsymbol{r}^{q,(k)}[1] \right\rangle, .., \left\langle \boldsymbol{\psi}_{n_1^*}^q, \boldsymbol{r}^{q,(k)}[M_2^q] \right\rangle \right)^T .$$

3. Find the common index of the best matching component of the selected rows of all $\tilde{\boldsymbol{P}}^q$:

$$n_2^* = \underset{n_2 \in [M2]}{\operatorname{argmax}} \sum_{q \in [Q]} w^q \left| \left\langle \boldsymbol{\phi}_{n_2}^q, \tilde{\boldsymbol{p}}_{n_1^*}^q \right\rangle \right|^2 .$$

4. Update the sparse matrices estimates $\hat{\boldsymbol{S}}^q$:

$$\hat{s}^{q,(k)}[n_1^*, n_2^*] = \frac{1}{M_1^q M_2^q} \left\langle \boldsymbol{\phi}_{n_2^*}^q, \tilde{\boldsymbol{p}}_{n_1^*}^q \right\rangle \qquad \forall\, q \in [Q] \,.$$

5. Remove the contribution of the selected atom:

$$\boldsymbol{R}^{q,(k+1)} = \boldsymbol{R}^{q,(k)} - \hat{s}^{q,(k)}[n_1^*, n_2^*] \boldsymbol{\psi}_{n_1^*}^q \boldsymbol{\phi}_{n_2^*}^{q\,T} \qquad \forall\, q \in [Q] \,.$$

6. $k \leftarrow k + 1$.

end

end

Chapter 3

Sparse Recovery Algorithms in the Radar Context

In this chapter, we apply and adapt the algorithms developed in Chap. 2 to the multistatic models developed in Chap.1. Therefore, we assume that the targets' locations and velocities are on known grids of target states and that the number K of targets to be detected is known in advance. Next, we study their behaviours with the inspection of a function called *Ambiguity Function*. This analysis enables the definition of conditions on the choice of location and velocity grids. Finally The effects of model mismatches are studied and compensated for by the introduction of an iterative correction.

1 Targets States Dection Algorithms

Thanks to the sparse models established in Chap. 1, the estimation problem of target parameters is translated to the reconstruction of the sparse vectors $\boldsymbol{s}^q$ for $q \in [Q]$, linked by a shared sparsity pattern. Let us remind that $\boldsymbol{s}^q := (s^q[1,1], .., s^q[1,N_v], s^q[2,1], .., s^q[N_x,N_v])^T$ for all $q \in [Q]$ where $s^q[n_x, n_v]$ was defined in (1.49). N_x and N_v are the respective sizes of the location grid Ω_X and velocity grid Ω_V defining our joint sparse representation. In all generality, the complete sparse model we are considering in this chapter is the multi-static MIMO model described in (1.54), the multi-static SISO model is simply a particular case of this model. We are also considering the simplified model proposed in (1.56). While using this model, the target parameters estimation problem becomes the reconstruction of the sparse matrices $\boldsymbol{S}^q$ for $q \in [Q]$, linked by a shared sparsity pattern. Let us remind that $\boldsymbol{S}^q \in \mathbb{C}^{N_x \times N_v}$ and the n_x-th element of its n_v-th column is $s^q[n_x, n_v]$. For both models, we identify, and eventually adapt, the appropriate algorithm amongst the ones proposed in Chap. 2 to perform the sparse signal reconstruction.

1 .1 Reconstruction from the complete multi-static model

The complete multistatic MIMO Radar sparse model described in (1.54) corresponds exactly to the multiple measurements vectors equation (2.9), for which we proposed an adaptation of BMP (Alg. 3). Looking at the dictionaries description proposed in (1.55), we notice that if there only one TX antenna in the q-th bistatic pair, then the q-th dictionary

is normalized such that[1],

$$\left\| \boldsymbol{d}^q_{n_x,n_v} \right\|_2^2 = N_R^q M, \quad \forall\, q \in [Q]\,, \quad \forall\, n_x \in [N_x]\,, \quad \forall\, n_v \in [N_v]\,, \tag{3.1}$$

where we remind that $M = M_s M_r$ is the number of acquired samples per antennas and $N = N_x N_v$ is the total number of target states. Therefore, as suggested by (2.14), the weights to use in BMP are

$$w^q = \frac{1}{\sum_{q\in[Q]} \frac{1}{\sigma_q^2 N_R^q M}} \frac{1}{\sigma_q^2 N_R^q M} = \frac{1}{\sum_{q\in[Q]} \frac{1}{\sigma_q^2 N_R^q}} \frac{1}{\sigma_q^2 N_R^q}, \tag{3.2}$$

If there is more than one TX antenna in a given bistatic pair, we may simply multiply each atom of the corresponding dictionary by an appropriate coefficient to get it normalized as above. Note that, by performing such a normalization, all indexes of $\boldsymbol{s}^q$ no longer have the same average power when occupied. This feature would also appear if we explitely added the path loss to our model. Because this feature may have impacts on the performance analysis that we have not yet studied, all the tests and simulations of this documents only consider systems with transmitting nodes composed of a single antenna, i.e., multi-static SIMO[2] systems. The BMP applied to our Radar model is described in Alg. 6.

Each iteration of Alg. 6 has a computational complexity in $\mathcal{O}\left(\sum_{q\in[Q]} N_R^q M N\right)$, which is directly linked to the sizes of the dictionaries, $\boldsymbol{D}^q \in \mathbb{C}^{N_R^q M \times N}$, involved in matrix products. In order to realize the meaning of this complexity, let us take a simple example where $N_R^q = 1$ for all $q \in [Q]$. In this example, the location grid result from the sampling of a square in the 2D location domain into 20×20 samples (i.e., $N_x = 400$). The set of velocity states result from the sampling in the 2D velocity domain into 20×20 samples (i.e., $N_v = 400$). As explained later in this chapter, $M_s = 20$ and $M_r = 20$ are typical values when using this set of target states. Therefore, the each dictionary has the huge size $400 \times 160,000$, which is hardly affordable. The complexity is in $\mathcal{O}\left(64,000,000Q\right)$.

1 .2 Reconstruction from the simplified multi-static model

The simplified model described in (1.56) corresponds to multiple measurements of signals having a *half-factored* model, i.e., as in (2.37). Therefore, we can apply the Factored Block Matching Pursuit to perform the reconstruction of the sparse matrices $\boldsymbol{S}^q$. However, the simplified model results from an approximation of the complete model. Therefore, in a given iteration of FBMP, even if we assume that it always correctly finds indexes n_x^* and n_v^* corresponding to a target, it removes an approximated contribution. However, we can correct it by using the same last steps ($2 \to 4$) as Alg. 6, without significant changes in the computational complexity. This adapted version of the second version of FBMP is proposed in Alg. 7. Note that we choose the second version because the iterative correction we introduce at the end of this chapter is only defined for this version.

[1] We dot not normalize all dictionaries to 1 because we want to keep the easy interpretation of the sparse vectors $\boldsymbol{s}^q$ in (1.49). Also, it keeps apparent the higher energy contained in a longer received signal for an **identical probability distribution of the non-zero elements' values** of all $\boldsymbol{s}^q$ for all $q \in [Q]$.

[2] SIMO stands for Single Input Multiple Output

The complexity of this algorithm is only $\mathcal{O}\left(\sum_{q\in[Q]} N_R^q M N_x + M_s N_v\right)$. For the same example we gave above, the complexity of each iteration is in $\mathcal{O}\left(168,000Q\right)$, i.e., 380 times faster than the application of BMP to the complete model. The comparison holds as both algorithms have the same number of iterations: K.

Algorithm 6: Block Matching Pursuit applied to multi-static Radars

Input : Number of targets K, measurement vectors $\boldsymbol{y}^q$.

Output : sparse vectors estimates $\hat{\boldsymbol{s}}^q$, for $q \in [Q]$.
Targets' locations estimates $\{\hat{\boldsymbol{x}}_k : k \in [K]\}$
Targets' velocities estimates $\{\hat{\boldsymbol{v}}_k : k \in [K]\}$.

begin

Initialization: $\boldsymbol{r}^{q,(0)} = \boldsymbol{y}^q, \quad \hat{\boldsymbol{s}}^{q,(0)} = \boldsymbol{0} \quad \forall\, q \in [Q]; \quad k = 1;$

While $k < K$ **:**

1. Find the common index of atoms matching with highest energy:

$$\{n_x^*, n_v^*\} = \underset{\substack{n_x\in[N_x]\\ n_v\in[N_v]}}{\operatorname{argmax}} \sum_{q\in[Q]} w^q \left|\left\langle \boldsymbol{d}^q_{n_x,n_v}, \boldsymbol{r}^{q,(k)}\right\rangle\right|^2.$$

2. Update the sparse vectors estimates and add an estimated target:

$$\hat{s}^{q,(k)}[n_x^*, n_v^*] = \frac{1}{N_R^q M_1 M_2}\left\langle \boldsymbol{d}^q_{n_x^*,n_v^*}, \boldsymbol{r}^{q,(k)}\right\rangle \qquad \forall\, q \in [Q]\,.$$

$$\hat{\boldsymbol{x}}_k = \Omega_X\left\{n_x^*\right\}, \quad \hat{\boldsymbol{v}}_k = \Omega_V\left\{n_v^*\right\}.$$

3. Remove the contributions of the selected atoms:

$$\boldsymbol{r}^{q,(k+1)} = \boldsymbol{r}^{q,(k)} - \hat{s}^{q,(k)}[n_x^*, n_v^*]\boldsymbol{d}^q_{n_x^*,n_v^*} \qquad \forall\, q \in [Q]\,.$$

4. $k \leftarrow k + 1$

end

end

Algorithm 7: Factored Block Matching Pursuit applied to multi-static Radars

Input : block-sparsity level K, signals $\boldsymbol{Y}^q$ for $q \in [Q]$ and weights w^q for $q \in [Q]$.

Output : sparse matrices estimates $\hat{\boldsymbol{S}}^q$ for $q \in [Q]$.
Targets' locations estimates $\{\hat{\boldsymbol{x}}_k : k \in [K]\}$
Targets' velocities estimates $\{\hat{\boldsymbol{v}}_k : k \in [K]\}$.

begin

Initialization: $\boldsymbol{R}^{q,(0)} = \boldsymbol{Y}^q$, $\hat{\boldsymbol{S}}^{q,(0)} = \boldsymbol{0} \quad \forall\, q \in [Q]; \quad k = 1;$

While $k < K$ **:**

1. Find the common first dimension index of atoms matching with highest energy:

$$n_x^* = \underset{n_x \in [N_x]}{\operatorname{argmax}} \sum_{q \in [Q]} w^q \sum_{m_r \in \left[M_r^q\right]} \left| \left\langle \boldsymbol{\psi}_{n_x}^q, \boldsymbol{r}^{q,(k)}[m_r] \right\rangle \right|^2 .$$

2. Select the n_1^*-th row of each $\tilde{\boldsymbol{P}}^q$ for $q \in [Q]$:

$$\tilde{\boldsymbol{p}}_{n_x^*}^q := \left(\left\langle \boldsymbol{\psi}_{n_x^*}^q, \boldsymbol{r}^{q,(k)}[1] \right\rangle, .., \left\langle \boldsymbol{\psi}_{n_x^*}^q, \boldsymbol{r}^{q,(k)}[M_r^q] \right\rangle \right)^T .$$

3. Find the common index of the best matching element of the selected rows of all $\tilde{\boldsymbol{P}}^q$:

$$n_v^* = \underset{n_v \in [M_r]}{\operatorname{argmax}} \sum_{q \in [Q]} w^q \left| \left\langle \boldsymbol{\phi}_{n_x^*, n_v}^q, \tilde{\boldsymbol{p}}_{n_x^*}^q \right\rangle \right|^2 .$$

4. Update the sparse matrices estimates $\hat{\boldsymbol{S}}^q$ and add an estimated target:

$$\hat{s}^{q,(k)}[n_x^*, n_v^*] = \frac{1}{N_R^q M_1 M_2} \left\langle \boldsymbol{d}_{n_x^*, n_v^*}^q, \operatorname{Vec}\left(\boldsymbol{R}^{q,(k)}\right) \right\rangle \qquad \forall\, q \in [Q] .$$

$$\hat{\boldsymbol{x}}_k = \Omega_X \left\{ n_x^* \right\}, \quad \hat{\boldsymbol{v}}_k = \Omega_V \left\{ n_v^* \right\} .$$

5. Remove the contribution of the selected atom:

$$\boldsymbol{R}^{q,(k+1)} = \boldsymbol{R}^{q,(k)} - \hat{s}^{q,(k)}[n_x^*, n_v^*] \operatorname{Mat}_{N_v}\left(\boldsymbol{d}_{n_x^*, n_v^*}^q\right) \qquad \forall\, q \in [Q] .$$

6. $k \leftarrow k + 1$.

end

end

2 Ambiguity Function

In the multi-static Radars signal processing framework, the Ambiguity Function (AF) is a four dimensional continuous function of the location and velocity, denoted by $\mathcal{X}(\boldsymbol{x}, \boldsymbol{v})$. This function relates the correlation between the theoretical acquired signals in the presence of K moving targets and the signals we would acquire if there was a single target located in $\boldsymbol{x}$, moving with a velocity $\boldsymbol{v}$ and characterized by scattering coefficients all equal to 1. According to this definition, the AF related to the q-th bistatic pair is defined,

$$\mathcal{X}^q(\boldsymbol{x}, \boldsymbol{v}) = \sum_{\substack{m_s \in [M_s] \\ m_r \in [M_r]}} y^q[m_s, m_r] d^q_{\boldsymbol{x}_k, \boldsymbol{v}_k}[m_s, m_r]. \tag{3.3}$$

The multi-static AF is derived from the corresponding optimal multi-static detector [5]. More precisely, in the first step of an iteration of Alg. 6, we evaluate the AF in all $\boldsymbol{x} \in \Omega_X$ and $\boldsymbol{v} \in \Omega_V$ and look for its maximum amongst the evaluated values, hence the multi-static AF is,

$$\mathcal{X}(\boldsymbol{x}, \boldsymbol{v}) := \sum_{q \in [Q]} w^q \left(\mathcal{X}^q(\boldsymbol{x}, \boldsymbol{v})\right)^2. \tag{3.4}$$

Alternatively, when we consider the simplified model of (1.45), we can separate this function into two functions: The location AF, denoted by $\mathcal{X}_X(\boldsymbol{x})$, associated to the simplified inner signal, hence only depending on $\boldsymbol{x}$, and the velocity AF, denoted by $\mathcal{X}_V(\boldsymbol{x}, \boldsymbol{v})$, associated to the outer signal, hence depending on both $\boldsymbol{x}$ and $\boldsymbol{v}$. We can interpret the actions of Alg. 7 with respect to the AFs. In the first step of an iteration of Alg. 7, we evaluate the location AF in all $\boldsymbol{x} \in \Omega_X$ and look for its maximum amongst the N_x evaluated values. Then, we select an index $n^*_x \in [N_x]$ and evaluate the velocity AF in $\boldsymbol{x} = \Omega_X \{n^*_x\}$, for all $\boldsymbol{v} \in \Omega_V$ to find an index $n^*_v \in [N_v]$.

In this section, we observe and mathematically describe the location and velocity AFs. According to all the above definitions, we have,

$$\mathcal{X}_X(\boldsymbol{x}) := \sum_{q \in [Q]} w^q \sum_{m_r \in [M_r]} \left(\sum_{m_s \in [M_s]} y^q[m_s, m_r] \psi^q_{\boldsymbol{x}}[m_s] \right)^2, \tag{3.5}$$

$$\mathcal{X}_V(\boldsymbol{v}) := \sum_{q \in [Q]} w^q \left(\sum_{m_r \in [M_r]} \left(\sum_{m_s \in [M_s]} y^q[m_s, m_r] \psi^q_{\boldsymbol{x}}[m_s] \right) \phi^q_{\boldsymbol{x}, \boldsymbol{v}}[m_r] \right)^2. \tag{3.6}$$

Applying the definition of the received signals and the definitions of the inner and outer signals provided in Sec. 3 .2 of Chap. 1, we get the following instrumental expressions,

$$\mathcal{X}_X(\boldsymbol{x}) \propto \sum_{k=1}^{K} \sum_{q \in [Q]} w^q |\alpha^q_k|^2 \left(\frac{\sin\left(\pi \frac{B}{c} \left(d^q_b(\boldsymbol{x}_k) - d^q_b(\boldsymbol{x})\right)\right)}{\sin\left(\pi \frac{B}{cM_s} \left(d^q_b(\boldsymbol{x}_k) - d^q_b(\boldsymbol{x})\right)\right)} \right)^2, \tag{3.7}$$

$$\mathcal{X}_V(\boldsymbol{x}, \boldsymbol{v}) \propto \sum_{k=1}^{K} \sum_{q \in [Q]} w^q |\alpha^q_k|^2 \left(\frac{\sin\left(\pi \frac{M_r f_0 T}{c} \left(v^q_b(\boldsymbol{x}_k, \boldsymbol{v}_k) - v^q_b(\boldsymbol{x}, \boldsymbol{v})\right)\right)}{\sin\left(\pi \frac{f_0 T}{c} \left(v^q_b(\boldsymbol{x}_k, \boldsymbol{v}_k) - v^q_b(\boldsymbol{x}, \boldsymbol{v})\right)\right)} \right)^2, \tag{3.8}$$

where $d_b^q(\boldsymbol{x})$ and $v_b^q(\boldsymbol{x}, \boldsymbol{v})$ are respectively defined in (1.43) and (1.44). Let us observe these functions for an arbitrary system with two TX nodes and two RX nodes located as in Tab. 3.1 and with the acquisition parameters of Tab. 3.2. All the nodes are composed of a single antenna. We evaluate $\mathcal{X}_X(\boldsymbol{x})$ in a location grid Ω_X obtained by uniformly sampling a square of length 30 meters, whose left-bottom corner is at the coordinate (1, -9), into a grid of 120x120 locations (See Fig. 3.1). We evaluate $\mathcal{X}_V(\boldsymbol{v})$ in a velocity grid Ω_V obtained by uniformly sampling a centered square of length 60 [m/s] into a grid of 120x120 velocities (See Fig. 3.2). There is a single *On-grid* target located in $\boldsymbol{x}_1 = (9.25, 2.50)$, moving with a velocity $\boldsymbol{v}_1 = (3.0, 0.0)$ and characterized by scattering coefficients $\alpha_1^q = 1$ for all $q \in [Q]$.

Nodes Locations	
TXs	**RXs**
(0, -2.5) [m]	(0, 2.5) [m]
(7.5, -10) [m]	(12.5, -10) [m]

Table 3.1: Locations of the transmit and receive nodes.

Parameter	B	f_0	T_s
Value	250MHz	24GHz	20μs
Parameter	M_s	M_r	T
Value	16	16	320μs

Table 3.2: Modulation and acquisition parameters.

Figure 3.1: Location grid of the example

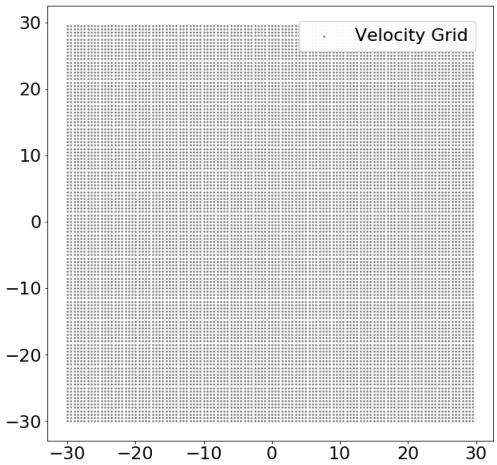


Figure 3.2: Velocity grid of the example

First, let us observe what happens if only TX1 is transmitting and only RX1 is receiving, i.e., we have a bistatic system. The corresponding location AF is displayed in Fig. 3.3. We observe several ellipses whose focal points are the locations of TX1 and RX1. It was expected by inspecting (3.12) because an ellipse with these focal points corresponds to a domain where $d_b^q(\boldsymbol{x})$ is constant. This location AF has a maximal value in $\boldsymbol{x} = \boldsymbol{x}_1$, as well as in every $\boldsymbol{x}$ such that $d_b^q(\boldsymbol{x}) = d_b^q(\boldsymbol{x}_1)$, which explains the main ellipse (i.e., on which the target is located). The equation (3.12) confirms the observed width of this ellipse, equal to $\frac{c}{B} = 1.2m$, this width is directly related to the resolution of the bistatic pair. It also confirms the secondary ellipses shown in the log scale figure. Finally, this equation also explains the repetition of the main ellipse. Indeed, because of the periodic nature of the function $\frac{\sin(x)}{\sin(Mx)}$, there is also a maximal value of the location AF in every $\boldsymbol{x}$ such that $d_b^q(\boldsymbol{x}) = d_b^q(\boldsymbol{x}_k) + aM_s\frac{c}{B}$, for all $a \in \mathbb{Z}$.

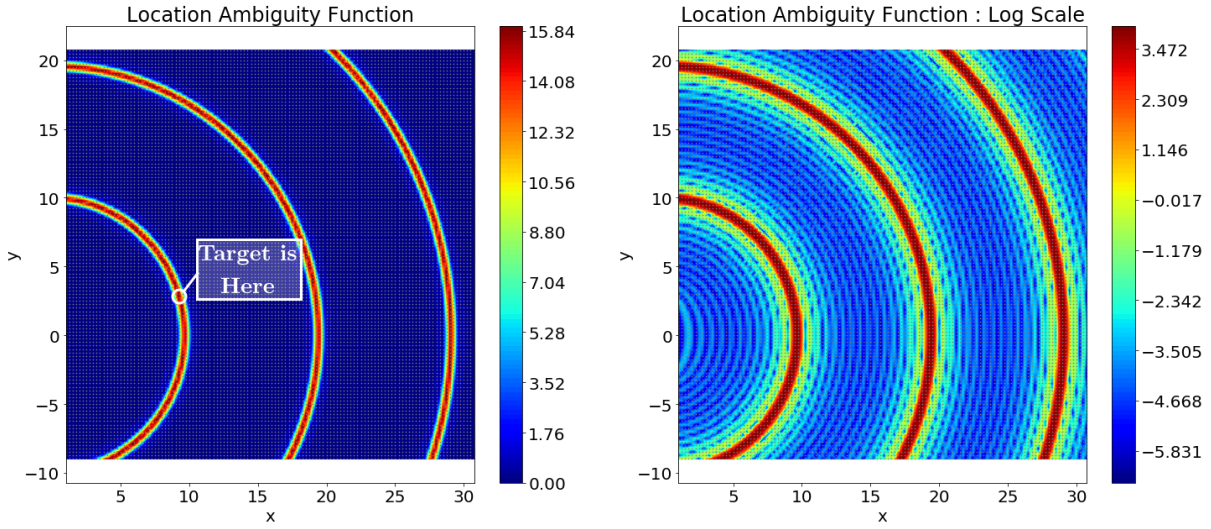


Figure 3.3: Location AF of a bistatic pair, evaluated in a discrete grid of locations defined by the uniform sampling of a square of length $L_x = 30m$ in the 2D location plane.

When the two TX nodes are transmitting and every RX nodes are listening to every TX nodes, yielding to $Q = 4$ pairs of communicating nodes, the location AF is a weighted sum of similar functions to the one shown in Fig. 3.3. Note that there is no noise, hence the weights w^q are all set to $\frac{1}{4}$, for all $q \in [Q]$. This location AF is shown in Fig. 3.4. In this example, the intersection of the ellipses almost enables the unambiguous location of the target. More precisely, we managed to locate the target in its ellipse, but the repetition of the main ellipses still generates ambiguity, such that there is tree maximums in our example. To avoid this problem, we have to select an *unambiguous set of location states*. While choosing an appropriate unambiguous set of bistatic range is a well known problem in the monostatic or bistatic Radars domain, the conditions for the definition an unambiguous location grid for multi-static systems is more challenging. Indeed, a grid may be ambiguous for one bistatic pair but unambiguous for the others, resulting in a globally unambiguous set. In this work, we decide to define a grid that is unambiguous for all bistatic pairs regardless of the location of the nodes, i.e., satisfying the following sufficient but unnecessary condition,

$$M_s \frac{c}{2B} \geq \max_{\boldsymbol{x},\boldsymbol{x}' \in \Omega_X} \|\boldsymbol{x} - \boldsymbol{x}'\|_2 \,. \tag{3.9}$$

Fig. 3.5 shows the same location AF evaluated in a location grid meeting the above condition. Finally, we can notice that the shape of the main lobe, which result from the intersection of ellipses, strongly depends on the way the ellipses intersect. Therefore, in addition to its dependence to $\frac{c}{2B}$, this lobe's width strongly depends on the location of the target and the locations of the TX and RX nodes, making difficult the establishment of a precise definition of the multi-static resolution. Still, the smallest possible lobe resulting from an ideal intersection has a circular shape of radius $\frac{c}{2B}$, suggesting that $\frac{c}{2B}$ is a lower bound of the resolution such as defined in the introduction. Moreover, because in practice the targets are not on grid, this circular shape of radius $\frac{c}{2B}$ suggests a minimal spacing between location states belonging to the location grid. For simplicity, from now in this document, we only consider uniformly sampled square grids. Let us assume the presence of

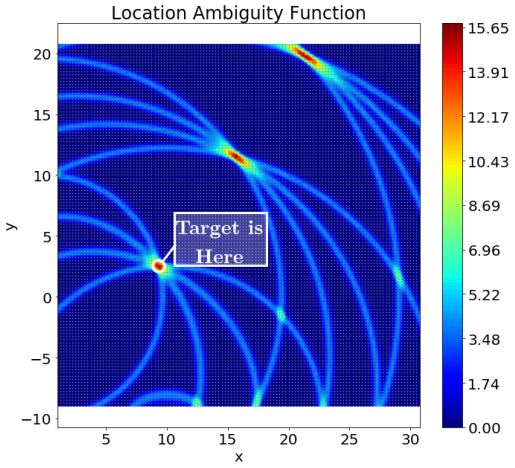


Figure 3.4: multi-static location Ambiguity Function evaluated in a discrete grid of locations defined by the uniform sampling of a square of length 30m in the 2D location plane. A single target is located in (9.25, 2.5).

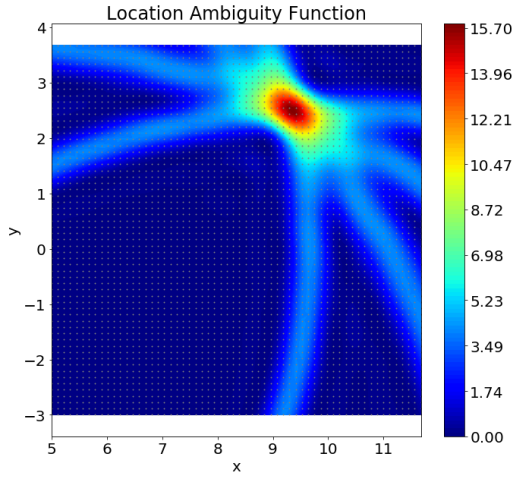


Figure 3.5: multi-static location Ambiguity Function evaluated in a canonical location grid.

a single *Off-Grid* target, if we want to be sure to detect the closest *On-Grid* location to the target's location, the spacing between two horizontally (or vertically) aligned consecutive locations of the grid needs to be smaller or equal to $\frac{c}{2\sqrt{2}B}$. From this conclusion and the condition (3.9), we introduce a *canonical location grid*. A canonical location grid is obtained from the uniform sampling of a square in the location 2D plane of length $L_x = M_s \frac{c}{2\sqrt{2}B}$ (called the canonical length of the location grid) into $M_s \times M_s$ locations. We call $\frac{M_s}{L_x}$ the canonical density of the location grid.

Fig. 3.6 shows the velocity AF evaluated for $\boldsymbol{x}$ fixed to the target's location. We observe straight lines instead of ellipses because the domain where $v_b^q(\boldsymbol{x}, \boldsymbol{v})$ is constant for all values of $\boldsymbol{v}$ and for $\boldsymbol{x}$ constant is a straight line. From (3.8), we know that these lines have a width equal to $\frac{c}{f_0 M_s T}$ m/s, yet similar difficulties as above prevent us to define a precise definition of the multi-static velocity resolution. Similarly to the location AF, we need to select an unambiguous velocity grid. In this work, we decide to define a grid satisfying the following sufficient but unnecessary condition,

$$\frac{c}{2f_0 T} \geq \max_{\boldsymbol{v}, \boldsymbol{v}' \in \Omega_V} \|\boldsymbol{v} - \boldsymbol{v}'\|_2 . \tag{3.10}$$

Fig. 3.7 shows the same velocity AF evaluated in velocities respecting the condition above. Similarly to the reasoning performed for the location grid, we introduce a *canonical velocity grid*. Such a grid is obtained from the uniform sampling of a square in the velocity 2D plane of length $L_v = \frac{c}{2\sqrt{2}f_0 T}$ (called the canonical length of the velocity grid) into $M_r \times M_r$ velocities. We call $\frac{M_r}{L_v}$ the canonical density of the velocity grid. Because there are no reason to give more attention to some directions of the velocity vector, the canonical velocity grid is always centered in $\boldsymbol{v} = (0, 0)$. An inherent property of this choice is

$$\|\boldsymbol{v}\|_2 \leq \frac{c}{4f_0 T}, \quad \forall\, \boldsymbol{v} \in \Omega_V . \tag{3.11}$$

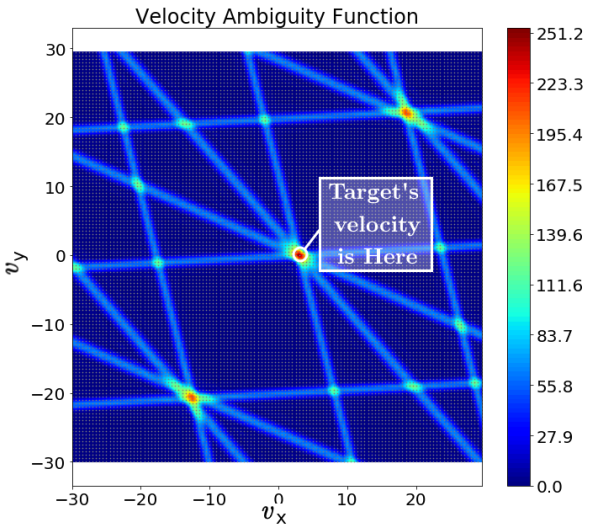


Figure 3.6: multi-static velocity Ambiguity Function evaluated in a discrete grid of velocities defined by the uniform sampling of a square of length 60m/s in the 2D velocity plane. A single target has a velocity: (3.0, 0.0) [m/s].

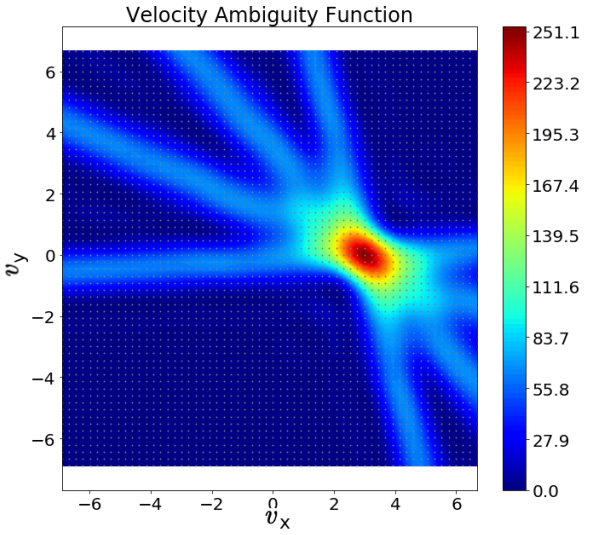


Figure 3.7: multi-static velocity Ambiguity Function evaluated in the canonical velocity grid

When the receivers are composed of a linear antenna array, the velocity AF remains unchanged, but the location AF is affected by a factor providing an angle of arrival resolution. As a rough estimation, we can express the multi-static SIMO location as follow:

$$\mathcal{X}_X(\boldsymbol{x}) \propto \sum_{k=1}^{K} \sum_{q\in[Q]} w^q |\alpha_k^q|^2 \left(\frac{\sin\left(\pi \frac{B}{c}\left(d_b^q(\boldsymbol{x}_k) - d_b^q(\boldsymbol{x})\right)\right)}{\sin\left(\pi \frac{B}{cM_s}\left(d_b^q(\boldsymbol{x}_k) - d_b^q(\boldsymbol{x})\right)\right)} \right)^2 \left(\frac{\sin\left(\pi \frac{f_0 N_R^q}{c} \delta_a^q(\boldsymbol{x})\right)}{\sin\left(\pi \frac{f_0}{c} \delta_a^q(\boldsymbol{x})\right)} \right)^2, \tag{3.12}$$

where,

$$\delta_a^q(\boldsymbol{x}) = \delta^q \frac{\langle \boldsymbol{A}_r^q, (\boldsymbol{x} - \boldsymbol{X}_r^q)\rangle}{\|\boldsymbol{x} - \boldsymbol{X}_r^q\|_2}, \tag{3.13}$$

where δ^q stands for the spacing between two consecutive antennas of the array, $\boldsymbol{A}_r^q$ is the second row of the 2D rotation matrix of angle β^q defined as in Fig. 3.8.

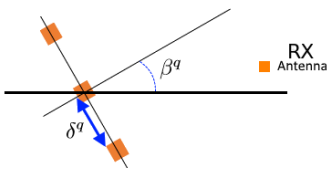


Figure 3.8: Schematic illustration of an antenna array's parameters

Let us consider that RX1 is now an array of three antennas aligned along the y-axis and RX2 is an array of three antennas aligned along the x-axis. In the presence of the same target as above, the location AF of the TX1-RX1 bistatic pair is depicted in Fig.

3.9, while the complete location AF is depicted in 3.10. Very interestingly, although the periodic nature of the location AF is still a matter for a given bistatic pair, it is no longer the case in the multi-static SIMO configuration, allowing us to observe much larger location domain than bistatic Radars can afford. This is a remarkable advantage of joint modeling over the disjoint model based methods for multi-static Radars. A full example of application of FBMP in the presence of 3 moving targets characterized by random scattering coefficients is provided in Appendix C .

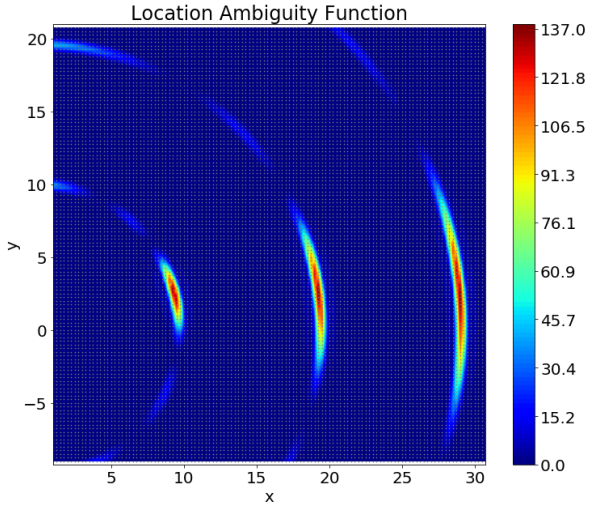


Figure 3.9: bistatic location Ambiguity Function evaluated in a discrete grid of locations defined by the uniform sampling of a square of length 30m in the 2D location plane. The receivers are arrays of three antennas.

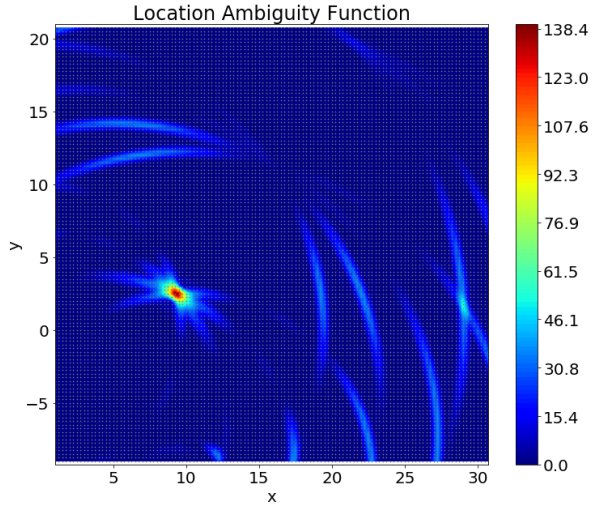


Figure 3.10: multi-static location Ambiguity Function evaluated in a discrete grid of locations defined by the uniform sampling of a square of length 30m in the 2D location plane. The receivers are arrays of three antennas.

3 Model Mismatches

In Chap. 1, we performed two simplifications. First, we neglected the coupling signal $\theta[m_s, m_r]$, this assumption has no significant effect because it is an extremely good approximation (see Appendix B). The second simplification, performed in (1.30), neglected the Doppler effect inside a ramp of the chirp modulation. This simplification has significant effects on the targets' states estimation, due to a mismatch between the received signals and the dictionaries used by our algorithms. More precisely, if we consider a the target located in $\boldsymbol{x}_1 = \Omega_X \{n_x^*\}$, and moving with a velocity $\boldsymbol{v}_1 = \Omega_V \{n_v^*\}$, the rows of the received matrices $\boldsymbol{Y}^q$ will not perfectly match $\boldsymbol{\psi}^q_{n_x^*}$. With respect to the simplified definition of the inner signal, the target will appear as if it was located in

$$\boldsymbol{x}_{shifted} = \boldsymbol{x}_1 + \frac{f_0 M_s T_s}{B} \boldsymbol{v}_1, \tag{3.14}$$

which is likely an *Off-Grid* location. To analyze the effect of this mismatch, let us consider the system we described above (Tab. 3.1 and Tab. 3.2) and the associated canonical grids shown in Fig. 3.12 and 3.13. In this section, we will observe the impact of the model mismatches on the estimation of the target location for different location

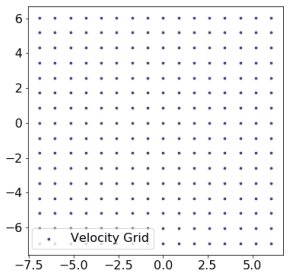


Figure 3.12: Canonical Velocity Grid. The grid density is canonical (16×16 samples)

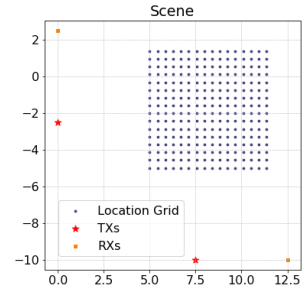


Figure 3.13: Canonical Location Grid. The grid density is canonical (16×16 samples)

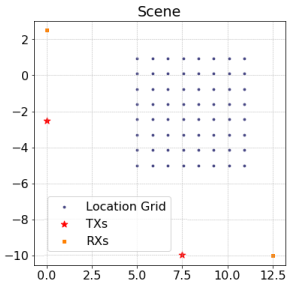


Figure 3.14: Location Grid of canonical length and with grid density equal to half the canonical one (8×8 samples)

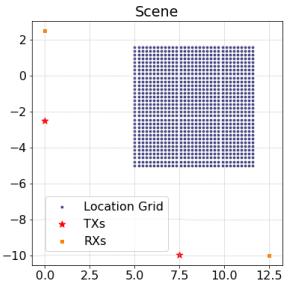


Figure 3.15: Location Grid of canonical length and with grid density equal to twice the canonical one (32×32 samples)

grid densities. More precisely, we select a square grid of canonical length and modify its density. Examples are given in Fig. 3.14 and 3.15. Because we chose the canonical velocity grid, we can define the bounds on the location shift:

$$\boldsymbol{x}_{shifted} - \boldsymbol{x}_1 \in \left[-\frac{M_s T_s}{T} \frac{c}{4\sqrt{2}B}, \; +\frac{M_s T_s}{T} \frac{c}{4\sqrt{2}B} \right]. \tag{3.15}$$

In our example, $M_s T_s = T$, therefore we have,

$$\boldsymbol{x}_{shifted} - \boldsymbol{x}_1 \in \left[-\frac{c}{4\sqrt{2}B}, \; +\frac{c}{4\sqrt{2}B} \right]. \tag{3.16}$$

This means that the shift will be observed, i.e., by picking a wrong bin of the location grid, only when we are using a location grid denser than the canonical one. This is illustrated in Fig. 3.11.

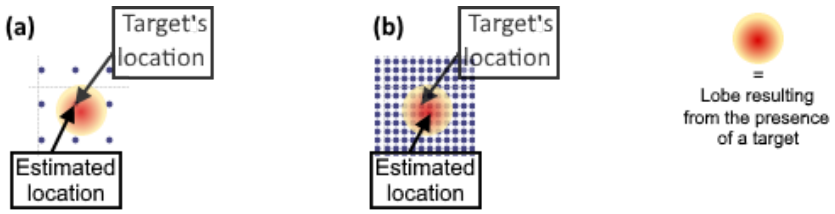


Figure 3.11: Rough Schematic illustration of the sampling of the location AF when a single targets generates a lobe in the AF. **(a)**, when the grid is of lower density than the canonical grid. **(b)**, when the grid is of higher density than the canonical grid.

In the following (noiseless) simulations, we tested several location grids, from a density of 4 to 40. For each location grid, we selected 20,000 times a single target with a random location in Ω_X and a random velocity in Ω_V. The scattering coefficients, hence the non-zero values of the sparse matrices $\boldsymbol{S}^q$ are all set to 1. In Fig. 3.16, we display the *Location Miss Rate*, i.e., the percentage of times the algorithm selected a wrong location against the density of the location grid. In Fig. 3.17, we display the *Average Location Estimation Error*. This metric is the average over all the 20,000 simulation of the location estimation error defined as,

$$\frac{\|\hat{\boldsymbol{x}}_1 - \boldsymbol{x}_1\|_2}{L_x}, \tag{3.17}$$

where $\hat{\boldsymbol{x}}_1$ is the estimate provided by Alg. 7. As we expected, the miss rate increases when the grid is denser. Indeed, denser grids manage to observe smaller shifts, hence this miss rate tends toward 100% when the grid density increases to infinity. The evolution of the average location estimation error with respect to the grid density clearly shows two regions: before the canonical density, where the shift is never seen an thus the target's location is always correctly estimated, and after the canonical density when the shift tends to always be observed.

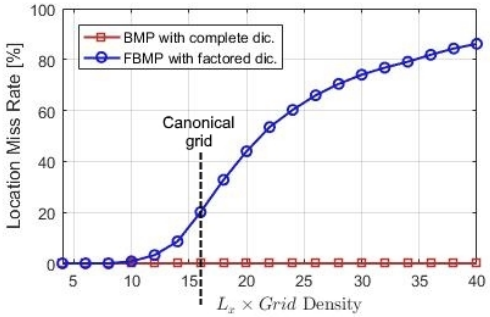


Figure 3.16: Location Miss Rate vs location grid density in the presence of single target.

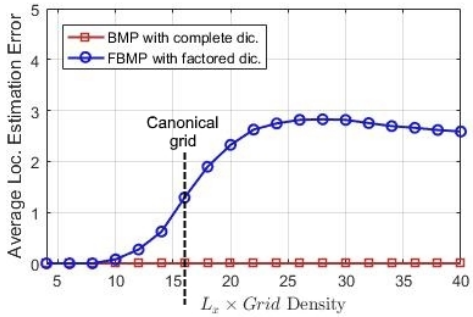


Figure 3.17: Average Location Estimation Error vs location grid density in the presence of single target.

Note that getting to the conclusion that choosing a low-density grid is better is foolish. Indeed, the above results are mainly due to the assumption of *On-Grid* targets. For low-density grids, this assumption is very strong and unrealistic. Moreover, remind that in practice, we should always select grids at least as dense as the canonical grid.

Fig. 3.18 and 3.19 show the same simulation where the Q scattering coefficients are selected such that α^q i.i.d $\sim \mathbb{C}\mathcal{N}(0,1)$ for $q \in [Q]$. This situation is much more consistent with the multi-static definition proposed in the introduction. In this situation, as shown by the miss rate curve, errors may also occur with low-density grids. This happens typically when one scattering coefficient has a much higher norm than the others. Such a situation is shown in Fig. 3.20 and 3.21, where the location grid is displayed on top of the continuous location AF. In this example, both high density grid and low density grid fail to located the target, but the higher density grids exhibit lower average location estimation errors. To sum up, in noiseless conditions, the model mismatches causes large but rare errors when we work with a location grid of density lower than the canonical density, and causes small but frequent errors when we work with a location grid of density higher than the canonical density.

4 Iterative Correction

The key simplification performed in (1.30) enables the factored expression of the model, but in return causes the model mismatches we studied in the previous section. However, it

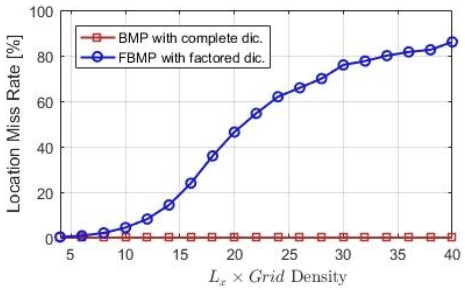


Figure 3.18: Location Miss Rate vs location grid density in the presence of single target. Scattering coefficients are randomly selected.

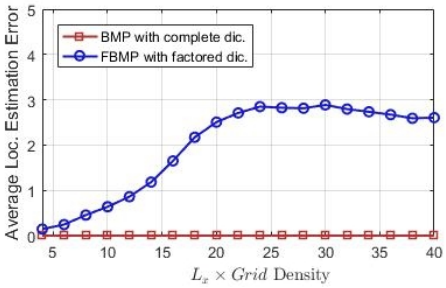


Figure 3.19: Average Location Estimation Error vs location grid density in the presence of single target. Scattering coefficients are randomly selected.

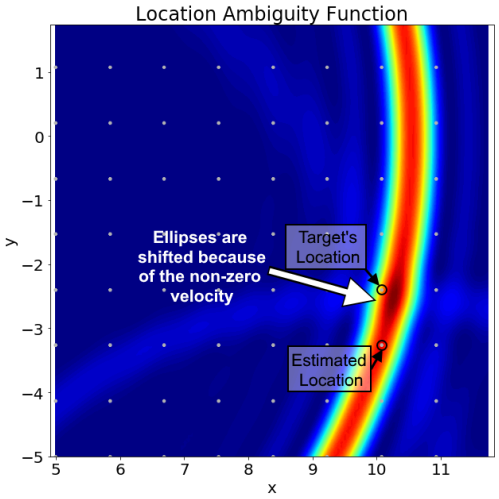


Figure 3.20: Continuous location AF sampled with a low-density location grid

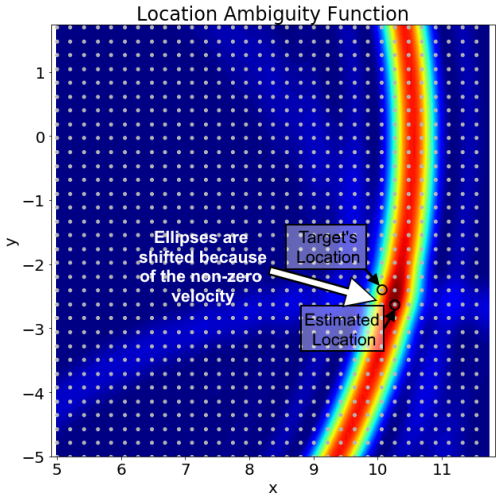


Figure 3.21: Continuous location AF sampled with a high-density location grid

is possible to counteract this mismatch with the addition of an iterative correction inside each iteration of Alg. 7. To understand how we can define such a correction, let us inspect the signals expression we would get without this simplification:

$$y^q[m_s, m_r] \simeq \sum_{k\in[K]} \alpha_k \psi^q_{\boldsymbol{x}_k,\boldsymbol{v}_k}[m_s]\phi^q_{\boldsymbol{x}_k,\boldsymbol{v}_k}[m_r] + e^q[m_s, m_r] \quad \forall\, q \in [Q]\,. \tag{3.18}$$

As done before, for each $q \in [Q]$, we can define a set of N_x different outer dictionaries: $\boldsymbol{\Phi}^q_{n_x} \in \mathbb{C}^{M_r \times N_v}$ whose n_v-th atom is $\phi_{\Omega_X\{n_x\},\Omega_V\{n_v\}}[m_r]$. Because we do not have the convenient simplification of the inner dictionary, let us similarly define a set of N_v different inner dictionaries: $\boldsymbol{\Psi}^q_{n_v} \in \mathbb{C}^{M_s \times N_x}$ whose whose n_x-th atom is $\psi_{\Omega_X\{n_x\},\Omega_V\{n_v\}}[m_s]$. The simplification of the inner dictionary we did in (1.30) corresponds to arbitrarily select $\boldsymbol{\Psi}^q = \boldsymbol{\Psi}^q_{n_v^0}$, where n_v^0 is the index of the $\boldsymbol{v} = \boldsymbol{0}$ in Ω_V, assuming that the zero velocity belongs to the velocity grid.

The idea is the following: After the third iteration of Alg. 7, we have a first guess of n_x^* and n_v^*. From this guess, we start over the iteration but using $\boldsymbol{\Psi}^q = \boldsymbol{\Psi}^q_{n_v^*}$ as the inner dictionary instead of the previously arbitrarily chosen inner dictionary. We can perform it over and over again until we reach a maximal number of try or until two consecutive guesses

of $\{n_x^*, n_v^*\}$ are identical. A schematic illustration of the FMP (i.e., with $Q = 1$, for the simplicity of the drawing) with the iterative correction is shown in Fig. 3.22. This algorithm described in this figure can easily be extended to $Q > 1$ similarly to what we did in Sec. 3 .6.

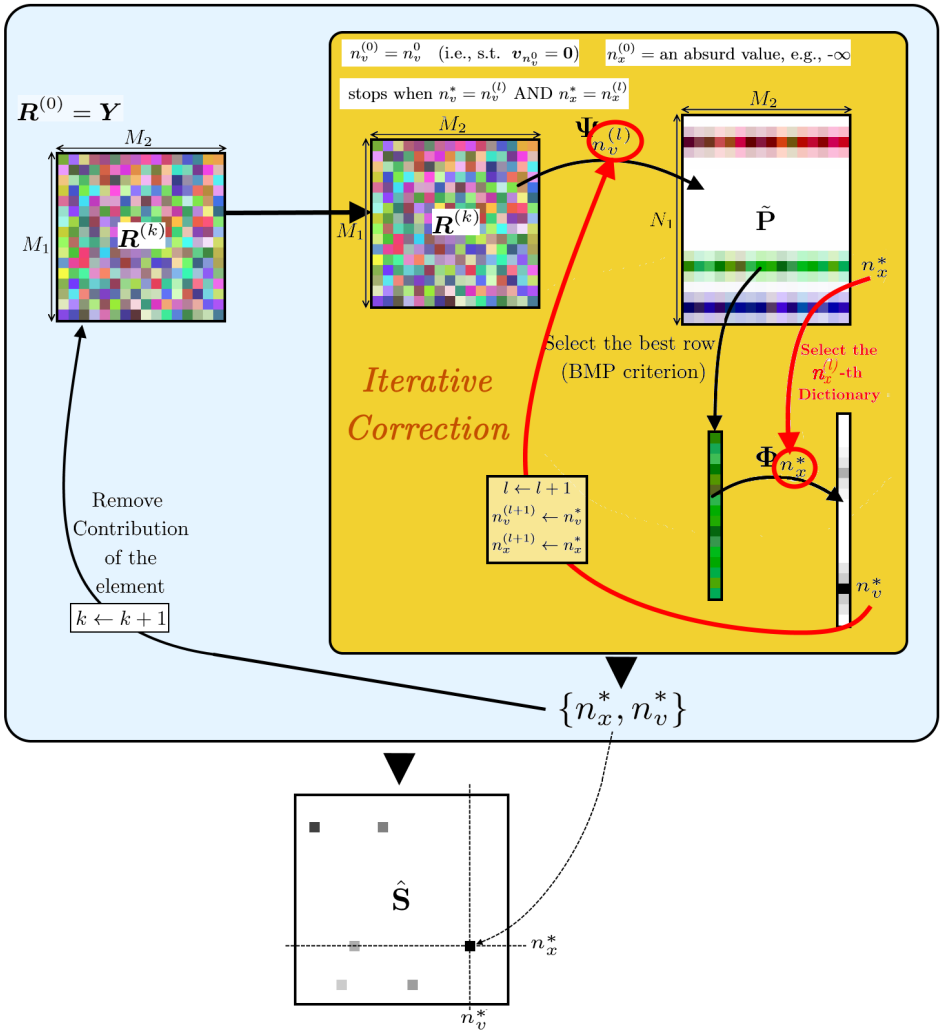


Figure 3.22: Schematic illustration of FMP with the iterative correction.

The figure 3.23 shows a comparative example of the discretized location AFs we obtain at the first (i.e., the first guess) and at the third corrective iteration. We clearly see that by processing the corrective iterations, the algorithm finds the correct dictionaries $\boldsymbol{\Psi}_{n_v}^q$ for $q \in [Q]$ matching the received signals.

Fig. 3.24 and 3.25 show, respectively, the location miss rate and the average location estimation error against the location grid density. In these figures we observe that a very few number of iterations is sufficient to completely counteract the model mismatches analyzed above. Finally, let us analyze the complexity of the FBMP with iterative correction (FBMP-ic). If N_{it} denotes the maximal number of corrective iterations we perform (counting the first guess of n_x^* and n_v^* as an iteration), then the complexity of one iteration of FBMP-ic is upper-bounded by $\mathcal{O}\left(N_{it}Q(M_rM_sN_x + M_rN_v)\right)$.

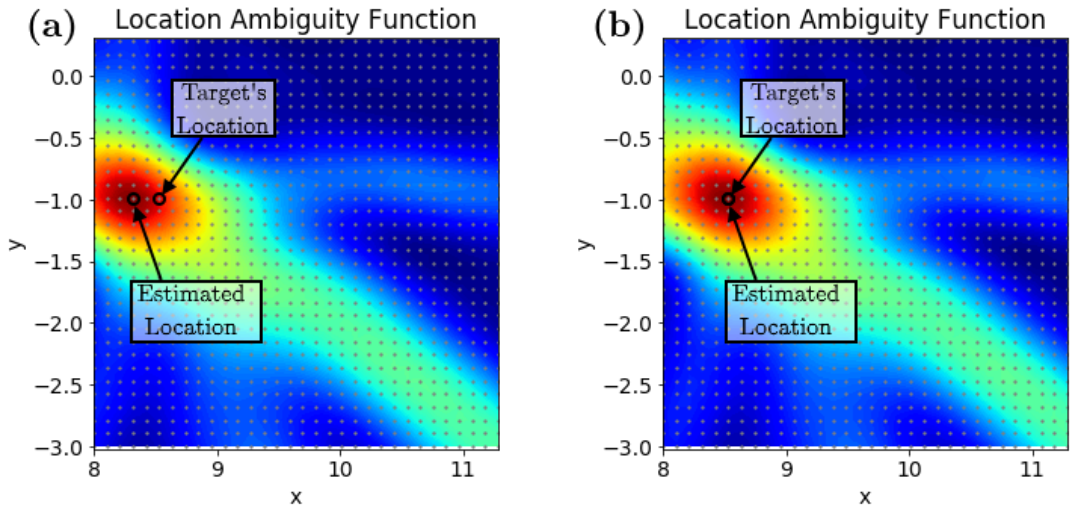


Figure 3.23: Example of location AF in the presence of a single target moving with a velocity (-6.9, 0). **(b)** First iteration of the correction, **(b)** third iteration of the correction.

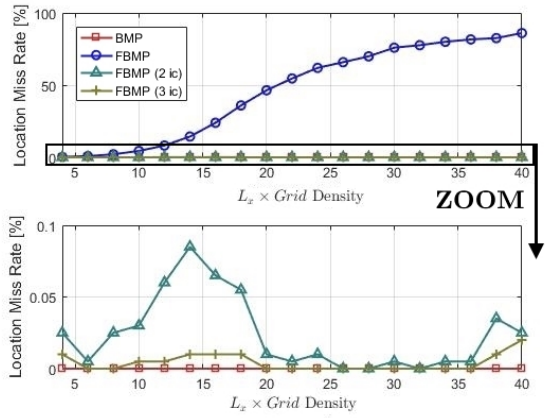


Figure 3.24: Location Miss Rate vs location grid density in the presence of single target. Comparing different number of iterative corrections.

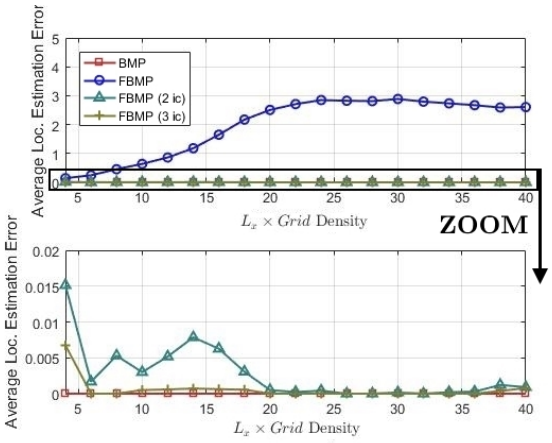


Figure 3.25: Average Location Estimation Error vs location grid density in the presence of single target. Comparing different number of iterative corrections.

Chapter 4

Performance Analysis

The final step of our work is the validation of the models the algorithms we established throughout this whole document. First, we study the algorithmic performances in terms of computational complexity and quality of target states estimation. This analysis is obtained by performing extensive simulations of multi-static Radar measurements. The purpose is to compare the algorithms between one another, but also to compare different choice of Radar configurations (typically, multi-static SISO, multi-static MIMO and bistatic SISO). More precisely, the computational complexity is evaluated against the grids' densities in order to validate the theoretical complexity we proposed, and the quality of estimation is evaluated using several metrics introduced below. Next we study the validity of our models and algorithms on actual multi-static Radars measurements. More precisely, we developed a simple measurement test bench acquired signals in the presence of one or several targets.

1 Theoretical Analysis

Unless otherwise stated, the system we simulated is composed of the nodes described in Tab. 4.1 and a modulation described in Tab. 4.2 (i.e., the same we used in Chap. 3). The default location and velocity grids we use are the canonical grids shown in Fig. 4.1 and 4.2. Most results shown in this chapter display the metrics described below against the Signal to Noise Ratio (SNR). Based on the complete multi-static signal model (1.54), when there are K targets (i.e., all $\boldsymbol{s}^q$ are K-sparse for $q \in [Q]$, we can define the SNR as

$$\mathrm{SNR} = \frac{1}{K} \frac{\mathbb{E}\left(\sum_{q\in[Q]} \|\boldsymbol{D}^q \boldsymbol{s}^q\|_2^2\right)}{\mathbb{E}\left(\sum_{q\in[Q]} \|\boldsymbol{e}^q\|_2^2\right)} = \frac{1}{K} \frac{\sum_{q\in[Q]} M^q \mathbb{E}\left(\|\boldsymbol{s}^q\|_2^2\right)}{\sum_{q\in[Q]} M^q \sigma_q^2}. \tag{4.1}$$

In this chapter, we always set the same noise power for all received signals, hence $\sigma_q^2 = \sigma_e^2$ for all $q \in [Q]$. Moreover, using the definition of $\boldsymbol{s}^q$ in (1.49), we know that $\mathbb{E}\left(\|\boldsymbol{s}^q\|_2^2\right) = K$ for all $q \in [Q]$, hence

$$\mathrm{SNR} = \frac{1}{\sigma_e^2}. \tag{4.2}$$

Finally, unless otherwise stated, every dot of every curve plotted in this chapter results from the average of the quantity being plotted over 20,000 simulations.

Nodes Locations	
TXs	**RXs**
(0, -2.5) [m]	(0, 2.5) [m]
(7.5, -10) [m]	(12.5, -10) [m]

Table 4.1: Locations of the transmit and receive nodes.

Parameter	B	f_0	T_s
Value	250MHz	24GHz	20μs
Parameter	M_s	M_r	T
Value	16	16	320μs

Table 4.2: Modulation and acquisition parameters.

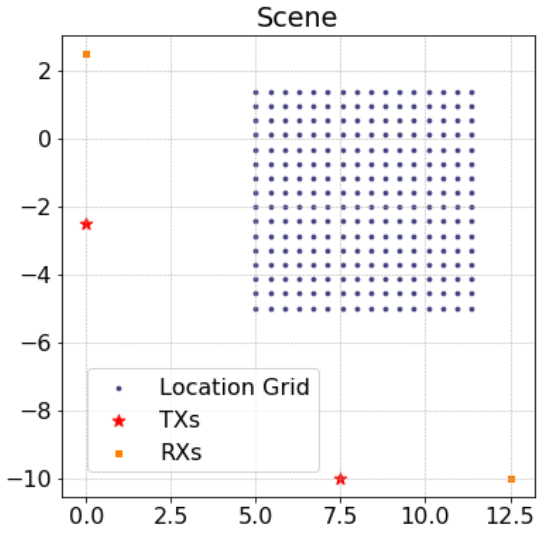


Figure 4.1: Canonical Location Grid. The grid density is canonical (16×16 samples)

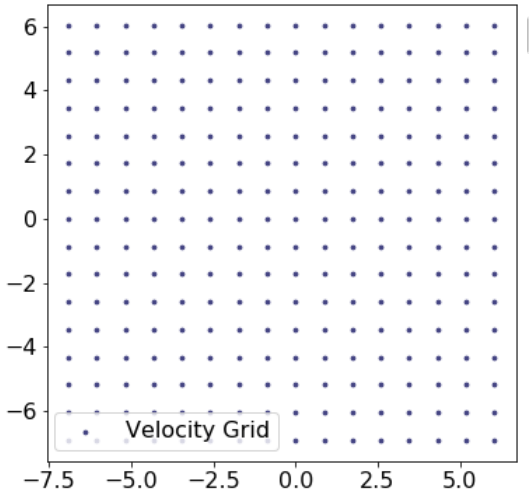


Figure 4.2: Canonical Velocity Grid. The grid density is canonical (16×16 samples)

1 .1 Definition of the performance metrics

To define the performances of the system, we use five metrics. Some of them have already been defined and used in Chap. 3, yet let us define them all here. The two first metrics are the location miss rate and the velocity miss rate. The location (resp. velocity) miss rate is defined as the ratio of simulated targets whose location (resp. velocity) has not been perfectly estimated. The miss rate is equal to 1 when the algorithm always fails and 0 if the algorithm always succeeds. Note that the miss rate is not always equal to 1 because we are assuming *On-Grid* targets.

The two next metrics are the average location estimation error and the average velocity estimation error. To compute these metrics, we compute the average over the 20,000 simulations of the location and velocity estimation errors defined as,

$$\begin{aligned} \text{Location estimation error} &= \frac{1}{K} \sum_{k \in [K]} \frac{\|\hat{\boldsymbol{x}}_k - \boldsymbol{x}_k\|_2}{L_x}, \\ \text{Velocity estimation error} &= \frac{1}{K} \sum_{k \in [K]} \frac{\|\hat{\boldsymbol{v}}_k - \boldsymbol{v}_k\|_2}{L_v}, \end{aligned} \tag{4.3}$$

Note that, when we are simulating multiple targets at the same time, in order to have a proper computation of these metrics, the estimators are sorted in a way that gives

the minimal value of the estimation errors, in order to somehow find which estimator corresponds to which target. Finally, the last metric is the reconstruction error, defined as,

$$\text{Reconstruction error} = \frac{1}{Q}\sum_{q\in[Q]}\frac{\|\boldsymbol{Y}^q - \boldsymbol{D}^q\hat{\boldsymbol{s}}^q\|_2^2}{\|\boldsymbol{Y}^q\|_2^2}. \tag{4.4}$$

1 .2 Comparison between BMP and FMP

Complexity and Memory

The first simulation aims to show the effectiveness of the application of FBMP (possibly with the iterative correction) to multi-static Radars with respect to the application of BMP, in term of computational complexity. We recorded the evolution of the average computation time of each algorithm with the sizes of the grids (i.e., N_x and N_v). More precisely, we set $N_x = N_v = \xi^2$. Therefore, BMP is expected to have complexity $\mathcal{O}(\xi^4)$ while FBMP/FBMP-ic is expected to have a complexity $\mathcal{O}(N_{it}\xi^2)$. These predictions are confirmed by the simulations depicted in Fig. 4.3. Note that all dictionaries were computed and stored before starting the simulations (i.e., all $\boldsymbol{D}^q$ for all $q \in [Q]$ for the BMP. All $\boldsymbol{\Psi}^q_{n_v}$ for all $n_v \in [N_v]$, all $\boldsymbol{\Phi}^q_{n_x}$ for all $n_x \in [N_x]$ for all $q \in [Q]$ for FBMP-ic). Also note that the memory requirement to store the $\boldsymbol{D}^q$ dictionaries is much higher the the memory requirement to store all the $\boldsymbol{\Psi}^q_{n_v}$ and $\boldsymbol{\Phi}^q_{n_x}$ combined, which is another advantage of FBMP.

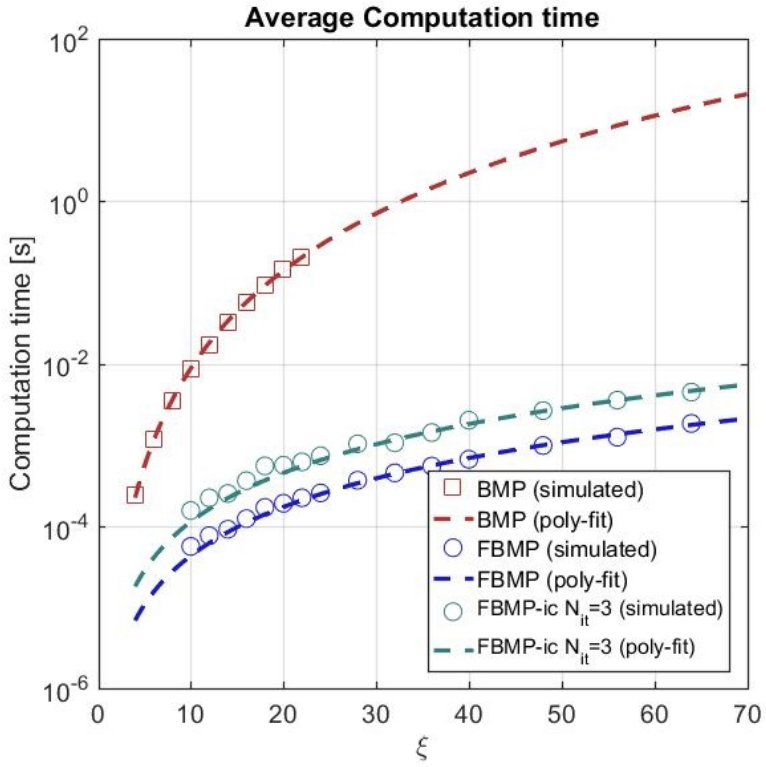


Figure 4.3: Average computation time of BMP, FMP and FMP-ic against the sizes of the grids. The polynomial fit of the BMP comp. time is $8.68e-07\xi^4$ [s]. The polynomial fit of the FBMP comp. time is $4.28e-07\xi^2$ [s]. The polynomial fit of the FBMP-ic comp. time is $1.12e-06\xi^2$ [s].

Robustness against noise

Fig. 4.4 shows the evolution of the five performance metrics with respect to the SNR in the presence of a single random target[1] and compares the FBMP, FBMP-ic and BMP. The reconstruction errors is of course improving when the SNR increases, but interestingly, there is no significant difference between the algorithms regarding this performance metric. The four other metrics, more related to the support recovery capability of the algorithm, exhibit significant difference between the algorithms. The miss rate and estimation error curves of FBMP quickly tend toward a fairly non-zero value for high values of SNR. This was expected because we saw in Chap. 3 that without the iterative correction, the model mismatches lead to significant estimation errors even in noiseless conditions. In the FBMP-ic curves, this saturation no longer occurs because the model mismatches have been compensated for, yet we still observe an SNR gap between the FBMP-ic curves and the FMP curve. This gap corresponds to the gap we studied in Sec. 3 .4 of Chap. 2, when we compared FMP and MP. This gap is the price to pay to benefit from the lower complexity demonstrated in Fig. 4.3.

Fig. 4.5 shows the same simulations performed in the presence of three random distinct targets. This time, we can see that all agorithms tend toward a non-zero value for high values of SNR, indicating the presence of errors in noiseless condition. Indeed, in the presence of multiple targets, in the location AF, the intersection of two ellipses coming from different targets may (yet unlikely) be stronger than the intersection of four ellipses. Fig. 4.6 shows a noiseless example when this problem does not occur while Fig. 4.7 shows the same noiseless example with different values of the scattering coefficients such that this problem occurs. This is an artefact of the non-orthogonality of the dictionaries. In fact, this is a perfect illustration of the non-guarantee of reconstruction of the MP algorithm we discussed in Sec. 2 .1 of Chap. 2. Developing *Orthogonal* versions of FBMP (i.e., based on the existing OMP and BOMP), is likely to provide far better results regarding this issue. Note that similar issue occurs in the trilateration/triangulation methods and is known as the *ghost targets* problem.

To sum up, the BMP and FBMP-ic have an equivalent reconstruction performance but BMP has a better support recovery performance. In the presence of several targets, their respective contributions to the location AF interfere with one another and may cause the detection of ghost targets.

1 .3 Comparison of radar configurations

In this section, we study the impact of the number of bistatic pairs Q, set to 4 up to now, and the number of receiving antennas in the RX nodes, set to $N_R^q = 1$ for all $q \in [Q]$ up to now, on the FBMP-ic algorithm. From now, the only algorithm we only simulate is FBMP-ic with $N_{it} = 3$. The evolution of the same five metrics is plotted against the SNR for different multi-static SIMO Radars configurations. Arbitrarily, we always set all the receivers with the same number of antennas (i.e., $N_R^q = N_R$ for all $q \in [Q]$. A configuration is referred to with the short name $Q(1 \times N_R)$. The nodes are the same as in Fig. 4.1.

[1]Selected uniformly on the grids and with scattering coefficients $\alpha^q \overset{\text{i.i.d.}}{\sim} \mathbb{C}\mathcal{N}(0,1) \quad \forall\, q \in [Q]$.

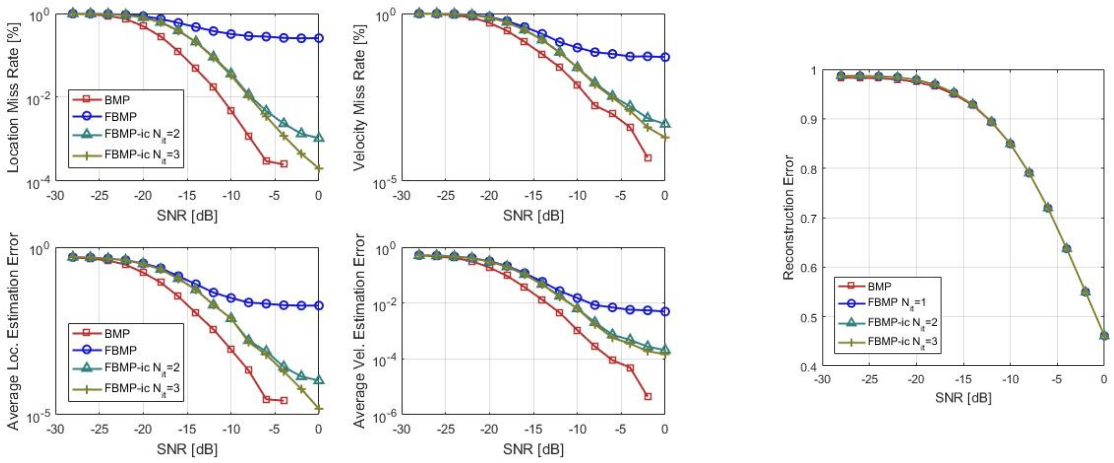


Figure 4.4: Evolution of the miss rates, estimation errors and reconstruction error with respect to the SNR, in the presence of a single target.

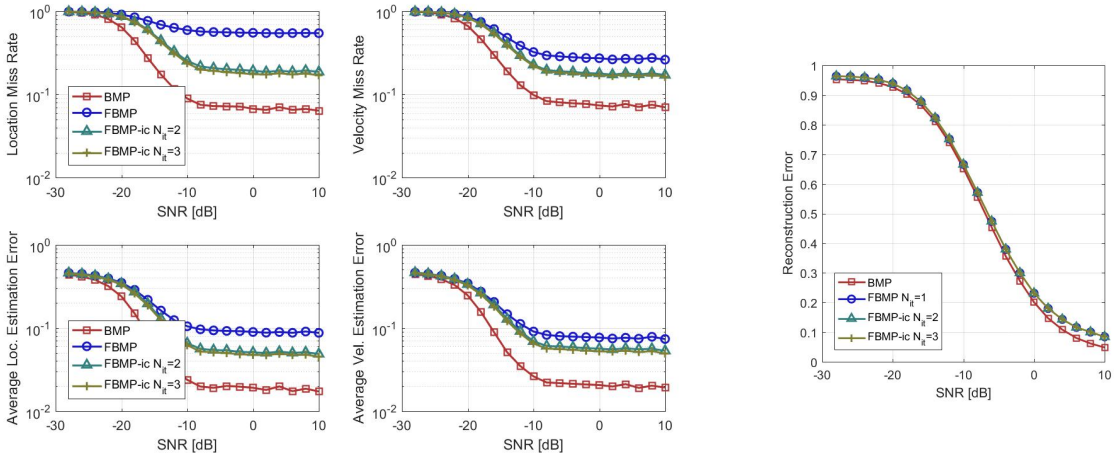


Figure 4.5: Evolution of the miss rates, estimation errors and reconstruction error with respect to the SNR, in the presence of three targets.

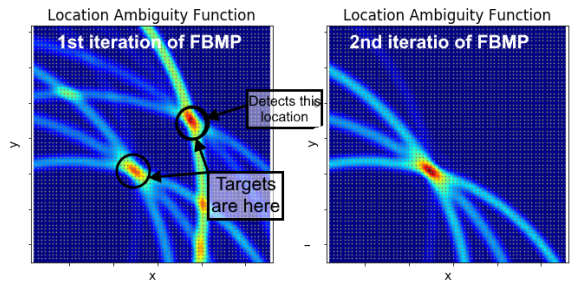


Figure 4.6: Schematic example of location AF obtained presence of 2 targets, with convenient scattering coefficients

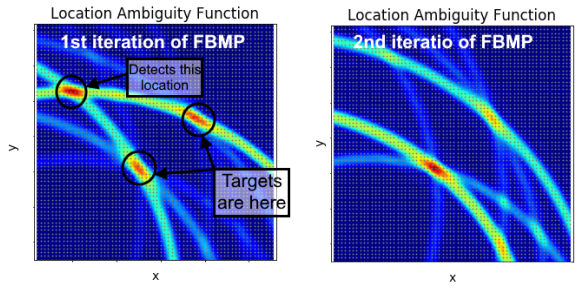


Figure 4.7: Schematic example of location AF obtained in the presence of 2 targets, with scattering coefficient such that an intersection of 2 ellipses is stronger than an intersection of 4 ellipses.

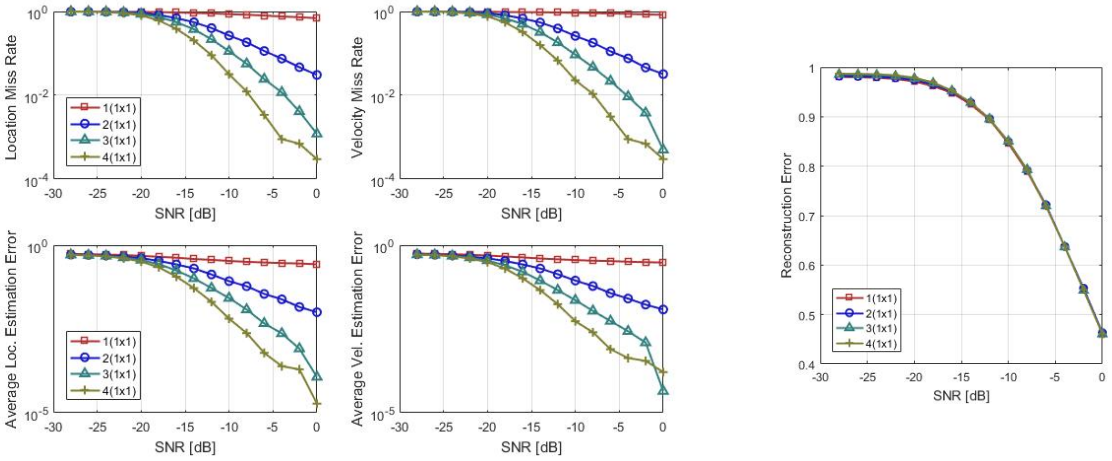


Figure 4.8: Evolution of the miss rates, estimation errors and reconstruction error with respect to the SNR, in the presence of a three targets. Comparison of different multi-static SISO configurations $Q(1 \times 1)$.

Fig. 4.8 compares multi-static configurations wtih different number of bistatic pairs, in the presence of a single target. When $Q = 2$ the 2 bistatic pairs are TX1-RX1 and TX1-RX2, when $Q = 3$, we add TX2-RX1. We see that the slope of the evolution of all metrics with respect to the noise increases with the number of bistatic pairs. This is the *Divsersity Gain* we get by viewing target through different angle. More precisely, this result comes from the fact that the Q scattering coefficients characterizing the target, α^q for all $q \in [Q]$, are all i.i.d. according to our model. Therefore, the probability that at least two of them[2] are large with respect to noise increases when the number of pair increases. Note that when $Q = 1$, the performances are terrible because the system is no longer able to see intersections of ellipses in the location AF. Actually, this is equivalent to the statement we gave in the introduction according to which a bistatic Radar is not able to detect more than one component of the location vector of a target.

Fig. 4.9 compares multi-static SIMO configurations with $Q = 4$ and different number of receiving antennas per RX node, in the presence of a single target. This time, we do not observe any change in the slope of the curves but we can see a shift to the left when the number of RX antenna per array increases. This is the *Array Gain* we get by viewing several times the target through the same angle. More precisely, this results comes from the assumption that all the scattering coefficients of the paths of a given MIMO pair are the same, allowing us to add the associated signals coherently. This shift is the same as if we acquired a higher number of samples M. However, one may wonder what role does the angular resolution provided by the antenna arrays play. When $Q = 4$ and in the presence of a single target, the answer in *none*, because we do not need this resolution to locate the intersection of the ellipses. Yet, in the presence of several target, the angular resolution greatly helps to avoid the ghost targets problem. Fig. 4.12 shows the same comparison in the presence of three targets. Not only the same shift as explained above occurs but also the high SNR values the metrics tend to are much smaller when we increase the number of RX antennas per node.

[2]At least two because we need an intersection of ellipses

Finally, one may wonder what is the best configuration to choose from a given number of antennas. We tried to answer this question by arbitrarily assuming we dispose of one TX antenna and 8 RX antennas. We simulated 4 configurations represented in Fig. 4.11. In the first configuration, we grouped all the 8 antennas in one RX node ($1(1 \times 8)$), in the second, we grouped it into 2 nodes of 4 antennas ($2(1 \times 4)$), in the third, we grouped it into 4 nodes of 2 antennas ($4(1 \times 2)$), finally the fourth one is composed of 8 nodes of 1 antenna ($8(1 \times 1)$). The results are displayed in Fig. 4.12. Interestingly, while the $4(1 \times 2)$ and $8(1 \times 1)$ configurations are almost equivalent regarding the estimation of target location while the $8(1 \times 1)$ exhibits better performances for the velocity estimation. That is, because the velocity vector needs to be seen from different angle to ensure a proper estimation, which is therefore less affected by the addition of co-located antennas.

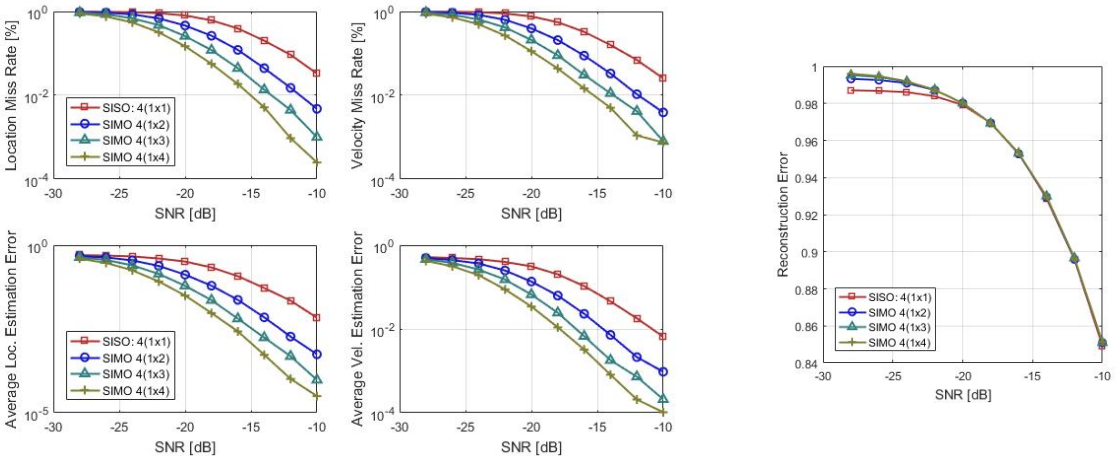


Figure 4.9: Evolution of the miss rates, estimation errors and reconstruction error with respect to the SNR, in the presence of a three targets. Comparison of different multi-static SIMO configurations $4(1 \times N_R)$.

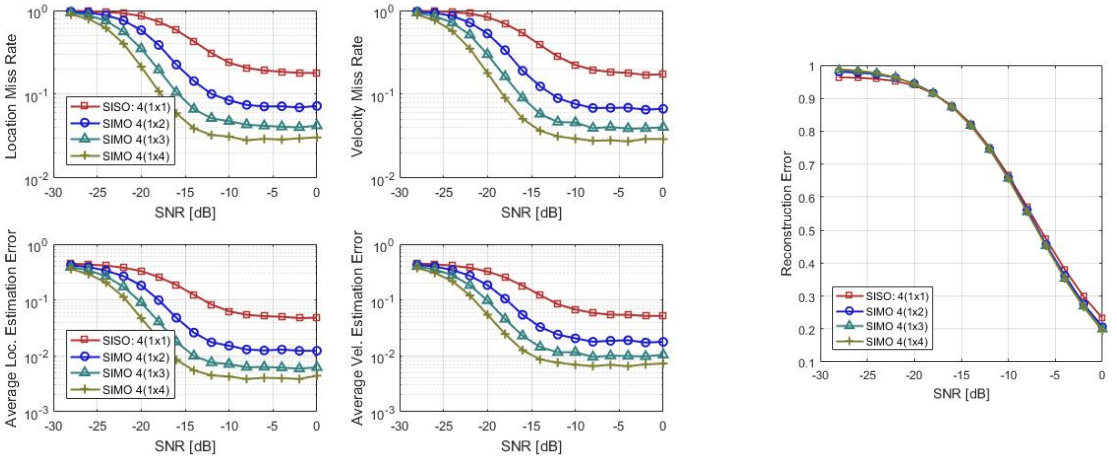
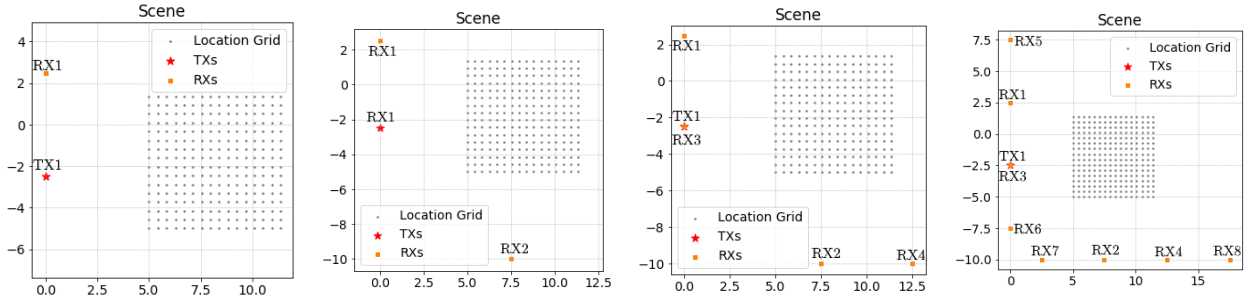


Figure 4.10: Evolution of the miss rates, estimation errors and reconstruction error with respect to the SNR, in the presence of a three targets. Comparison of different multi-static SIMO configurations $4(1 \times N_R)$.



(a) $1(1 \times 8)$ configuration (the RX node is an array of 8 antennas).

(b) $2(1 \times 4)$ configuration (each RX node is an array of 4 antennas).

(c) $4(1 \times 2)$ configuration (each RX node is an array of 2 antennas).

(d) $8(1 \times 1)$ configuration (each RX node contains one antenna).

Figure 4.11: Locations of the nodes for 4 different multi-static SIMO Radar configurations.

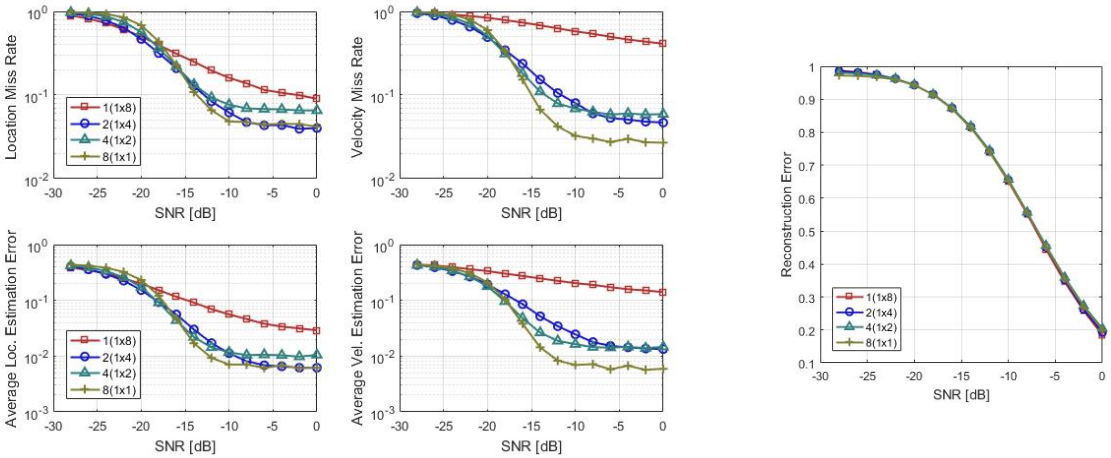


Figure 4.12: Evolution of the miss rates, estimation errors and reconstruction error with respect to the SNR, in the presence of a three targets. Comparison of different multi-static SIMO configurations $4(1 \times N_R)$.

2 Experimental Validation

2 .1 Test bench description

In order to validate our models and the algorithms for practical situations, we set up a very simple test bench. The system we created is composed of two monostatic pairs as shown in Fig. 4.13 ($Q = 2$). Each RX node is composed of two receiving antennas ($N_R = 2$), hence the system could have the short name $2(1 \times 2)$. Because we used only identical Radars (i.e., working on the same frequency band), to obtain this multi-static system, we can either turn on two monostatic Radars and apply shifted modulations as in Fig. 1.4, or we can make a full acquisition with one Radar, then move it to another location and make the other acquisition. The second strategy is much simpler to establish, yet not very realistic for actual applications. Also, this second method does not enable meaningful velocity measurements. Still we used this second method to create the system described in Tab. 4.3 and Tab. 4.4. Before making the actual acquisitions, we first made background measurements for calibration, i.e., we made an acquisition when no target was

present to get $\boldsymbol{y}_{bg}^{q}$ for $q = 1, 2$, then we acquired the actual data $\boldsymbol{y}_{data}^{q}$ for $q = 1, 2$ and entered the following in the FBMP algorithm,

$$\boldsymbol{y}^{q} = \boldsymbol{y}_{data}^{q} - \boldsymbol{y}_{bg}^{q}, \quad q = 1, 2. \tag{4.5}$$

Nodes Locations	
TXs	**RXs**
(0, 0) [m]	(0, 0) [m]
(13, -13) [m]	(13, -13) [m]

Table 4.3: Locations of the transmit and receive nodes.

Parameter	B	f_0	T_s
Value	156.5MHz	24.124GHz	20μs
Parameter	M_s	M_r	T
Value	32	16	640μs

Table 4.4: Modulation and acquisition parameters.

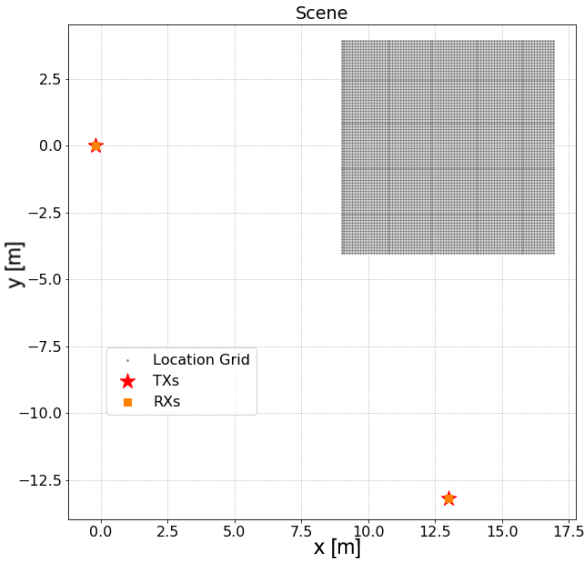


Figure 4.13: Radar configuration and location grid used in our measurements.

2 .2 Results

For the first set of measurements, we defined a location grid by uniformly sampling a square of length 8m into 80×80 samples, as shown in Fig. 4.13. We started with measurement of a single target. For each scenario, we made 100 measurements. Fig. 4.14 shows one location AF we obtained with a single static target located in $\boldsymbol{x} = (13, 0)$. The portions of ellipses of width equal to $\frac{c}{B} = 1.92m$ we observe confirm the theoretical model we established. Fig. 4.15 shows a mapping of the location we estimated over the 100 measurements. Most of the time, we estimate the exact location, or very close to it (always less than 50cm of location estimation error). Fig. 4.17 and 4.16 show another measurement with a target located in $\boldsymbol{x} = (11, 0)$. In this result, we can see that there are many ellipses that come from the ground reflections and calibration artefacts, that are discussed in Chap. 5.

We also made measurements in the presence of two targets, located in $\boldsymbol{x}_1 = (14, 1)$ and $\boldsymbol{x}_2 = (11, -2)$. The location AF obtained for one of the 100 measurements we performed

for this scene and at each iteration of FBMP is displayed in Fig. 4.18. We can see that the algorithm succeeds to first detect the target in $\boldsymbol{x}_2$ then removes properly its contribution for the next iteration, during which it detects properly $\boldsymbol{x}_1$. In Fig. 4.19, we can see the mapping of the estimated locations. There are two main zones corresponding to $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, and one secondary zone, corresponding to some failure, i.e., when the ghost target is stronger (see Fig. 4.20), and therefore detected instead of a correct target. Amongst the 200 targets estimated (2 for each of the 100 measurements), this wrong detection occurred 4 times.

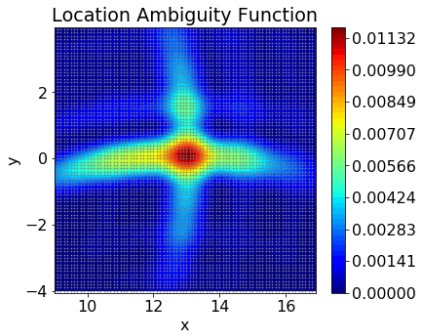


Figure 4.14: One of the 100 measurement location AF

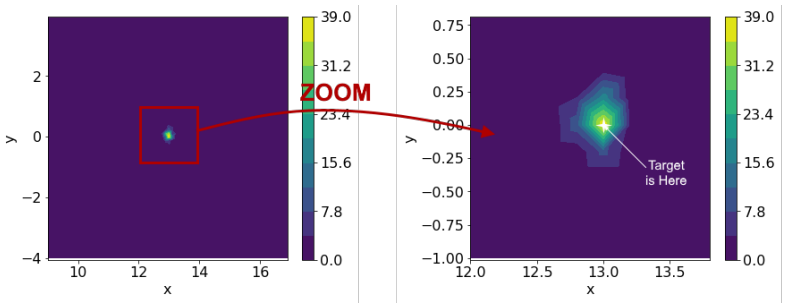


Figure 4.15: Mapping of the 100 results of estimated location.

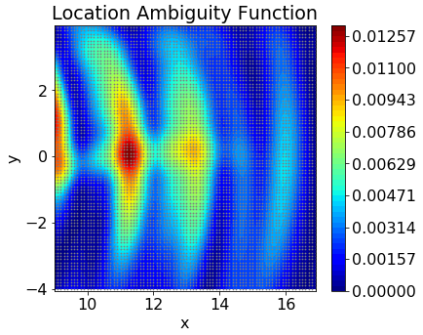


Figure 4.16: One of the 100 measurement location AF

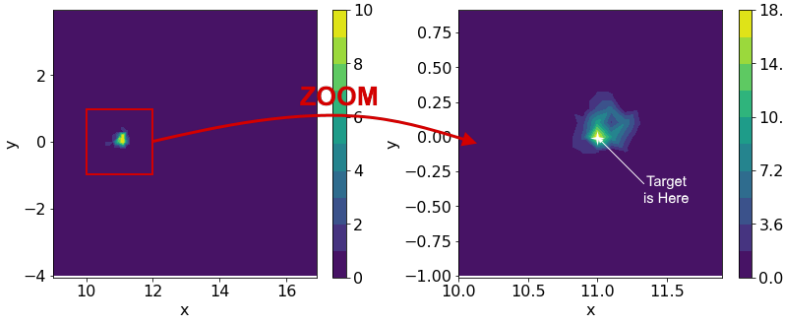


Figure 4.17: Mapping of the 100 results of estimated location. The correct location is $\boldsymbol{x} = (11, 0)$.

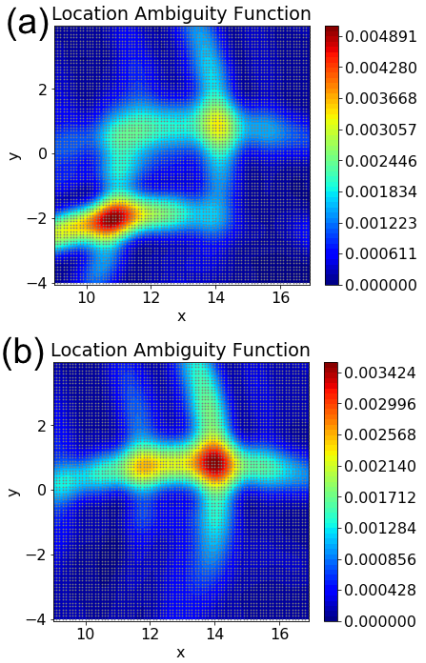


Figure 4.18: One of the 100 measurement location AF. **(a)** First iteration of FBMP. **(b)** Second iteration of FBMP.

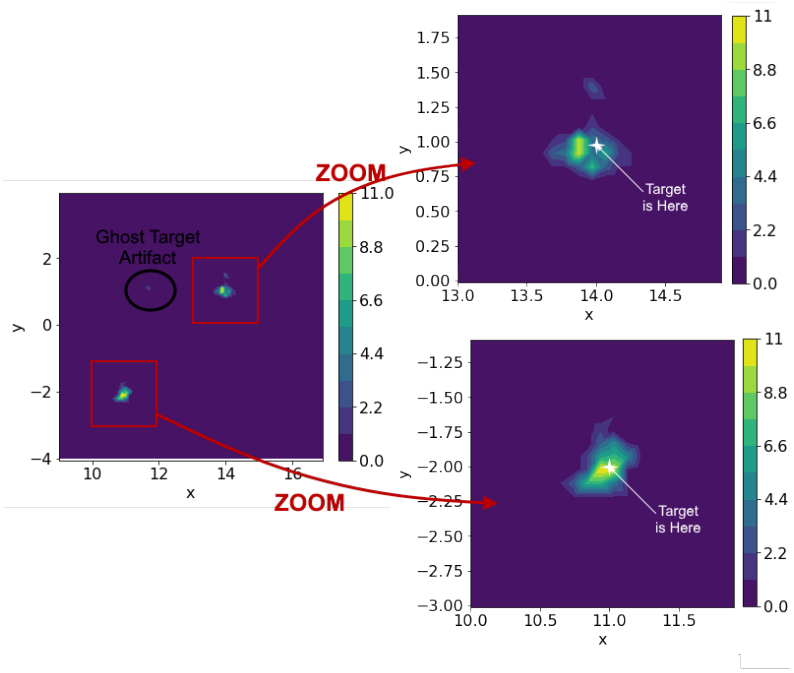


Figure 4.19: Mapping of the 100 results of estimated location (hence mapping of 200 dots because we estimated 2 locations for each measurement). The correct locations are $\boldsymbol{x}_1 = (14, 1)$ and $\boldsymbol{x}_2 = (11, -2)$.

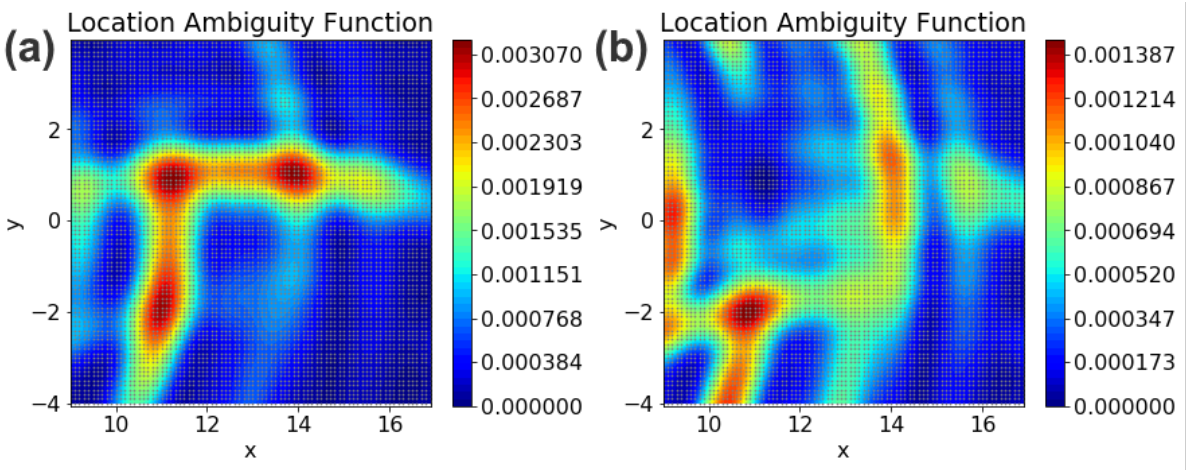


Figure 4.20: One of the 100 measurement location AF in the rare situation when the ghost target is detected. **(a)** First iteration of FBMP. **(b)** Second iteration of FBMP.

Chapter 5

Discussions and Perspectives for Future Works

In this chapter, we bring a critical view to the algorithms we studied in this document and we sum up all the questions that have been raised by this work and still need to be investigated. This chapter is separated into three section. First we discuss the validity of our measurements and models compared to real applications. We also discuss the effectiveness of joint models and associated algorithms with respect to disjoint models and algorithms (i.e., such as triangulation and trilateration). Next we enumerate many possible improvements to the sensing model and optimizations to FBMP that could be studied in order to lower its weaknesses and take advantage of its strengths. Finally, our models and algorithms have opened the path to some advanced topics that we also enumerate and briefly describe.

1 Discussions

In the analysis performed in Chap. 3 and 4, we only considered the joint models. This analysis led to the conclusion that the BMP applied to the complete, non-simplified model gives the best performances in terms of accuracy and robustness against the noise. On the other hand, FBMP applied to the simplified model is much faster but suffers from model mismatches that can be compensated for by FBMP-ic. Yet, FBMP-ic still exhibits lower robustness against noise than BMP and is more affected by the ghost targets' detection problem than the BMP.

1 .1 Models and algorithms: joint against disjoint

In multi-static Radars, other possibilities that have not been treated in these chapters are methods adapted to disjoint models. More precisely, each bistatic pair is treated by an algorithm suited for bistatic Radar systems, hence providing a set of estimators of the locations and velocities for each bistatic pair. Then, another algorithm tries to associate the estimators that correspond to the same target to draw a better estimation of the location and velocity. In the future, a deep analysis comparing the joint methods and the disjoint methods should be done. Yet we can already try to anticipate some results by the means of the following discussion.

Let us first consider multi-static SISO systems. In disjoint methods, we no longer need to have a joint sparsity domain for all the received signals. Therefore, the estimators provided by each bistatic pair are usually estimation of the bistatic range and bistatic speeds, hence allowing to use the model described in (1.38). In that case, FFT can be used in the algorithms. Because the location and velocity grids have a very different definition in the disjoint model, it is hard to define a relevant comparison of the joint and disjoint methods. Yet, as a rule of thumb, if we want to define target states spaced with a density equivalent to the canonical density used for joint methods, the complexity of algorithms suited for disjoint models of multi-static SISO Radar systems is $\mathcal{O}\left(\sqrt{N_x}\sqrt{N_v}\log_2(M_rM_s)\right)$ against $\mathcal{O}\left(N_xM_sM_r + N_vM_s\right)$ for FBMP applied to the joint model. Therefore, the algorithms for disjoint model are much faster. If we consider multi-static SIMO systems, each bistatic pair can also estimate the angle of arrival, giving a model such a shown in Fig. 6 in the Introduction. In that case, as a rule of thumb, the complexity is $\mathcal{O}\left(N_x\sqrt{N_v}\log(N_RM_sM_r)\right)$, which is usually still faster than FBMP applied to the joint model. In multi-static SIMO systems both methods are expected to suffer equivalently of the ghost target issue. However, the disjoint method is expected to be slightly less affected by this problem in multi-static SIMO systems.

However, the joint method offers many advantages. To obtain the low complexity of disjoint methods provided above, very specific grids must be used, while joint models offers a much better flexibility on the shape of the grids, allowing us to look only in a very specific region of the 2D location plane (or velocity plane, although it is less relevant in practical applications). Note that this is note really an advantage of the joint method because we can also use arbitrary grids on disjoint models, however doing so removes its advantage of lower complexity.

The joint model enables the use of location grids far larger than the maximal non-ambiguous bistatic range. This was observed in Fig. 3.10. This is only possible for joint models because to make a decision about one target's location, the algorithm gathers the information of all the bistatic pairs, while in the disjoint method, each bistatic pair makes its decision on its own, which, as shown in Fig. 3.9 is ambiguous for large grids. Actually, this collective decision-making of the joint method is its key advantage. This feature allows the joint method to be much more robust against many non-idealities such as ground reflection and target occlusion. A measurement we realized tends to confirm this statement, see Appendix D .

1 .2 Validity of the simulations and measurements

To obtain the result of our simulation, we set a lot of non-realistic assumptions. These assumptions have impact on the relevance of the comparison of different Radar configurations, especially the comparison of Fig. 4.12. Indeed, some configurations are likely to perform better with respect to the others regarding the ground reflections and target occlusion. Also, the location of the TX node and the RX nodes, as well as the location grid were pure arbitrary choices that impacted the results.

The measurements performed were very simple but allowed us to show that the theory we developed and confirmed with simulations also fits the reality. However, these measurements are still both not realistic because of the non-simultaneous acquisition we performed and not yet compelling for real practical applications. Indeed, the conditions were not optimal because the Radars we used were not perfectly calibrated, i.e., we had no perfect guarantee of the frequencies we transmitted, as well as we had no perfect guarantee of the perfect phase coherence between the received signals of the two RX antennas. Better performance could be obtained with a more reliable test bench. We are already planning to set up better measurements in the near future.

2 Optimizations

2 .1 Exactitude and optimization of the sensing model

The assumption of the complex normal random distribution of the scattering coefficient is theoretical and intuitive. In practice, this coefficient depends on many parameters that could be taken into account in our model. For example, the path loss and the antenna gains, which are injected into this coefficient in our model may be accounted for by assuming that the n-th component of the sparse vector $\boldsymbol{s}^q$ no longer has the same probability distribution as the n'-th component of the same vector with $n \neq n' \in [N]$. Similarly, the n-th component of the sparse vector $\boldsymbol{s}^q$ no longer has the same probability distribution as the n-th component of $\boldsymbol{s}^{q'}$, with $q \neq q' \in [Q]$. Such considerations impact the algorithms. Notably because the ML approach presented in Appendix A no longer holds and should be replaced by a Maximum A Posteriori (MAP) approach, more challenging to compute in this framework. Moreover, if we assume a prior knowledge of the environment where measurements are performed, we may be able to anticipate the ground reflection and other similar scattering, and thus, modify the content of the dictionaries to take account of these undesired effects.

Multistatic radars offer the additional degree of freedom of choosing and even optimizing the locations of the TX and RX nodes. In this work, we always set arbitrary locations, yet this is a system parameter that can be optimized according to several criteria. For instance, the maximal size of a non-ambiguous location grid is strongly related to this parameter and could be one criterion for the optimization. As an extend to this reasoning, in the future, we will study the effect of several signal and system parameters on well known and relevant estimation metrics. In the context of information geometry, the Cramer-Rao bound (CRB), as well as the Ziv-Zakai bound (ZZB) which is more relevant at low-SNR regimes, determine the smallest achievable estimation error variances; hence characterizing the best reachable accuracy of the radar system independently of any particular algorithm. The derivation of such bounds will provide criteria and guidelines for the optimization of the sensing parameters.

Another metric of interest is the self-coherence of the dictionaries. For a given dictionary $\boldsymbol{\Psi}$, the self-coherence $\mu(\boldsymbol{\Psi})$ is the maximal correlation between two distinct atoms of $\boldsymbol{\Psi}$, i.e., with our normalization convention,

$$\mu = \frac{1}{M}\max_{i\neq j} \left|\left\langle \boldsymbol{\psi}_i, \boldsymbol{\psi}_j \right\rangle\right|. \tag{5.1}$$

A small value $\mu(\boldsymbol{\Psi})$ of makes $\boldsymbol{\Psi}$ closer to an orthonormal basis and improves the estimation the sparse vector $\boldsymbol{s}$. Unfortunately, in the Radar framework, the coherence the dictionaries we use ($\boldsymbol{D}^q$, $q \in [Q]$ for the complete model, $\boldsymbol{\Psi}^q$ and $\boldsymbol{\Phi}^q$, $q \in [Q]$ for the factored model) increases if the densities of Ω_X and/or Ω_V increase. In this work, we always used very simple uniform square location and velocity grids. However, this choice is most likely not optimal regarding the coherence of the dictionaries resulting from it. Finally, links between CRB/ZZB and the coherence may be unveiled in future work.

2 .2 Algorithmic optimizations

All the algorithms we explained and developed in this document are based on MP. However, MP has no guarantee to provide a perfect reconstruction of the sparse vectors $\boldsymbol{s}^q$. On the other hand, the Orthogonal Matching Pursuit guarantees the recovery of up to $K \leq \frac{1}{2}\left(\frac{1}{\mu}+1\right)$ atoms, i.e., in the framework of Radar applications, targets [19]. The derivation and application of similar algorithms to the one presented in this work, yet based on the OMP instead of MP, is expected to provide much higher robustness against undesired effects due to the interference between the targets, such as the detection of ghost targets.

In practice, when we define a location grid Ω_X and a velocity grid Ω_V, this usually means that we are looking for targets such that there is at most one target located in a given location in Ω_X, hence with a single velocity. In other terms, in FBMP, when we select a location index $n_x^* \in [N_x]$, we are expecting that the n_x^*-th row of the matrices $\boldsymbol{S}^q$ are 1-sparse for all $q \in [Q]$. However, in real measurement, it may happen that a single target location contains multiple active velocities. Indeed, imagine a location grid with states spatially spaced by 1m with respect to one another. If a person is standing still in a location perfectly matching the location grid, yet spinning around, this location will appear to have multiple velocities. If we desire to avoid selecting multiple times the same location[1], we can replace the fifth step of Alg. 7 by the following:

$$\boldsymbol{R}^{q,(k+1)} = \boldsymbol{R}^{q,(k)} - M_r \boldsymbol{s}^{q,(k)}[n_x^*, n_v^*]\boldsymbol{\psi}^q_{n_x^*}\left(\tilde{\boldsymbol{p}}^q_{n_x^*}\right)^T. \tag{5.2}$$

This way, we remove the whole contribution of the n_x^*-th location instead of the specific contribution of the n_v^*-th velocity at the n_x^*-th location. Actually, this adaptation is inspired by the first version of FMP we proposed in Fig. 2.1, because we are forcing a more constraining sparsity pattern where each row of $\boldsymbol{S}^q$ is at most 1-sparse.

[1]It is a desire because selecting multiple times the same location to find all the velocities it contains is not a mistake. However, for some reason, we may want to detect only targets on distinct locations in Ω_X.

3 Advanced Topics

3 .1 Large targets

When the targets are no longer considered as point targets but large and shaped targets, the definitions of grids we proposed in this document is no longer relevant because the targets are no longer strictly located in a precise location. A better way to interpret it is to suppose that the targets are not only defined by a location and a velocity but also by an additional bunch of new features such as the shape, the scale, the orientation or even the spin. As we did for the location and the velocity, we can define a discrete set of possible values for each feature. For instance, we may define a set of possible orientation angles and a set of possible shapes. Each additional feature adds a dimension to our dictionaries, and hence dramatically increases its size. Therefore, a whole research topic may consist of finding new ways to factor these additional dimensions to perform an order-$D > 2$ FBMP. These features are thus additional target parameters to estimate, enabling target recognition. In Fig. 5.1, we can see an example in which we want to detect pedestrians and cars. Pedestrians are modeled as disks while cars are modeled as rectangles, possibly rotated. Therefore, we may add the shape feature, with two possible states and the rotation feature with N_θ possible rotations angles. To build the dictionaries $\boldsymbol{D}^q$ for all $q \in [Q]$ corresponding to this set of features, we should derive a new model taking account of the shape and angles of the targets, possibly by integrating the waveform derived in Chap. 1, yet, it is expected to be trickier because of the non-constant scattering properties along the shape (for instance, if the target is composed of several materials). The derivation of target signatures is a vast topic. Finally, the multi-static configuration is expected to provide better performances in terms of target recognition due to its ability to observe the targets from different angles [6].

Figure 5.1: Example of scenario with three distinct large targets.

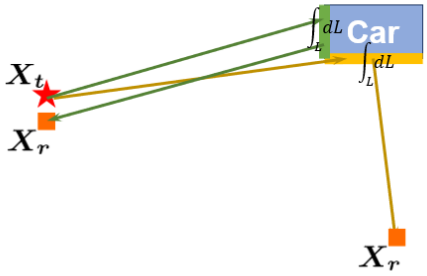


Figure 5.2: Illustration of the signal transmission in the presence of a large target.

3 .2 Compressed Sensing

As stated in the introduction, CS shows how to sample a *low-complexity* signal (typically, with a low sparsity level) from general random linear observations in order to be able to acquire much less samples than required by the classical Shannon-Nyquist theorem while

still being able to reconstruct the few information contained in the signal. In the framework of the multi-static Radars as viewed by this document, we simply assumed a classical Shannon-Nyquist sampling. Indeed, in Chap 3, we established the properties of the grids defining the sparsity domain, hence defining the sparse signals $\boldsymbol{s}^q$' properties that we are able to properly recover. Notably using (3.9) and (3.10), we saw how such grids should be defined with respect to the number of acquired samples, i.e., M_s and M_r. Equivalently, these equations directly result from the application of the Shannon-Nyquist sampling theory, according to which $f_{max} - f_{min} \leq \frac{1}{T_s}$, where f_{max} and f_{min} are the maximal and minimal frequencies of the sampled signal. That is, the condition given in (3.9) results from the following condition,

$$\frac{c}{BM_sT_s}\left(\max\left(d_b^q(\boldsymbol{x})\right) - \min\left(d_b^q(\boldsymbol{x})\right)\right) = f_{max} - f_{min} \leq \frac{1}{T_s} \quad \forall\, q \in [Q]\,. \tag{5.3}$$

If M_s increases, we are able to recover sparse vectors $\boldsymbol{s}^q$ resulting from a larger grid while keeping the same grid density, hence the same self-coherence and cross-coherence of the dictionaries $\boldsymbol{D}^q$, $q \in [Q]$. As highlighted by the simulation of Fig. 4.12, The robustness against inter target interference is strongly linked to the coherence. In a nutshell, according to the Shannon-Nyquist sampling the analysis of this work are based on, for a given desired performance we can establish a relation between the sparsity domain we manage to reconstruct and the required number of samples[2], i.e.,

$$\{K, \Omega_X, \Omega_V\} \longleftrightarrow \{M_s, M_r\}\,. \tag{5.4}$$

CS theory suggests that we may still be able to recover sparse vectors defined by $\{K, \Omega_X, \Omega_V\}$ with a number of acquired samples $M' = M'_s M'_r$ much smaller than $M = M_s M_r$ when using random sampling strategies. Inspired by previous approaches so far restricted to monostatic and bistatic radar systems [10], future work intends to apply CS theory to multi-static FMCW Radar systems as a way to efficiently sample, quantize and process the received signals by investigating the possibility of applying random time subsampling of received signals. To illustrate it, let us consider a single bistatic pair, i.e., $Q = 1$. In this work we are performing the sensing illustrated in Fig. 5.3, while the application of CS described above is represented in Fig. 5.4, with $\boldsymbol{R}$ standing for the random subsampling matrix. An additional application of CS to Radars is the addition of the quantization operator in the sensing model, as depicted in Fig. 5.5. In future work, we may consider various quantization scenarios. This model will be used for tackling the 1-bit quantization question keeping only the signs of the signals. When the sensing described in this figure is applied to all bistatic pairs, we obtain the multi-static compressive Radar sensing scheme.

[2]We assume here that the Radars' properties (B, f_0) are constant parameters

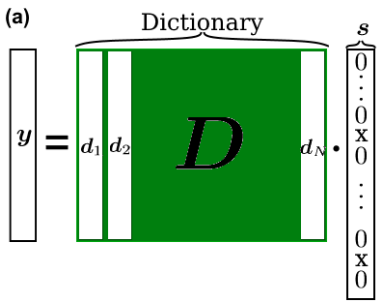


Figure 5.3: Shannon-Nyquist sensing.

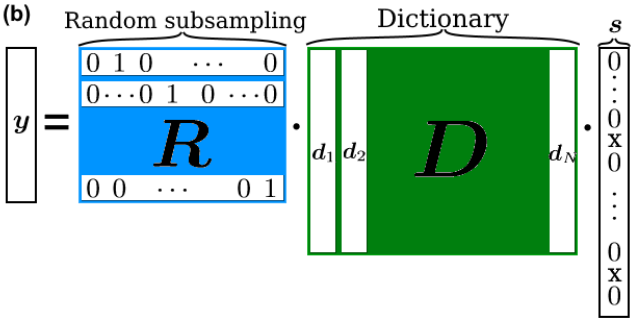


Figure 5.4: Random subsampling from compressive radar. On a single bistatic pair, a random time selection operator $\boldsymbol{R}$, corresponding to random rows of the identity matrix, is inserted in the sensing model. According to CS theory, under certain conditions on the sampling rate, the sparsity of $\boldsymbol{s}$ enables the possibility to reconstruct it from $\boldsymbol{y}$. This constitutes an application of Compressive Sensing in the context of radar systems.

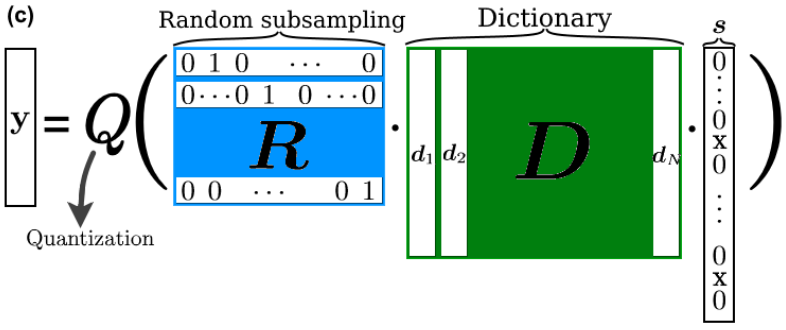


Figure 5.5: Random subsampling from compressive radar with quantization operator denoted by $\mathcal{Q}$, on a single bistatic pair.

Conclusion

The study of multi-static Radars is a recent yet growing research field. The spatial diversity they provide is a key advantage for a robust detection of moving targets, but also for target recognition and tracking. Yet, while most researches in the field are studying Pulse Doppler Radars, this thesis intended to extend these contributions to the FMCW Radars framework, more suited for short range applications such as the smart cities or the automotive world, instead of the military domain generally targeted by the PDRs. This thesis goes even further by considering so-called multi-static MIMO systems, i.e., systems composed of widely separated nodes, each composed of close and aligned antennas.

The application of the sparse signals recovery theory in Radar applications has also recently grown in interest and further led to the development of compressive Radar systems. However, such researches have mainly been applied to monostatic and bistatic Radars, and hence such that most works proposing sparse representations for multi-static Radar systems are treating them as a composition of multiple bistatic Radars, each to be modeled and processed independently. Nonetheless, working with a joint acquisition model for the whole multi-static Radar system provides a higher robustness against many non-idealities over the mentioned above disjoint modeling. That is, because the decision of target detection is performed by directly collecting a common "opinion" of all Radars. The few works that tackled this joint modeling and processing quickly faced the issue of a huge dimensionality of the problem to be solved. Inspired by some hints of solution to this dimensionality problem proposed in PDRs framework, this thesis intended to deeply study the possibility of dimensionality reduction of the sparse signal recovery adapted to the joint modeling of FMCW multi-static Radars.

This study required to leave away the abstraction of general dictionaries and precisely define what they are made of. Chapter 1 reached this purpose with the derivation of a complete joint model of the received signals of a FMCW multi-static Radar system and goes even further with the extension to multi-static MIMO Radar systems. While known simplifications of the bistatic Radars signal models enable to reduce its expression to the DFT-2 of a sparse vector, and hence defining a factored dictionary, we saw in this chapter that reaching the same expression for a multi-static joint model is not possible. Despite this difficulty, we managed to reach a so-called *half factored* expression of the received signals.

In the second chapter we considered several sparse reconstruction algorithms regardless of any application. From existing algorithms, we successfully derived a fast algorithm for recursively reconstructing each dimension of a sparse multi-dimensional signal from

a multi-dimensional measurement signal. This new algorithm, called Factored Matching Pursuit (FMP) was shown to be much faster and less memory demanding than the application of the known matching pursuit to factored dictionaries, yet at the price of a lower robustness against the noise. Still, we also showed that FMP is extremely efficient for highly redundant dictionaries both in terms of complexity and robustness against noise. However, much more theoretical studies are still to be done to fully characterize the impact of FMP and system parameters on the robustness againt the noise. Next, we showed that FMP is suited for the *half factored* kinds of dictionaries resulting from the modeling of Chap 1. Finally, inspired by the Block Matching Pursuit found in the literature on sparse reconstruction of joint sparse vectors, we extended FMP to the situation of multiple measurement vectors to build the Factored Block Matching Pursuit, perfectly suited for measurement obtained with multi-static Radars.

When we applied the algorithms developed to the derived multi-static MIMO Radar model, we noticed that the simplifications performed to enable the use these algorithms cause model mismatches leading to significant errors when estimating the locations velocities of moving targets. In Chap. 3, with the analysis of the ambiguity functions, we managed to fully characterize these mismatches. This enabled the introduction of adaptations and of an iterative correction for our algorithms to compensate for the mismatches. Chap. 4 showed the comparative results of many extensive simulations confirming the great success of the adaptations, but also showed the weaker robustness against noise of Factored Block Matching Pursuit (FBMP), even with the adaptations, with respect to BMP. The last simulations proposed a comparison of system compositions and seem to show that for a fixed amount of RX antenna, creating a fully multi-static SISO system performs better than a multi-static SIMO hybrid system. However, these results must be interpreted carefully because of the intrinsic assumptions we made on the definition of multi-static and MIMO, but also because of the arbitrary choice of antennas' locations. Finally, we also showed measurement results that validate the effectiveness of all the models and algorithms developed in this thesis. Yet, the measurements performed were very simple and not really realistic to be sufficiently compelling for real applications. In the future, a much more rigorous and complete measurement campaign will be performed.

At the end of the day, this work is a huge open door for future contributions. We have still a lot to learn and optimize regarding the sensing model, from the definition of a more accurate signal model to the extension to the case of large targets, from opitmization of the Radars' location to the optimization of the grids defining sparsity domain. The derivation of fundamental performance bounds and criterion such as the Cramer Rao bounds and the coherence of the dictionaries must be studied in the future to enable such optimizations. Moreover, the algorithms we developed are prone to receive many improvement and further developments, notably by deriving *Orthogonal* enhanced versions of them. Finally, the natural next step of development of the sparse modeling we proposed is the application of Compressed Sensing theory. Therefore, the contribution of this thesis paves the way for the development of efficient compressive multi-static FMCW Radars possibly leading to a new generation of low-cost sensing systems for smart-cities and automotive worlds of applications.

Bibliography

[1] Mario Figueiredo Elad, Michael and Yi Ma. On the role of sparse and redundant representations in image processing. *Proceedings of the IEEE*, (6):972–982, 2010.

[2] M. G. Amin F. Ahmad. Noncoherent approach to through-the-wall radar localization. *IEEE transactions on aerospace and electronic systems*, 42(4), October 2006.

[3] A. P. Petropulu B. Li. Distributed mimo radar based on sparse sensing: Analysis and efficient implementation. *IEEE Transactions on Aerospace and Electronics Systems*, 51(4), January 2015.

[4] K. Woodbridge T.E. Derham, S. Doughty and C.J. Baker. Design and evaluation of a low-cost multistatic netted radar system. *IET Radar Sonar Navig*, (1):362–368, 2007.

[5] K. Woodbridge T.E. Derham, S. Doughty and C.J. Baker. Ambiguity functions for spatially coherent and incoherent multistatic radar. *IEEE Transactions on aerospace and electronic systems*, 46(1):362–368, January 2010.

[6] F. Gini M. La Manna P. Stinco, M. S. Greco. Non-cooperative target recognition in multistatic radar systems. *IET Radar Sonar Navig*, (4):396–405, 2014.

[7] T. Kirubarajan A. A. Gorji, R. Tharmarasa. Widely separated mimo versus multistatic radars for target localization and tracking. *IEEE Transactions on Aerospace and Electronics*, 49(4), 2013.

[8] Mario Figueiredo Elad, Michael and Yi Ma. On the role of sparse and redundant representations in image processing. *Proceedings of the IEEE*, 98.6:972–982, 2010.

[9] Emmanuel J. Candès and Michael B. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 25(2):21–30, 2008.

[10] K. Jamil F. E. Abd El-Samie M. A. Hadi, S. Alshebeili. Compressive sensing applied to radar systems: an overview. *Signal, Image and Video Processing, 2015*, 9(1):25–39, December 2015.

[11] T. Strohmer M. A. Herman. High-resolution radar via compressed sensing. *IEEE Transaction Signal Processing*, 57(6), 2009.

[12] A. Nehorai S. Gogineni. Target estimation using sparse modeling for distributed mimo radar. *IEEE Transaction Signal Processing*, 59(11), November 2011.

[13] XizhangWei Bin Sun, Haowen Chen and Xiang Li. Multitarget direct localization using block sparse bayesian learning in distributed mimo radar. *International Journal of Antennas and Propagation*, pages 11–23, 2015.

[14] Helmut Bölcskei Yonina C. Eldar, Patrick Kuppinger. Block-sparse signals: Uncertainty relations and efficient recovery. *IEEE Transaction Signal Processing*, 58(6), 2010.

[15] Jos´e M. F. Moura Christian R. Berger. Noncoherent compressive sensing with application to distributed radar. *Annual Conference on Information Sciences and Systems*, 45(4), 2011.

[16] Balas Kausik Natarajan. Sparse approximate solutions to linear systems. *SIAM journal on computing*, 24(2):227–234, 1995.

[17] Stéphane G. Mallat and Zhifeng Zhang. Matching pursuits with time-frequency dictionaries. *Signal Processing, IEEE Transactions*, 41(12):3397–3415, 1993.

[18] Joel A. Tropp. Just relax: Convex programming methods for subset selection and sparse approximation. *IEEE Transactions on Information Theory - TIT*, 04-04, 01 2004.

[19] H. Rauhut S. Foucart. *A Mathematical Introduction to Compressive Sensing.* 2013.

Appendix

A Maximum Likelihood Approach for Block Matching Pursuit

Let us define the space of K-sparse vectors with length N: $\Sigma_K^N := \left\{ \boldsymbol{s} \in \mathbb{C}^N : \|\boldsymbol{s}\|_0 \leq K \right\}$. For the simplicity of notations, only here we define $\{\boldsymbol{s}^q\} := \{\boldsymbol{s}^q : q \in [Q]\}$, $\{\hat{\boldsymbol{s}}^q\} := \{\hat{\boldsymbol{s}}^q : q \in [Q]\}$ and $\{\boldsymbol{y}^q\} := \{\boldsymbol{y}^q : q \in [Q]\}$. Let us define the set of sets of Q K-sparse vectors sharing a sparsity pattern:

$$\Sigma_{K,Q}^N := \left\{ \{\boldsymbol{s}^q\} : \boldsymbol{s}^q \in \mathbb{C}^N, \|\boldsymbol{s}^q\|_0 \leq K \quad \forall\, q \in [Q]\,, \quad \mathrm{Supp}\left(\boldsymbol{s}^q\right) = \mathrm{Supp}\left(\boldsymbol{s}^{q'}\right) \quad \forall\, q, q' \in [Q] \right\}.$$

We compute here the Maximum Likelyhood approach applied to multiple measurement vectors under AWGN conditions,

$$\boldsymbol{y}^q = \boldsymbol{\Psi}^q \boldsymbol{s}^q + \boldsymbol{e}^q \quad \forall\, q \in [Q]\,,$$

where $\boldsymbol{e}^q \overset{\text{i.i.d.}}{\sim} \mathbb{C}\mathcal{N}\left(0, \sigma_q^2\right)$. The ML approach reads,

$$\begin{aligned}
\{\hat{\boldsymbol{s}}^q\} &= \underset{\{\boldsymbol{s}^q\} \in \Sigma_{K,Q}^N}{\mathrm{argmax}} \ \mathrm{P}\left(\{\boldsymbol{y}^q\} \,|\, \{\boldsymbol{s}^q\}\right) \\
&= \underset{\{\boldsymbol{s}^q\} \in \Sigma_{K,Q}^N}{\mathrm{argmax}} \ \prod_{q \in [Q]} \frac{1}{\sqrt{2\pi\sigma_q^2}} \prod_{m \in [M^q]} \exp\left(- \frac{\left| y^q[m] - \sum_n \psi_n^q[m] s^q[n] \right|^2}{2\sigma_q^2} \right),
\end{aligned}$$

where $\mathrm{P}\left(\{\boldsymbol{y}^q\} \,|\, \{\boldsymbol{s}^q\}\right)$ stands for the conditional probability of measuring $\{\boldsymbol{y}^q\}$ if its sparse representation is $\{\boldsymbol{s}^q\}$. Computing the Log-likelihood,

$$\begin{aligned}
\hat{\boldsymbol{s}} &= \underset{\{\boldsymbol{s}^q\} \in \Sigma_{K,Q}^N}{\mathrm{argmin}} \ \sum_{q \in [Q]} \frac{1}{\sigma_q^2} \sum_{m \in [M^q]} \left| y^q[m] - \sum_n \psi_n^q[m] s^q[n] \right|^2 \\
&= \underset{\{\boldsymbol{s}^q\} \in \Sigma_{K,Q}^N}{\mathrm{argmin}} \ \left\| \boldsymbol{y}^q - \boldsymbol{\Psi}^q \boldsymbol{s}^q \right\|_2^2 \\
&= \underset{\{\boldsymbol{s}^q\}}{\mathrm{argmin}} \ \left\| \boldsymbol{y}^q - \boldsymbol{\Psi}^q \boldsymbol{s}^q \right\|_2^2 \ \text{ s.t. } \ \{\boldsymbol{s}^q\} \in \Sigma_{K,Q}^N.
\end{aligned}$$

And we reach the sparse recovery problem formulation for multiple measurement vectors. Going further, we can write,

$$\hat{\boldsymbol{s}} = \underset{\{\boldsymbol{s}^q\}\in\Sigma^N_{K,Q}}{\operatorname{argmin}} \sum_{q\in[Q]} \frac{1}{\sigma_q^2} \left(\sum_{m\in[M^q]} |y^q[m]|^2 - 2\Re\left(\sum_{m\in[M^q]} \sum_{n\in[N]} \left(y^q[m]\right)^* \psi_n^q[m] s^q[n] \right) \right.$$
$$\left. + \sum_{n\in[N]} \sum_{n'\in[N]} \left(\sum_{m\in[M^q]} \psi_n^q[m]\psi_{n'}^q[m] \right) s^q[n]s^q[n'] \right).$$

The first term in constant with respect to the argument of the minimization, therefore, we can remove it from the expression to maximize. Moreover, if we neglect the cross correlation between atoms in the last term (in radar application, neglect the inter targets interference), we can write,

$$\hat{\boldsymbol{s}} \simeq \underset{\{\boldsymbol{s}^q\}\in\Sigma^N_{K,Q}}{\operatorname{argmin}} \sum_{q\in[Q]} \frac{1}{\sigma_q^2} \left(-2\Re\left(\sum_{m\in[M^q]} \sum_{n\in[N]} \left(y^q[m]\right)^* \psi_n^q[m] s^q[n] \right) + \sum_{n\in[N]} \left(\sum_{m\in[M^q]} |\psi_n^q[m]|^2 |s^q[n]|^2 \right) \right).$$

Let us remind that $\sum_{m\in[M^q]} |\psi_n^q[m]|^2 = M^q$ by definition of our normalization in (2.2). Moreover, we may also add the following artifice,

$$\hat{\boldsymbol{s}} \simeq \underset{\{\boldsymbol{s}^q\}\in\Sigma^N_{K,Q}}{\operatorname{argmin}} \sum_{q\in[Q]} \frac{1}{\sigma_q^2} \left(\frac{1}{M^q} \sum_{n\in[N]} \left| \sum_{m\in[M^q]} y^q[m]\psi_n^q[m] \right|^2 - 2\sum_{n\in[N]} \Re\left(\sum_{m\in[M^q]} \left(y^q[m]\right)^* \psi_n^q[m] s^q[n] \right) \right.$$
$$\left. + \sum_{n\in[N]} M^q |s^q[n]|^2 \right).$$

Therefore,

$$\hat{\boldsymbol{s}} \simeq \underset{\{\boldsymbol{s}^q\}\in\Sigma^N_{K,Q}}{\operatorname{argmin}} \sum_{q\in[Q]} \frac{1}{\sigma_q^2 M^q} \sum_{n\in[N]} \left| M^q s^q[n] - \sum_{m\in[M^q]} \left(y^q[m]\right)^* \psi_n^q[m] \right|^2.$$

In one iteration of BMP, we solve the local optimization problem of finding the best set of 1-sparse signals, i.e., $\{\boldsymbol{s}^{\boldsymbol{q}q}\} \in \Sigma^N_{1,Q}$. Therefore, we compute,

$$n^* = \underset{\{\boldsymbol{s}^q\}\in\Sigma^N_{K,Q}}{\operatorname{argmax}} \sum_{q\in[Q]} \frac{1}{\sigma_q^2 M^q} \sum_{n\in[N]} \left| \sum_{m\in[M^q]} \left(y^q[m]\right)^* \psi_n^q[m] \right|^2.$$

This gives the values to give to the weights:

$$w^q \propto \frac{1}{\sigma_q^2 M^q}.$$

B Proof of the Model Simplifications

Neglecting the coupling signal

The purpose if to show that $\theta_{\boldsymbol{x},\boldsymbol{v}}[m_s, m_r] \simeq 1$ for all $m_s \in [M_s]$, $m_r \in [M_r]$, where,

$$\begin{aligned}\theta_{\boldsymbol{x},\boldsymbol{v}}[m_s, m_r] \;:=\; & \exp\left(-j2\pi \tfrac{1}{c}\tfrac{B}{M_s} v_b(\boldsymbol{x},\boldsymbol{v})(m_r T + m_s T s) m_s\right) \\ & \exp\left(j2\pi \tfrac{B}{M_s T_s}\tfrac{1}{c^2}(d_b(\boldsymbol{x}) v_b(\boldsymbol{x},\boldsymbol{v})(m_r T + m_s T_s))\right). \\ & \exp\left(j2\pi \tfrac{B}{2M_s T_s}\tfrac{1}{c^2}(v_b(\boldsymbol{x},\boldsymbol{v})(m_r T + m_s T_s))^2\right).\end{aligned} \tag{5.5}$$

We will show that the factor multiplying $j2\pi$ in eah of the exponential factors is very small with respect to 1. We assume a velocity grid Ω_V of canonical size, and therefore

$$v_b(\boldsymbol{x},\boldsymbol{v}) \leq \frac{c}{2f_0 T} \quad \forall\, \boldsymbol{x} \in \mathbb{R}^2, \quad \forall\, \boldsymbol{v} \in \Omega_V.$$

For the first exponential factor:

$$\begin{aligned}\frac{1}{c}\frac{B}{M_s} v_b(\boldsymbol{x},\boldsymbol{v})(m_r T + m_s T s) m_s &\leq \frac{1}{c}\frac{B}{M_s}\frac{c}{2f_0 T}(M_r T + M_s T s) M_s \\ &\leq \frac{B}{2f_0 T}(M_r T + M_s T s) \\ &\leq \frac{B}{2f_0}(M_r + 1)\end{aligned}$$

And $\frac{B}{2f_0}(M_r+1) \ll 1$ if $B(M_r+1) \ll f_0$.

For the second exponential factor, due to Hyp. 4, we must have a location grid Ω_X such that $d_b(\boldsymbol{x}) < cT_s$ for all $\boldsymbol{x} \in \Omega_X$, therefore:

$$\begin{aligned}\frac{B}{M_s T_s}\frac{1}{c^2}(d_b(\boldsymbol{x}) v_b(\boldsymbol{x},\boldsymbol{v})(m_r T + m_s T_s)) &< \frac{B}{M_s T_s}\frac{1}{c^2}\left(cT_s \frac{c}{2f_0 T}(M_r T + M_s T_s)\right) \\ &< \frac{B}{2f_0}\frac{M_r+1}{M_s}.\end{aligned}$$

And $\frac{B}{2f_0}\frac{M_r+1}{M_s} \ll 1$ if $B \ll f_0$, which is the case for the narrowband signals we are working with.

For the third exponential factor:

$$\begin{aligned}\frac{B}{2M_s T_s}\frac{1}{c^2}\left(v_b(\boldsymbol{x},\boldsymbol{v})(m_r T + m_s T_s)\right)^2 &\leq \frac{B}{2M_s T_s}\frac{1}{c^2}\left(\frac{c}{2f_0 T}(M_r T + M_s T_s)\right)^2 \\ &\leq \frac{B}{8f_0^2 M_s T_s}(M_r+1)^2 \\ &\ll 1\end{aligned}$$

Neglecting the contribution of the velocity in the inner signal

In the definition of the inner signal $\psi_{\boldsymbol{x},\boldsymbol{v}}[m_s]$,

$$\begin{aligned}\psi_{\boldsymbol{x},\boldsymbol{v}}[m_s] := \exp\left(-j2\pi\tfrac{f_0}{c}d_b(\boldsymbol{x})\right)\exp\left(j2\pi\tfrac{B}{2M_sT_s}\tfrac{1}{c^2}(d_b(\boldsymbol{x}))^2\right)\\ \exp\left(-j2\pi\tfrac{1}{c}\left(\tfrac{B}{M_s}d_b(\boldsymbol{x})+f_0T_sv_b(\boldsymbol{x},\boldsymbol{v})\right)m_s\right),\end{aligned} \tag{5.6}$$

we can re-express the last exponential factor as,

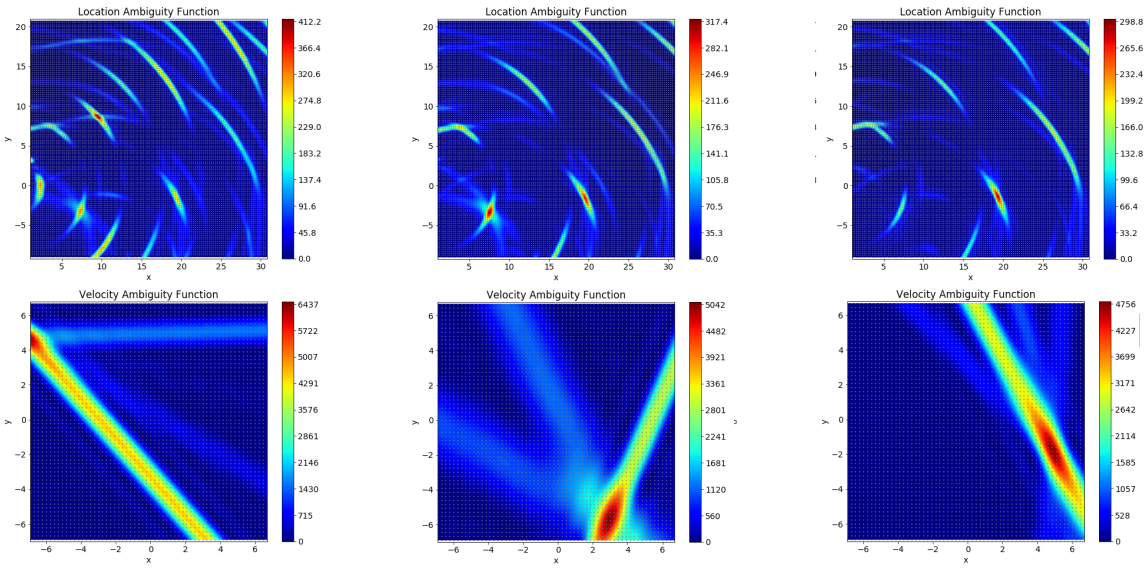
$$\exp\left(-j2\pi\frac{1}{c}\frac{B}{M_s}\left(d_b(\boldsymbol{x})+\frac{M_s}{B}f_0T_sv_b(\boldsymbol{x},\boldsymbol{v})\right)m_s\right).$$

Therefore, the purpose is to show that $\frac{M_s}{B}f_0T_sv_b(\boldsymbol{x},\boldsymbol{v})$ is much smaller than the range of values of $d_b(\boldsymbol{x})$, which according to the definition of the canonical location grid is $M_s\frac{c}{2B}$.

$$\begin{aligned}\frac{M_s}{B}f_0T_sv_b(\boldsymbol{x},\boldsymbol{v}) &\leq \frac{M_s}{B}f_0T_s\frac{c}{2f_0T}\\ &\leq \frac{M_sT_s}{T}\frac{c}{2B}\\ &\leq \frac{c}{2B}\\ &\ll M_s\frac{c}{2B}.\end{aligned}$$

C Example of Application of FBMP/FBMP-ic in Noiseless conditions

Example of application of FBMP



(a) AFs computed in the first iteration of FBMP

(b) AFs computed in the second iteration of FBMP

(c) AFs computed in the third iteration of FBMP

Figure 5.6: Location and Velocity AFs computed during the computation of FBMP applied for the noiseless simulation of three random targets. The system simulated is the one described in Tab. 3.1 and Tab. 3.2, with three RX antennas per RX node.

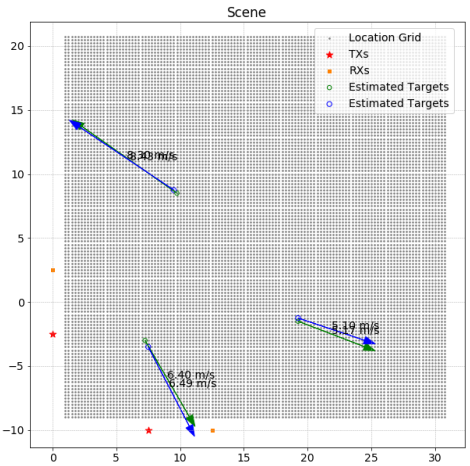
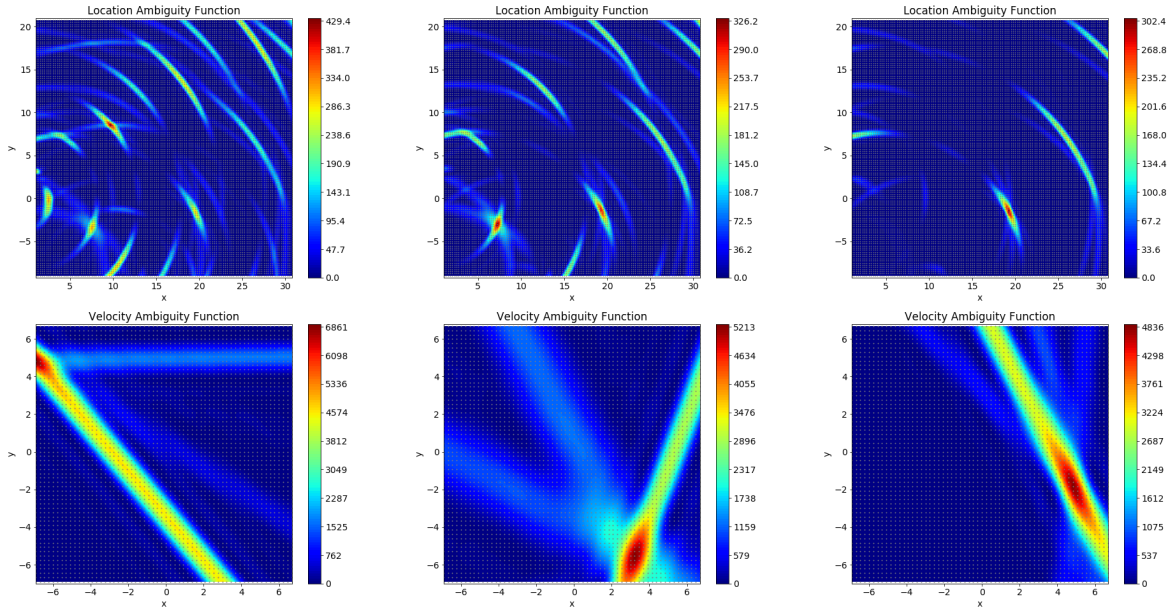


Figure 5.7: Estimation results of the computation of FBMP applied as above.

Example of application of FBMP-ic



(a) AFs computed in the third corrective iteration of the first iteration of FBMP-ic $N_{it} = 3$

(b) AFs computed in the third corrective iteration of the second iteration of FBMP-ic $N_{it} = 3$

(c) AFs computed in the third corrective iteration of the third iteration of FBMP-ic $N_{it} = 3$

Figure 5.8: Location and Velocity AFs computed during the computation of FBMP-ic $N_{it} = 3$ applied for the noiseless simulation of three random targets. The system simulated is the one described in Tab. 3.1 and Tab. 3.2, with three RX antennas per RX node.

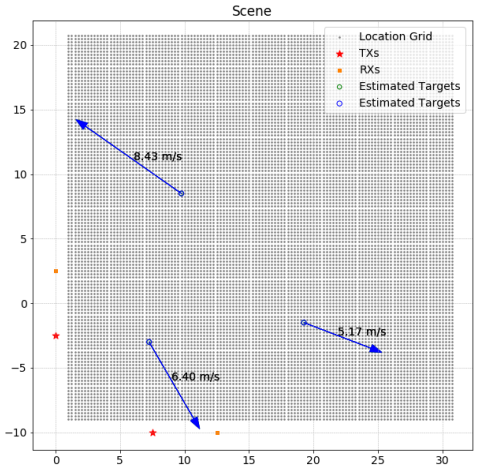
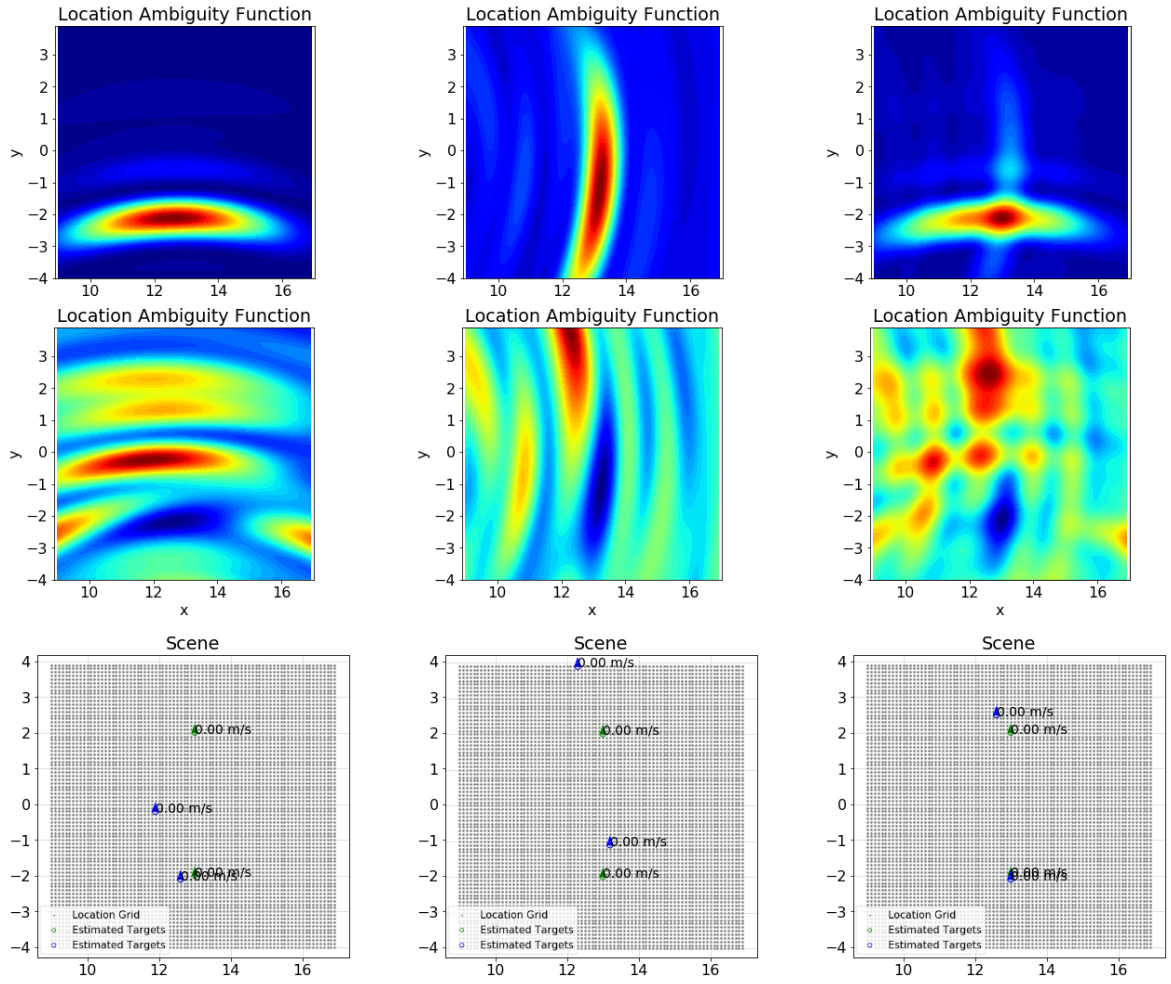


Figure 5.9: Estimation results of the computation of FBMP-ic $N_{it} = 3$ applied as above.

D Relevant Measurements to Show the Effectivness of the Joint Method



(a) FMP on only the first Radar's received signal.

(b) FMP on only the second Radar's received signal.

(c) BFMP on only both Radars' received signals.

Figure 5.10: Location AFs computed during the computation of FBMP. In each sub-figure, the top image is the AF computed during the first iteration of FBMP, the middle image is the AF computed during the second iteration of FBMP and the bottom image shows the estimated locations of the two targets. We see that each Radar is completely lost when performing estimation alone. Therefore, disjoint or "two-steps" procedures would not work. While the joint procedure succeeds to estimate the (almost) correct location.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
École polytechnique de Louvain
Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/epl**