UCLouvain
Institute for Information
and Communication Technologies,
Electronics and Applied Mathematics

Accelerated Sparse Signal Decomposition for Single-Step Estimation in Multistatic Radars and Sensor Networks

Gilles Monnoyer de Galland

Thesis submitted in partial fulfillment
of the requirements for the degree of
Docteur en Sciences de l'Ingénieur

Dissertation committee:

Prof. Luc Vandendorpe (UCLouvain, advisor)
Prof. Laurent Jacques (UCLouvain, advisor)
Prof. Jérôme Louveaux (UCLouvain)
Prof. François Horlin (ULB, Belgium)
Dr. Thomas Feuillen (University of Luxembourg, Luxembourg)
Dr. Cédric Herzet (INRIA, Rennes, France)
Prof. Christophe Craeye (UCLouvain, chair)

Version of September 22, 2023.

Abstract

With advancements in electronics and computer science, new research possibilities have emerged, enabling the resolution of previously inaccessible complex problems. Multiple sensor network applications have benefited from this aspect, particularly regarding real-time feasibility for low-cost implementation. Our main objective is to introduce accelerated adaptations of robust algorithms based on sparse coding theory, previously overlooked due to their heavy computational requirements. Built upon this observation, this thesis explores accelerated adaptations of robust sparsity-driven estimation techniques, explicitly focusing on multistatic radars and extending to sensor networks sharing similar characteristics. By leveraging existing research on algorithmic accelerations, we provide insights to replace conventional standard real-time approaches with more robust algorithms.

The results of this thesis are divided into three targeted acceleration contexts, progressively studying our objective from an abstract point of view toward application-specific considerations. First, we investigate the acceleration of robust sparse signal decomposition algorithms, aiming to solve challenging NP-hard sparse decomposition problems efficiently. Our novel approach called "Peeling" successfully accelerates standard Branch-and-Bound procedures for the sparse decomposition task. Second, we focus on decomposing continuous sparse signals in the estimation and sensing context. In particular, we developed a novel interpolation-based continuous version of the Orthogonal Matching Pursuit (OMP) algorithm that is faster with respect to state-of-the-art propositions and more efficient than conventional discrete approaches, which rely on gridding the observed scene. We also provide extensions that are specifically tailored for radar systems. Lastly, we address the single-step estimation processes in sensor networks by introducing a generalized acceleration framework for single-step methods called "Grid Hopping". Bridging a gap between distinct research communi-

ties, we formally study existing trade-offs between (i) two-step procedures, commonly used for their simplicity and low computational requirements, and (ii) single-step methods offering more robust results but suffering from high computational complexity.

Throughout this thesis, we present detailed motivations and descriptions of our contributions to each research question. The newly designed algorithms are evaluated through extensive Monte-Carlo simulations and real-world applications in the context of Multistatic Radar (MSR). The evaluation includes analyzing numerical results and using measurements recorded by our active multistatic radar system tailored for this research, hence assessing the acceleration provided by our propositions compared to existing algorithms.

Remerciements

L'écriture de cette thèse ne signifie pas seulement l'achèvement de ces quatre années de doctorat, mais aussi la conclusion de ces neuf années passées en tant qu'étudiant à l'UCLouvain. Cette aventure a démarré pour moi comme un long fleuve tranquille, auquel sont d'abord venus s'ajouter des remous, jusqu'à devenir un torrent qui aurait pu m'emporter. Mais une chose est restée constante : Tout au long de cette aventure, j'ai été entouré de gens exceptionnels. Ce bref mot est pour moi l'occasion d'enfin leur rendre un petit peu de ce qu'ils m'ont apporté.

En premier lieu, je souhaite bien entendu remercier mes promoteurs, Laurent Jacques et Luc Vandendorpe, sans qui cette thèse n'aurait jamais existé. En particulier, pour la confiance qu'ils ont eue en moi toutes ces années, notamment en me laissant la liberté travailler selon des horaires, disons, parfois peu conventionnels (Team chouettes-nocturnes ! Coucou François).

Laurent est une véritable encyclopédie vivante. Quiconque a déjà travaillé avec lui a déjà dû l'entendre dire des dizaines de fois la phrase : "Ce que tu me racontes, ça m'évoque tel papier écrit par tels auteurs en telle année". Je tiens à le remercier spécifiquement pour la rigueur mathématique que, je pense, il a réussi à me transmettre.

Je tiens aussi à souligner le rôle crucial que Luc a eu dans le déroulement de cette thèse. Tout d'abord pour son véritable œil de lynx lors de la relecture des nombreux documents que j'ai rédigé, mais aussi pour son expertise et son expérience pour gérer les différents projets, encadrement de cours et autres conférences. Je n'aurais jamais pu profiter d'une telle diversité d'expériences sans son appui.

Ensuite, je souhaite remercier mon jury de thèse : Jérôme Louveaux, François Horlin, Christophe Craeye, Cédric Herzet et Thomas Feuillen, pour leurs précieux commentaires et conseils. En particulier, travailler

avec Cédric, Théo et Clément lors de mon séjour à l'INRIA Rennes fût un réel plaisir ! Trois rencontres inoubliables, associées à de super souvenir, autant tous ensemble à s'acharner sur une expression duale d'un problème d'optimisation récalcitrant qu'autour d'une (demi-douzaine de) bière(s) au Hibou ! Évidemment, je ne pourrais manquer d'évoquer à quel point Thomas a été un collègue de travail d'exception. Mon troisième promoteur non-officiel, un mentor, et un ami. Toujours là pour répondre à mes questionnements ou me remettre les idées en place par rapport aux angoisses existentielles que je m'invente sur des futilités. Je te remercie aussi tout simplement de m'avoir formé durant ces 7 dernières années.

En parlant de personnes qui ont eu une contribution des plus notables dans ma formation, je souhaiterais mentionner Charles Wiame et Claude Oestges d'avoir été respectivement un si bon tuteur et un si bon prof, et ensuite, de chouettes collègues.

Ensuite, quel plaisir cela aura été de travailler avec Maxime Drouget, tant nos compétences sont complémentaires. Cette collaboration m'aura permis de sortir un peu du monde cérébral et théorique des maths pour faire un peu de pratique, sortir le tournevis et mettre au point tout le hardware dont j'avais besoin. Sans lui, nombre de mes résultats n'auraient simplement pas vu le jour. Je voudrais également remercier Martin Willame, d'avoir repris à mes côtés l'encadrement du projet ELEC après le départ de Thomas, et ce, avec beaucoup de professionnalisme. Je pense en particulier à tout le travail que tu as pris sur tes épaules pendant mon rush de rédaction.

Et bien sûr, je compte bien mentionner tous les potes et collègues de bureau qui m'ont accompagné dans cette aventure : Antoine, Emre, Simon, Florian, Guillaume, François, Jérôme, Charles-Henry, Victoria, Adrian, Karim, Thibault, Nicolas, Balazs, Olivier, Rémi, Benoit, Jean, Modeste, Khaldoun, et quiconque aurait été oublié.

Il me tient à cœur d'étendre ces remerciements au-delà de mes collègues. C'est pourquoi je démarrerai cette partie avec, ... un collègue, mais qui est surtout mon frère : Charles. La camaraderie, la solidarité et l'entraide qui soudent notre fraternité est extrêmement précieuse pour moi. Merci de toujours répondre présent quand j'ai besoin de soutien, de conseils ou de passer un moment en bonne compagnie ! Je remercie bien sûr aussi Raphaël, mon autre frère, toujours prêt à venir me soutenir dans les moments difficiles. Il en va de même pour mes parents, qui m'ont toujours épaulé.

Mon parrain, Vincent, a aussi occupé une place primordiale tout au long de cette aventure, et même bien au-delà. Nos échanges, tes enseignements,

ton soutien moral et les bons moments que nous passons autour d'un bon repas, d'une bonne bière et/ou d'un bon film, sont et continueront de m'être inestimables.

Enfin, je tiens tout particulièrement à dire un immense Merci à Violette. Ton soutien est ce qui m'aura été donné de plus précieux tout au long de ce parcours. Tu as subi et géré toutes les émotions que cette thèse a généré en moi. Tu t'es toujours efforcée de me tirer vers l'équilibre, vers mon bien être, malgré la résistance que j'ai pu y opposer. Tu as su, maintes fois, me relever à la seule force de tes mots et de tes bras.

Et en dernier lieu, il reste une personne que je ne saurai jamais si je dois remercier ou détester; il s'agit de Gilles-du-passé, pour avoir pris la décision de démarrer cette thèse. Maintenant au bout du voyage, je ne ressens pas de regret de l'avoir entamé, et achevé. Je ne pourrai cependant jamais savoir ce que j'aurais été si j'avais pris la décision inverse. Sans doute moins fier et moins riche de ces innombrables apprentissages, mais sans doute aussi moins tourmenté.

Une petite note supplémentaire semble tout de même nécessaire pour remercier mes héros du quotidien : les chats, le potager (sauf le mildiou, toi, tu peux aller te faire ***), les poules et les hérissons du jardin !

Contents

List of Figures

Glossary

ADC Analog to Digital Converter. 24, 32

AE Average Error. 135

AHE Average Hit Error. xiv, 92–95, 100, 102, 103, 135, 136

AWGN Additive White Gaussian Noise. 22

BnB Branch-and-Bound. i, xiii, 5, 6, 10, 37, 42, 43, 46, 47, 50–52, 55–57, 59, 61–63, 65, 179, 181, 183

BOMP Block Orthogonal Matching Pursuit. xvi, xvii, 120, 121, 127, 128, 130, 150, 154, 156, 157, 172

COMP Continuous OMP. xiv, 44, 52, 74, 76, 88–97, 99, 100, 102, 104, 151, 181

CRLB Cramèr-Rao Lower Bound. 11, 183

DAC Digital to Analog Converter. 23, 32

DFT Discrete Fourier Transform. 124

DoA Direction of Arrival. xv–xvii, 111, 116, 121, 122, 125, 126, 128–131, 165

DTFT Discrete Time Fourier Transform. 29, 124

F-COMP Factorized COMP. xiv, 96, 97, 99, 100, 102–104

F-OMP Factorized OMP. xiv, 96, 100, 102–104

FFT Fast Fourier Transform. 30, 98, 135

FM Frequency Modulation. 17

FMCW Frequency-Modulated Continuous Wave. xi, xii, xiv, 8, 10, 13, 15–17, 20, 21, 23–25, 36, 39, 74, 89, 95, 97, 99–101, 112, 121, 126, 127, 132, 180, 182

GCC Generalized Cross-Correlation. 122, 124, 151

GH Grid Hopping. xvii, 8, 11, 83, 122, 130, 138, 149–157, 159–170, 172, 174, 180, 182, 183

LS Least-Squares. xvii, 80, 83, 95, 163, 164

MR Miss Rate. xiv, 92–95, 100, 102–104, 135, 136

MSR Multistatic Radar. ii, viii, xvi, xvii, 3, 4, 6–8, 10, 11, 13, 17, 19, 22, 29, 32, 36, 111, 112, 121, 122, 126, 128, 130–138, 140, 142, 149, 150, 155, 160, 163, 164, 166, 172, 179, 180, 182, 184, 185

NN Nearest Neighbour. 83, 87, 164

OMP Orthogonal Matching Pursuit. i, xiii, 5, 6, 10, 31, 37, 43, 44, 47, 51, 52, 56, 73, 74, 82, 88, 92, 94, 96, 100, 103, 104, 118, 120, 156, 183

PDR Pulse Doppler Radars. 15, 16

PPF Parameter Projection Function. xv, 113–115, 121, 124, 126, 127, 132, 140–142, 157, 168, 182, 184

RCS Radar Cross Section. 27

RF Radio Frequency. 17–20, 24, 111, 126

RX receiver. xv, xvi, 17–20, 24, 27, 122, 123, 126, 127, 133, 152

SL Source Localization. xvi, 4, 7, 8, 130, 131, 149, 150

SNR Signal-to-Noise Ratio. xii, 31, 135, 136, 163, 167, 184

SRP Steered Response Power. 150, 151

TDoA Time Difference of Arrival. xvi, 111, 121–124, 126, 128–131

TX transmitter. xv, xvi, 17–20, 24, 27, 123, 126, 128, 152

UWB Ultra-Wide Bandwidth. 123, 149

VCO Voltage Controlled Oscillator. 17, 18, 21, 23, 24

WDO Window-Disjoint Orthogonality. 125

Notations and Symbols

Notation	
$\mathbb{R}$	Real Domain
$\mathbb{C}$	Complex Domain
j	Imaginary symbol for $\sqrt{-1}$
$\boldsymbol{a}$	Vectors are denoted by bold lowercase symbols
$\boldsymbol{A}$	Matrices are denoted by bold uppercase symbols
$\boldsymbol{A}^\top$	Transpose of $\boldsymbol{A}$
$\boldsymbol{A}^\mathrm{H}$	Hermitian (conjugate and transpose) of $\boldsymbol{A}$
$\cdot^*$	Marker for ground truth quantities
$\hat{\cdot}$	Marker for estimated quantities
$\tilde{\cdot}$	Marker for approximated quantities
$\bar{\cdot}$	Marker for quantities taken from a grid
$\otimes$	Outer product
$\times_\ell$	ℓ-mode tensor product
$\langle \cdot, \cdot \rangle$	Scalar product
$\langle \cdot, \cdot \rangle_\mathrm{F}$	Frobenius scalar product
$\| \cdot \|_0$	ℓ_0 pseudo-norm (counting non-zero components)

$\|\cdot\|_p$	ℓ_p norm ($p \geq 1$)
$\|\cdot\|_{\mathrm{F}}$	Frobenius norm
$\mathbb{B}_p(r)$	ℓ_p ball of radius r centered in $\mathbf{0}$
$\|a\|$	Absolute value of the value a
$\|\mathcal{S}\|$	Cardinality of the set $\mathcal{S}$
$[N]$	Set of integers $\{1, \cdots N\}$
$\lfloor\cdot\rfloor$	Floor integer rounding
$\lceil\cdot\rceil$	Ceil integer rounding
$\lfloor\cdot\rceil$	Rounding to the nearest integer
$\eta(\cdot)$	Indicator function
$\cdot \longleftarrow \cdot$	Affecting a new value in a variable (Algorithm context)
$\angle(\cdot)$	Angle operator
$\mathbf{1}$	Vector of ones
$\mathbf{0}$	Vector of zeros

Part I

Introduction and Background

Chapter 1

Introduction

The ever-increasing advancements in electronics and computer science have recently opened up new research possibilities by enabling the resolution of problems that were previously considered inaccessible due to their complexity. Many applications relying on networks of multiple sensors have greatly benefited from these developments. In particular, estimation techniques that were once impractical to be solved in real-time have now become feasible and attainable in low-cost applications. Nonetheless, this increase in the computational power reachable within low-cost constraints may only be exploited with careful attention and dedicated efforts to adapting algorithms initially designed for non-real-time contexts to the real-time domain. In this research thesis, we aim to take a step in that direction by introducing accelerated adaptations of robust algorithms that were previously left apart due to their heavy computational requirements. By building upon existing research on algorithmic accelerations in the sensor network context, we intend to provide new insights to replace conventional real-time approaches with more robust algorithms.

Among sensor networks, our particular interest lies in Multistatic Radar (MSR) systems. As depicted by Fig. 1.1, a multistatic radar comprises one or more transmitters and at least two receivers positioned at widely spread locations around the observed scene. These systems find increasing utilization in diverse applications, including low-cost monitoring and automotive systems [LLC+16, SN16, SCM18]. The particular multistatic geometry enables the estimation of the location and velocity of one or multiple targets with an increased diversity by viewing the targets from different angles [HBC08, SGGM13, SCWL15]. While passive bistatic and multistatic

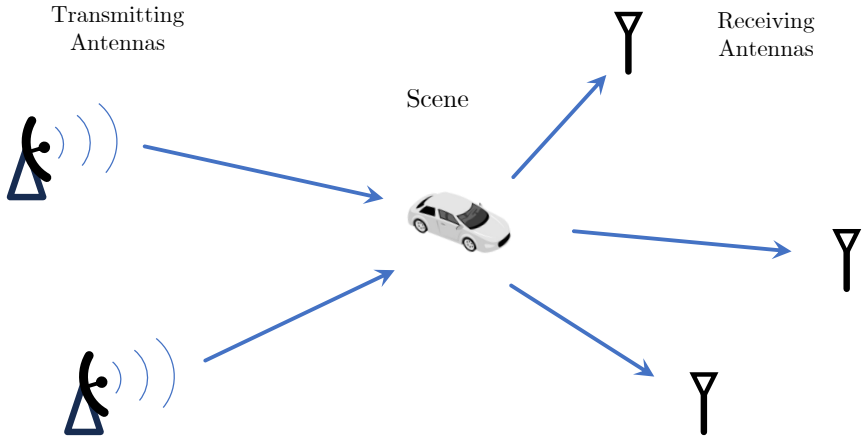

Figure 1.1: Representation of a multistatic radar with two transmitters and three receivers.

radars are vastly studied in the literature, active multistatic radars have more recently emerged in shorter-range applications [DDWB07].

Despite our specific focus on MSR systems, the scope of this research encompasses a broader domain of sensor networks that share common characteristics with this system. In particular, we focus on systems for which the estimation of multiple features from the scene amounts to recovering a *joint* sparse decomposition of the signals recorded by the system in an application-dependent vector space. In addition to MSR [GN11, YGD13, LP15], this is for instance the case of sound Source Localization (SL) from arrays of microphones [PGPM13, DLJ+22].

Presented from general to particular, our research question is articulated around accelerating three specific contexts currently known for their high computational requirements when addressed using state-of-the-art methods.

(RQ. 1) Accelerating the "strong" sparse signal decomposition defined from the minimization of an ℓ_0-regularized problem,

(RQ. 2) Accelerating the decomposition into *continuous* sparse signals in the estimation context of sensor networks and

(RQ. 3) Accelerating the "single-step" estimation processes for sensor networks.

Each context is ultimately particularized for use in MSR systems.

1.1 Targeted Acceleration Contexts

Before detailing the outline of this dissertation with the actual descriptions of our contributions, let us deepen the motivation of the three research questions listed above.

Regardless of the applicative context, finding a sparse representation of a given measurement $\boldsymbol{y}$ consists of identifying a small subset of an extensive collection of vectors, called "atoms", along with weights forming a linear combination approximating $\boldsymbol{y}$. Introducing a few notations, with $\boldsymbol{A}$ being a matrix gathering in its columns the collection of atoms, this task is equivalent to finding a *sparse* vector $\boldsymbol{s}$ —*i.e.,* with a few of non-zero components— such that $\boldsymbol{y} \simeq \boldsymbol{A}\boldsymbol{s}$. This task can be accomplished in many ways, ranging from potentially very fast heuristics to solving intensive non-convex optimization problems. The raw or "strong" formulations of the decomposition problems involve the inconvenient use of the "ℓ_0 pseudonorm operator". These have been proven to define NP-hard problems and are condemned to be solved within a combinatorial computational complexity. Standard approaches to solve them, like Branch-and-Bound (BnB) algorithms [Cpl09, BNCM15, BKM16, BCWP21, HMS22, GHE22], thus maintain the complexity of an exhaustive search in the worst case scenario, rendering them not suited for real-time sensing scenarios.

To address this challenge, numerous alternative formulations emerged over the past few decades, offering tractable procedures with polynomial computation time able to recover approximate solutions of the raw decomposition problem [FR13, Ch. 3]. Among these, the Orthogonal Matching Pursuit (OMP) [PRK95] stands as one of the fastest heuristics and is famous even within specific sensor network research communities for its effectiveness in recovering "very sparse" decomposition. Although OMP and other methods successfully recover the solutions of the raw formulation in "easy" setups, they usually fall short for more challenging instances of the problem.

Addressing the problem formulated in RQ. 1, we initially investigated the possible enhancement of OMP by incorporating tests originating from BnB procedures into its iterations, potentially leading to greedy algorithms that align more closely with the exact solution of the raw sparse decomposition problem. However, defining the intended enhanced OMP appeared to be a more challenging task than initially expected since BnB procedures rely on convex relaxations which are conventionally defined too loosely for potential extensions to the OMP context. In collaboration with the team of Dr.

Cédric Herzet during my stay at INRIA (Rennes), we thus decided to concentrate our effort on a correlated topic: **Accelerating the BnB method tailored to solve the raw problem.** Our acceleration operates by iteratively tightening the convex relaxation internally exploited by the BnB procedures. The resulting technique, coined "peeling" straightforwardly applies to many scientific fields that find interest in solving the raw ℓ_0 problem, such as statistics, machine learning, or inverse problems [CT05, DET05, TW10]. It is additionally hoped to pave the way for future work on new generations of greedy algorithms for sparse signal decomposition.

Next, RQ. 2 extends beyond the discrete framework considered in RQ. 1. While this discrete framework characterized by a finite number of atoms in $\boldsymbol{A}$ has maintained the focus in the field of sparse signal representation, many applications strive to recover a sparse representation that describes a physical reality in a continuous domain. One such example is localization sensor systems, whose objective is to represent the scene as a sparse signal parameterized by the spatial domain. The resulting sparse signal's support corresponds to a map indicating the locations of the observed targets. Since locations are taken from a continuous domain, applying the conventional discrete approaches amounts to defining a gridding of the observed scene. Although the grid approach is widely employed, it exhibits limitations resulting in low estimation performance when using a coarse grid or excessive computational time when employing a finer grid.

Alternatively, existing decomposition algorithms relying on continuous parametric dictionaries of atoms have emerged with multiple formulations, each commonly extending classic *on-the-grid* problems [CFG14, DDP17, SRHG18, DHD19, TAL19, KTTG17]. Most of them involve solving one or multiple highly non-convex optimization problems, often making them impracticable for real-time sensing applications. From this observation, multiple contributions suggesting approximations of the continuous parametric domain by means of interpolation strategies have been proposed over the past decade [ETS11, FDD13, KYHP15, CH19]. In particular, the authors in [KYHP15] have proposed an interpolation-based extension of OMP. Although promising in terms of complexity, their proposition computationally remains out of reach for real-time low-cost applications such as MSR systems. Thus, a focus of our work pertaining to RQ. 2 is to provide a faster adaptation of their proposition that is suitable for such applications. Additionally, we pursued careful efforts to develop a continuous OMP that explicitly matches the particularities of radars, such as the inherent factorization properties of its signal model.

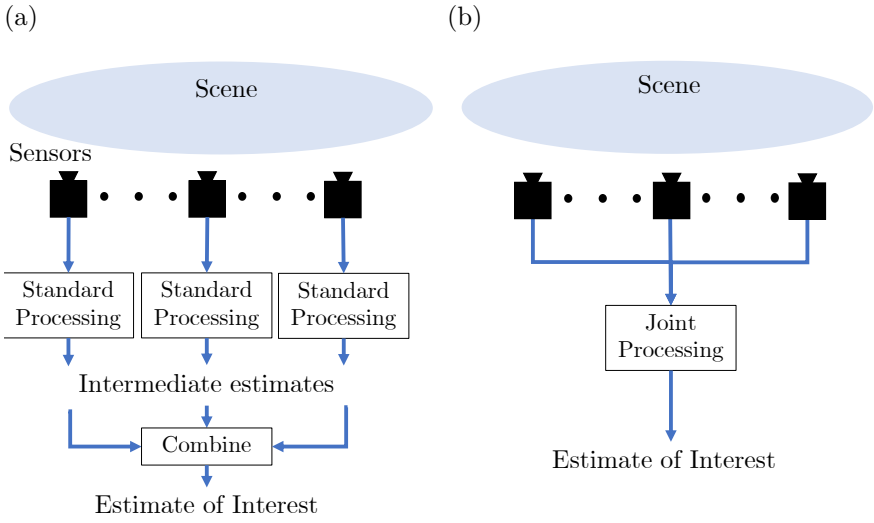

Figure 1.2: Simplified representation of the **(a)** two-step and the **(b)** single-step procedures.

The third research question tackled by this thesis is specific to the sensor network architecture. In multistatic radars, as well as other sensor network applications, the algorithms designed for estimating features from the observed scene are commonly classified into two categories: the "two-step" and the "single-step" procedures. The two-step methodology is commonly employed in sensor network estimation applications because of its simplicity of implementation and low computational requirement. As shown in Fig. 1.2(a), two-step procedures consist of each sensor applying first the same signal processing technique one would use for the sensor alone and, next, combining the resulting intermediate estimates to obtain those of interest.

Single-step methods have long been left aside in many applications (including MSR) because of their potentially very high computational complexity. They are, however, known to provide more robust results against many practical non-idealities in the measurement processes [DSB01, P.03, ZFZ08, BKY16]. In specific cases such as sound SL applications, the single-step method has maintained interest and led this community to develop some application-specific accelerated single-step algorithms [CML11, DDSvW21]. Our contribution associated with RQ. 3 is, therefore, to fill the gap between

research communities by extending acceleration approaches to a generalized framework that encompasses multiple sensor networks sharing common properties with MSR and SL contexts. We named our generalized acceleration framework "Grid Hopping (GH)". This technique relies on the interpolation strategies studied in RQ. 2 to accelerate evaluations of objective functions on a grid of interest from a more practical gridding of the sensing space.

In the next section, we provide an overview of this document's organization and explain its contributions in connection with our research questions.

1.2 Outline and Contributions

Fig. 1.3 describes the overall organization of this thesis. The subsequent chapters are organized into four distinct parts, each represented by a color in this diagram. The first part is depicted by gray blocks. It provides no contribution related to our three research questions but instead focuses on providing the reader with the theoretical background required to go through the following chapters. More specifically, basic concepts on radar systems are introduced in Chapter 2 and will serve for the validation of our methods proposed in Chapters 5 to 7. The sparse coding theory is another crucial ingredient of this research and standard concepts are explained in Chapter 3.

The second part of this document is represented by the blue blocks in Fig. 1.3. That part focuses on acceleration strategies for the sparse decomposition of a *single* measurement (typically, given a single sensor). We respectively address RQ. 1 and RQ. 2 in Chapters 4 and 5.

The third part is represented by green blocks in Fig. 1.3 and tackles acceleration procedures designed for sparse decomposition in the sensor network context. While Chapter 6 is more of a transition chapter from single sensor to sensor networks, our main contributions regarding RQ. 3 are given in Chapter 7.

The last part, shown by the white block, provides discussions on the contributions and concludes this work. In detail, the role and contributions of each chapter are as follows.

- In **Chapter 2**, we introduce essential concepts on radar systems. More specifically, an emphasis is set on Frequency-Modulated Continuous Wave (FMCW) radars, which fit low-range and low-cost applications well. In that chapter, we explain the derivation of a simplified

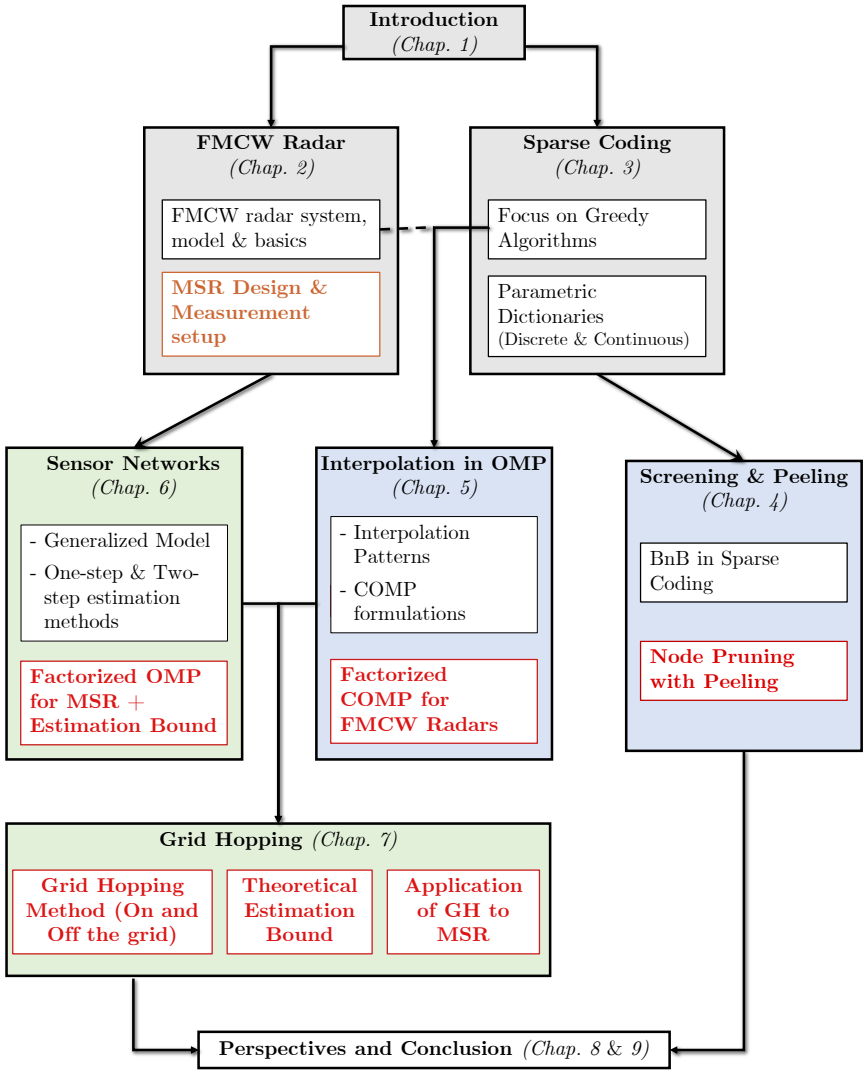

Figure 1.3: Outline of the thesis. Gray blocks represent the first part: "Introduction and Background". The blue and green blocks stand for the chapters that respectively belong to Part II and Part III. The white block corresponds to the conclusion. White sub-blocks represent descriptions of existing concepts in the literature, and red texts are actual contributions. The orange text is a minor side contribution we briefly overview in this document.

model for the signals acquired by such a system. We also introduce signal processing techniques that will later be assimilated into the sparse coding theory. Finally, this chapter presents a side and minor contribution (represented by the orange box). It consists of the design of an active MSR system that I conducted with Maxime Drouguet. Since the challenges associated with this design lie outside the more theoretical scope of this thesis, we focus on a comprehensive description of the system and explain the measurement campaign we performed with the system. The data acquired by those measurements will later be used to assess our accelerated algorithms in Chapter 7.

- **Chapter 3** is another introductive chapter focusing on more theoretical considerations. In substance, the general problem of decomposing a signal into a linear combination of a few waveforms taken from a dictionary is presented in both discrete and continuous formulations. The raw NP-hard problem involving the ℓ_0 norm is presented, and the standard BnB approach to solving it is introduced. This chapter also presents OMP as well as simple continuous extensions of this greedy algorithm. All the concepts from that chapter are extensively used in Chapter 4. Chapter 5 and 7 also rely on concepts provided in Chapter 3 since they intend to extend OMP.

- With **Chapter 4**, we introduce the first contribution of this thesis, addressing RQ. 1. This contribution is the result of a collaboration with Théo Guyard, Clément Elvira, and Cédric Herzet during my stay at INRIA Rennes and presented at EUSIPCO 2023 [GMEH23]. In that chapter, we present our novel methodology, coined peeling, to tighten the convex relaxation used in BnB procedures. This results in a faster version of this algorithm, allowing a quicker exploration of the decision tree intrinsically built by BnB.

- **Chapter 5** moves on to RQ. 2 and starts with a review of interpolation techniques in sparse coding. Next, we describe existing formulations of continuous OMP algorithms. Studying them leads to the definition of a new interpolation-based continuous version of OMP that best matches the requirement of targeted real-time applications. In particular, we extend our algorithm to the factorized context of the FMCW radar signal, described in Chapter 2. These latest results have been presented at iTwist'20 [dGVJ20] and ICASSP'21 [dGFVJ21].

- **Chapter 6** acts as a transition toward our third research question. In that chapter, we gathered multiple sensor network applications into a single framework. The resulting unified model is thereby used to rewrite the principle of two-step and single-step methods. In particular, this chapter offers a contribution in the scope of RQ. 3 by extending a work proposed by my Master thesis on the specific application of single-step procedures to the factorized model resulting from the context of MSR. Indeed, we provide a novel theoretical bound enabling us to quantify the precision loss due to the approximations required for the factorization of the single-step procedures. Using Monte-Carlo simulations, we also demonstrate superior robustness against the noise of the single-step methods over their two-step counterparts in the MSR context.

- In **Chapter 7**, we introduce and study our main contribution in the sensor network field: Grid Hopping (GH). The technique relies on the concepts intensively described in Chapters 5 and 6. After a detailed connection with similar techniques in the literature, we present the general formulation of Grid Hopping (GH). Next, we evaluate our technique's performance through Monte Carlo simulations and with a novel theoretical error bound on its output in a simplified scenario. Finally, specific concerns about the practical implementations of GH to the MSR context are addressed. We end the chapter by applying the three methods (two-step, single-step, and GH) to the measurement acquired with the system that we designed and described in Chapter 2. We recently presented the application of GH to multistatic radars at the ISCS 2023 conference [MFD+23].

- In **Chapter 8** we discuss the contributions of this thesis. We investigate new possibilities to extend our work and connect it to other fields and concepts such as machine learning, target tracking, and the Cramèr-Rao Lower Bound (CRLB).

In the next section, we list the publications produced by the research presented in those chapters.

1.3 List of Publications

1.3.1 Journal Papers

- *(to be Submitted in IEEE Transaction on Signal Processing)* G. Monnoyer, T. Feuillen, L. Vandendorpe, L. Jacques, *Grid Hopping in Sensor Networks: Acceleration Strategies for One-Step Estimation Algorithms*.

- *(Published, Dec. 2022)* M. Bodehou, G. Monnoyer, M. Drouguet, K. Al Khalifeh, L. Vandendorpe, C. Craeye, *Metasurface Antennas for FMCW Radar*, in IEEE Antennas and Wireless Propagation Letters, vol. 22, no. 5, pp. 1040-1044, May 2023.

1.3.2 Conference Papers

- *(Sep. 2023)* T. Guyard, G. Monnoyer, C. Elvira, C. Herzet, *Safe peeling for l0-regularized least-squares*, in Proceedings of European Signal Processing Conference 2023 (EUSIPCO 2023).

- *(Jun. 2023)* G. Monnoyer, T. Feuillen, M. Drouguet, L. Jacques, L. Vandendorpe, *Grid Hopping: Accelerating Direct Estimation Algorithms for Multistatic FMCW Radar*, in Proceedings of the International Symposium on Computational Sensing 2023 (ISCS2023).

- *(Jun. 2021)* G. Monnoyer, T. Feuillen, L. Vandendorpe, L. Jacques, *Sparse Factorization-Based Detection of Off-the-Grid Moving Targets Using FMCW Radars*, in Proceedings of the 2021 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2021).

- *(May 2021)* G. Monnoyer, T. Feuillen, L. Vandendorpe, L. Jacques, *Going Below and Beyond Off-the-Grid Velocity Estimation from 1-bit Radar Measurements*, in Proceedings of 2021 IEEE Radar Conference (RadarConf21).

- *(Jun. 2020)* G. Monnoyer, L. Vandendorpe, L. Jacques, *Factorization over interpolation: A fast continuous orthogonal matching pursuit*, in Proceedings of International Traveling Workshop on Interactions between low-complexity data models and Sensing Techniques (i'TWIST20).

- *(Dec. 2019)* G. Monnoyer, T. Feuillen, L. Jacques, L. Vandendorpe, *Sparsity-Driven Moving Target Detection in Distributed Multistatic FMCW Radars* in Proceedings of 2019 IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP).

1.4 Additional Contributions

In parallel with the work performed on the thesis, I actively contributed to the "LUMINET" Walloon region's project. This project aimed to investigate new possibilities to take advantage of the lampposts structure in cities to incorporate traffic monitoring systems based on the collaborative working of opportunistic wireless source localization and multistatic radar systems. Participating in this project provided me with specific feedback on my theoretical work from researchers whose focuses lie on more practical concerns, hence increasing the interdisciplinary aspect of my work. Moreover, it provided a favorable context, in terms of available resources, for the design of our active multistatic radar working in the K-band. While this system is explained in Chapter 2, we take a brief instant at the end of this introduction to mention a corollary work on that topic: the development of the software that I called "Picoradar". It is open source and can be found at `https://gitlab.com/gilles.monnoyer.dg/Picoradar`.

The name Picoradar refers to its initial implementation intended to interface the "Picoscope" acquisition board. Still, I ultimately extended the software to interface both this board and the one from *National Instrument* mentioned in Chapter 2. Picoradar provides an efficient user-friendly graphical interface tailored for FMCW multistatic radar applications. It encapsulates all the interfacing procedures to translate from the high-level radar signal point of view toward the low-level electronic considerations of the acquisition board. In other words, it enables the efficient application of all the electronic modifications in the acquisition board from a signal-oriented graphical interface in a real-time fashion. This software was therefore of great help in developing and debugging the MSR hardware as well as efficiently performing the measurement. The easy-to-handle interface offered by my software also makes it an excellent pedagogical tool that we have been using for the past three years since its early development in project courses intended to introduce radar systems to third-year bachelor students. Picoradar enabled the student to observe in real-time the results of their

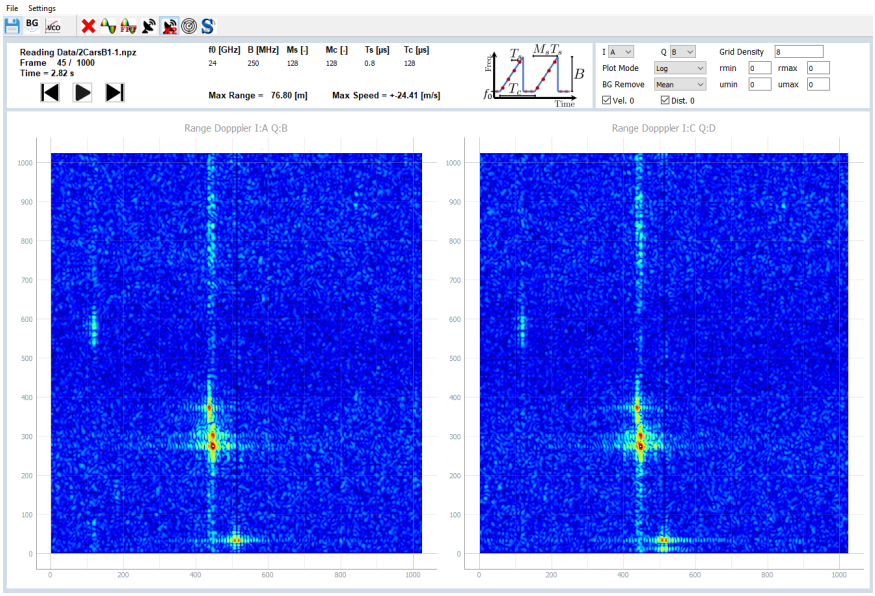

Figure 1.4: graphical interface of Picoradar in "read mode".

algorithms while acquiring data and to directly observe the impact of any change in the radar modulation's properties. Indeed, the software offers a short and simple kind of "API" to easily include new custom processing tabs, which is a convenient feature for the students, as well as during hardware development where one desires to switch in real-time from one processing to another.

Fig. 1.4 shows an overview of the graphical interface offered by Picoradar. The software has two modes: (i) an "acquisition mode" for which the acquisition board must be connected to the computer to offer all the real-time features mentioned above and (ii) a read mode enabling the simulation of the record mode by loading data recorded during an instance of acquisition mode.

Chapter 2

Background on FMCW Radar Systems

This chapter provides the reader with the background on the FMCW radar technology. We begin with the derivation of a simplified signal model that will be used in subsequent chapters to particularize our approaches to the radar context. A specific emphasis is set on the assumptions and approximations required for the derivation of this standard model, as well as insight into the potential impacts they have on the techniques proposed in the following chapters. Standard targets' parameter estimation algorithms are also presented and illustrated.

We finally describe the principle of the multistatic radar. We overview the conception of an experimental setup designed in collaboration with Maxime Drouguet.

Thereby, this chapter's purpose is twofold: Introduce the reader to the main concepts of FMCW radar systems and describe the experimental setup from which we recorded real-world data that we use in the next chapters to validate the algorithms developed throughout this work.

2.1 Introduction to Radar Systems

In civilian, monitoring, or military applications, radar systems have been widely considered to offer low-cost solutions. In practice, radar systems offer a wide range of possible designs, each with its unique properties and advantages. Two fundamental classes of radars are Pulse Doppler Radars

(PDR) and FMCW Radars. While Pulse Doppler Radars (PDR)s use sequences of short pulses separated by long breaks, FMCW radars employ continuous modulated signals. While this chapter is focused on the design and the theory of FMCW radars, understanding first the ideas behind the PDR enables us to understand better the specificity of FMCW radars.

PDRs represent the most intuitive design for a radar. In brief, the short pulses they transmit $s_{\mathrm{T}}(t)$ are waveforms modulated around a carrier frequency f_0. At the reception, the distance information is encoded in the delay of the received signal, which is the echo of the transmitted one that the target has reflected. Speed information can be recovered by inspecting the Doppler effect on the received waveform compared to the known properties of the transmitting one. Alternately, the phase shift between successive pulses can be used to determine the speed. The resolution of the range estimation is directly connected to the pulse duration T_{P} by $\frac{\mathsf{c}T_{\mathrm{P}}}{2}$, where c stands for the speed of light. Since the pulse duration is approximately the inverse of the pulse's bandwidth, it results by the Shannon-Nyquist sampling theorem that given the sampling period, T_{s} must be smaller than T_{P}. This implies that PDRs require a high sampling frequency to reach a fine resolution. For example, the multistatic PDR developed in [DDWB07] uses a 100MHz sampling frequency for a 6m resolution. PDRs are therefore more suited for long-range applications, such as aircraft detection, where such a broad resolution makes sense. This is all the more the case as PDRs enables to concentrate a low average power into brief and powerful pulses able to reach long distances and compensate for the resulting path loss.

On the other hand, FMCW radars transmit uninterrupted signals, with an instantaneous carrier frequency $f_{\mathrm{c}}(t)$, modulated around a central carrier. The received signal is demodulated by mixing it with the currently transmitted signal. The distance information is thus contained in the difference between the received frequency and the one that is currently being transmitted. Since the demodulated signal directly contains this difference in frequency, the resolution no longer depends on the sampling frequency. This resolution is instead given by

$$\text{Range Resolution} = \frac{\mathsf{c}}{2B}, \tag{2.1}$$

where B denotes the bandwidth over which the modulation is performed. For a given modulation, the sampling frequency instead imposes a maximal unambiguous range. This aspect will be clarified in the next sections of this chapter. Typical sampling frequencies used by FMCW radars are thus

around a few MHz at most [FMV16], which makes them suited for low-cost short-range applications. Considering this aspect as well as the expertise of the UCLouvain research team in telecommunications, we therefore focused on the design of an active FMCW radar in order to validate our works on accelerated estimation algorithms for sensor networks. The subsequent sections develop the FMCW radar theory and we end this chapter by describing our Multistatic Radar (MSR) system.

2.2 FMCW Radar System Description

With this section, we present the practical implementation of a simple FMCW radar composed of a single transmitter and receiver. The main hardware components used by this system are represented in Fig. 2.1. The scheme on this figure is valid for both the monostatic configuration, where the transmitter (TX) and the receiver (RX) are colocated, and the bistatic configuration in which the receiver is remotely located.

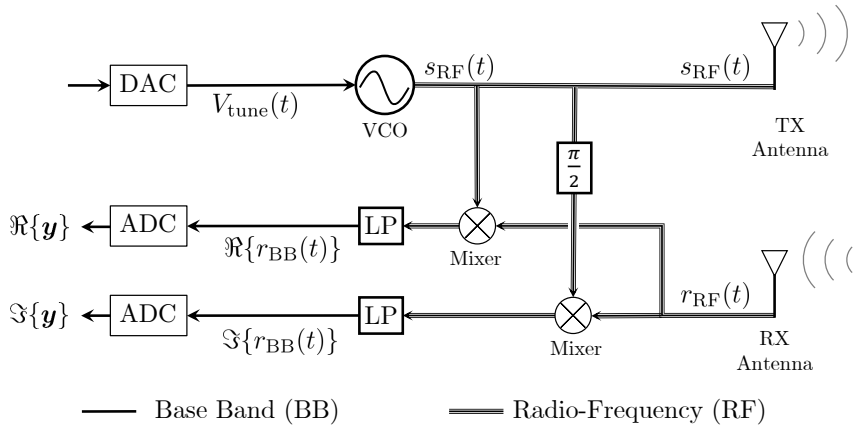

Figure 2.1: Simplified schematic representation of the main hardware components in the IQ demodulation of a FMCW radar composed of a single transmitting antenna and a single receiving antenna.

The radar system shown in that figure is controlled by the V_{tune} signal which is generated according to the desired modulation and the frequency response of the Voltage Controlled Oscillator (VCO). This signal inputs the VCO which performs the Frequency Modulation (FM) process by outputting the Radio Frequency (RF) signal $s_{\text{RF}}(t)$, itself broadcasted by the

TX antenna. At the receiving side, the signal $r_{\mathrm{RF}}(t)$ provided by the RX antenna is first amplified before feeding two "mixers" for the demodulation process. Each mixer multiplies $r_{\mathrm{RF}}(t)$ either by the transmitted signal $s_{\mathrm{RF}}(t)$ or by a $\pi/2$ phase-shifted version of it. After the application of a low pass filter, the real and imaginary parts of the baseband signal $r_{\mathrm{BB}}(t)$ are respectively obtained before being sampled at a rate that depends on the performance targeted by the system. The sampling process finally provides the measurement $\boldsymbol{y}$ whose expression is derived in the next section.

We end this section by drawing attention to two aspects of the system which need to be carefully taken into account in the actual design of the multistatic radar presented at the end of this chapter.

The first aspect concerns the properties of the VCO. This component outputs a signal whose instantaneous frequency is driven by the value $V_{\mathrm{tune}}(t)$ following its frequency response $f_{\mathrm{VCO}}(\cdot)$. Thereby, assuming a unit amplitude for $s_{\mathrm{RF}}(t)$ without loss of generality, its ideal expression reads

$$s_{\mathrm{RF}}(t) = \cos\left(2\pi \int_0^t f_{\mathrm{VCO}}\big(V_{\mathrm{tune}}(t')\big)\mathrm{d}t' + \phi_0\right), \tag{2.2}$$

where ϕ_0 stands for the initial phase of the VCO. For reasons beyond the scope of this work, the response f_{VCO} is commonly non-linear and temperature-dependent. Therefore, given a desired modulation $f_{\mathrm{c}}(t)$ of the transmitted frequency, the feeding signal V_{tune} needs to be carefully calibrated. When this calibration is imperfect —and such that $f_{\mathrm{VCO}}\big(V_{\mathrm{tune}}(t)\big)$ differs from the targeted modulation $f_{\mathrm{c}}(t)$— mismatches between the actual received signal and the model from which the estimation algorithms are derived happen. These mismatches grow stronger as the delay between the transmission of $s_{\mathrm{RF}}(t)$ and the reception of $r_{\mathrm{RF}}(t)$ at the mixer's level is higher. Therefore, this non-ideality becomes a stronger concern in bistatic —and hence multistatic— configurations where at least one RF signal must inevitably be carried across long distances, inducing extra delays in the received signal with respect to the transmitted one.

As a second aspect, we highlight that the delay from which distances are estimated is captured in $\boldsymbol{y}$ at the mixer level. Said differently, the measurement $\boldsymbol{y}$ is parameterized by the delay as measured between the output of the VCO and the reception of $r_{\mathrm{RF}}(t)$ in the mixer. It is therefore commonly necessary to correct the estimated distance by subtracting a distance that approximately corresponds to the length of the RF cables. A simple calibration procedure, consisting of the observation of a target with a known

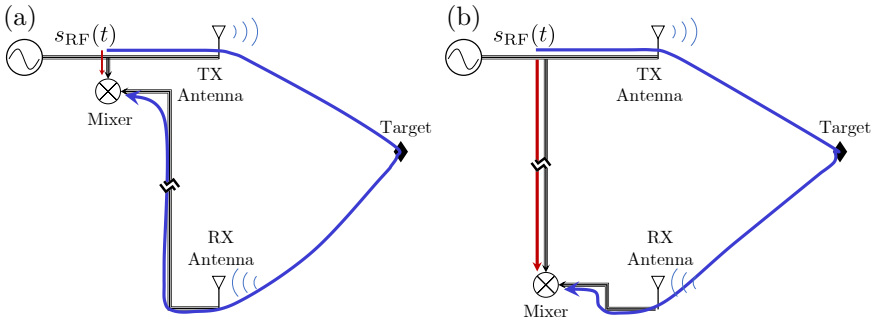

Figure 2.2: Two possible configurations for the mixer's location in a bistatic radar configuration. The resulting delay corresponds to the difference between the distances of the blue path minus the red path.

location, is enough to determine the value of the corrective constant term. While this aspect is often anecdotal in monostatic configurations, bistatic —and hence, multistatic— radars must consider it carefully in its design.

In short, the mixers of a multistatic radar can be either colocated with the transmitter or the receiver, as shown in Fig. 2.2. When the mixer is located with the transmitter as in (a), then the delay caused by the long RF connection between the nodes is positively added to the delay of interest (the one from the path TX-target-RX). On the other hand, the delay of the connection is subtracted in the second configuration. Since the connection commonly follows the shortest path from TX to RX, the second configuration typically does not cause the observation of negative delays (*i.e.,* the blue delay is typically always longer than the red one). We explained before that some non-idealities, such as an imperfectly calibrated V_{tune} signal, show a stronger impact on the signal's quality when affecting longer delays. The second configuration should therefore be preferred to optimize the quality of measurement. In practice however, we still considered the first configuration in the design of our MSR system because placing all the mixers associated with all receiving nodes in the same location enabled an easier management of power supplies, and hence a reduced cost as well as a better stability of the system with less power amplifiers.

The third element affecting the performance of radar measurement is the clutter. The clutter is a generic term used in radar systems to name all the undesired and non-constant signals that cannot be considered as noise or interference. The three most common sources of clutter are (i)

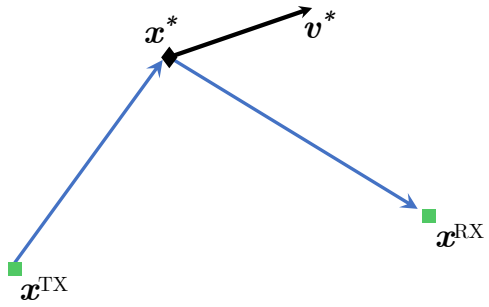

Figure 2.3: Representation of the simple estimation scenario of a FMCW bistatic radar, graphically showing the main geometrical quantities used in the model. The blue arrows depict the path followed by the wave in the propagation process.

the internal leakage within the RF circuit, (ii) the direct transmission from the TX antenna toward the RX antenna, and (iii) reflections on the static "background" of the observed scene. Although many techniques exist to remove clutter from radar measurements, either by means of an enhanced hardware design or with more sophisticated signal processing techniques, we restrict our application to an aggressive pre-processing step. In brief, we erase from the signal all contributions from non-moving targets.

All in all, these three aspects constitute the most important sources of differences between the theory studied in this whole work and the results from the experimental measurements.

2.3 FMCW Radar Signal Model

We can now derive a simplified model for the measurement $\boldsymbol{y}$ received from the FMCW radar represented in Fig. 2.1. The standard radar system is composed of one transmitting antenna located in $\boldsymbol{x}^{\mathrm{TX}}$ and one receiving antenna located in $\boldsymbol{x}^{\mathrm{RX}}$. This couple forms a *bistatic pair*, which is a generic building block for all radar configurations. Indeed, the monostatic radar is a particular case of bistatic radar with $\boldsymbol{x}^{\mathrm{TX}} = \boldsymbol{x}^{\mathrm{RX}}$, and a multistatic radar is a collection of multiple bistatic "transmitter-receiver" pairs.

We begin our derivation by considering the observation of a scene composed of a single point target located in $\boldsymbol{x}^*$ and moving with a constant velocity vector $\boldsymbol{v}^*$, respectively known to belong to domains of interest $\mathcal{X}$ and $\mathcal{V}$, as depicted by Fig. 2.3. By the end of this section, we extend the

model to the observation of K targets with the k-th target's location and velocity respectively denoted by $\boldsymbol{x}_k^*$ and $\boldsymbol{v}_k^*$.

Let us emphasize that, for the sake of simplicity in the notations throughout this presentation, we adopt the following normalization.

Definition 2.1 (Normalization of the Target Parameters)**.** All locations provided in a radar context are expressed with dimensionless quantities, normalized with respect to the radar's range resolution given by $\frac{\mathsf{c}}{2B}$. Similarly, the velocities are expressed in a normalized manner with respect to the radar's speed resolution given by $\frac{\mathsf{c}}{2f_0T_{\mathrm{PF}}}$, for a carrier frequency f_0 and a processing frame duration T_{PF}.

The expressions used for the resolutions in the normalization originate from the reasoning given at the end of this section. To make this normalization clearer, consider the FMCW radar with a bandwidth $B = 250$MHz, a carrier $f_0 = 24$GHz, and acquiring processing frames of duration $T_{\mathrm{PF}} = 10$ms. This radar's range and speed resolution are given by .6 meters and .625 meters per second. Thus a target located in $[12, 6]$ in meters and moving following a velocity vector $[5, 7.5]$ in m/s is modelled by the unitless vectors $\boldsymbol{x}^* = [20, 10]$ and $\boldsymbol{v}^* = [8, 12]$. In addition to simplifying the signal model, this normalization allows for a direct assessment of the efficiency of the algorithms proposed in this presentation. This is because the estimation performance on $\boldsymbol{x}^*$ and $\boldsymbol{v}^*$ are independent of the radar's internal properties, which are determined by the choice of its modulation parameters.

2.3.1 Signal Model Derivation

After the reflection of the transmitted wave on a single point target modeled by $(\boldsymbol{x}^*, \boldsymbol{v}^*)$, the signal received by the system is modeled as a noisy version of the transmitted signal $s_{\mathrm{RF}}(t)$ affected by a delay-Doppler term τ. Assuming V_{tune} to be perfectly calibrated to get the desired modulation $f_{\mathrm{c}}(t)$ at the output of the VCO, the received signal is modelled by

$$r_{\mathrm{RF}}(t) = As_{\mathrm{RF}}(t-\tau) + w(t) \tag{2.3}$$

$$= A\cos\left(2\pi\int_{\tau}^{t} f_{\mathrm{c}}(t'-\tau)\mathrm{d}t' + \phi_0\right) + w(t) \tag{2.4}$$

$$= A\cos\left(2\pi\int_{0}^{t-\tau} f_{\mathrm{c}}(t')\mathrm{d}t' + \phi_0\right) + w(t), \tag{2.5}$$

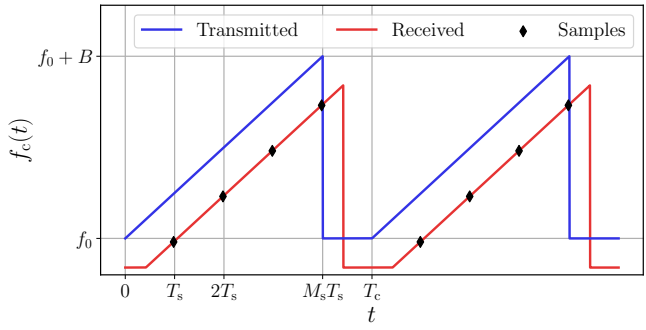

Figure 2.4: Schematic representation of the linear chirps frequency modulation on both the transmitter and receiver sides. This modulation is characterized by a bandwidth B, a main carrier f_0, and a chirp repetition period T_c.

where A is an attenuation factor. In short, this factor integrates the combined effects on the received power of the path loss of the transmission, the gains of the antennas, and the reflectivity of the target. We shall discuss this aspect in more detail by the end of this section. The Additive White Gaussian Noise (AWGN) term $w(t)$ amounts to the noise that mainly originates from the system's hardware components. After the coherent demodulation, the resulting baseband signal is modeled by

$$r_{\mathrm{BB}}(t) = \bar{\alpha} \exp\left(\mathrm{j}2\pi\Big(\int_0^t f_c(t')\mathrm{d}t' - \int_0^{t-\tau} f_c(t')\mathrm{d}t'\Big)\right) + w(t), \tag{2.6}$$

where $\bar{\alpha} \in \mathbb{C}$ is a complex coefficient whose amplitude is proportional to A and phase depends on the demodulation process.

The modulation that has the focus in our work — since we are employing it in our MSR system— is the linear chirp function represented in Fig. 2.4. This modulation is characterized by a bandwidth B, a main carrier f_0, and a chirp repetition period T_c. The chirp duration is given by $M_sT_s \leq T_c$ where M_s and T_s respectively stand for the number of samples acquired per chirp and the sampling frequency used in the acquisition process. A processing frame is composed of M_c successive chirps, hence resulting in $T_{\mathrm{PF}} = M_cT_c$.

In Fig. 2.4, the distance information is encoded in the horizontal shift, which is in turn translated into a constant frequency difference between the transmitted and the received frequencies in a given linear chirp. Since the demodulation can be interpreted as the extraction of this difference of

frequency, estimating range from a FMCW radar using linear chirp amounts to determine a constant frequency in the demodulated signal. The speed information is translated into successive phase shifts from one chirp to the next in the received signal. It thereby produces a constant frequency on the signal built by taking one corresponding sample in each chirp from the demodulated signal.

We underline that while forcing the equality $M_\mathrm{s}T_\mathrm{s} = T_\mathrm{c}$ appears as a natural choice, the respective settling times of the Digital to Analog Converter (DAC) and VCO constrain us to insert breaks between subsequent linear chirps. Given m_c and m_s respectively denoting a chirp index and a sample index inside the chirp, we must ensure that the $(m_\mathrm{c}M_\mathrm{s} + m_\mathrm{s})$-th acquired sample always provides a sample in the echo of the m_c-th transmitted linear chirp. To this end, we additionally assume that $\tau < T_\mathrm{s}$. Still, this constraint can be relaxed at the cost of a loss in the effective bandwidth if one discards the first acquired samples of each received chirp.

For a given chirp index m_c, the modulated frequency $f_\mathrm{c}(t)$ with the repetition period T_c represented in Fig. 2.4 is thus expressed by

$$f_\mathrm{c}(t) = \begin{cases} f_0 + \frac{B}{M_\mathrm{s}T_\mathrm{s}}(t - m_\mathrm{c}T_\mathrm{c}) & \text{if } t \in [m_\mathrm{c}T_\mathrm{c},\ m_\mathrm{c}T_\mathrm{c} + M_\mathrm{s}T_\mathrm{s}] \\ f_0 & \text{otherwise.} \end{cases} \tag{2.7}$$

With this expression, one reformulates (2.6) by applying

$$\int_0^t f_\mathrm{c}(t')\mathrm{d}t' = f_0 t + \tfrac{1}{2}BM_\mathrm{s}T_\mathrm{s}m_\mathrm{c} + \tfrac{1}{2}\frac{B}{M_\mathrm{s}T_\mathrm{s}}(t - m_\mathrm{c}T_\mathrm{c})^2. \tag{2.8}$$

The demodulated signal corresponding to the m_c-th chirp is then written as

$$r_\mathrm{BB}(t) = \bar{\alpha}\exp\Big[-\mathrm{j}2\pi\big(f_0\tau + \frac{B}{M_\mathrm{s}T_\mathrm{s}}\,\tau\,(t - m_\mathrm{c}T_\mathrm{c}) - \tfrac{1}{2}\frac{B}{M_\mathrm{s}T_\mathrm{s}}\tau^2\big)\Big] + w(t). \tag{2.9}$$

For the next step of this derivation, we link the delay-Doppler term τ to the target's parameters $\boldsymbol{x}^*$ and $\boldsymbol{v}^*$. To this end, let us notice that this term encompasses the delay corresponding to the path from $\boldsymbol{x}^\mathrm{TX}$ to $\boldsymbol{x}^*$ and coming back to $\boldsymbol{x}^\mathrm{RX}$ as well as the Doppler shift due to the target's velocity. Those are commonly modeled using the concepts of *bistatic range*

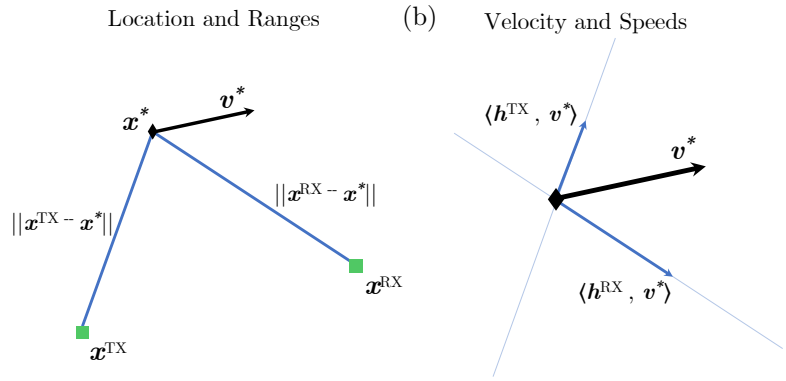

Figure 2.5: Representation of the simple estimation scenario of a FMCW bistatic radar, connecting the location to the bistatic range and the velocity to the bistatic speed.

and *bistatic speed* which are respectively defined by

$$r^* := \tfrac{1}{2}\left(\|\boldsymbol{x}^* - \boldsymbol{x}^{\mathrm{TX}}\|_2 + \|\boldsymbol{x}^* - \boldsymbol{x}^{\mathrm{RX}}\|_2 + r_{\mathrm{cable}}\right) \quad \text{and} \tag{2.10}$$

$$u^* := \tfrac{1}{2}\left\langle \boldsymbol{h}^{\mathrm{TX}} + \boldsymbol{h}^{\mathrm{RX}}, \boldsymbol{v}^* \right\rangle, \tag{2.11}$$

where we define unit directional vectors by

$$\boldsymbol{h}^{\mathrm{TX}} := \frac{\boldsymbol{x}^* - \boldsymbol{x}^{\mathrm{TX}}}{\|\boldsymbol{x}^* - \boldsymbol{x}^{\mathrm{TX}}\|_2} \quad \text{and} \quad \boldsymbol{h}^{\mathrm{RX}} := \frac{\boldsymbol{x}^* - \boldsymbol{x}^{\mathrm{RX}}}{\|\boldsymbol{x}^* - \boldsymbol{x}^{\mathrm{RX}}\|_2}. \tag{2.12}$$

Those quantities are graphically shown in Fig. 2.5. The $\frac{1}{2}$ factors are conventions, thanks to which, for a monostatic configuration, these concepts reduce to the range and the radial speed of the target. The constant term r_{cable} models the additional delay caused by the length of the RF circuitry connecting the VCO to the TX antenna and the RX antenna to the mixer. All in all, given c to denote the speed of light and taking into account the normalization from Def. 2.1, the delay-Doppler term reads

$$\tau = \tfrac{2}{\mathsf{c}}\left(\left(\tfrac{\mathsf{c}}{2B}\right)r^* + \left(\tfrac{\mathsf{c}}{2f_0 M_{\mathrm{c}} T_{\mathrm{c}}}\right)u^* t\right). \tag{2.13}$$

We finally derive the expression of the sampled measurement $\boldsymbol{y} \in \mathbb{C}^{M_{\mathrm{c}} M_{\mathrm{s}}}$. The Analog to Digital Converter (ADC) samples the signal at a rate $1/T_{\mathrm{s}}$ and discards the samples corresponding to breaks between chirps. There-

fore, we construct $\boldsymbol{y}$ by setting its $(m_{\mathrm{c}}M_{\mathrm{s}}+m_{\mathrm{s}})$-th component by

$$y_{m_{\mathrm{c}}M_{\mathrm{s}}+m_{\mathrm{s}}} := r_{\mathrm{BB}}(m_{\mathrm{c}}T_{\mathrm{c}}+m_{\mathrm{s}}T_{\mathrm{s}}). \tag{2.14}$$

Injecting (2.13) and (2.9) in the above ultimately leads to the FMCW complete signal model

$$\boldsymbol{y} = \alpha \boldsymbol{a}^{\mathrm{fmcw}}(r^*, u^*) + \boldsymbol{w}, \tag{2.15}$$

where α is a phase-shifted version of $\bar{\alpha}$ discussed in Sec. 2.3.2 and the $(m_{\mathrm{c}}M_{\mathrm{s}}+m_{\mathrm{s}})$-th component of the FMCW radar *atom* is written as a triple product of complex exponential signals that read

$$\begin{aligned} a^{\mathrm{fmcw}}_{m_{\mathrm{c}}M_{\mathrm{s}}+m_{\mathrm{s}}}(r^*, u^*) := \exp\big[-\mathrm{j}2\pi((r^*+\sigma_{\mathrm{u}}u^*)\tfrac{m_{\mathrm{s}}}{M_{\mathrm{s}}})\big] \\ \exp\big[-\mathrm{j}2\pi u^*\tfrac{m_{\mathrm{c}}}{M_{\mathrm{c}}}\big]\exp\big[-\mathrm{j}2\pi D\big]. \end{aligned} \tag{2.16}$$

In the above, we define the shift factor by $\sigma_{\mathrm{u}} := \frac{M_{\mathrm{s}}T_{\mathrm{s}}}{M_{\mathrm{c}}T_{\mathrm{c}}}$. Moreover, D is a distortion term which non-linearly couples the indices m_{s} and m_{c} and which is sometimes neglected [MFJV19, dGVJ20]. For the study presented in Chapter. 5, we only require to point out that under the assumption $\tau < T_{\mathrm{s}}$, the following inequality holds for all $m_{\mathrm{c}} \in [M_{\mathrm{c}}]$ and all $m_{\mathrm{s}} \in [M_{\mathrm{s}}]$,

$$|D| < \tfrac{B}{2f_0}u^*. \tag{2.17}$$

The third exponential factor can thus safely be neglected for low-speed targets since by definition of a FMCW radar, $B \ll f_0$ and hence $|D| \ll 1$. When a radar is designed for the estimation of higher speeds, the algorithms built upon the model simplified by neglecting D can suffer from model mismatch which deteriorates their results. This effect will be highlighted in the validation of the *off-the-grid* algorithms proposed in Chapter 5.

When the simplification is applied, we can rewrite $\boldsymbol{y}$ in factorized fashion by reshaping it into the matrix $\boldsymbol{Y} \in \mathbb{C}^{M_{\mathrm{c}}\times M_{\mathrm{s}}}$ such that $Y_{m_{\mathrm{c}},m_{\mathrm{s}}} := y_{m_{\mathrm{c}}M_{\mathrm{s}}+m_{\mathrm{s}}}$. The factorization results from the following observation. In (2.16), the first exponential factor expresses a frequency along the dimension indexed by m_{s}. Since this corresponds to samples within the chirps, we refer to this part of the model as the *inner signal*. The second exponential similarly expresses a frequency along m_{c} and is called the *outer signal*. It follows

$$\boldsymbol{Y} = \alpha \boldsymbol{a}^{\mathrm{out}}(u^*) \otimes \boldsymbol{a}^{\mathrm{in}}(r^*+\sigma_{\mathrm{u}}u^*) + \boldsymbol{W}, \tag{2.18}$$

where $\otimes$ denotes the outer vector product. In the above, $\boldsymbol{W}$ corresponds

to the noise vector $\boldsymbol{w}$ which has similarly been reshaped. The inner and outer radar *atoms* are specific cases of *Fourier* atoms. Their expressions are respectively given by $\boldsymbol{a}^{\text{in}}(\cdot) := (a_1^{\text{in}}(\cdot), \cdots, a_{M_\text{s}}^{\text{in}}(\cdot))^\top$ and $\boldsymbol{a}^{\text{in}}(\cdot) := (a_1^{\text{out}}(\cdot), \cdots, a_{M_\text{c}}^{\text{out}}(\cdot))^\top$ with

$$a_{m_\text{s}}^{\text{in}}(r) := \exp\big[-\text{j}2\pi\, r\, m_\text{s}/M_\text{s}\big], \tag{2.19}$$

$$a_{m_\text{c}}^{\text{out}}(u) := \exp\big[-\text{j}2\pi\, u\, m_\text{c}/M_\text{c}\big]. \tag{2.20}$$

Since the range and the speed are both related to a frequency to be estimated from the discretized measurement, the sampling theory imposes that these quantities must be a-priory known to belong to intervals of values with respective widths equal to M_s and M_c. This is required to guarantee an unambiguous estimation. With the normalization from Def. 2.1, the conventional prior knowledge on these are

$$r^* \in \mathcal{R} := [r_{\min},\ r_{\min} + M_\text{s}[\quad \text{and} \tag{2.21}$$

$$u^* \in \mathcal{U} := [-M_\text{c}/2,\ M_\text{c}/2[, \tag{2.22}$$

where $r_{\min}$ is an arbitrary minimal range of interest which is chosen depending on the practical application. To ensure that we meet the above conditions on the range and the speed, we will always assume that the following sufficient conditions are met in the radar context. First, the location domain $\mathcal{X}$ is such that for all $\boldsymbol{x} \in \mathcal{X}$,

$$\begin{aligned} r_{\min} \leq \|\boldsymbol{x} - \boldsymbol{x}^{\text{TX}}\|_2 \leq r_{\min} + M_\text{c} \quad \text{and} \\ r_{\min} \leq \|\boldsymbol{x} - \boldsymbol{x}^{\text{RX}}\|_2 \leq r_{\min} + M_\text{c}. \end{aligned} \tag{2.23}$$

Also, the velocity domain $\mathcal{V}$ is such that for all $\boldsymbol{v} \in \mathcal{V}$,

$$\|\boldsymbol{v}\|_2 \leq \tfrac{M_\text{c}}{2}. \tag{2.24}$$

The reader may have noticed that the frequency of the inner atom is related to a linear combination of the range and the speed, and not the range alone. This has however no effect on the definition of $\mathcal{R}$ since the sampling theory only imposes its span. Moreover, combining $\sigma_\text{u} = \frac{M_\text{s} T_\text{s}}{M_\text{c} T_\text{c}} \leq \frac{1}{M_\text{c}}$ with (2.22) yields $|\sigma_\text{u} u^*| \leq 0.5$. In other words, the shift caused by the velocity in the frequency of $\boldsymbol{a}^{\text{in}}$ is bounded by half the range resolution. While this coupling has no effect on the estimation of ranges and speed from a monostatic (or bistatic) radar, we will show in Chapter 6 that it is

detrimental to the performance of algorithms directly estimating $\boldsymbol{x}^*$ in a single step from the collection of signals provided by a multistatic radar. We will also demonstrate that the further away the receivers are located from the observed scene, the smaller the detrimental effect.

Thanks to the normalization, the models do not depend on the choice of bandwidth, carrier frequency, sampling frequency, and chirp duration. The effect of these parameters' values is gathered in the translation from the normalized to the actual values. Thus, the validation results obtained for a given set of values for $M_\mathrm{s} - M_\mathrm{c}$ are valid regardless of the choice of other parameters.

2.3.2 Modelling the Scattering Coefficient

Let us add a few notes on the model we use in this work for the scattering coefficient α. Just like the range r^* and the speed u^*, the scattering coefficient α is an unknown variable that can be estimated. The coefficient α is related to $\bar{\alpha}$ by

$$\alpha := \bar{\alpha} \exp\big(- \mathrm{j}2\pi \tfrac{f_0}{B} r^* \big). \tag{2.25}$$

Since we consider a K-band signal, the phase shift applying in (2.25) is highly sensitive to the exact value of r^*. Considering a typical value of $f_0 = 24\mathrm{GHz}$, the phase increases by 2π every increment of 6.25 millimeters in the range. As the sizes of the targets themselves are usually significantly larger than 1cm, the phase of α is supposed to be a random variable that follows a uniform distribution between 0 and 2π.

We remind here that the amplitude α is proportional to the attenuation term A introduced in the expression of the received signal (2.3). Therefore, the value of $|\alpha|^2$ is given by the received power P_R, which correlates with the target's location by the "*Radar Equation*". This equation provides an expression of the received power as a function of the transmitted power P_T by

$$P_\mathrm{R} = P_\mathrm{T} \frac{G_\mathrm{T} G_\mathrm{R} \lambda^2}{(4\pi)^3 r_\mathrm{T}^2 r_\mathrm{R}^2} \sigma. \tag{2.26}$$

In the above, G_T and G_R respectively denote the gain of the TX and the RX antennas, λ is the wavelength of the transmitted signal, r_T and r_R are respectively the distances from the TX and RX antennas to the target and σ is a variable called as the Radar Cross Section (RCS). This variable corresponds to an equivalent surface of the target, which is highly dependent on its geometry, material, and the frequency band used by the radar. Due to

this extreme sensitivity, the value of σ is commonly modeled as stochastic with an exponential distribution. Putting all together, it results that α follows a centered complex normal distribution whose variance depends on the target's location following the value of P_{R}.

In all the simulations performed in the subsequent chapters, we however assume that $\alpha \sim \mathbb{C}\mathcal{N}(0,1)$ independently of the target's location. This assumption enables us to focus our result on the conclusions directly drawn from the algorithms' performance. This assumption is motivated by a work that I performed in collaboration with Dr. Modeste Bodehou [BdGD+23]. In this contribution, we demonstrated the effectiveness of low-cost antennas that exhibit radiation patterns that compensate for the path loss, given a known geometry of the radar system. Those antennas were shown to provide signals with a distribution of amplitudes that is independent of the target's location.

2.3.3 Observation of Multiple Targets

We conclude the section by extending the measurement model (2.15) to the observation of multiple targets. Given the observation of K point targets, if k-th of them is modelled by the location-velocity couple $(\boldsymbol{x}_k^*, \boldsymbol{v}_k^*)$, then using the relations (2.10) and (2.11), one can define the associated range-speed couple (r_k^*, u_k^*). Since the waves coming back from the different targets add to one another, the model (2.15) is straightforwardly extended with the linear combination

$$\boldsymbol{y} = \sum_{k=1}^{K} \alpha_k \boldsymbol{a}^{\mathrm{fmcw}}(r_k^*, u_k^*) + \boldsymbol{w}. \tag{2.27}$$

In the above, the scattering coefficient α_k associated to the k-th target is considered to be independent to $\alpha_{k'}$ for all $k' \neq k$. Using the same approximations as explained in Sec. 2.3.1, the model above is factorized into

$$\boldsymbol{Y} = \sum_{k=1}^{K} \alpha_k \boldsymbol{a}^{\mathrm{out}}(u_k^*) \otimes \boldsymbol{a}^{\mathrm{in}}(r_k^* + \sigma_{\mathrm{u}} u_k^*) + \boldsymbol{W}. \tag{2.28}$$

The model (2.27) will extensively be used in the next chapters to generate synthetic radar signals. The atom function $\boldsymbol{a}^{\mathrm{fmcw}}$ will also constitute a building block for algorithms taking the distortions into account. Those algorithms are mainly implemented in Chapter 5 to serve as a comparison standard for the algorithms based on the factorized model (2.28).

In subsequent chapters, both formulations will be connected to the sparse coding theory. More precisely, apart from the noise, the measurement $\boldsymbol{y}$ appears to be decomposable as a sparse signal in the continuous domain $\mathcal{R} \times \mathcal{U}$. The decomposition is expressed either through the single continuous dictionary parameterized by the range-speed couple with the atom function $\boldsymbol{a}^{\text{fmcw}}$, or in a factorized fashion through a pair of inner and outer dictionaries respectively parameterized by the range and the speed.

2.4 Estimation of Target Parameters

Estimating the target's location and velocity vectors requires the use of more than one bistatic pair, *i.e.,* a multistatic radar. This introductive chapter focuses on basic radar theory, and hence we restrict this section to concepts for the estimation of ranges and speeds from one bistatic pair. We leave the estimation algorithm designed for MSR to the Chapters 6 and 7, which are dedicated to sensor networks.

In most radar applications, the factorized model given by (2.18) is used for the design of estimation algorithms. When a single target needs to be located, a conventional maximum likelihood approach leads to the maximization of the 2D Discrete Time Fourier Transform (DTFT) of $\boldsymbol{Y}$. This amounts to a particularization of a correlation between the received signal and the continuous dictionaries given by the parametric atoms $\boldsymbol{a}^{\text{in}}$ and $\boldsymbol{a}^{\text{out}}$. Mathematically, since the scattering coefficient α is unknown and modeled as a complex normal random variable, the range and speed estimates $\hat{r}$ and $\hat{u}$ are computed from

$$(\hat{r}, \hat{u})^\top := \underset{r \in \mathcal{R},\, u \in \mathcal{U}}{\arg\max}\ \tfrac{1}{2} \left| \left[\boldsymbol{a}^{\text{in}}(r + \sigma_{\text{u}} u) \right]^\top \boldsymbol{Y}^H\, \boldsymbol{a}^{\text{out}}(u) \right|^2. \tag{2.29}$$

Remark 2.1. It is important to note that the periodicity of the output of the DTFT implies that the above can be solved by first ignoring the constraint $r \in \mathcal{R},\ u \in \mathcal{U}$ and then projecting the resulting estimates onto the domain $\mathcal{R} - \mathcal{U}$ with an appropriate modulo operation. Moreover, doing so enables us to ignore the term $\sigma_{\text{u}} u$ in the objective function of (2.29) and subtract $\sigma_{\text{u}} \hat{u}$ from the resulting preliminary range estimate.

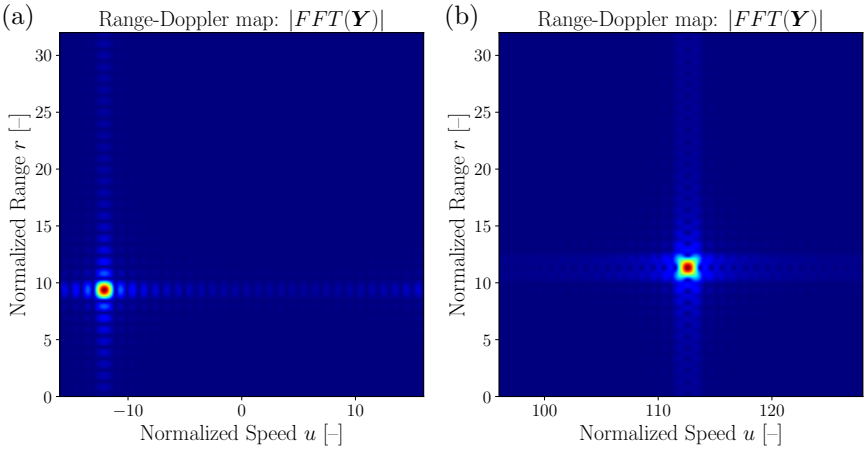

Figure 2.6: Range-Doppler maps of simulated signals following the model (2.15) with one target characterized by **(a)** $r^* = 10$ and $u^* = -12$ and with $M_\mathrm{s} = M_\mathrm{c} = 32$, and **(b)** $r^* = 10$ and $u^* = 112$ and with $M_\mathrm{s} = 32$, $M_\mathrm{c} = 256$, zoomed around the target's value.

The problem (2.29) is commonly solved by an *on-the-grid* strategy which discretizes the space $\mathcal{R}\times\mathcal{U}$. Given a uniform discretization with respectively N_r and N_u range and speed bins, the use of the Fast Fourier Transform (FFT) algorithm yields the computational complexity $O(N_\mathrm{r}N_\mathrm{u}\log(M_\mathrm{s}))$.

Two examples of the output of the 2D FFT of the measurement, called a *range-Doppler map*, are given in Fig. 2.6. In Fig. 2.6(a), $\boldsymbol{y}$ corresponds to the noiseless observation of single target with $r^* = 10$ and $u^* = -12$. Since we used $M_\mathrm{s} = M_\mathrm{c} = 32$, $B = 250\mathrm{MHz}$, and $f_0 = 24\mathrm{GHz}$, the distorsion is negligible because $|D| < 0.17$ for all $u^* \in \mathcal{U}$. The range-Doppler thus corresponds to a 2D Dirichlet kernel shifted according to values of r^* and u^*. More precisely, this figure shows a negligibly distorted version of the function

$$\begin{aligned}|\alpha|\left|\left\langle \boldsymbol{a}^{\mathrm{in}}(r),\, \boldsymbol{a}^{\mathrm{in}}(r^* + \sigma_\mathrm{u}u^*)\right\rangle\right| \left|\left\langle \boldsymbol{a}^{\mathrm{out}}(u),\, \boldsymbol{a}^{\mathrm{out}}(u^*)\right\rangle\right| \\ = |\alpha|\left|\frac{\sin(\pi(r^* + \sigma_\mathrm{u}u^* - r))}{\sin(\pi(r^* + \sigma_\mathrm{u}u^* - r)/M_\mathrm{s})}\right|\left|\frac{\sin(\pi(u^* - u))}{\sin(\pi(u^* - u)/M_\mathrm{c})}\right|, \quad (2.30)\end{aligned}$$

evaluated in all $r \in \mathcal{R}_\mathrm{G}$ and $u \in \mathcal{U}_\mathrm{G}$ where $\mathcal{R}_\mathrm{G}$ and $\mathcal{U}_\mathrm{G}$ respectively denote discretizations of $\mathcal{R}$ and $\mathcal{U}$. The width of the main lobe of the kernel cor-

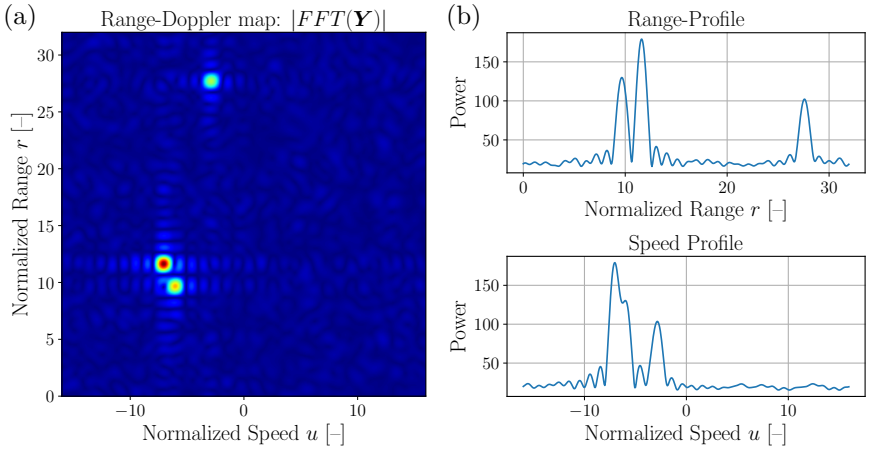

Figure 2.7: **(a)** Range-Doppler map corresponding to a synthetic measurement $\boldsymbol{y}$ following the model (2.27) with $K = 3$ targets simulated and a SNR of 15dB. The radar properties are $M_\mathrm{s} = M_\mathrm{c} = 32$. **(b)** Range and Speed Profiles were obtained by taking the norms respectively of rows and columns of the range-Doppler map.

responds to the radar's resolution. By definition of the normalization given in Def. 2.1, it always equals 1. Also, note that the peak appears in a range of 9.7 instead of 10 because of the shift caused by $\sigma_\mathrm{u} u^*$. In Fig. 2.6(b), the choice $M_\mathrm{c} = 256$ enables $\mathcal{U}$ to include higher values for the normalized speed. Thus, since the target depicted in this figure is characterized by $u^* = 112$, we have a non-negligible distortion, and hence, the Dirichlet kernel deteriorates. This observation has an impact on the translation-invariance of the actual kernel. Thereby, many algorithms, including the ones presented in this work, exhibit lower performance when M_c is higher.

Finally, Fig. 2.7 shows an example of a range-Doppler map obtained from a simulated measurement with $K = 3$ targets under a 15dB Signal-to-Noise Ratio (SNR) scenario. This is the kind of range-Doppler map one can expect from actual experimental results.

To estimate the parameters from multiple targets, we will consider iterative strategies that sequentially estimate and clear the contribution of each target in the measurement. Formally, we will apply *off-the-grid* extensions of the OMP algorithm to radars in Chapters 5 and 7. Background on sparse coding theory and heuristics like OMP are provided in the next chapter.

2.5 Multistatic Radar: System Description

A minor contribution of this thesis is the design of an active Multistatic Radar composed of one transmitter and 8 receiving antennas. The radar system is built around the BGT24 family of radio frequency transceivers from Infineon. Fig. 2.8 shows a simplified schematic of the designed hardware, and Fig. 2.9 is a picture of the transceiver board. The *base station* includes a computer, the acquisition board, one transmitter, and a four antennas array. Three *remote stations* are connected to the base station through coaxial cables of length up to 24 m and are equipped with respectively a two-antennas array and twice a single antenna. A general-purpose data acquisition board (NI 6378 from National Instruments) outputs voltage ramps toward the voltage-controlled oscillator (BGT24MTR11) to generate chirps of duration $T_c = 128\mu$s around $f_0 = 24$GHz, with $B = 250$MHz. This signal feeds the transmitting antenna and is also split to feed the 16 mixers required for the demodulation of the signals coming from all the receivers. The receivers are IQ demodulators (BGT24MR2) followed by baseband amplification and filtering circuits. The remote stations have a local low-noise amplifier to compensate for the losses in the cables. The baseband signals are sampled by the acquisition board sending processing frames of $M_c \times M_s$ samples, with $M_s = M_c = 128$, to the computer to perform the detection of targets.

The chosen NI acquisition board can transmit and sample all 16 channels at a rate of 2MHz synchronously by connecting all the ADCs and the DAC to the same internal clock. This is a crucial aspect since any time shift between those components potentially results in additional distortions in the signal when they grow beyond $1/B$. We also emphasize that, according to the remark given in Sec. 2.2, we located all the mixers in the transceiver board. This potentially leads to degradation in the received signals since the V_{tune} signal is unlikely to be perfectly calibrated at each instant because of its strong temperature dependency.

To validate the different algorithms developed in Chapters 6 and 7, we performed an extensive measurement campaign with this multistatic radar system in the framework of the LUMINET project. A picture of the whole system as used during the measurement is given in Fig. 2.10. We considered four distinct scenarios in which one or two cars moved along known paths represented in Fig. 2.11. For each scenario, we recorded 1000 processing frames of 128×128 samples each, for an approximate total of

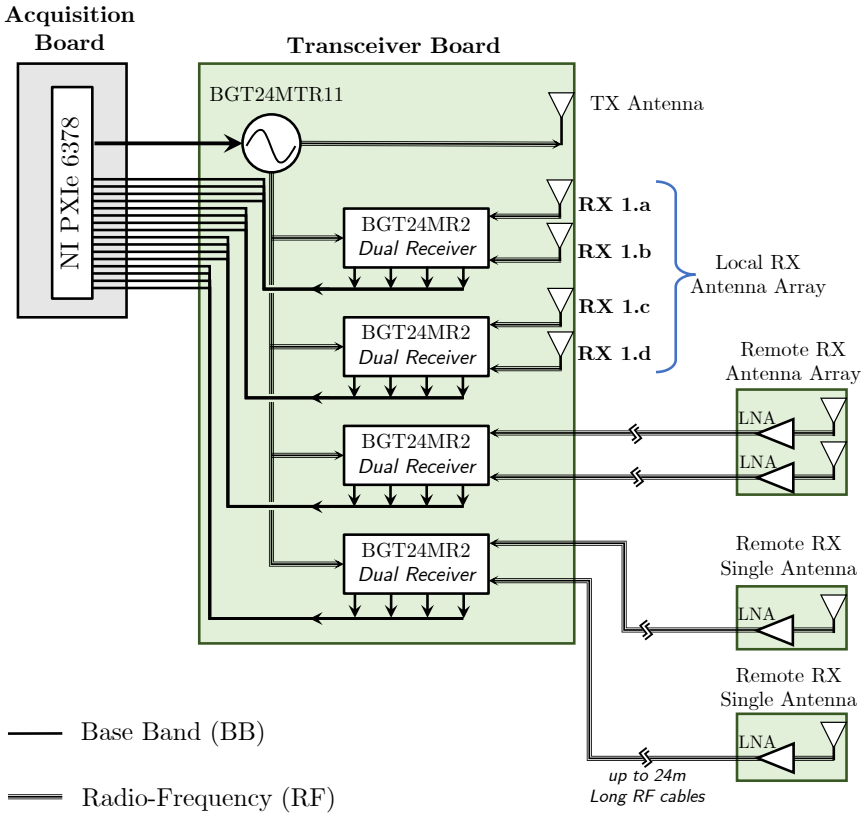

Figure 2.8: Simplified schematic representation of our active multistatic radar design.

60 seconds. We repeated each scenario five times, resulting in a total of 20,000 recorded processing frames for 22GB of raw data. The dataset can be found in open access at `https://gitlab.com/gilles.monnoyer.dg/msr-dataset-22-08-2022`. While results of the application of our algorithms on a selection of this data set will be given in Chapter 7, we already show in Fig. 2.12 an example of range-Doppler maps obtained from the simultaneous observation of two cars by the system. We applied the preprocessing mentioned in Sec. 2.2 to remove the static clutter. More precisely, we remove from each chirp the average of all the chirps, resulting in a cancellation of all the static contributions in the received signal. We see in this figure the main lobes corresponding to each target, confirming the proper working of our system.

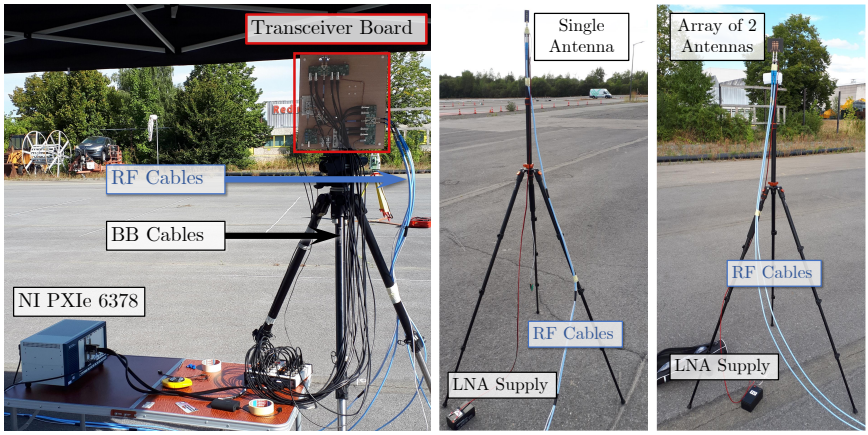

Figure 2.9: Photo of the transceiver board used for the LUMINET project

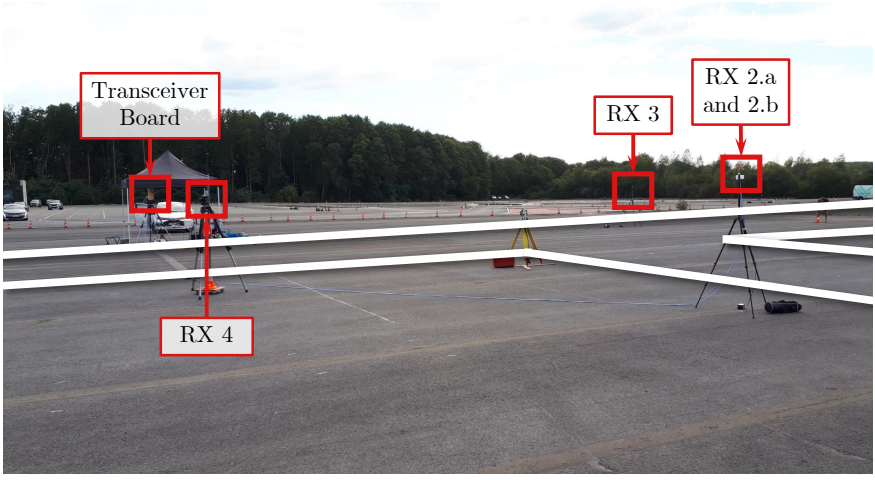

Figure 2.10: Picture of the whole multistatic radar. The white lines represent the delimitation of the road over which the cars moved during the measurement.

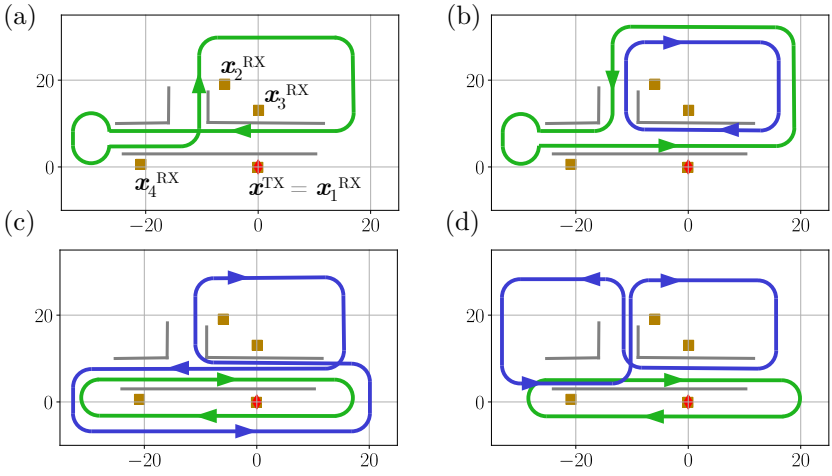

Figure 2.11: Representation of the four paths used in the measurement campaign.

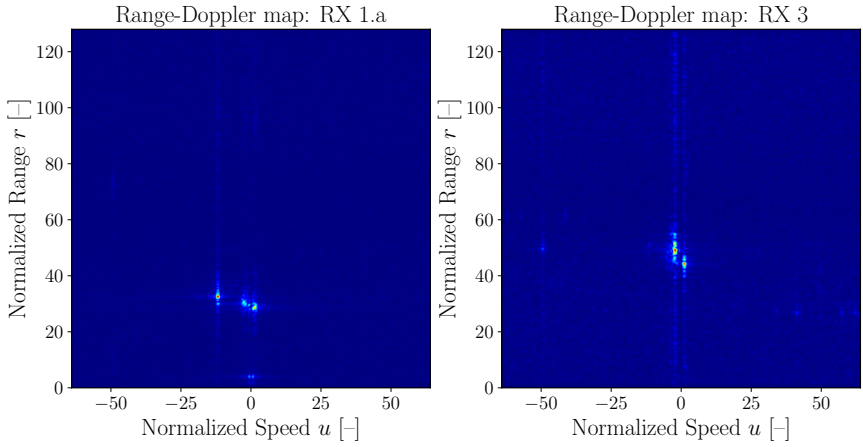

Figure 2.12: Example of the range-Doppler maps obtained from our measurement campaign while simultaneously observing two cars. We restrict the example to one antenna from the monostatic node and one remote antenna.

2.6 Summary

This chapter introduced the basics of FMCW radar technologies regarding both their hardware conception and the standard estimation algorithms. We highlighted a few non-idealities resulting from imperfect hardware. When applying to Multistatic Radar the algorithms we propose in Chapters 5, we will particularly study the impact of these specific limitations on the resulting performance.

While deriving the FMCW radar signal model, an emphasis was put on the factorization of the model and the simplifications this process requires. Since the inner and outer atoms $\boldsymbol{a}^{\text{in}}$ and $\boldsymbol{a}^{\text{out}}$ are particular cases of Fourier atoms, they will be extensively to illustrate concepts in subsequent chapters chapter.

We finally presented a side contribution: the conception of an active FMCW MSR tailored to validate our algorithms. Chapter 7 will show a sample of processed data resulting from our measurement. Still, the reader can find the whole dataset at `https://gitlab.com/gilles.monnoyer.dg/msr-dataset-22-08-2022`.

Chapter 3

Introduction to Sparse Coding

Throughout this dissertation, we describe multiple propositions of algorithms that share a common property: they all amount to extensions of existing procedures defined in the sparse-coding framework. We start with the definition of the general problem of decomposing a signal into a linear combination of a few waveforms taken from a continuous dictionary. Next, we particularize in Sec. 3.2 this problem to a discrete framework. This enables us to present multiple standard formulations and corresponding solvers from the literature. More precisely, the Branch-and-Bound (BnB) approaches are described in Sec. 3.4 and will serve as a starting point for the proposed "peeling" method that we introduce in Chapter 4.

In Sec. 3.3 we describe the Orthogonal Matching Pursuit (OMP), a fast heuristic that intends to approximately solve the problem which is exactly solved by the BnB approaches. We formulate it in both the continuous and the discrete setups. This algorithm will be extended multiple times in Chapters 5 to 7. We end this chapter with an intuitive comparison of the respective ideas behind BnB and OMP, hence motivating the connection between Chapter 4 and the next ones.

3.1 Sparse Representation of a Signal

This chapter is articulated around a generalization of the measurement $\boldsymbol{y}$ provided in the previous chapter for radar applications. The model (2.27) hinted that a radar signal can be decomposed into a linear combination of a few waveforms which can, under a few assumptions, be assimilated to functions called Fourier *atoms*. In this chapter, $\boldsymbol{y} \in \mathbb{C}^M$ denotes a generic measurement which possesses, up to some additive noise $\boldsymbol{w}$, a K-sparse decomposition in a sensing domain Θ. More precisely, we assume that a measurement $\boldsymbol{y}$ such that there exists a collection of K parameters $\theta_1^*, \cdots, \theta_K^* \in \Theta$ from which it decomposes into

$$\boldsymbol{y} = \sum_{k=1}^{K} \alpha_k \, \boldsymbol{a}(\theta_k^*) + \boldsymbol{w}, \tag{3.1}$$

where $\boldsymbol{a}$ denotes the *atom function* mapping any parameter taken from Θ onto a waveform in $\mathbb{C}^M$ which acts as a building block for the measurement $\boldsymbol{y}$. Without loss of generality and for the sake of simplicity, we always define the atom such that $\|\boldsymbol{a}(\theta)\|_2 = 1$ for all $\theta \in \Theta$. The coefficients α_k model unknown complex coefficients randomly taken from a complex normal distribution, in accordance with the conclusions of Sec. 2.3.2. Equivalently, $\boldsymbol{y}$ is said to be decomposable by the continuous parametric dictionary $\mathcal{A}$, defined following Def. 3.1 [Gri01, JV08].

Definition 3.1 (Parametric Dictionary)**.** Given an *atom* function $\boldsymbol{a}(\theta) : \Theta \mapsto \mathbb{C}^M$, the continuous parametric dictionary $\mathcal{A}$ is defined as the set

$$\mathcal{A} := \{\boldsymbol{a}(\theta) : \theta \in \Theta\}. \tag{3.2}$$

The sparse decomposition problem is formulated as the recovery of the ground truth parameters $\theta_1^*, \cdots, \theta_K^*$ from $\boldsymbol{y}$. We restrict our work to dictionaries such that *any* measurement $\boldsymbol{y}$ following the model (3.1) with *no noise* has a unique K-sparse decomposition in $\mathcal{A}$. In [DE03] Donoho and Elad provided a necessary and sufficient condition for the above to hold when working with a *discrete* dictionary gathering a finite set of atoms. Following extensions of their conclusion for continuous dictionaries [YX14], we impose the Assumption 3.2 below on the dictionary.

Assumption 3.2 (Uniqueness Condition)**.** The sparsity K is such that $K \leq \frac{M}{2}$ and the parametric dictionary $\mathcal{A}$ is such that any set of $2K$ distinct atoms[a] taken from it forms a linearly independent set of vectors.

[a] In practice, a minimal separation between the parameters $\{\theta_k^*\}_{k=1}^K$ is required to ensure numerical stability.

Let us highlight that the assumption above implies that the parametric atom function $\boldsymbol{a}$ is injective over Θ and such that for all $\theta' \neq \theta \in \Theta$ there exists no scalar $\alpha \in \mathbb{C}$ such that $\boldsymbol{a}(\theta) = \alpha \boldsymbol{a}(\theta')$.

We also emphasize that unless otherwise stated, we assume throughout the whole document that we only use dictionaries exhibiting a property of *tranlation-invariance* [Den18, CH21]. This concept is defined as below.

Definition 3.3. The dictionary $\mathcal{A}$ is said *translation-invariant* if the inner products of its atoms obey the following invariance property for all $\theta, \theta' \in \Theta$,

$$\langle \boldsymbol{a}(\theta), \boldsymbol{a}(\theta') \rangle = \kappa(\theta - \theta'). \tag{3.3}$$

This property naturally holds for Fourier dictionaries, which we will extensively use for illustration purposes in Chapter 5 on fast *off-the-grid* sparse decomposition algorithms.

Example 3.1 (Fourier Dictionary)*.* In subsequent chapters, we will often illustrate properties of dictionaries and algorithms using a Fourier dictionary with atom length $M = 16$ and $\Theta = [0, M[$. The corresponding atom function is characterized by its m-th component, which reads

$$a_m(\theta) := \exp\left(\mathrm{j}2\pi\theta\tfrac{m}{M}\right). \tag{3.4}$$

Let us consider the measurement example $\boldsymbol{y} = \alpha \boldsymbol{a}(\theta^*)$ with $\alpha = 1 + .5\mathrm{j}$ and $\theta^* = 5.3$. The correlation $|\langle \boldsymbol{a}(\theta), \boldsymbol{y} \rangle|^2 = |\alpha|^2 |\kappa(\theta - \theta^*)|^2$ corresponds to a "*Fejér*" kernel [Hof07], which is translation invariant and shown in Fig. 3.1.

We remind that the inner and outer atom functions from a radar measurement, respectively given in the equations (2.19) and (2.20), are particular cases of Fourier atom function. Therefore, Ex. 3.1 notably corresponds to a simplified FMCW radar signal modulated with a single chirp and 16

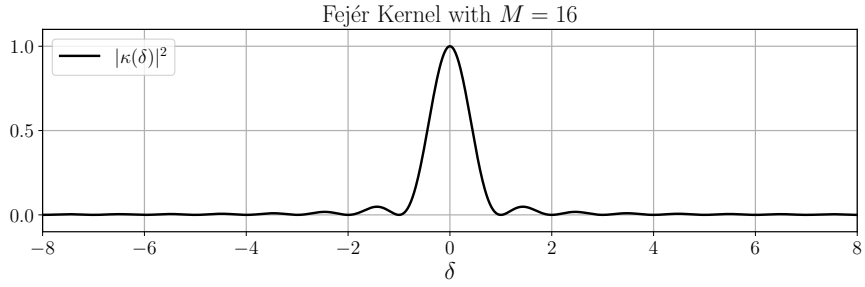

Figure 3.1: Plot of a Féjér kernel corresponding to the function $|\kappa(\cdot)|^2$ associated to the atom function from Ex. 3.1.

samples and resulting from the observation of a static target located in a normalized distance $r = 5.3$.

We now express the decomposition problem associated with the continuous model (3.1). The literature on the *continuous* sparse decomposition of signals presents many different formulations for this problem [JV08, TBSR13, TA18, KTTG17, EGSH21]. For the sake of simplicity, we restrict ourselves to the following,

$$\min_{\boldsymbol{\theta}\in\Theta^K} \min_{\boldsymbol{\alpha}\in\mathbb{C}^K} \left\| \boldsymbol{y} - \sum_{k=1}^{K} \alpha_k \, \boldsymbol{a}(\theta_k) \right\|_2^2, \tag{3.5}$$

where $\boldsymbol{\alpha} := (\alpha_1, \dots \alpha_K)^\top$ and $\boldsymbol{\theta} := (\theta_1, \dots, \theta_K)^\top$. The above is mathematically similar to the problem addressed by P-J. Bénard et al. in [TA18, TAL19] without the minimal separation constraint between the parameters, and is most likely NP-hard. Addressing this problem, however, intrinsically assumes the prior knowledge of the sparsity level K. When this information is neither known nor arbitrarily defined by the application, one can use regularized formulations that we detail for discrete cases in the next section.

3.2 Discretized Problem

Although estimation problems —such as those derived for radar applications in the previous chapter— are more naturally connected to the field of sparse coding by means of a continuous dictionary, the field has initially been more widely studied with discrete formulations. In fact, the aforementioned estimation problems are also commonly reduced in practice to a discretized version. Moreover, most solvers, either exact or heuristics, for

the problem (3.5) are derived by extending existing techniques designed for its discretized counterparts.

Definition 3.4 (Discretized Parametric Dictionary)**.** Given the dictionary $\mathcal{A}$ parameterized over the domain Θ, with $\Theta_{\mathrm{G}} := \{\bar{\theta}_n\}_{n=1}^N$ discretizing Θ, discretized parametric dictionary is defined by the matrix $\boldsymbol{A} \in \mathbb{C}^{M\times N}$ such that

$$\boldsymbol{A} := \big[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_N\big], \tag{3.6}$$

where $\boldsymbol{a}_n := \boldsymbol{a}(\bar{\theta}_n)$.

Using the discrete dictionary $\boldsymbol{A}$, we reformulate the decomposition problem (3.5) with the restriction to approach the measurement $\boldsymbol{y}$ by a linear combination of the atoms gathered in the columns of $\boldsymbol{A}$ only. With $\boldsymbol{s}$ denoting the sparse vector to construct in the domain $\mathbb{C}^N$, this problem reads

$$\min_{\boldsymbol{s}\in\mathbb{C}^N} \ \tfrac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{s}\|_2^2. \quad \text{such that} \quad \|\boldsymbol{s}\|_0 \leq K, \tag{3.7}$$

where the operator $\|\cdot\|_0$ denotes the ℓ_0-pseudonorm which counts the number of non-zero elements in its argument. In brief, solving (3.7) amounts to finding a sparse vector with no more than K non-zero components whose locations correspond to the parameters $\hat{\theta}_1, \cdots, \hat{\theta}_K$ (via the gridding Θ_{G}). We emphasize that these problems are similarly used in contexts involving non-parametric dictionaries, *i.e.,* potentially any matrix $\boldsymbol{A}$ of size $M \times N$, typically (but not necessarily) with $N \geq M$.

In the literature, a variation of this decomposition problem is often studied. This alternative imposes the sparsity of its solution using a regularization instead of the constraint used in (3.7). With $\lambda > 0$ denoting a regularization parameter, it reads

$$p^\star = \min_{\boldsymbol{s}\in\mathbb{C}^N} \ \mathrm{P}(\boldsymbol{s}) := \tfrac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{s}\|_2^2 + \lambda\|\boldsymbol{s}\|_0. \tag{3.8}$$

Note that problems (3.7) and (3.8) have been shown to be *partially* equivalent [Nik16]. We summarized this partial equivalence with the schematic example in Fig. 3.2. Considering first a "very" high value for λ (in the far right of the figure), the optimizer of (3.8) is the zero vector. Progressively decreasing λ will modify this optimizer per stage. More precisely, at some point, λ is sufficiently small to afford one non-zero element in the optimal

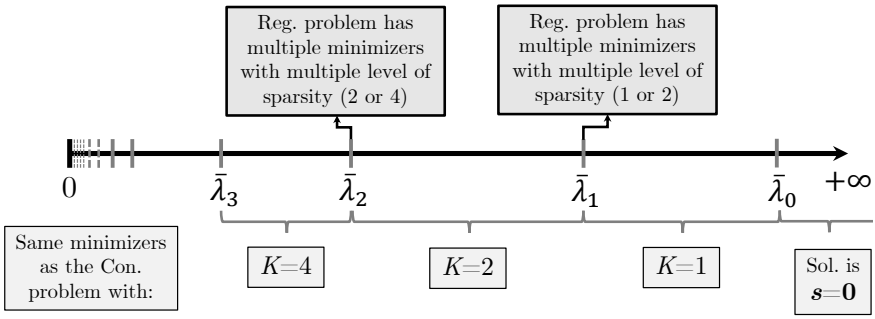

Figure 3.2: Schematic example of the connection between the regularized problem (3.8) and the constrained problem (3.8). The values $\bar{\lambda}_i$ represent critical values for the regularizer for which the equivalence with a single constrained formulation does not hold.

value of $\boldsymbol{s}$. Then, there is a smaller value of λ such that the solution of (3.8) can afford being 2-sparse, and so on. However, some values of sparsity level K may have no corresponding stage in the partition of the possible values for λ. Moreover, the critical values of the regularizer between subsequent stages yield regularized problems containing multiple optimal sets of the constrained problems. Nonetheless, this partial equivalence is generally sufficient to translate most conclusions drawn for one formulation into equivalent or adapted conclusions for the other.

Solving (3.8) is of paramount interest in many scientific fields such as statistics, machine learning, or inverse problems [CT05, DET05, TW10]. Unfortunately, this problem also turns out to be NP-hard [Nat95, CGWY12]. This means that an exhaustive search among all the possible subsets of the column of $\boldsymbol{A}$ is required to find the solution of (3.8).

The challenging complexity of these problems has led researchers of the last decades to propose a wide range of computationally tractable procedures able to recover approximate solutions of either (3.7) or (3.8). Among these fast alternatives, convex relaxations and greedy algorithms have become very popular, see [FR13, Ch. 3]. Although these procedures often successfully recover the actual solutions of (3.7), they are usually unsuited to tackle challenging instances of the problem. This observation, combined with recent advances in integer optimization and hardware performance, has revived the interest in methods solving (3.8). A standard approach is to use a BnB algorithm that solves (3.8), see [Cpl09, BNCM15, BKM16, BCWP21, HMS22, GHE22]. The approach provides a potentially faster

solving procedure of (3.8) than the exhaustive search. Still, the NP-Hard nature of this problem remains, and hence, BnB amounts in a theoretical —yet very unlikely— worst-case scenario to the exhaustive search.

In the next section, we will focus on the standard greedy algorithm known as Orthogonal Matching Pursuit (OMP), which is a fast heuristic to approximately solve (3.7) or (3.5), respectively in discrete and continuous setups. Next, Sec. 3.4 is dedicated to introducing BnB procedures. In an attempt to leverage ingredients from BnB methodology to improve the performance of OMP, we will present in Chapter 4 a novel approach, coined "Peeling", which extends beyond the standard formulation of BnB. This peeling strategy potentially offers the possibility of future work on enhancing OMP by injecting, in its procedure, some tests inspired by the BnB philosophy. These tests may enable future work to guide OMP by discarding or validating atoms known to match the decomposition provided by the exact solution of the ℓ_0 problems given in this section.

3.3 Orthogonal Matching Pursuit

Greedy algorithms offer a fast alternative to solving the NP-hard problem presented in the previous section. The most famous of these algorithms is the Orthogonal Matching Pursuit (OMP). We formulated it in its standard discrete formulation in Alg. 1, with the operator $\leftarrow$ denoting the affectation of a new value in a variable. By design, OMP's philosophy can be interpreted as an attempt to solve the constrained formulation (3.7) approximately. Indeed, from the knowledge of a targeted sparsity K for the resulting vector $\boldsymbol{s}$, OMP greedily minimizes (3.7) by iteratively solving

$$\min_{\boldsymbol{s}\in\mathbb{C}^N,\ \|\boldsymbol{s}\|_0=1} \tfrac{1}{2}\|\boldsymbol{r}-\boldsymbol{A}\boldsymbol{s}\|_2^2, \tag{3.9}$$

and subsequently updating the residue $\boldsymbol{r}$. With different words, each iteration looks for the atom from $\boldsymbol{A}$ that best fits the residue and next, removes from $\boldsymbol{y}$ an orthogonal projection of this signal onto the space generated by the set of selected atoms at the current iteration.

Remark 3.2. OMP can be adapted to match the philosophy of the regularized problem. To this end, one can replace the pre-determined number of iterations with a stopping criterion. Indeed, given the iteration k, when

Algorithm 1: Orthogonal Matching Pursuit (OMP)

Input : $\boldsymbol{y}$, $\boldsymbol{A}$, K

Output: $\boldsymbol{s}$

begin

Initialization: $\boldsymbol{r} \longleftarrow \boldsymbol{y}$;

for $k = 1$ **to** K **do**

1. Selection Step

$$\hat{n}_k = \arg\max_{n \in [M]} |\langle \boldsymbol{a}_n, \boldsymbol{r} \rangle| \tag{3.10}$$

2. Fitting Step

$$\hat{\boldsymbol{\alpha}} \longleftarrow \arg\min_{\boldsymbol{\alpha} \in \mathbb{C}^k} \frac{1}{2} \left\| \boldsymbol{y} - \left[\boldsymbol{a}_{\hat{n}_1}, \cdots, \boldsymbol{a}_{\hat{n}_k} \right] \boldsymbol{\alpha} \right\|_2^2 \tag{3.11}$$

3. Residue Update

$$\boldsymbol{r} \longleftarrow \boldsymbol{y} - \left[\boldsymbol{a}_{\hat{n}_1}, \cdots, \boldsymbol{a}_{\hat{n}_k} \right] \hat{\boldsymbol{\alpha}} \tag{3.12}$$

$\boldsymbol{s} = \left[\boldsymbol{a}_{\hat{n}_1}, \cdots, \boldsymbol{a}_{\hat{n}_K} \right] \hat{\boldsymbol{\alpha}}$

$\frac{1}{2}\|\boldsymbol{r}\|_2^2 < \lambda k$, one stops the procedure since any additional atom selection would increase of cost function of (3.8).

When applied to the parameter estimation context with a parametric dictionary, OMP is by definition restricted to the grid Θ_{G}. Many propositions of greedy heuristics extending the OMP's principle have been studied in the literature, each with its own characteristics [KYHP15, EGSH21, KTTG17]. In this chapter, we maintain the focus on the most standard approach to defining Continuous OMP (COMP), described in Alg. 2. This algorithm follows the exact same principle as the standard discrete OMP by searching for atoms in $\mathcal{A}$ instead of $\boldsymbol{A}$. The collection of N scalar products in the selection step of OMP is thus replaced by a usually non-convex optimization problem defined in the space Θ (see equation (3.14)). In fact, this COMP formulation iteratively solves

$$\arg\min_{\theta \in \Theta} \min_{\alpha \in \mathbb{C}} \|\boldsymbol{r} - \alpha \boldsymbol{a}(\theta)\|_2^2, \tag{3.13}$$

Algorithm 2: "Standard" Continuous OMP

Input : $\boldsymbol{y}$, $\mathcal{A}$, K

Output: $\hat{\theta}_1, \cdots, \hat{\theta}_K$

begin

Initialization: $\boldsymbol{r} \longleftarrow \boldsymbol{y}$;

for $k = 1$ **to** K **do**

1. Selection Step

$$\hat{\theta}_k = \arg\max_{\theta \in \Theta} |\langle \boldsymbol{a}(\theta), \boldsymbol{r} \rangle| \tag{3.14}$$

2. Fitting Step

$$\hat{\boldsymbol{\alpha}} \longleftarrow \arg\min_{\boldsymbol{\alpha} \in \mathbb{C}^k} \frac{1}{2} \left\| \boldsymbol{y} - \left[\boldsymbol{a}(\hat{\theta}_1), \cdots, \boldsymbol{a}(\hat{\theta}_k) \right] \boldsymbol{\alpha} \right\|_2^2 \tag{3.15}$$

3. Residue Update

$$\boldsymbol{r} \longleftarrow \boldsymbol{y} - \left[\boldsymbol{a}(\hat{\theta}_1), \cdots, \boldsymbol{a}(\hat{\theta}_k) \right] \hat{\boldsymbol{\alpha}} \tag{3.16}$$

which is reduced to (3.14) since we only consider atoms of unit power. This problem can, for instance, be solved by means of a gradient ascent initialized by a grid search [TA18, TAL19], or approximately with an interpolation strategy [KYHP15, CH19].

Alternate and more advanced approaches suggest reinforcing the fitting step (3.15) with a global gradient descent, yielding more robust results with a potentially very high computational requirement (see *e.g.,* [KTTG17]). Formally, Alg. 2 can be enhanced (or "refined") by adding, after the step (3.15), a global gradient descent aiming to solve

$$\hat{\alpha}, \hat{\boldsymbol{\theta}} \longleftarrow \underset{\boldsymbol{\alpha} \in \mathbb{C}^k,\, \boldsymbol{\theta} \in \Theta^k}{\arg\min} \left\| \boldsymbol{y} - \left[\boldsymbol{a}(\theta_1), \cdots, \boldsymbol{a}(\theta_k) \right] \boldsymbol{\alpha} \right\|_2^2 \tag{3.17}$$

and typically initialized with the current values of $\hat{\alpha}$ and $\hat{\boldsymbol{\theta}} := (\hat{\theta}_1, \cdots, \hat{\theta}_k)$. This additional step enables the greedy method to iteratively adjust decisions taken on $\hat{\boldsymbol{\theta}}$ during the previous selection steps. Formulated differently, this extension amounts to solving the initial problem (3.5) with a gradient descent whose initialization is iteratively constructed by the greedy pro-

cedure. In our work, we use the name "Refined-COMP" to refer to the algorithm extending Alg. 2 by including the iterative global gradient descent (3.17). This enables us to distinguish it from the "Standard-COMP" that exactly corresponds to Alg. 2.

3.4 Principles of BnB Methods for ℓ_0 Sparse Decomposition

Most of the content of this section is taken from our publication presented at EUSIPCO 2023. "T. Guyard, G. Monnoyer, C. Elvira, C. Herzet, *Safe peeling for l0-regularized least-squares*, in Proceedings of European Signal Processing Conference 2023 (EUSIPCO 2023)." [GMEH23]

In this section, we introduce the main principles of BnB procedures that we use to solve the ℓ_0 regularized problem in a discretized context, *i.e.,* the problem (3.8). We only review the elements of interest to introduce the proposed peeling method. We refer the reader to [Wol20, Ch. 7] for an in-depth treatment of the subject. Also, we formulate it in the context of a dictionary $\boldsymbol{A}$ and sparse signals $\boldsymbol{s}$ defined over $\mathbb{R}$ instead of $\mathbb{C}$. This is done without loss of generality since any complex problem can be reformulated in a real context by splitting the real and the imaginary parts, hence leading to a problem in which all dimensions are doubled. The targetted problem thus amounts to finding one or multiple vectors $\boldsymbol{s} \in \mathbb{R}^N$ minimizing (3.8).

We remind the reader that the problem (3.8) is NP-hard. Finding its minimizers requires, in the worst case, to perform the whole exhaustive search over the 2^N possible supports for the solution (each index of $\boldsymbol{s}^*$ can either be zero or non-zero, leading to 2^N possibilities). The BnB procedures offer a strategy to explore the set of possible supports efficiently.

In this section, it is convenient to use the indicator operator η to write the constraints on the optimization problems. This operator returns 0 if the condition in its input is met, and $+\infty$ otherwise.

3.4.1 Pruning

The principle of BnB methods is illustrated in Fig. 3.3. Its crux consists in identifying and discarding some subsets of $\mathbb{R}^N$ which do not contain a minimizer of (3.8). To do so, one constructs a decision tree in which each

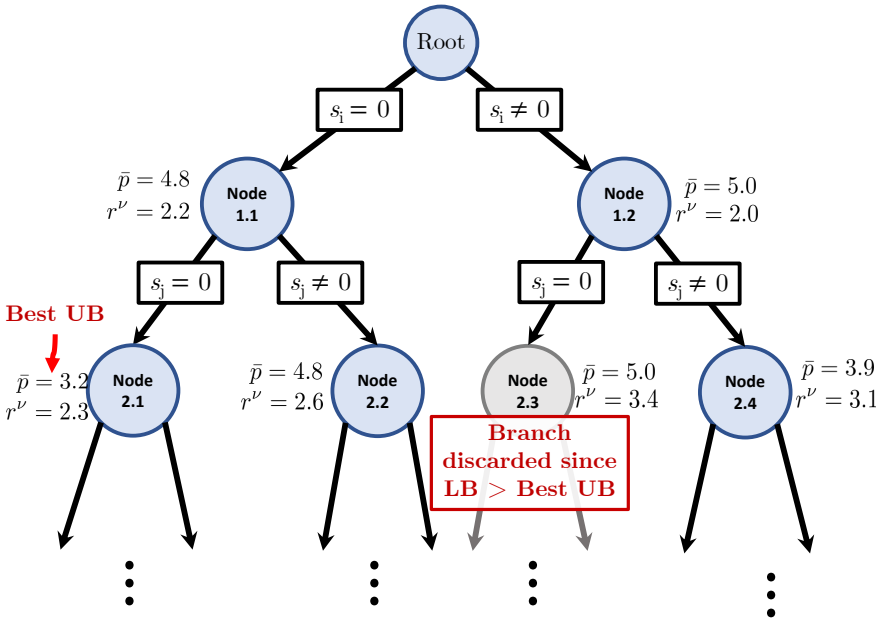

Figure 3.3: Schematic representation of an example of a tree constructed by a BnB procedure. At each node, an upper bound on the root problem is computed with a fast heuristic (*e.g.*, OMP), and a lower bound on the problem restricted to this node is computed from a convex relaxation. The lower bound is given at each node ν by the value of r^ν instead of p^ν for reasons detailed in Sec. 3.4.2.

node corresponds to a particular subset of $\mathbb{R}^N$. In our context, a tree node is identified by two disjoint subsets of $[N]$, say ν_0 and ν_1. The goal at node $\nu := (\nu_0, \nu_1)$ is to detect whether a solution of (3.8) can be attained within

$$\mathcal{S}^\nu := \big\{ \boldsymbol{s} \in \mathbb{R}^N \colon \boldsymbol{s}_{\nu_0} = \boldsymbol{0},\ \boldsymbol{s}_{\nu_1} \neq \boldsymbol{0} \big\}, \tag{3.18}$$

where $\boldsymbol{s}_{\nu_k}$ denotes the restriction of $\boldsymbol{s}$ to its elements in ν_k. For example, the node 2.2 in Fig. 3.3 is characterized by $\nu_0 = \{i\}$ and $\nu_1 := \{j\}$, for some $i \neq j \in [N]$ arbitrarily chosen to construct the tree.

In particular, let $\mathcal{S}^\star$ be the non-empty set of minimizers of (3.8). Given some known upper bound $\bar{p}$ on the optimal value $p^\star$, if we let

$$p^\nu := \min_{\boldsymbol{s} \in \mathbb{R}^N} \mathrm{P}^\nu(\boldsymbol{s}) \tag{3.19}$$

with $\mathrm{P}^\nu(\boldsymbol{s}) := \mathrm{P}(\boldsymbol{s}) + \eta(\boldsymbol{s} \in \mathcal{S}^\nu)$, we obtain the implication

$$p^\nu > \bar{p} \implies \mathcal{S}^\nu \cap \mathcal{S}^\star = \emptyset. \tag{3.20}$$

Indeed, since $\bar{p} \geq p^\star$, the above implies $p^\nu > p^\star$ and hence, $\mathcal{S}^\nu$ cannot contain the minimizer of the root problem (3.8). In other words, if the minimal objective of the decomposition problem restricted to ν is higher than a known upper bound $\bar{p}$ on the initial problem, then the subset $\mathcal{S}^\nu$ does not contain any solution of (3.8). It can therefore be discarded from the search space of the optimization problem. This operation is usually referred to as "*pruning*".

3.4.2 Bounding and Relaxing

Making a pruning decision at node ν requires the knowledge of $\bar{p}$ and p^ν. On the one hand, finding $\bar{p}$ is an easy task since the value of the objective function in (3.8) at any feasible point constitutes an upper bound on $p^\star$. Typically, we evaluate a new feasible point each time we reach a node, providing a candidate for our upper bound. We keep track along the tree of the best upper bound that has been met so far. On the other hand, evaluating p^ν is NP-hard. This issue can nevertheless be circumvented by finding a tractable lower bound r^ν on p^ν and relaxing (3.20) as

$$r^\nu > \bar{p} \implies \mathcal{S}^\nu \cap \mathcal{S}^\star = \emptyset. \tag{3.21}$$

One ubiquitous approach in the literature [BMBN22, BCWP21, BNCM15] to find such a lower bound consists in:

i) Adding an extra term "$\eta(\boldsymbol{s} \in [\boldsymbol{l}, \boldsymbol{u}])$" to the cost function of (3.8), for some *well-chosen* bounds $\boldsymbol{l} \in \mathbb{R}^N_-$ and $\boldsymbol{u} \in \mathbb{R}^N_+$.[1] In particular, the new constraint "$\boldsymbol{s} \in [\boldsymbol{l}, \boldsymbol{u}]$" must lead to a problem fully equivalent to (3.8), that is

$$\forall \boldsymbol{s}^\star \in \mathcal{S}^\star : \ \boldsymbol{s}^\star \in [\boldsymbol{l}, \boldsymbol{u}]. \tag{3.22}$$

ii) Exploiting the convex relaxation of the function $\|\cdot\|_0$ on the bounded

[1] This additional constraint usually takes the form "$-M \leq s_i \leq M, \ \forall i$" with $M > 0$ and is known as *"Big-M"* constraint, see [BNCM15, Sec. 3]

set $\mathcal{S}^\nu \cap [\boldsymbol{l}, \boldsymbol{u}]$, given by

$$\|\boldsymbol{s}\|_0 \geq |\nu_1| + \sum_{i \in \nu_\bullet} \frac{[s_i]_+}{u_i} - \frac{[-s_i]_+}{l_i}, \tag{3.23}$$

where $[s]_+ = \max(s, 0)$ refers to the positive-part function and $|\cdot|$ denotes the cardinality of a set. in the above, $\nu_\bullet := [N] \setminus (\nu_0 \cup \nu_1)$ and we use the convention "$0/0 = 0$".

On the one hand, item *i)* implies that the pruning test (3.20) involves the following quantity (rather than p^ν):

$$p^\nu(\boldsymbol{l}, \boldsymbol{u}) = \inf_{\boldsymbol{s} \in \mathbb{R}^N} \mathrm{P}^\nu(\boldsymbol{s}; \boldsymbol{l}, \boldsymbol{u}) \tag{3.25}$$

where $\mathrm{P}^\nu(\boldsymbol{s}; \boldsymbol{l}, \boldsymbol{u}) := \mathrm{P}^\nu(\boldsymbol{s}) + \eta(\boldsymbol{s} \in [\boldsymbol{l}, \boldsymbol{u}])$ is the objective function of the decomposition problem restricted to the node ν endowed with the extra term "$\eta(\boldsymbol{s} \in [\boldsymbol{l}, \boldsymbol{u}])$". On the other hand, a lower bound $r^\nu(\boldsymbol{l}, \boldsymbol{u})$ on $p^\nu(\boldsymbol{l}, \boldsymbol{u})$ can be obtained by using (3.23) and solving

$$r^\nu(\boldsymbol{l}, \boldsymbol{u}) = \min_{\boldsymbol{s} \in \mathbb{R}^N} \mathrm{R}^\nu(\boldsymbol{s}; \boldsymbol{l}, \boldsymbol{u}) \tag{3.27}$$

where

$$\mathrm{R}^\nu(\boldsymbol{s}; \boldsymbol{l}, \boldsymbol{u}) := \tfrac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{s}\|_2^2 + \lambda \sum_{i \in \nu_\bullet} \tfrac{[s_i]_+}{u_i} - \tfrac{[-s_i]_+}{l_i} + \lambda|\nu_1| + \eta(\boldsymbol{s}_{\nu_0} = \boldsymbol{0}) + \eta(\boldsymbol{s} \in [\boldsymbol{l}, \boldsymbol{u}]).$$

We note that (3.27) is a convex problem and can be solved efficiently to good accuracy via numerous polynomial-time numerical procedures, see *e.g.,* [Bec17, Ch. 10].

In practice, the choice of $\boldsymbol{l}$ and $\boldsymbol{u}$ must respect two conflicting imperatives. First, the new constraint "$\boldsymbol{s} \in [\boldsymbol{l}, \boldsymbol{u}]$" should not modify the solution of our target problem (3.8) and condition (3.22) must therefore be verified. Since $\mathcal{S}^\star$ is obviously not accessible beforehand, this suggests that the entries of $\boldsymbol{l}$ and $\boldsymbol{u}$ should be chosen "large enough" in absolute values. Some heuristics are commonly used in the literature to select proper values of the bounds, see [BNCM15, Sec. V.B], [GHE22, Sec. 5.1] or [MBM+20, Sec. 4]. Second, the tightness of $r^\nu(\boldsymbol{l}, \boldsymbol{u})$ with respect to $p^\nu(\boldsymbol{l}, \boldsymbol{u})$ degrades with the spread of the set $[\boldsymbol{l}, \boldsymbol{u}]$. This impairment pertains to a large class of mixed-integer problems and is well known in the literature, see *e.g.*, [CRT90]. In

particular, the right-hand side of (3.23) tends to $|\nu_1|$ when $\boldsymbol{l} \ll \boldsymbol{s}$ and $\boldsymbol{s} \ll \boldsymbol{u}$. Therefore, setting the entries of $\boldsymbol{l}$ and $\boldsymbol{u}$ with too large absolute values is likely to degrade the effectiveness of the relaxed pruning decision (3.21).

In Chapter 4, we present a collaborative work performed with Théo Guyard under the supervision of Dr. Cédric Herzet in which we propose a solution to address this problem by deriving a methodology that locally tightens the constraint $\boldsymbol{s} \in [\boldsymbol{l}, \boldsymbol{u}]$ at each node of the decision tree while preserving the correctness of the BnB procedure.

3.4.3 Enhancing BnB with Screening

In more recent contributions [AG20, GHE21], the authors suggested an acceleration of BnB procedures. This acceleration adapts to the BnB context a technique called "screening" that was initially defined to accelerate solvers tailored for convex relaxations (*e.g.,* the LASSO problem [Tib96]) of the sparse decomposition problem [BERG15, XWR17, NFGS16]. Understanding this method motivates the connection between BnB and greedy algorithms and thus motivates the research on the peeling methodology that we describe in Chapter 4.

In a nutshell, the screening technique in the BnB context enables the procedure to anticipate, with almost no additional computation time, lower bounds on nodes that have not yet been explored in the tree.

To understand the principle of this technique, let us assume what happens in BnB when reaching a node ν. Arriving at node ν, the procedure potentially updates the upper bound kept in track and solves the relaxed problem (3.27) to obtain a lower bound for this node, denoted r^ν. Next, if the procedure has not pruned the node, it must select a new index, say i, that has not yet been affected to either ν_0 or ν_1 to keep descending along the tree. Let us denote by ν_i one of the two resulting child nodes[2]. Reaching ν_i, a conventional BnB method will repeat the same operations as for node ν. With the screening, instead of directly solving the relaxation on the node ν_i to obtain r^{ν_i}, we first quickly evaluate a lower bound on r^{ν_i}. With d_i denoting this lower bound, we directly conclude by construction that

$$d_i > \bar{p} \implies r_{\nu_i} > \bar{p}. \tag{3.28}$$

If the above is observed, the knowledge of d_i, instead of r_{ν_i}, is sufficient to prune the node ν_i.

[2] it does not matter which of the two nodes it refers to

Exploiting the weak duality theorem [BV04, Chapter 5], the screening technique succeeds in providing multiple values for lower bound d_i that can be computed from the knowledge of r^ν with no additional computation time. In short, this theorem states that there exists a dual function $\mathrm{D}^{\nu_i}(\boldsymbol{w})$ such that for all $\boldsymbol{w} \in \mathbb{R}^M$

$$\mathrm{D}^{\nu_i}(\boldsymbol{w}) \leq r^{\nu_i}. \tag{3.29}$$

In [AG20, GHE21], the authors showed that the conventional solvers used to compute r^ν typically evaluate ingredients required to evaluate $\mathrm{D}^{\nu_i}(\boldsymbol{w})$ for all i and for as many values of $\boldsymbol{w}$ as the solver went through during its iterations. Therefore, during the process of solving the relaxation at node ν, we simultaneously evaluate values for $\mathrm{D}^{\nu_i}(\boldsymbol{w})$ and for each index i, we keep track of the best (higher) evaluation encountered. All in all, given the distinction ν_i^0 and ν_i^1 between the node in which i has, respectively, been added to ν_0 or to ν_1, we can directly remove from the problem restricted to ν *all* the atom indices that meet the condition

$$\mathrm{D}^{\nu_i^0}(\boldsymbol{w}) > \bar{p}. \tag{3.30}$$

Similarly, we can directly validate that any atom satisfying

$$\mathrm{D}^{\nu_i^0}(\boldsymbol{w}) > \bar{p} \tag{3.31}$$

is selected in the solution of the ν-restricted problem (3.19).

3.5 Hints of Screening in OMP

The Orthogonal Matching Pursuit algorithm described in Alg. 1 can be interpreted as a simple exploration of the BnB tree of nodes. That is, each selection step corresponds to a branching decision, choosing which index to add into the set ν_1, which is initialized to the empty set at the beginning of the algorithm. Therefore, at a given iteration of OMP, one can decide to solve (3.27) either exactly or approximately to obtain a value for a dual variable $\boldsymbol{w}$ from which the tests (3.30) and (3.31) can be performed. If the test (3.30) passes for some indices i, then the corresponding atoms can be removed from the dictionary $\boldsymbol{A}$, hence dynamically reducing the dimension of the problem to solve, which amounts to reducing the computation time of the next iterations of OMP. If on the other hand the test (3.31) passes for an index i, then this index can directly be selected in the solution set

of selected indices. This index is not only part of the solution of the greedy algorithm, but also guaranteed to belong to the solution of the ℓ_0 NP-hard problem restricted to the node defined by the current iteration of OMP.

This strategy potentially enables the acceleration of OMP and provides guarantees that some indices of the resulting solution match the solution of the initial NP-hard problem that OMP intends to approximately solve. This was the initial goal of my collaboration with the team of Dr. Cédric Herzet at the INRIA research center in Rennes. In practice, however, we have not yet succeeded in successfully implementing this technique in a way that enables us to effectively pass the abovementioned tests in the OMP context. We suspected that this is caused by bad conditioning of the bounds $\boldsymbol{l}$ and $\boldsymbol{u}$ and thus, motivated us to concentrate on the problem of improving the evaluation of the best possible value for these bounds. This leads us to the peeling strategy that we present in Chapter 4.

3.6 Summary

In this chapter, we introduced multiple formulations for finding sparse signal representations. We focused on the parametric dictionaries we exemplified with a decomposition in the Fourier domain.

Distinguishing discrete and continuous situations, we provided the standard formulation of a famous greedy algorithm: the Orthogonal Matching Pursuit. In the continuous configuration, the name Continuous OMP have been used in the literature to denote multiple possible extension of the OMP. We highlighted two of them that we named the “Standard-COMP” and the ”Refined-COMP”. The latest is expected to provide stronger results than the first, with an increased computational complexity. In Chapter 5, we will explore other alternatives extending these two algorithms using interpolation.

This chapter also described the basic concepts of the BnB algorithms. These concepts act as the starting point for the next chapter.

Part II

Accelerated Signal Decomposition

Chapter 4

Accelerating Sparse Decomposition with Peeling

The results of this chapter have been accepted to appear in the proceedings of the IEEE conference EUSIPCO 2023. This is a work performed in collaboration with Théo Guyard (INSA Rennes), Dr. Clément Elvira (CentraleSupelec Rennes), and Dr. Cédric Herzet (INRIA Rennes). The numerical results presented in this chapter have been produced by Théo Guyard and are available in `https://github.com/TheoGuyard/BnbPeeling.jl`. The mathematical development and theory of the peeling technique, as well as the interpretation of the results, is a collaborative work of all the authors.

Most of the content of this chapter is taken from our publication presented at EUSIPCO 2023. "T. Guyard, G. Monnoyer, C. Elvira, C. Herzet, *Safe peeling for l0-regularized least-squares*, in Proceedings of European Signal Processing Conference 2023 (EUSIPCO 2023)." [GMEH23]

Accelerating the solving process of the ℓ_0 sparse decomposition problem (3.8) is by itself of high interest in many scientific fields that we listed in the previous chapter. We emphasized in that chapter that Branch-and-Bound (BnB) procedures require the definition of some bounding constraint on the non-zeros values of the reconstructed

sparse vector. The arbitrary nature of adding this constraint comes with two issues we detailed in the last chapter. First, since the optimal vector is unknown before solving the problem, these bounds are commonly set with high absolute values to avoid altering the solution. Higher values, however, reduce the ability of BnB to prune node and hence to provide acceleration with respect to the brutal exhaustive search. Second, we concluded before that a loosely conditioned relaxation (*i.e.,* with a high absolute value of these bounds) may prevent us from the possibility of adapting the node pruning principle to greedy algorithms such as OMP. Since this latest extension of OMP could lead in the future to significant improvement in the performance of greedy algorithms, which have the focus in Chapter 5 to 7, it is of paramount interest to find a way to improve the conditioning of those bounding constants.

This chapter thus proposes a new strategy, dubbed *"safe peeling"*, to enable more aggressive pruning in the BnB tree by refining those bounding values. In a nutshell, our contribution is a computationally simple test applied at each node of the BnB decision tree to identify some intervals of $\mathbb{R}^N$ which cannot contain a solution of (3.8). Our numerical experiments show that the proposed method significantly reduces the solving time compared to state-of-the-art concurrent methods. The name *"safe peeling"* comes from the fact that the proposed method enables to reduce (or in more figurative terms, "to peel") of the feasible set of the problem at each node of the decision tree while safely preserving the correctness of the BnB procedure.

Proofs of the results presented in this chapter are postponed to Sec. 4.4.

4.1 The Peeling Principle

In this section, we introduce our proposed peeling procedure. As an initial assumption, we suppose that some interval $[\boldsymbol{l}, \boldsymbol{u}]$ verifying condition (3.22) is known. This assumption will be relaxed later on in Sec. 4.1.3.

Our goal is to find a new interval $[\boldsymbol{l}', \boldsymbol{u}']$ such that

$$\forall \boldsymbol{s} \in [\boldsymbol{l}, \boldsymbol{u}] \setminus [\boldsymbol{l}', \boldsymbol{u}'] : \ \mathrm{P}^{\nu}(\boldsymbol{s}; \boldsymbol{l}, \boldsymbol{u}) > \bar{p} \tag{4.1a}$$

$$[\boldsymbol{l}', \boldsymbol{u}'] \subseteq [\boldsymbol{l}, \boldsymbol{u}]. \tag{4.1b}$$

These requirements imply that the pruning decision (3.20) made at node ν remains unchanged when replacing constraint "$\boldsymbol{s} \in [\boldsymbol{l}, \boldsymbol{u}]$" by "$\boldsymbol{s} \in [\boldsymbol{l}', \boldsymbol{u}']$" in (3.25). More specifically, the result from the following lemma holds.

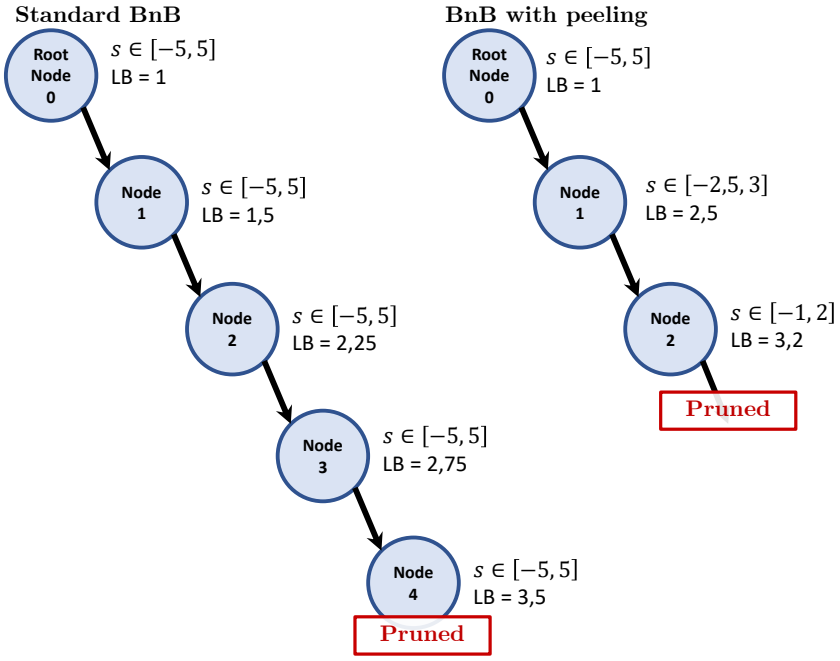

Figure 4.1: Schematic representation of an example of tree exploration (left) with no peeling and (right) with our peeling methodology.

Lemma 4.1. *Assume* $[\boldsymbol{l}, \boldsymbol{u}]$ *and* $[\boldsymbol{l}', \boldsymbol{u}']$ *verify* (4.1a)-(4.1b), *then*

$$p^{\nu}(\boldsymbol{l}', \boldsymbol{u}') > \bar{p} \iff p^{\nu}(\boldsymbol{l}, \boldsymbol{u}) > \bar{p}. \tag{4.2}$$

A proof of this result is available in Sec. 4.4. A consequence of preserving the pruning decision is that taking the new constraint "$\boldsymbol{s} \in [\boldsymbol{l}', \boldsymbol{u}']$" into account at node ν does not alter the output of the BnB procedure. In particular, it still correctly identifies the solutions of (3.8). The second requirement (4.1b) implies that $r^{\nu}(\boldsymbol{l}', \boldsymbol{u}')$ can possibly be larger than $r^{\nu}(\boldsymbol{l}, \boldsymbol{u})$ since the lower bound in (3.23) is tightened by considering lower absolute values for $\boldsymbol{l}$ and $\boldsymbol{u}$. Overall, any choice of $[\boldsymbol{l}', \boldsymbol{u}']$ verifying (4.1a)-(4.1b) thus keeps unchanged the output of the BnB procedure while allowing for potentially more aggressive pruning decisions. This aspect is illustrated in Fig. 4.1, showing how the peeling potentially enables the BnB procedure to prune a node earlier than its standard counterpart with no peeling.

In the rest of this section, we describe a strategy to find some interval $[\boldsymbol{l}', \boldsymbol{u}']$ satisfying (4.1a)-(4.1b). Because of the symmetry of the problem at stake, we only focus on the construction of the upper bound $\boldsymbol{u}'$. The identification of a lower bound $\boldsymbol{l}'$ can be done along the same lines.

4.1.1 Target Peeling Strategy

Given some index $j \in \nu_\bullet$ and $\alpha > 0$, we consider the following perturbed versions of (3.25):

$$p_\alpha^\nu(\boldsymbol{l}, \boldsymbol{u}) := \inf_{\boldsymbol{s} \in \mathbb{R}^N} \mathrm{P}^\nu(\boldsymbol{s}; \boldsymbol{l}, \boldsymbol{u}) + \eta(s_j > \alpha). \tag{4.3}$$

Problem (4.3) corresponds to (3.25) where s_j is additionally constrained to be strictly greater than α. The following lemma then trivially follows from the definition of $p_\alpha^\nu(\boldsymbol{l}, \boldsymbol{u})$:

Lemma 4.2. *If* $\alpha \in [0, u_j[$ *and*

$$p_\alpha^\nu(\boldsymbol{l}, \boldsymbol{u}) > \bar{p}, \tag{4.4}$$

then (4.1a)-(4.1b) *hold with*

$$u_i' = \begin{cases} \alpha & \text{if } i = j \\ u_i & \text{otherwise.} \end{cases} \tag{4.5}$$

This result thus states that any $\alpha \in [0, u_j[$ verifying (4.4) enables to construct some $\boldsymbol{u}'$ automatically fulfilling (4.1a)-(4.1b). Unfortunately, evaluating (4.4) involves the same computational burden as solving (3.25). This problem can nevertheless be circumvented by finding some proper lower bound on $p_\alpha^\nu(\boldsymbol{l}, \boldsymbol{u})$ as described in the next section.

4.1.2 Tractable Implementation

In Sec. 4.4, we leverage Fenchel-Rockafellar duality for problem (3.27) to show that for any $\boldsymbol{w} \in \mathbb{R}^M$, the following lower bound on $p_\alpha^\nu(\boldsymbol{l}, \boldsymbol{u})$ holds:

$$p_\alpha^\nu(\boldsymbol{l}, \boldsymbol{u}) \geq \mathrm{D}^\nu(\boldsymbol{w}; \boldsymbol{l}, \boldsymbol{u}) + \psi_j(\boldsymbol{w}; \boldsymbol{l}, \boldsymbol{u}) + \alpha[-\boldsymbol{a}_j^\top \boldsymbol{w}]_+, \tag{4.6}$$

where

$$\begin{aligned}\mathrm{D}^{\nu}(\boldsymbol{w};\boldsymbol{l},\boldsymbol{u}) &:= \tfrac{1}{2}\|\boldsymbol{y}\|_2^2 - \tfrac{1}{2}\|\boldsymbol{y}-\boldsymbol{w}\|_2^2 + \lambda|\nu_1| \\ &\quad - \sum_{i\in\nu_1}\mu_{0,i}(\boldsymbol{a}_i^\top\boldsymbol{w}) - \sum_{i\in\nu_\bullet}\mu_{\lambda,i}(\boldsymbol{a}_i^\top\boldsymbol{w}) \\ \psi_j(\boldsymbol{w};\boldsymbol{l},\boldsymbol{u}) &:= \mu_{\lambda,j}(\boldsymbol{a}_j^\top\boldsymbol{w}) - u_j[\boldsymbol{a}_j^\top\boldsymbol{w}]_+ + \lambda\end{aligned}$$

and $\mu_{\rho,i}(v) := [u_i v - \rho]_+ + [l_i v - \rho]_+$.

Using this result, condition (4.4) can be relaxed as

$$\mathrm{D}^{\nu}(\boldsymbol{w};\boldsymbol{l},\boldsymbol{u}) + \psi_j(\boldsymbol{w};\boldsymbol{l},\boldsymbol{u}) + \alpha[-\boldsymbol{a}_j^\top\boldsymbol{w}]_+ > \bar{p}. \tag{4.7}$$

Hence, choosing any $\alpha \in [0, u_j[$ verifying (4.7) for some $\boldsymbol{w} \in \mathbb{R}^M$ defines a new valid constraint via (4.5), in the sense of (4.1a)-(4.1b). Interestingly, the left-hand side of (4.7) depends linearly on α, thus allowing to precisely characterize the range of possible values satisfying the strict inequality (4.7). This leads us to the main result of this section.

Proposition 4.3. *Let $\boldsymbol{w} \in \mathbb{R}^M$. If $\boldsymbol{a}_j^\top\boldsymbol{w} \geq 0$ and*

$$\mathrm{D}^{\nu}(\boldsymbol{w};\boldsymbol{l},\boldsymbol{u}) + \psi_j(\boldsymbol{w};\boldsymbol{l},\boldsymbol{u}) > \bar{p}, \tag{4.8}$$

then $\boldsymbol{u}'$ defined as in (4.5) with $\alpha = 0$ fulfills (4.1a)-(4.1b). Moreover, if $\boldsymbol{a}_j^\top\boldsymbol{w} < 0$, then $\boldsymbol{u}'$ defined as in (4.5) with any $\alpha \in [0, u_j[$ verifying

$$\alpha > \bar{\alpha} := \frac{\bar{p} - \mathrm{D}^{\nu}(\boldsymbol{w};\boldsymbol{l},\boldsymbol{u}) - \psi_j(\boldsymbol{w};\boldsymbol{l},\boldsymbol{u})}{[-\boldsymbol{a}_j^\top\boldsymbol{w}]_+} \tag{4.9}$$

fulfills (4.1a)-(4.1b).

Our next result shows that Prop. 4.3 can be applied to all indices $j \in [N]$ either sequentially or in parallel, while preserving the correctness of the BnB procedure:

Lemma 4.4. *Let $[\boldsymbol{l}', \boldsymbol{u}']$ and $[\boldsymbol{l}'', \boldsymbol{u}'']$ be two intervals satisfying (4.1a)-(4.1b). Then, the interval $[\boldsymbol{l}', \boldsymbol{u}'] \cap [\boldsymbol{l}'', \boldsymbol{u}'']$ also fulfills (4.1a)-(4.1b).*

A proof is available in Sec. 4.4.

We note that in terms of complexity the parallel application of Prop. 4.3 to all indices $j \in [N]$ requires the computation of the inner products

$\{\boldsymbol{a}_i^\top \boldsymbol{w}\}_{i=1}^N$ and *one single* evaluation of $\mathrm{D}^\nu(\boldsymbol{w}; \boldsymbol{l}, \boldsymbol{u})$. Interestingly, these inner products are already computed in most numerical procedures solving (3.27) and are thus usually available at no additional cost, see *e.g.,* [GHE22, Sec. 4.3]. The overhead complexity of applying in parallel our proposed peeling strategy thus scales as $\mathcal{O}(N + M)$.

4.1.3 Propagating Peeling Down the Tree

In this section, we emphasize that any interval $[\boldsymbol{l}', \boldsymbol{u}']$ verifying (4.1a)-(4.1b) at node ν can be used as a starting point to apply our peeling procedure at the child nodes of ν. More specifically, the following result holds:

Lemma 4.5. *Let $[\boldsymbol{l}', \boldsymbol{u}']$ be some interval verifying (4.1a)-(4.1b) at node ν and let ν' be some child node of ν. Assume that the peeling procedure defined in Prop. 4.3 is applied at node ν' with $[\boldsymbol{l}', \boldsymbol{u}']$ as input, rather than $[\boldsymbol{l}, \boldsymbol{u}]$, to generate a new interval $[\boldsymbol{l}'', \boldsymbol{u}'']$. Then we have*

$$\forall \boldsymbol{s} \in [\boldsymbol{l}, \boldsymbol{u}] \setminus [\boldsymbol{l}'', \boldsymbol{u}''] : \mathrm{P}^{\nu'}(\boldsymbol{s}; \boldsymbol{l}, \boldsymbol{u}) > \bar{p} \tag{4.10a}$$

$$[\boldsymbol{l}'', \boldsymbol{u}''] \subseteq [\boldsymbol{l}, \boldsymbol{u}]. \tag{4.10b}$$

A proof of this result is available in Sec. 4.4. In other words, Lem. 4.5 states that any peeled interval $[\boldsymbol{l}', \boldsymbol{u}']$ computed at node ν can be used as a starting point to apply a new peeling step at any child node ν'. This allows to propagate the peeled interval $[\boldsymbol{l}', \boldsymbol{u}']$ down the decision tree to hopefully improve sequentially the tightness of the convex relaxation (3.27).

4.2 Numerical Results

This section reports an empirical study demonstrating the effectiveness of the proposed peeling procedure to accelerate the resolution of (3.8) on a synthetic dataset. We refer the reader to [BKM16, HTT17] for an in-depth study of the statistical properties of the optimizer obtained from this problem.

4.2.1 Experimental Setup

We consider instances of problem (3.8) with dimensions $(M, N) = (100, 150)$. For each trial, new realizations of $\boldsymbol{A}, \boldsymbol{y}$ and λ are generated as follows. Each

row of the dictionary $\boldsymbol{A}$ is drawn from a multivariate normal distribution with zero mean and covariance matrix $\boldsymbol{K} \in \mathbb{R}^{N\times N}$. The (i,j)th entry of $\boldsymbol{K}$ is defined as $K_{ij} = \rho^{|i-j|}$, $\forall i,j \in [N]$, with $\rho = 0.1$. Each realization of $\boldsymbol{y}$ is generated in two steps. We first create a K-sparse vector $\boldsymbol{s}^\dagger \in \mathbb{R}^N$ with evenly-distributed non-zero components, where $K = 5$. The non-zero entries are defined as $s_i^\dagger = \text{sign}(r_i) + r_i$ where r_i is an independent realization of a zero-mean Gaussian with variance σ^2. We then set $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{s}^\dagger + \boldsymbol{w}$ for some zero-mean white Gaussian noise $\boldsymbol{w}$. The variance of the noise is adjusted so that the SNR $:= 10\log_{10}(\|\boldsymbol{A}\boldsymbol{s}^\dagger\|_2^2/\|\boldsymbol{w}\|_2^2)$ is equal to 15dB. The parameter λ is calibrated for each instance of (3.8) using the cross-validation tools of the L0Learn package [HM20] with the default parameters. More specifically, we call the `cv.fit` procedure that takes $\boldsymbol{y}$ and $\boldsymbol{A}$ as inputs and returns couples $(\boldsymbol{s}_{\lambda_i}, c_{\lambda_i})$ from a grid of values $\{\lambda_i\}_{i\in\mathbb{N}}$ selected data-dependently by the package. The vector $\boldsymbol{s}_{\lambda_i}$ is an approximate solution of (3.8) where the weight of the ℓ_0-norm is set to λ_i and c_{λ_i} is an associated cross-validation score computed on 10 folds of $M/10$ randomly sampled rows in $\boldsymbol{A}$ and entries in $\boldsymbol{y}$. We then set

$$\lambda = \arg\min_{\lambda_i} \; c_{\lambda_i} \text{ s.t. } \|\boldsymbol{s}_{\lambda_i}\|_0 = \left\|\boldsymbol{s}^\dagger\right\|_0. \tag{4.11}$$

4.2.2 Competing Procedures

We consider the following numerical solvers addressing (3.8): *i)* Cplex [Cpl09], a generic mixed-integer problem solver; *ii)* L0bnb [HMS22], a standard BnB procedure using a "breadth-first search" exploration strategy, see [HMS22, Sec. 3.3]; *iii)* Sbnb, a standard BnB procedure using a "depth-first search" exploration strategy, see [MBM+20, Sec. 2.2]; *iv)* Sbnb-N, corresponding to Sbnb enhanced with additional "node-screening" techniques, see [GHE22]; *v)* Sbnb-P, corresponding to Sbnb enhanced with the peeling strategy presented in this paper. L0bnb, Sbnb, Sbnb-N and Sbnb-P all use the same solving procedure for relaxed problem (3.27), namely a coordinate descent method [Wri15]. We use the C++ implementation of Cplex[1] and the Python implementation of L0bnb.[2] Sbnb, Sbnb-N and Sbnb-P are implemented in Julia.

For Sbnb-P, peeling is applied at each iteration of the numerical procedure solving the relaxed problem (3.27). We use the current iterate, say $\boldsymbol{s}^{(k)}$, to define $\boldsymbol{w} := \boldsymbol{y} - \boldsymbol{A}\boldsymbol{s}^{(k)}$ and apply the peeling rules defined in Prop. 4.3

[1] `https://github.com/jump-dev/CPLEX.jl`
[2] `https://github.com/hazimehh/L0Learn`

in parallel, *i.e.,* simultaneously for all the components of $\boldsymbol{s}$. The value of α satisfying (4.9), if any, is chosen as $\alpha = \bar{\alpha} + 10^{-16}$. The peeled intervals are propagated through the decision tree as described in Sec. 4.1.3.

All the solving procedures are provided with the initial bounds $\boldsymbol{l} = -M\mathbf{1}$ and $\boldsymbol{u} = M\mathbf{1}$ for some proper value of M. This corresponds to the standard *"Big-M"* constraint commonly considered in the literature [MBM+20, BMBN22, BCWP21, BNCM15]. As far as our random simulation setup is concerned, it can be shown that (3.8) admits a unique minimizer $\boldsymbol{s}^\star$ with probability one. We thus choose $M = \gamma\|\boldsymbol{s}^\star\|_\infty$ for some $\gamma \geq 1$ in our simulations. This requires to solve (3.8) once beforehand to identify $\boldsymbol{s}^\star$. This operation is here only done for the sake of comparing the sensibility of the solving methods to the choice of γ.

In practice, we obtain $\boldsymbol{s}^\star$ by solving a sequence of problems with an increasing value of M in the *Big-M* constraint. More specifically, letting $\boldsymbol{s}^\star_M$ denote the solution of (3.8) with the additional constraint "$-M\mathbf{1} \leq \boldsymbol{s} \leq M\mathbf{1}$", we are guaranteed that $\boldsymbol{s}^\star = \boldsymbol{s}^\star_M$ as soon as the strict inequality $\|\boldsymbol{s}^\star_M\|_\infty < M$ holds. We thus compute the solution $\boldsymbol{s}^\star_M$ for a sequence of M of the form $\{\eta^i M_0\}_{i\in\mathbb{N}}$ with $\eta = 1.1$ and for some $M_0 > 0$ and stop as soon as $\|\boldsymbol{s}^\star_M\|_\infty < M$.

4.2.3 Computational Gains

Fig. 4.2 presents the performance of the considered solving procedures. All results are averaged over 50 problem instances. Experiments were run on one Intel Xeon E5-2660 v3 CPU clocked at 2.60 GHz with 16 GB of RAM. The left column in Fig. 4.2 represents the average solving time of each procedure as a function of γ (top) and σ (bottom); the right column illustrates the gain allowed by the proposed method in terms of solving time (solid) and number of nodes explored (dashed) as compared to its best competitor, that is SBNB-N.

We note that SBNB-P leads to the smallest running time in all the considered setups. Since the latter corresponds to SBNB where peeling has been added, the spacing between the red and green curves materializes the gain provided by peeling. As far as our simulation setup is concerned, we see that the proposed method enables an acceleration of almost one order of magnitude with respect to SBNB. It is noticeable that this acceleration occurs even if $\gamma = 1$, that is, the *Big-M* constraint is perfectly tuned to the problem at hand. This is due to the fact that peeling can refine *individually* each component of the initial bounds $\boldsymbol{l}$ and $\boldsymbol{u}$ at *each node* of the BnB

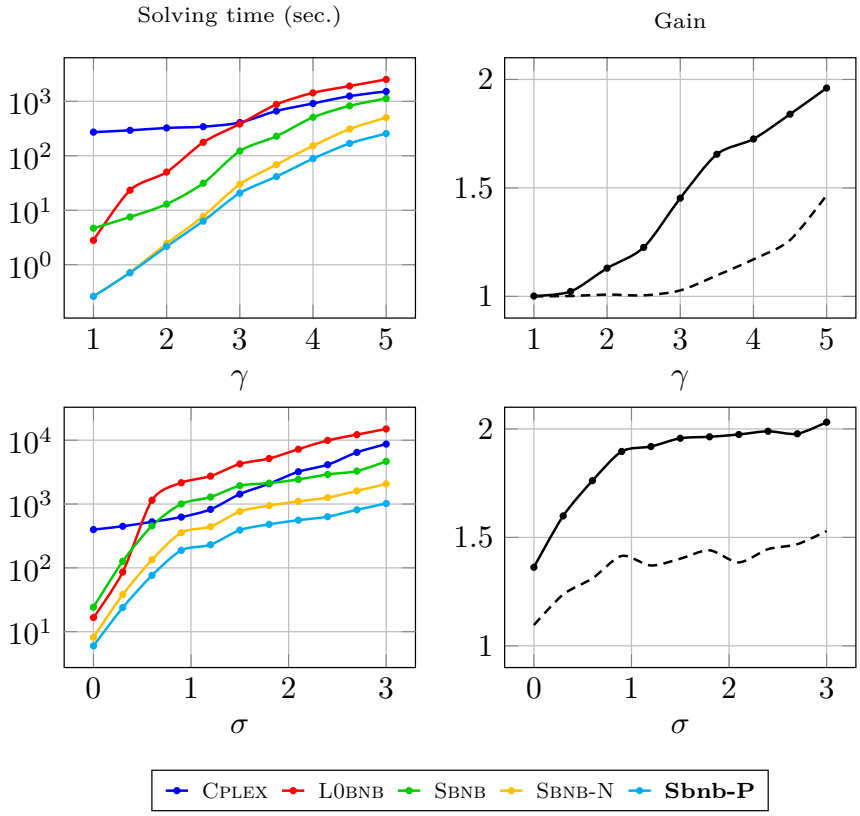

Figure 4.2: Left: Solving time as a function of γ (top, $\sigma = 1$) and σ (bottom, $\gamma = 5$). Right: gain in terms of solving time (solid) and number of nodes explored (dashed) with respect to SBNB-N.

decision tree to fit the local geometry of the problem.

We also notice that SBNB-P improves over SBNB-N, which can be seen as another acceleration of SBNB. In particular, SBNB-P performs always as well as SBNB-N as emphasized by the gains in the right-hand side of Fig. 4.2. We note in particular the gain provided by peeling in terms of number of nodes processed by the BnB procedure: as expected, peeling allows for more aggressive pruning and thus reduces the number of nodes to be explored.

For now, we only varied the parameters that may directly influence the *Big-M* constraint. To give a broader overview of the performance allowed by our peeling methodology, we have selected three different types of instances of (3.8) with increasing hardness. We still fix M, N and SNR but we vary

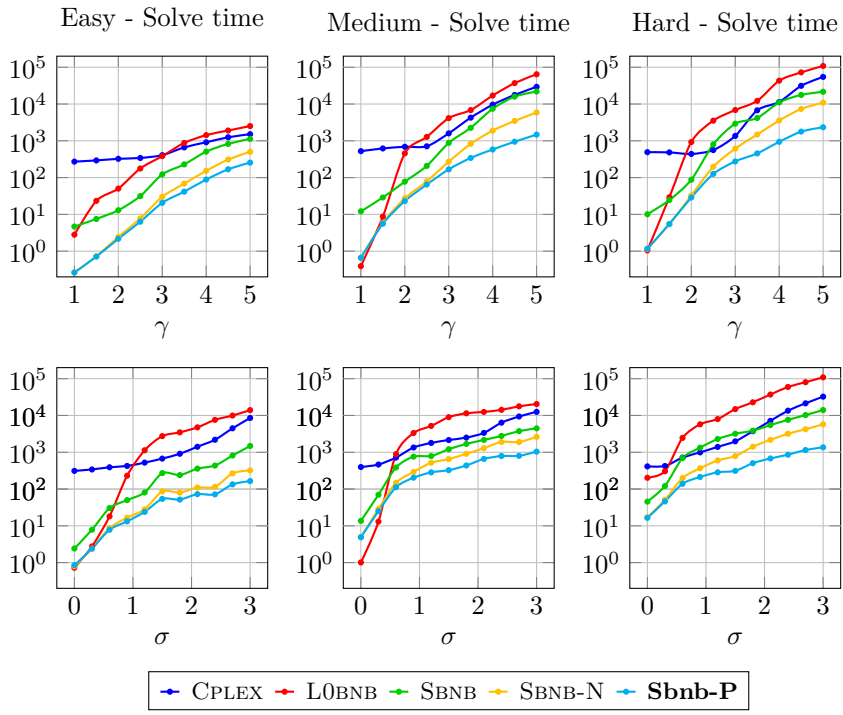

Figure 4.3: Solving time as a function of γ (top, $\sigma = 1$) and σ (bottom, $\gamma = 3$).

K and ρ to control the instance's inherent complexity. On the one hand, larger K leads to problems with a more complex combinatorial nature. On the other hand, increasing ρ leads to more correlation between the columns of $\boldsymbol{A}$, which brings the local minima of (3.8) closer to the global ones. Both of these effects make the problem harder to solve. Our three setups are constructed with the following combination of the parameters:

	K	M	N	ρ	SNR
Easy	5	100	150	0.1	15
Medium	7	100	150	0.1	15
Hard	7	100	150	0.8	15

We note that the maximum solution time allowed for the solver is 30 hours, that is just above 10^5 seconds. For particularly hard instances, some solvers

hit this time limit. Since it only concerns a minority of the trials, we have decided not to remove them, but this may have slightly lower their mean solution time.

Fig. 4.3 shows the solution time of the different methods for the three different types of instances when varying γ and σ. We roughly observe the same behaviors as in Fig. 4.2. A notable remark is that the more difficult the instance, the larger the gains permitted by our peeling methodology. Our empirical analysis is that on harder instances, the BnB tree has to be explored more extensively. Tightening the bounds at a given node will therefore allow a greater number of nodes to benefit from the strengthening of the relaxations. The solving time will therefore be further reduced. When varying γ, the gains of the method implementing our peeling strategy (SBNB-P) again its best concurrent (SBNB-N) almost reach a factor 2 on EASY instances. On MEDIUM instances, it reaches a factor 4 and on HARD instances, it reaches a factor 5.

4.3 Discussion

In this chapter, we presented a tractable strategy, named "*peeling*", to tighten the box constraints used in a BnB procedure tailored to ℓ_0-regularized least-squares problems. Unlike the standard approach, which imposes *one global* constraint to the problem, our strategy aims to locally refine the box constraints *at each node* of the decision tree. This refinement enables to strengthening the convex relaxations used in the pruning decisions made by the BnB procedure and can lead to significant improvements in terms of solving time, as emphasized by our simulation results.

Following the comments from Sec. 3.5, the results of this chapter potentially open up new ways for future research projects to extend the peeling and screening techniques to enhance greedy algorithms. Indeed, the stronger convex relaxation the peeling provides may enable screening tests to pass in the branches explored by greedy approaches. Achieving this task is expected to require additional research on identifying "good" suboptimal values of the dual variable from which the evaluation of the screening tests is more likely to pass successfully.

4.4 Appendix: Proofs

4.4.1 Proof of Lemma 4.1

We first have

$$p^\nu(\boldsymbol{l}', \boldsymbol{u}') \geq p^\nu(\boldsymbol{l}, \boldsymbol{u})$$

since $[\boldsymbol{l}', \boldsymbol{u}'] \subseteq [\boldsymbol{l}, \boldsymbol{u}]$ from (4.1b) and the left-hand side corresponds to the minimum value of a problem more constrained than the right-hand side. The reverse implication in (4.2) is thus always verified.

The direct implication results from the following observations. If $p^\nu(\boldsymbol{l}, \boldsymbol{u}) \leq \bar{p}$, then we have

$$p^\nu(\boldsymbol{l}, \boldsymbol{u}) = p^\nu(\boldsymbol{l}', \boldsymbol{u}')$$

since (4.1a) ensures that the minimizers of $\mathrm{P}^\nu(\boldsymbol{s}; \boldsymbol{l}, \boldsymbol{u})$ do not belong to $[\boldsymbol{l}, \boldsymbol{u}] \setminus [\boldsymbol{l}', \boldsymbol{u}']$ and are therefore also minimizers of $\mathrm{P}^\nu(\boldsymbol{s}; \boldsymbol{l}', \boldsymbol{u}')$. We thus obtain the direct implication by contraposition.

4.4.2 Proof of (4.6)

We first note that

$$\begin{aligned}
p^\nu_\alpha(\boldsymbol{l}, \boldsymbol{u}) &= \inf_{\boldsymbol{s} \in \mathbb{R}^N} \mathrm{P}^\nu(\boldsymbol{s}; \boldsymbol{l}, \boldsymbol{u}) + \eta(s_j > \alpha) && \text{(4.12a)} \\
&\geq \inf_{\boldsymbol{s} \in \mathbb{R}^N} \mathrm{R}^\nu_j(\boldsymbol{s}; \boldsymbol{l}, \boldsymbol{u}) + \eta(s_j > \alpha) && \text{(4.12b)} \\
&\geq \inf_{\boldsymbol{s} \in \mathbb{R}^N} \mathrm{R}^\nu_j(\boldsymbol{s}; \boldsymbol{l}, \boldsymbol{u}) + \eta(s_j \geq \alpha) && \text{(4.12c)} \\
&= \min_{\boldsymbol{s} \in \mathbb{R}^N} \mathrm{R}^\nu_j(\boldsymbol{s}; \boldsymbol{l}, \boldsymbol{u}) + \eta(s_j \geq \alpha) && \text{(4.12d)}
\end{aligned}$$

where

$$\begin{aligned}
\mathrm{R}^\nu_j(\boldsymbol{s}; \boldsymbol{l}, \boldsymbol{u}) := \tfrac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{s}\|_2^2 + \lambda \sum_{i \in \nu_\bullet \setminus j} \left(\tfrac{[s_i]_+}{u_i} - \tfrac{[-s_i]_+}{l_i} \right) \\
+ \lambda(|\nu_1| + 1) + \eta(\boldsymbol{s}_{\nu_0} = \boldsymbol{0}) + \eta(\boldsymbol{s} \in [\boldsymbol{l}, \boldsymbol{u}]) \quad \text{(4.13)}
\end{aligned}$$

and the last equality follows from the fact that the objective function is lower-semi-continuous and its domain is closed and bounded.

The lower bound in (4.6) then stems from the Fenchel-Rockafellar dual

problem of (4.12d). More specifically, we first notice that (4.12d) can be rewritten as

$$\min_{\boldsymbol{s}\in\mathbb{R}^N} f(\boldsymbol{A}\boldsymbol{s}) + g(\boldsymbol{s}) \tag{4.14}$$

with

$$f(\mathbf{z}) = \tfrac{1}{2}\|\boldsymbol{y}-\mathbf{z}\|_2^2 \tag{4.15a}$$
$$g(\boldsymbol{s}) = g_j(s_j) + \sum_{i\neq j} g_i(s_i) \tag{4.15b}$$

and

$$g_j(s) = \eta(s\in[\alpha,u_j]) + \lambda$$
$$g_i(s) = \begin{cases} \eta(s=0) & \text{if } i\in\nu_0 \\ \eta(s\in[l_i,u_i]) + \lambda & \text{if } i\in\nu_1 \\ \eta(s\in[l_i,u_i]) + \lambda(\frac{[s]_+}{u_i} - \frac{[-s]_+}{l_i}) & \text{if } i\in\nu_\bullet\backslash j. \end{cases}$$

Second, we have from standard results of convex optimization [Bec17, Ch. 12] that the Fenchel-Rockafellar dual problem of (4.14) reads

$$\max_{\boldsymbol{w}\in\mathbb{R}^M} \; -f^\star(-\boldsymbol{w}) - g^\star(\boldsymbol{A}^\top\boldsymbol{w}), \tag{4.16}$$

where $f^\star$ and $g^\star$ denote the convex conjugates of f and g (see [Bec17, Def. 4.1]). In particular, we have $\forall \boldsymbol{s}\in\mathbb{R}^N, \boldsymbol{w}\in\mathbb{R}^M$:

$$f(\boldsymbol{A}\boldsymbol{s}) + g(\boldsymbol{s}) \geq -f^\star(-\boldsymbol{w}) - g^\star(\boldsymbol{A}^\top\boldsymbol{w}) \tag{4.17}$$

as a consequence of weak duality applied to the pair (4.14)-(4.16). Our lower bound in (4.6) corresponds to a particularization of the right-hand side of (4.17) to (4.15a)-(4.15b). In particular, applying the definition of the convex conjugate to the function f, one easily obtains:

$$f^\star(\boldsymbol{w}) = -\tfrac{1}{2}\|\boldsymbol{y}\|_2^2 + \tfrac{1}{2}\|\boldsymbol{y}+\boldsymbol{w}\|_2^2. \tag{4.18}$$

Invoking Lem. 4.6 in Sec. 4.4 with $(b,l,u) = (\lambda,\alpha,u_j)$, we obtain

$$g_j^\star(v) = u_j[v]_+ - \alpha[-v]_+ - \lambda. \tag{4.19}$$

Applying again Lem. 4.6 with

$$(b, l, u) = \begin{cases} (0, 0, 0) & \text{if } i \in \nu_0 \\ (\lambda, l_i, u_i) & \text{if } i \in \nu_1, \end{cases}$$

leads to

$$g_i^\star(v) = \begin{cases} 0 & \text{if } i \in \nu_0 \\ [u_i v]_+ + [l_i v]_+ - \lambda & \text{if } i \in \nu_1 \end{cases} \tag{4.20}$$

where we have used the fact that for all $i \in \nu_1$, $u_i[v]_+ = [u_i v]_+$ (resp. $l_i[-v]_+ = -[l_i v]_+$) since $u_i \geq 0$ (resp. $l_i \leq 0$). Similarly, using Lem. 4.7 in Sec. 4.4 with $(a, l, u) = (\lambda, l_i, u_i)$ we obtain:

$$g_i^\star(v) = [u_i v - \lambda]_+ + [l_i v - \lambda]_+ \quad \text{if } i \in \nu_\bullet \backslash j. \tag{4.21}$$

Finally, combining (4.18)-(4.21) leads to the desired result.

4.4.3 Proof of Lemma 4.4

Let $[\boldsymbol{l}', \boldsymbol{u}']$ and $[\boldsymbol{l}'', \boldsymbol{u}'']$ be two intervals satisfying (4.1a)-(4.1b), and defines $\mathcal{S} = [\boldsymbol{l}', \boldsymbol{u}'] \cap [\boldsymbol{l}'', \boldsymbol{u}'']$. We first note that $[\boldsymbol{l}', \boldsymbol{u}'] \cap [\boldsymbol{l}'', \boldsymbol{u}'']$ can be rewritten as

$$\mathcal{S} = \bigtimes_{i=1}^{N} [\max(l_i', l_i''), \min(u_i', u_i'')]$$

where $\bigtimes$ denotes the Cartesian product of sets, and therefore defines an interval. It thus remains to show that $\mathcal{S}$ fulfills (4.1a)-(4.1b).

First, the inclusion $\mathcal{S} \subset [\boldsymbol{l}, \boldsymbol{u}]$ follows from the fact that both $[\boldsymbol{l}', \boldsymbol{u}']$ and $[\boldsymbol{l}'', \boldsymbol{u}'']$ verify (4.1b). Second, the fact that $\mathcal{S}$ fulfills (4.1a) is a consequence of the following set equality:

$$[\boldsymbol{l}, \boldsymbol{u}] \setminus \mathcal{S} = \big([\boldsymbol{l}, \boldsymbol{u}] \setminus [\boldsymbol{l}', \boldsymbol{u}']\big) \cup \big([\boldsymbol{l}, \boldsymbol{u}] \setminus [\boldsymbol{l}'', \boldsymbol{u}'']\big).$$

Hence, if $\boldsymbol{s} \in [\boldsymbol{l}, \boldsymbol{u}] \setminus \mathcal{S}$ then $\boldsymbol{s} \in \big([\boldsymbol{l}, \boldsymbol{u}] \setminus [\boldsymbol{l}', \boldsymbol{u}']\big)$ or $\boldsymbol{s} \in \big([\boldsymbol{l}, \boldsymbol{u}] \setminus [\boldsymbol{l}'', \boldsymbol{u}'']\big)$. Assume without loss of generality that $\boldsymbol{s} \in \big([\boldsymbol{l}, \boldsymbol{u}] \setminus [\boldsymbol{l}', \boldsymbol{u}']\big)$. Since $[\boldsymbol{l}', \boldsymbol{u}']$ fulfills (4.1a), we then have $\mathrm{P}^\nu(\boldsymbol{s}; \boldsymbol{l}, \boldsymbol{u}) > \bar{p}$. One concludes the proof by noting that the latter rationale holds for all $\boldsymbol{s} \in [\boldsymbol{l}, \boldsymbol{u}] \setminus \mathcal{S}$.

4.4.4 Proof of Lemma 4.5

If the peeling procedure defined in Prop. 4.3 is applied at node ν' with $[\boldsymbol{l}', \boldsymbol{u}']$ as input to generate a new interval $[\boldsymbol{l}'', \boldsymbol{u}'']$, we have by construction:

$$\forall \boldsymbol{s} \in [\boldsymbol{l}', \boldsymbol{u}'] \setminus [\boldsymbol{l}'', \boldsymbol{u}''] : \mathrm{P}^{\nu'}(\boldsymbol{s}; \boldsymbol{l}', \boldsymbol{u}') > \bar{p} \tag{4.22a}$$

$$[\boldsymbol{l}'', \boldsymbol{u}''] \subseteq [\boldsymbol{l}', \boldsymbol{u}']. \tag{4.22b}$$

We note that (4.22a) is equivalent to

$$\forall \boldsymbol{s} \in [\boldsymbol{l}', \boldsymbol{u}'] \setminus [\boldsymbol{l}'', \boldsymbol{u}''] : \mathrm{P}^{\nu'}(\boldsymbol{s}; \boldsymbol{l}, \boldsymbol{u}) > \bar{p} \tag{4.23}$$

since $\boldsymbol{s} \in [\boldsymbol{l}', \boldsymbol{u}']$ and $[\boldsymbol{l}', \boldsymbol{u}'] \subseteq [\boldsymbol{l}, \boldsymbol{u}]$ by hypothesis.

Moreover, we have

$$\mathrm{P}^{\nu'}(\boldsymbol{s}; \boldsymbol{l}, \boldsymbol{u}) \geq \mathrm{P}^{\nu}(\boldsymbol{s}; \boldsymbol{l}, \boldsymbol{u}).$$

since $\mathcal{S}^{\nu'} \subset \mathcal{S}^{\nu}$ for any child node ν' of ν. Hence, any interval $[\boldsymbol{l}', \boldsymbol{u}']$ verifying (4.1a) at node ν obviously also fulfills (4.1a) at ν', that is

$$\forall \boldsymbol{s} \in [\boldsymbol{l}, \boldsymbol{u}] \setminus [\boldsymbol{l}', \boldsymbol{u}'] : \mathrm{P}^{\nu'}(\boldsymbol{s}; \boldsymbol{l}, \boldsymbol{u}) > \bar{p}. \tag{4.24}$$

Combining (4.22b), (4.23) and (4.24), we finally obtain the result.

We finally derive the expressions of two conjugate functions appearing in the derivation of (4.6). The results are encapsulated in Lem. 4.6 and 4.7 below.

Lemma 4.6. *Let* $g \colon \mathbb{R} \to \mathbb{R}$ *be defined for all* $s \in \mathbb{R}$ *as*

$$g(s) = b + \eta(s \in [l, u])$$

for $b \in \mathbb{R}$ *and* $l \leq u$. *Then,*

$$g^\star(v) = u[v]_+ - l[-v]_+ - b.$$

Proof. By definition of the convex conjugate (see [Bec17, Def. 4.1]), one has

$\forall v \in \mathbb{R}$:

$$\begin{aligned} g^\star(v) &= \sup_{s\in\mathbb{R}} \; vs - g(s) \\ &= \max_{l\le s\le u} \; vs - b \\ &= u[v]_+ - l[-v]_+ - b \end{aligned}$$

□

Lemma 4.7. *Let* $g\colon \mathbb{R} \to \mathbb{R}$ *be defined for all* $s \in \mathbb{R}$ *as*

$$g(s) = \eta(s \in [l, u]) + \tfrac{a}{u}[s]_+ - \tfrac{a}{l}[-s]_+$$

for $a \ge 0$ *and* $l < 0 < u$. *Then,*

$$g^\star(v) = [uv - a]_+ + [lv - a]_+.$$

This expression remains valid for $l \le 0 \le u$ *as long as we use the convention "*$0/0 = 0$*".*

Proof. Prior to proving the lemma, let us first show that the following result holds true: for all scalars $\beta \in \mathbb{R}, \mu \ge 0$, we have

$$\max_{0\le s\le \mu} \; \beta s - \tfrac{a}{\mu} s = [\mu\beta - a]_+. \tag{4.26}$$

Note that the maximum in (4.26) is attained since the optimization problem amounts to maximizing a linear function over a (nonempty) compact set. Now, if $\mu = 0$ we have

$$\max_{0\le s\le \mu} \; \beta s - \tfrac{a}{\mu} s = 0$$

by using the convention "$0/0 = 0$". Moreover, if $\mu > 0$:

$$\begin{aligned} \max_{0\le s\le \mu} \; \beta s - \tfrac{a}{\mu} s \; &= \max_{0\le s'\le 1} \; s'(\mu\beta - a) \\ &= [\mu\beta - a]_+. \end{aligned}$$

Overall, one can summarize the two cases as

$$\max_{0\le s\le \mu} \; \beta s - \tfrac{a}{\mu} s = [\mu\beta - a]_+ \tag{4.28}$$

since $a \geq 0$.

We now turn to the proof of Lem. 4.7. By definition of the convex conjugate (see [Bec17, Def. 4.1]), one has $\forall v \in \mathbb{R}$:

$$g^\star(v) = \sup_{s \in \mathbb{R}} \; vs - g(s).$$

The above problem can be equivalently expressed as

$$g^\star(v) = \max\left\{g_1^\star(v); g_2^\star(v)\right\} \tag{4.29}$$

where

$$\begin{aligned} g_1^\star(v) &:= \sup_{s \geq 0} \; vs - g(s) \\ g_2^\star(v) &:= \sup_{s \leq 0} \; vs - g(s). \end{aligned}$$

On the one hand, $g_1^\star(v)$ can be rewritten as

$$g_1^\star(v) \; = \sup_{0 \leq s \leq u} \; vs - \tfrac{a}{u} s.$$

By applying (4.26) with $(\beta, \mu) = (v, u)$, we then obtain:

$$g_1^\star(v) = [uv - a]_+. \tag{4.32}$$

On the other hand, we have

$$\begin{aligned} g_2^\star(v) \; &= \sup_{l \leq s \leq 0} \; vs - \tfrac{a}{l} s \\ &= \sup_{0 \leq s' \leq -l} \; -vs' - \tfrac{a}{-l} s'. \end{aligned}$$

Hence, using (4.26) with $(\beta, \mu) = (v, -l)$ leads to

$$g_2^\star(v) = [lv - a]_+. \tag{4.34}$$

Finally, plugging (4.32)-(4.34) into (4.29) yields

$$\begin{aligned} g^\star(v) &= \max\left\{[uv - a]_+; [lv - a]_+\right\} \\ &= [uv - a]_+ + [lv - a]_+ \end{aligned}$$

where the last equality follows from the fact that

$$uv - a > 0 \implies lv - a \leq 0.$$

Indeed, if $uv - a > 0$ then necessarily $u > 0$ since $a \geq 0$. We thus have $v > \frac{a}{u} \geq 0$. This implies that $lv - a \leq 0$ since $l \leq 0$ and $a \geq 0$. □

Chapter 5

Leaving the Grid with Interpolation

EXTENDING the sparse coding theory to sensor applications with both low-cost and real-time requirements, such as radars, is a challenging task that can only be addressed by means of fast heuristics. Indeed, these circumvent the challenge of solving the initial NP-hard problem which requires too much computation power. The famous Orthogonal Matching Pursuit (OMP) algorithm presented in Chapter 3 for parametric dictionaries is known for its efficiency in the decomposition of highly sparse signals (*i.e.,* with a very small number of atoms in the decomposition). This observation however only holds for its discrete version, restricted to a grid of parameters. Reducing the errors caused by the gridding can be achieved either by (i) using a denser grid or (ii) solving a highly non-convex optimization problem [TA18]. Both cases result in a deterioration of the speed of the decomposition procedure. This chapter investigates alternatives based on a principle of "atom interpolation", recently introduced in multiple contributions [ETS11, FDD13, KYHP15, LZL+18, CH19]. We leverage this technique to derive extensions of OMP, able to leave the grid quickly.

In more detail, the progression of this chapter can be explained in three parts, each with one contribution. This first part theoretically introduces the interpolation procedure (Sec. 5.1) and details the different patterns (Sec. 5.2). It provides then, as a contribution, a theoretical bound on the application of interpolation in the first iteration of OMP (Sec. 5.3). The second part (Sec. 5.4) compares different practical implementations of an

interpolation-based Continuous OMP (COMP) with a novel proposition. Our proposition, built upon the conclusions drawn in the first part, demonstrates superior performance compared to existing propositions from the literature when the ground truth is subjected to some separation constraints. The third and last part starts with the introduction, in Sec. 5.5, of our "Factorization over Interpolation" strategy that we presented in [dGVJ20]. Our implementation of COMP is thereby extended to encompass factorized models such as encountered in the context of FMCW radars. We finally show and discuss the numerical results of applying our factorized algorithm compared with other methods.

All the proofs of the theorems and lemmas given in this chapter are gathered in the last section of this chapter.

5.1 Interpolation-based Atom Selection

The atom selection is the key step of greedy algorithms such as OMP. In the continuous formulation given in Alg. 2, this step reads

$$\hat{\theta} = \arg\max_{\theta \in \Theta} \big| \langle \boldsymbol{a}(\theta), \boldsymbol{r} \rangle \big|, \tag{5.1}$$

given $\boldsymbol{r}$ to denote the residual signal at the beginning of a certain iteration. This problem is commonly highly non-convex and hence, the available methods to solve it often appear as computationally too demanding for real-time low-cost applications. This chapter thereby focuses on accelerated alternatives, built upon interpolation of atoms, to approximately compute $\hat{\theta}$.

The presented technique enables us to approximate any atom taken from the continuous set $\mathcal{A}$ with a linear combination of a finite set of atoms, gathered in the discrete dictionary $\boldsymbol{A}$ given by Def. 3.4. More precisely, the interpolation is performed piecewise from a partition of the domain Θ. To maintain the simplicity of notations, we impose the partition to provide exactly one subset of Θ around each grid point from $\Theta_{\mathrm{G}} := \{\bar{\theta}_n\}_{n=1}^{N}$. The resulting partition is schematically illustrated in Fig. 5.1. This yields, for $n \in [N]$, the subsets denoted by Θ_n such that $\Theta = \cup_{n=1}^{N} \Theta_n$. Formally, with I denoting the interpolation order standing for the number of interpolants, we gather in $\boldsymbol{A}_n$ the atoms from which the interpolation in Θ_n is performed. More precisely, if $\bar{\theta}_{n,i}$ denotes the i-th interpolant grid component taken around $\bar{\theta}_n$ in Θ_{G}, it reads

$$\boldsymbol{A}_n := \big[\, \boldsymbol{a}(\bar{\theta}_{n,1}), \cdots, \boldsymbol{a}(\bar{\theta}_{n,I}) \,\big]. \tag{5.2}$$

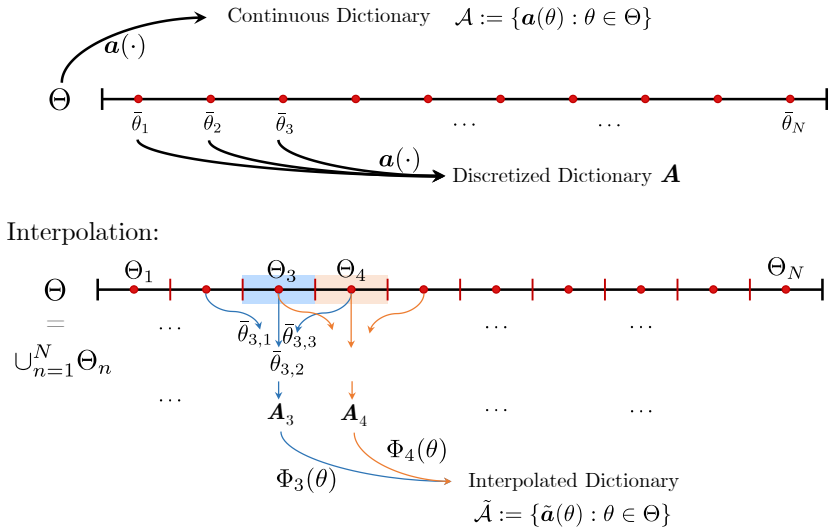

Figure 5.1: Schematic representation of the partition of the parameter domain Θ that is used to generate the interpolated dictionary $\tilde{\mathcal{A}}$ defined in Def. 5.1.

Typically, a 1-D and uniform grid Θ_G with a spacing Δ between adjacent bins gives $\bar{\theta}_{n,i} = \bar{\theta}_n + (i - \lfloor \frac{I-1}{2} \rfloor)\Delta$. The approximated dictionary $\tilde{\mathcal{A}}$ is then defined by the following.

Definition 5.1 (Approximated Continuous Dictionary)**.** The approximated dictionary $\tilde{\mathcal{A}}$ reads

$$\tilde{\mathcal{A}} := \{\tilde{\boldsymbol{a}}(\theta) : \theta \in \Theta\}, \tag{5.3}$$

where $\tilde{\boldsymbol{a}}$ approximates the atom function $\boldsymbol{a}$ through the piecewise interpolation reading for $\theta \in \Theta_n$,

$$\tilde{\boldsymbol{a}}(\theta) = \boldsymbol{A}_n \boldsymbol{c}. \tag{5.4}$$

In the above, we obtain the I interpolating weights by $\boldsymbol{c} = \Phi_n(\theta) \in \mathbb{R}^I$ where the expressions for the interpolating functions $\{\Phi_n\}_{n=1}^N$ are driven by the choice of interpolation pattern. By definition, those are defined such that the approximation is exact on the grid and hence

$$\tilde{\boldsymbol{a}}(\theta) = \boldsymbol{a}(\theta) \quad \text{for all} \quad \theta \in \Theta_G \tag{5.5}$$

We can now accelerate the atom selection by searching for the best matching atom in $\tilde{\mathcal{A}}$ instead of $\mathcal{A}$. This idea was notably hinted at in [CH19].

Using the decomposition (5.4), the approximated selection step writes

$$\max_{\theta\in\Theta} |\langle \tilde{\boldsymbol{a}}(\theta),\, \boldsymbol{y}\rangle| = \max_{n\in[N]} \max_{\theta\in\Theta_n} |\langle \Phi_n(\theta),\, \boldsymbol{A}_n^H \boldsymbol{y}\rangle|. \tag{5.6}$$

In practice, this selection computationally amounts to first evaluating the whole *on-the-grid* correlation $\boldsymbol{A}^H\boldsymbol{y}$ and next, for each index n, extract a set of I elements from the resulting vector to perform the inner maximization. This latest typically requires less computation time to solve than (5.1) because we commonly choose (i) an interpolation order I such that $I \ll M$ and (ii) interpolating functions Φ_n with simple expressions. As shown in the next section, the commonly employed interpolation patterns come with closed-form expressions for the inner maximizer of (5.6). This results in a complexity for the above which is dominated by the evaluation of $\boldsymbol{A}^H\boldsymbol{y}$.

Remark 5.1. The selection (5.6) is only valid when $\tilde{\mathcal{A}}$ contains atoms with constant ℓ_2 norm. Otherwise, one should instead inject the approximation into (3.13), which results in replacing the inner maximization of (5.6) by

$$\max_{\theta\in\Theta_n} \frac{|\langle \Phi_n(\theta),\, \boldsymbol{A}_n^H \boldsymbol{y}\rangle|}{\|\boldsymbol{A}_n \Phi_n(\theta)\|_2}. \tag{5.7}$$

With the above denominator entailing the norm of an M-length vector, the computational complexity of solving (5.7) is anticipated to exhibit a computational complexity that scales proportionally to M.
To guarantee a negligible computation time for this step, we suggest in this work the following *approximation* when the interpolated atoms have a non-uniform power over Θ. First, The non-normalized selection step (3.13) is relaxed as

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}\in\mathbb{C}^I} \tfrac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}_n\boldsymbol{\beta}\|_2^2, \tag{5.8}$$

and next, the minimizer $\hat{\boldsymbol{\beta}}$ of the above is projected onto the space of feasible interpolating coefficients through

$$\max_{\theta\in\Theta_n} \frac{|\langle \Phi_n(\theta),\, \hat{\boldsymbol{\beta}}\rangle|}{\|\Phi_n(\theta)\|_2}. \tag{5.9}$$

In Sec. 5.4 and 5.5, we also leverage this relax-and-project approach to propose a version of COMP that is faster than existing implementations from the literature.

5.2 Interpolation Patterns

In Def. 5.1, we stated that the interpolating function Φ_n depends on the chosen interpolation pattern. This aspect is clarified in this section, which summarizes several patterns from the literature and reformulates them to fit the framework given in the previous section. To this end, we restrict the range of interpolating functions to those matching the decomposition below. Given a function $\Psi_n(\theta)$,

$$\Phi_n(\theta) := \boldsymbol{U}_n \, \Psi_n(\theta), \tag{5.10}$$

for some matrix $\boldsymbol{U}_n \in \mathbb{C}^{I \times I}$.

To maintain lighter notations, we drop the index n from the interpolating function. We also impose the resulting function Ψ to be expressed as a function of the *off-the-grid* shifts $\delta := \theta - \bar{\theta}_n$, thereby restricting the study to interpolation strategies that are uniform over the partitioning of Θ. Interestingly, actual implementations of interpolation-based methods commonly meet these restrictions. Still, the reader can straightforwardly extend the patterns described below to their non-uniform version by inserting back the indexes n accordingly. This simplification yields, for all $\theta \in \Theta_n$ and given $n \in [N]$,

$$\tilde{\boldsymbol{a}}(\theta) = \boldsymbol{A}_n \, \boldsymbol{U} \, \Psi(\theta - \bar{\theta}_n). \tag{5.11}$$

The following lemma provides sufficient conditions for the approximation (5.11) to define an interpolation from the grid Θ_{G}. Its proof is provided in Sec. 5.8.

Lemma 5.1 (Interpolation Constraint)**.** *Given* $\tilde{\Theta} := \cup_{n=1}^{N} \{\theta - \bar{\theta}_n : \theta \in \Theta_n\}$, *if* $\{\Psi(\delta) : \delta \in \tilde{\Theta}\} > I$ *is such that any set of* I *vectors taken from it forms a set of linearly independent vector, then the approximation* (5.11) *is an interpolation in the sense of the constraint* (5.5)*, with*

$$\boldsymbol{U} := \begin{bmatrix} \big[\Psi(\bar{\theta}_{n,1} - \bar{\theta}_n)\big]^\top \\ \vdots \\ \big[\Psi(\bar{\theta}_{n,I} - \bar{\theta}_n)\big]^\top. \end{bmatrix}^{-1}, \tag{5.12}$$

The most commonly used interpolation patterns —that we present in the next section— all meet the condition provided by Lem. 5.1. They are thus fully explained by their corresponding expression for Ψ. Still, in the literature, the matrix $\boldsymbol{U}$ is either intrinsically encapsulated in the interpolating function Ψ [CH21] or in the definition of interpolant atoms [ETS11, FDD13, KYHP15]. In the next section, we reformulate the patterns described from those contributions to fit (5.11).

We now provide the expressions of Ψ corresponding to different existing patterns.

5.2.1 Polynomial and Taylor Interpolations

The simplest interpolation is obtained when the interpolating function provides a polynomial basis. For simplicity, we provide it for a 1-D domain Θ, *i.e.,*

$$\Psi(\delta) = \left[1,\, \delta,\, \cdots,\, \delta^{I-1}\right]. \tag{5.13}$$

Following the same idea, the authors in [ETS11] suggested an interpolation based on a Taylor expansion of the dictionary. Precisely, this pattern does not exactly fit the model (5.11) because $\boldsymbol{A}_n$ should contain the atom $\boldsymbol{a}(\theta)$ and its derivatives evaluated in $\bar{\theta}_n$. This in turn multiplies by I the number of discrete dictionaries to store in memory. In practice, this also multiplies by I the number of scalar products to perform prior to solving the inner maximizations in (5.6). Yet, we point out that the polynomial interpolation here above follows the same philosophy as the Taylor one because the product $\boldsymbol{A}_n \boldsymbol{U}$ yields approximations of the derivatives of $\boldsymbol{a}$, through finite differences. This can be observed in the simple example below.

Example 5.2. Considering a uniform grid Θ_{G} with spacing Δ and the first order polynomial interpolation, *i.e.,* with $I = 2$, it results

$$\boldsymbol{U} = \begin{bmatrix} 1 & 0 \\ 1 & \Delta \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ \frac{1}{\Delta} & -\frac{1}{\Delta} \end{bmatrix}. \tag{5.14}$$

In the search for computational efficiency, one should prefer the polynomial over the Taylor approximation since it requires I times fewer operations to perform for asymptotically similar results as Θ_{G} get denser.

Remark 5.3. We emphasize that first-order polynomial interpolation ($I = 2$) delivers a closed form expression for (5.9) since setting the derivative of its objective function to zero amounts to determine the roots of the following second-order polynomial in δ,

$$\Re\{\hat{\boldsymbol{\beta}}_1\hat{\boldsymbol{\beta}}_2^*\}(1-\delta^2) + (|\hat{\boldsymbol{\beta}}_2|^2 - |\hat{\boldsymbol{\beta}}_1|^2)\delta = 0. \tag{5.15}$$

Note that when $\hat{\boldsymbol{\beta}}$ is real, the above is solved by $\delta = \hat{\boldsymbol{\beta}}_2/\hat{\boldsymbol{\beta}}_1$.

5.2.2 Least-Squares Optimal Interpolation

While the polynomial interpolation is simple and versatile, it is known to result in limited performance when compared to other existing patterns. One can thus search for the best pattern by minimizing the squared distance between the actual atom and its approximation. This is equivalent to orthogonally projecting each atom $\boldsymbol{a}(\theta)$ onto a vectorial space V of dimension I. Since we aim for an approximation performing an interpolation, the unique possibility is for V to match the space generated by the columns of $\boldsymbol{A}_n$, which directly leads to

$$\tilde{\boldsymbol{a}}(\theta) = \boldsymbol{A}_n\boldsymbol{A}_n^\dagger\boldsymbol{a}(\theta). \tag{5.16}$$

For translation invariant dictionaries, when Θ_{G} defines a 1-D uniform grid, the above fits the model (5.11) with

$$\Psi(\delta) = \left[\, \kappa(\delta - \lfloor \tfrac{I-1}{2} \rfloor \Delta), \cdots, \kappa(\delta + \lfloor \tfrac{I}{2} \rfloor \Delta) \,\right]^\top, \tag{5.17}$$

where the kernel κ is defined as in Def. 3.3. Note that the optimal interpolation proposed here is a narrower definition than the optimal approximation mentioned in [CH19] which is not restricted to being an interpolation.

Although this pattern provides the best approximation $\tilde{\boldsymbol{a}}$, it is often computationally inefficient in practice for multiple reasons. First, the function κ might have no closed-form expression or a complicated one. Additionally, If $\mathcal{A}$ is defined over $\mathbb{C}$, then the interpolating coefficients resulting from (5.11) are likely to be complex numbers. Computationally speaking, this results in an interpolation with an apparent order $2I$.

5.2.3 Polar and Raised-Cosine Interpolations

In seeking a compromise between the polynomial and the least-squares optimal interpolations, the polar interpolation has been introduced in the context of sparse coding by E. P. Simoncelli et al. in [ETS11] and then applied to Fourier dictionaries in [FDD13]. These papers describe this interpolation in what we shall call here its "classic" formulation, intrinsically restricted to applications with a 1-D uniform grid Θ_{G} and with $I = 3$. This pattern has shown particularly good performance for translation invariant dictionaries. In short, it exploits the fact that the atoms from $\mathcal{A}$ define a manifold that lies on the sphere of radius 1 in $\mathbb{C}^M$. Therefore, the set $\mathcal{A}_n$ can be efficiently approximated by the unique arc of a circle that passes through three atoms as represented in Fig. 5.2(a). In our formalism, the interpolating function reads

$$\Psi(\delta) = \left[1,\ R\cos(\tfrac{\phi}{\Delta}\delta),\ R\sin(\tfrac{\phi}{\Delta}\delta)\right]^\top, \tag{5.18}$$

where the radius R and the angle ϕ are as shown in Fig. 5.2(c). Since the product $\boldsymbol{A}_n\boldsymbol{U}$ projects the interpolant atoms $\boldsymbol{A}_n$ onto the basis $(\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2, \hat{\boldsymbol{e}}_3)$ depicted in Fig. 5.2(c), it directly yields approximations $\tilde{\boldsymbol{a}}(\theta)$ belonging to the circle shown in this figure, given $\theta \in \Theta_n$. In practice, R and ϕ are computed with simple trigonometric identities from the angles (see Fig. 5.2(b))

$$\psi_1 := \Re\{\langle \boldsymbol{a}(\bar{\theta}_n),\, \boldsymbol{a}(\bar{\theta}_n - \Delta)\rangle\}, \tag{5.19}$$

$$\psi_2 := \Re\{\langle \boldsymbol{a}(\bar{\theta}_n + \Delta),\, \boldsymbol{a}(\bar{\theta}_n - \Delta)\rangle\}, \tag{5.20}$$

More recently F. Champagnat and C. Herzet have shown that the polar interpolation reaches the least-square optimality for dictionaries that exhibit raised-cosine-shaped kernels [CH19]. From this observation, they proposed an extension of this pattern that goes beyond the restrictions for which the initial polar interpolation applies [CH21]. In this sense, their proposition can be assimilated as a higher-order polar interpolation. In practice, this technique is similar to the Least-Squares (LS)-optimal interpolation by replacing the exact kernel κ by a raised cosine approximation

$$\tilde{\kappa}(\delta) = \lambda_0 + \textstyle\sum_{i=1}^{\lfloor \frac{I}{2} \rfloor} \lambda_i \cos(\omega_i \delta), \tag{5.21}$$

with some variables λ_i and ω_i to be determined to meet the constraint

$$\tilde{\kappa}(\bar{\theta}_{n,i} - \bar{\theta}_{n,i'}) = \kappa(\bar{\theta}_{n,i} - \bar{\theta}_{n,i'}) \tag{5.22}$$

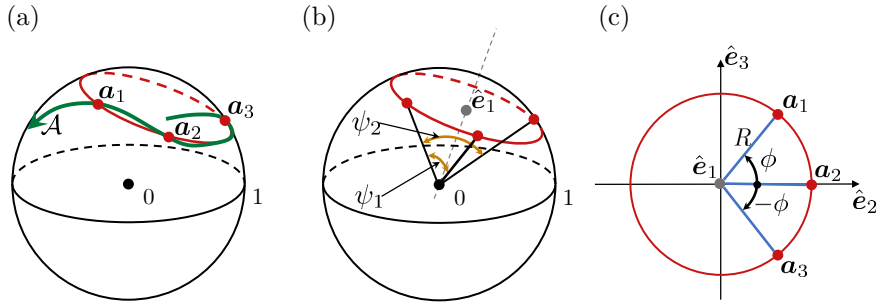

Figure 5.2: Schematic representation of the polar interpolation principle. The notation $\boldsymbol{a}_i$ is a shortcut for the interpolant atoms $\boldsymbol{a}(\bar{\theta}_{n,i})$. The manifold described by $\mathcal{A}$ is shown in (a) and is approached by the red circle whose arc between $\boldsymbol{a}_1$ and $\boldsymbol{a}_3$ represents $\mathcal{A}_n$.

for all $i, i' \in [I]$. When $I = 3$ and Θ_{G} is a 1-D uniform grid, the above is equivalent to the classic polar interpolation. All in all, this pattern offers the advantage of preserving in $\tilde{\mathcal{A}}$ both the normalized and the translation-invariant natures of $\mathcal{A}$. Moreover, the classic polar interpolation offers a simple expression for the inner maximization (5.6). This expression is provided by the following lemma, whose proof can be found in Sec. 5.8.2.

Lemma 5.2 (Polar Selection Step)**.** *Given $\boldsymbol{g} \in \mathbb{C}^I$ and the interpolating function Ψ defined as in (5.18), the optimizer of*

$$\max_{\delta \in [-\frac{\Delta}{2}, \frac{\Delta}{2}]} \tfrac{1}{2} |\langle \Psi(\delta), \boldsymbol{g} \rangle|, \tag{5.23}$$

is given by

$$\hat{\delta} = \Delta \operatorname{clip}_{[-\frac{1}{2}, \frac{1}{2}]} \left(\phi^{-1} \arccos \left(\frac{\Re\{g_2/\zeta\}}{\sqrt{\Re\{g_2/\zeta\}^2 + \Re\{g_3/\zeta\}^2}} \right) \right), \tag{5.24}$$

where $\zeta \in \mathbb{C}$ obeys

$$\zeta = g_1 + R \, \frac{\Re\{g_2/\zeta\} g_2 + \Re\{g_3/\zeta\} g_3}{\sqrt{\Re\{g_2/\zeta\}^2 + \Re\{g_3/\zeta\}^2}} \tag{5.25}$$

Remark 5.4. When $\boldsymbol{g} \in \mathbb{R}^I$, all the denominators ζ cancel out in Lem. 5.2, and hence the above particularizes into Lemma 1 from [CH19], which directly provides a closed form expression for $\hat{\delta}$. Otherwise, we evaluate $\hat{\delta}$ quickly by initializing $\zeta = g_1$ and iteratively updating it according to (5.25). Note that a complex dictionary like in radar applications typically yields $\boldsymbol{g} \in \mathbb{C}^I$.

5.2.4 Interpolation Gap

We complete our tour of the interpolation patterns with a comparison of their intrinsic performance when applied to the Fourier dictionary from Ex. 3.1. First, for illustrative purpose, Fig. 5.3 shows the objective function $|\langle \tilde{\boldsymbol{a}}(\theta), \boldsymbol{y} \rangle|^2$ resulting from the application of different patterns, in comparison with the exact squared correlation $|\langle \boldsymbol{a}(\theta), \boldsymbol{y} \rangle|^2$, where $\boldsymbol{y}$ corresponds to the measurement signal from the same example.

The intrinsic precision of each interpolation technique, relative to a given parametric dictionary $\mathcal{A}$, is formally evaluated by the *interpolation gap* explained in Def. 5.2 below. The performance of interpolation-based sparse decomposition algorithms is intrinsically related to this quantity, as studied in the next section.

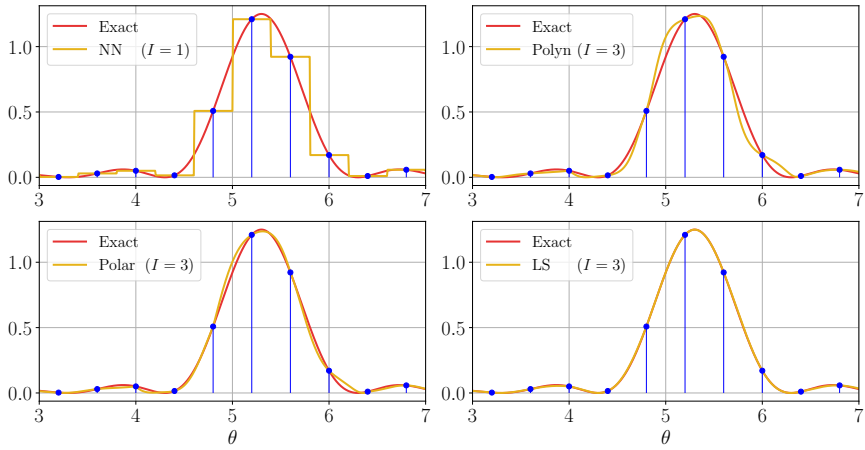

Figure 5.3: Application of multiple interpolation patterns to the selection step of the single iteration of continuous extensions of OMP, applied to the signal given in Ex. 3.1. The resulting approximated objective functions for the selection step are compared with the exact one.

Definition 5.2 (Interpolation Gap)**.** Given an interpolation strategy generating $\tilde{\mathcal{A}}$ according to Def. 5.1, the interpolation gap is the worst-case distance between any atom from $\mathcal{A}$ and its correspondence in $\tilde{\mathcal{A}}$. It writes

$$\xi := \max_{\theta \in \Theta} \|\tilde{\boldsymbol{a}}(\theta) - \boldsymbol{a}(\theta)\|_2 \tag{5.26}$$

In substance, the value of ξ is driven by three elements: (i) the nature of the interpolation, (ii) its order I, and (iii) the density of Θ_{G} that we characterize by N_Θ/M, reminding that $N_\Theta := |\Theta_{\mathrm{G}}|$. This effect is observed in Fig. 5.4, showing the gaps corresponding to all the patterns previously presented when applied to Ex. 3.1. Note that the Nearest Neighbour (NN) technique corresponds to a unique possible pattern exhibiting $I = 1$, which amounts to no interpolation by taking the closest grid component from Θ_{G}. Since the considered grid Θ_{G} is uniform, N_Θ/M is proportional to the inverse spacing Δ^{-1} . Commencing from a density sufficiently high to capture the curvature of the manifold of $\mathcal{A}$, we empirically observe that the interpolation gap decreases following the standard rule

$$\xi = h\, \Delta^I, \tag{5.27}$$

for some pattern-dependent constant h. This behavior is a well-known result for polynomial interpolations [SM03]. As expected, this constant is the smallest for the LS-optimal pattern while the polar interpolation, suited for the translation-invariant dictionary $\mathcal{A}$, exhibits better performance than the polynomial one. We can now mathematically relate the interpolation to the performance of the selection step in a simple scenario.

5.3 Bound on the Selection Process

The first contribution of this chapter is the derivation of a bound on the estimation error between the ground truth θ^* and its estimation $\hat{\theta}$ being the maximizer of the interpolated selection (5.6). We consider a noiseless measurement $\boldsymbol{y}$ that follows the model (3.1) with $K = 1$. This bound is later extended for the Grid Hopping method presented in Chapter 7.

The bound on $\|\hat{\theta} - \theta^*\|_2$ is derived using two concepts: (i) the interpolation gap and (ii) the lobe property of the *squared* kernel $|\kappa(\delta)|^2$ that we define here below.

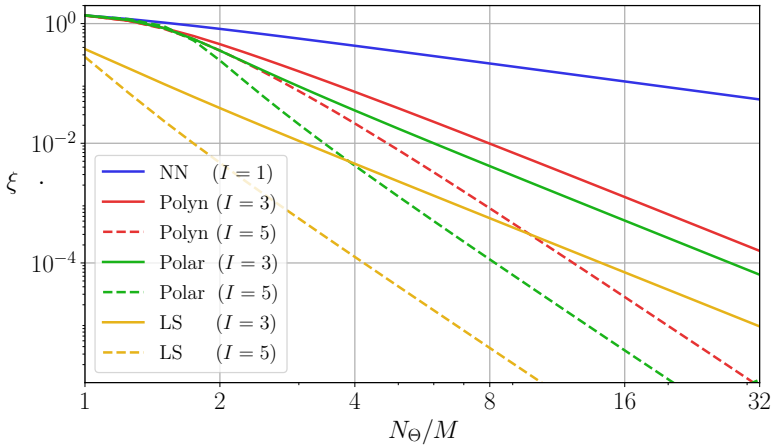

Figure 5.4: Interpolation gaps of multiple interpolation patterns applied to approximate the dictionary from Ex. 3.1, performed from a uniform grid of N_Θ bins.

Definition 5.3 (lobe-r function)**.** Given $r > 0$ and a domain $\mathcal{Z} \subset \mathbb{R}^n$, a function $f : \mathcal{Z} \mapsto \mathbb{R}$ is said to be lobe-r if there is $R \geq r$ such that for all $\boldsymbol{z}^{\text{in}} \in \mathbb{B}_2(r)$ and for all $\boldsymbol{z}^{\text{out}} \in \mathcal{Z} \backslash \mathbb{B}_\infty(R)$, we have

$$f(\boldsymbol{z}^{\text{in}}) > f(\boldsymbol{z}^{\text{out}}), \tag{5.28}$$

and if there are positive lobe-bounding constants $0 < U \leq L$ such that for all $\boldsymbol{z} \in \mathbb{B}_\infty(R)$, the following *lobe-bounding property* holds

$$1 - L\|\boldsymbol{z}\|_2^2 \leq f(\boldsymbol{z}) \leq 1 - U\|\boldsymbol{z}\|_2^2. \tag{5.29}$$

In particular, all the above implies that $f(\boldsymbol{z})$ has a unique global maximizer at $\boldsymbol{z} = \boldsymbol{0}$ with $f(\boldsymbol{0}) = 1$.

We illustrate the concept of lobe-r function with a simple example that follows up Ex. 3.1.

Example 5.5 (Féjér Kernel as a Lobe-r function). Given a Fourier dictionary whose atom expression is given in (3.4), the resulting kernel $|\kappa(\delta)|^2$ defines a Fejér kernel, which is lobe-r as depicted in Fig. 5.5.

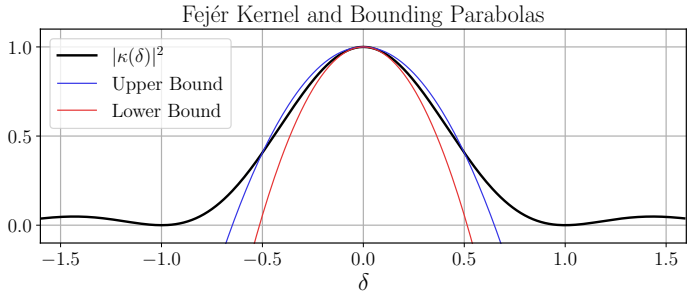

Figure 5.5: Representation of the squared kernel (zoomed in around zero) of the Fourier dictionary from Ex. 3.1 and the optimal choice, according to (5.30), of lower and upper bounding parabolas given the choice $r = .5$.

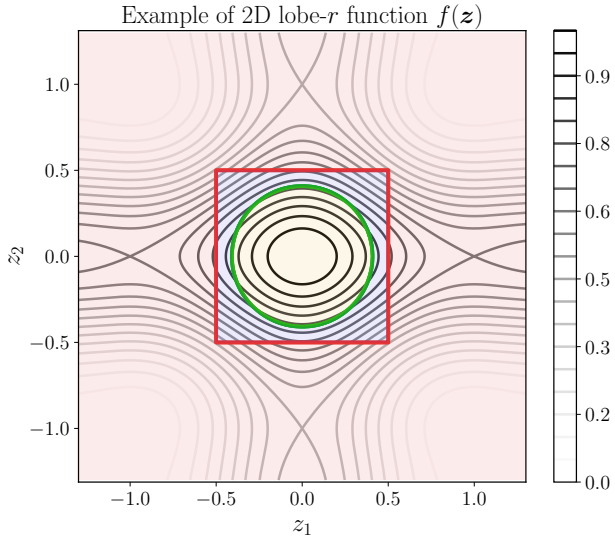

Figure 5.6: Example of 2D lobe-r function with a color mapping for the inner lobe (in green), the outer lobe (in blue), and the exterior (in red).

The features of lobe functions can be better observed from the 2D example, obtained by an outer summation of Fejér kernels, shown in Fig. 5.6 from which three distinct regions can be observed. First, the terms *Inner Lobe* and *Exterior* denote respectively the domains $\mathbb{B}_2(r)$ and $\mathcal{Z}\backslash\mathbb{B}_\infty(R)$, respectively shown in green and red in Fig. 5.6. The equation (5.28) implies

that all evaluations of f inside the inner lobe yield a higher value than all evaluations from the exterior. Practically speaking, the inner radius r of the squared kernel $|\kappa(\cdot)|^2$ in a specific application is closely related to the concept of resolution of the given sensor.

The infinite ball is used to define the limitation of the exterior because of its invariance properties in concatenations. This choice will come in handy when the bound presented here will be extended to multi-sensor domains where outer summations of lobe functions are involved.

The blue region is the *Outer Lobe*, inside which no such hierarchy can be established on the evaluations of f. Nonetheless, the lobe-r function is known to conform to the quadratic decreasing properties given by (5.29). The reader may notice that there is commonly no unique choice for the bounding constants U and L, and thereby Fig. 5.5 only shows one arbitrary choice for them. In practice, for the remainder of this chapter, given an inner radius r, we set the values of these constants following the rule

$$\begin{aligned} L &= \tfrac{1}{2}|f''(0)|, \\ U &= \tfrac{1}{r^2}(1 - f(r)). \end{aligned} \tag{5.30}$$

Under some conditions, the values for U and L given by (5.30) may be their tightest values (respectively highest and lowest admissible value). However, this aspect still needs to be proven.

Using the notion of lobe functions, we provide the bounding theorem.

Theorem 5.3 (Interpolated Selection Error Bound)**.** *Given*

1. *a translation-invariant parametric dictionary $\mathcal{A}$ such that $|\kappa(\delta)|^2$ is lobe-r with upper and lower bounding constants U and L,*
2. *an interpolation strategy yielding an interpolation gap ξ, and*
3. *the noiseless 1-sparse measurement $\boldsymbol{y} = \boldsymbol{a}(\theta^*)$,*

if

$$\xi \leq 2^{-\frac{3}{2}} U r^2, \tag{5.31}$$

then the estimate $\hat{\theta}$ optimizing (5.6) exhibits the bound

$$\|\hat{\theta} - \theta^*\|_2 \leq \tfrac{\sqrt{L}+\sqrt{L+2U}}{U}\, \xi. \tag{5.32}$$

The proof of Thm. 5.3 is provided in Sec. 5.8.

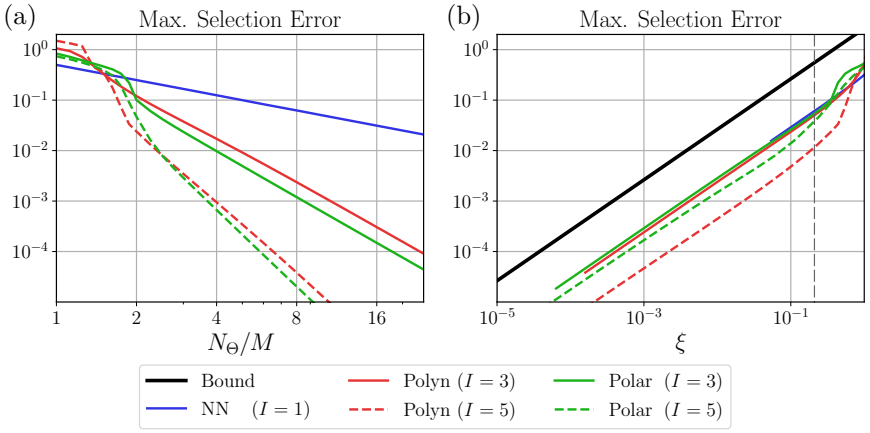

Figure 5.7: Performance of the interpolation-based selection step applied to Ex. 3.1 with multiple interpolation patterns. The results shown in (b) with respect to the interpolation gap and compared with the theoretical bound from Thm. 5.3.

We end this section by evaluating this bound's tightness through Monte-Carlo simulations. To this end, we drew 10,000 realizations of random 1-sparse noiseless measurements $\boldsymbol{y}$ with the dictionary from Ex. 3.1. The worst-case performance of the approximated selection step is represented in Fig. 5.7(a). In Fig. 5.7(b), we displayed these same results with respect to the value of the interpolation gap ξ (see Def. 5.2 and Fig. 5.4) and compared them with the bound from Thm. 5.3. We computed the bound using the bounding constants shown in Fig. 5.5. We remark that the maximum error resulting from the NN "interpolation" corresponds to the discretization error, directly due to the gridding Θ_G, since no interpolation is actually performed. The other interpolation pattern exhibits wort-case errors that properly fit the tendency given by our bound.

We emphasize that the bound could be improved by adaptively selecting the best inner radius r according to the value of ξ by activating the constraint (5.31). This would lead to a ξ-dependent value of U in (5.32). Still, when the squared kernel is closely upper bounded by a parabola, the value of U, given by (5.30), fluctuates only slightly with r (or equivalently, with ξ). Consequently, a fixed value of U is sufficient to provide a bound that captures the linear link between the worst-case selection error and the interpolation gap, as depicted in Fig. 5.7(b). Finally, we inform the reader that to avoid redundancy, we do not show the values for the average selection error here since it follows, up to some constant factor, the same evolution as the maximum error shown in Fig. 5.7.

5.4 Interpolation in Continuous OMP

While the previous sections focused on the application of atom interpolation to the decomposition of 1-sparse signals, we now return to the initial problem of representing K-sparse signals, given $K > 1$, with fast heuristics. To this end, we presented in Sec. 3.3 the Continuous OMP (COMP) in its most natural version and explained that the literature on continuous versions of OMP can be ambiguous since many distinct approaches, yet relying on different principles, have been assimilated this name [KYHP15, EGSH21, FDD13, CH19, KTTG17]. The second contribution of this chapter is thus the presentation of a continuous extension of OMP that is built upon the above-mentioned propositions to best fit the computational complexity requirements of real-time sensing applications such as radars. In substance, this section compares the Standard-COMP formulation, given by Alg. 2, with two competing alternatives for interpolation-based continuous OMP.

5.4.1 Implementations of Interpolation-based COMP

The first interpolation-based alternative was briefly exposed in the introduction of this chapter. It simply extends the Standard-COMP by replacing the exact selection step (3.14) with the approximation (5.6). One may, however, notice that solving the inner maximization of (5.6) for all $n \in [N]$ can be computationally demanding, even with a low-order interpolation. We circumvent this issue by remarking that, given a grid Θ_{G} sufficiently dense, the following approach can safely substitute the selection of $\hat{\theta}$,

$$\hat{n} = \arg\max_{n\in[N]} |\langle \boldsymbol{a}(\bar{\theta}_n),\, \boldsymbol{r}\rangle|, \tag{5.33}$$

$$\hat{\theta} = \max_{\theta\in\Theta_n} |\langle \Phi_{\hat{n}}(\theta),\, \boldsymbol{A}_n^H \boldsymbol{r}\rangle|, \tag{5.34}$$

to output the same results as (5.6). Let us emphasize that this simplification follows the same principle as implementing the selection step from Standard-COMP with a gradient descent initialized by an *on-the-grid* search, instead of many descents from all possible initializations taken from Θ_{G}. Additionally, let us point out that for interpolation patterns providing approximations of the non-uniform norm over Θ, the maximization (5.34) is replaced by the relax-and-project procedure explained in Rem. 5.1. We label "Interp-COMP" the version of COMP that modifies Alg. 2 by substituting selection step (3.14) with the approximation (5.6) as explained here above.

More robust results can be obtained if one works from the Refined-COMP instead of the Standard-COMP. In Sec. 3.3, we explained how this refined version provides better results by iteratively refining the parameter estimates using the global gradient descent (3.17). The proposition of interpolation-based COMP proposed by K. Knudson et al. in [KYHP15] can be interpreted as an extension of this Refined-COMP. Mathematically, their proposition amounts to (i) applying the interpolated selection step explained here above and (ii) approximating the global gradient descent (3.17) using the interpolation. This approximation can be expressed using $\boldsymbol{\Psi}$ to denote the space of feasible interpolating coefficients (*i.e.,* the image space of the interpolating function Ψ), up to a multiplicative complex scalar. With $\hat{\boldsymbol{\beta}} := [\hat{\boldsymbol{\beta}}_1, \cdots, \hat{\boldsymbol{\beta}}_k]$ with $\hat{\boldsymbol{\beta}}_i \in \mathbb{C}^I$ for all i, it reads

$$\hat{\boldsymbol{\beta}} \longleftarrow \underset{\boldsymbol{\beta}_i \in \boldsymbol{\Psi},\ \forall i \in [k]}{\arg\min} \ \tfrac{1}{2} \big\| \boldsymbol{y} - \textstyle\sum_{i=1}^{k} \boldsymbol{A}_{\hat{n}_i} \boldsymbol{U} \boldsymbol{\beta}_i \big\|_2^2. \tag{5.35}$$

By construction, it results that $\hat{\boldsymbol{\beta}}_i$ carries information about both $\hat{\alpha}_k$ and $\hat{\delta}_i := \hat{\theta}_i - \bar{\theta}_{\hat{n}_i}$ that can be recovered by applying subsequently

$$\hat{\delta}_i = \underset{\delta \in \Theta_{\hat{n}_i} - \bar{\theta}_{\hat{n}_i}}{\arg\max} \ \frac{|\langle \Psi(\delta),\, \hat{\boldsymbol{\beta}}_i \rangle|}{\|\Psi(\delta)\|_2} \tag{5.36}$$

and the basic closed form expression (3.15) to recover the coefficients α_i. We call "Knudson-COMP" the resulting algorithm extending the Strong-COMP by using the above instead of (3.17).

In the next section, we show by means of Monte Carlo simulation that the Knudson-COMP indeed enables to approach the performance of the Strong-COMP in a lower computation time. However, this reduction in the computation time is unsatisfying to enable its use for real-time applications such as FMCW radars. Our contribution to the topic is thus to propose an additional alternative based on Knudson's proposition. Our method, detailed in Alg. 3, widens the application of the relax-and-project idea (see Rem. 5.1) to the joint computation of matching coefficients $\hat{\boldsymbol{\beta}}$, performed from *all* the interpolant atoms by (5.38). It thus amounts to drop the "$\boldsymbol{\beta}_i \in \boldsymbol{\Psi}$" constraint from (5.35), hence leading to the step (5.38). This change still allows our algorithm to progressively tune both $\hat{\boldsymbol{\alpha}}$ and the off-grid shifts $\hat{\delta}_i$, while reducing the computational complexity of the algorithm since (5.38) comes with a well-known closed-form expression. For its connection with the Knudson-COMP, we name our proposition the "Relaxed-COMP".

Algorithm 3: Relaxed COMP

Input : $\boldsymbol{y}$, $\boldsymbol{A}$, K
Output: $\{\hat{\theta}_k\}_{k=1}^{K}$
begin
Initialization: $\boldsymbol{r} \longleftarrow \boldsymbol{y}$.
for $k = 1$ **to** K **do**

$$\hat{n}_k = \arg\max_{n\in[N]} |\langle \boldsymbol{a}(\bar{\theta}_n), \boldsymbol{r}\rangle| \tag{5.37}$$

$$\hat{\boldsymbol{\beta}} \longleftarrow \arg\min_{\boldsymbol{\beta}\in\mathbb{C}^{kI}} \tfrac{1}{2}\left\|\boldsymbol{y} - \left[\boldsymbol{A}_{\hat{n}_1}, \cdots, \boldsymbol{A}_{\hat{n}_k}\right]\boldsymbol{\beta}\right\|_2^2 \tag{5.38}$$

$$\boldsymbol{r} \longleftarrow \boldsymbol{y} - \left[\boldsymbol{A}_{\hat{n}_1}, \cdots, \boldsymbol{A}_{\hat{n}_k}\right]\hat{\boldsymbol{\beta}} \tag{5.39}$$

for $k = 1$ **to** K **do**

$$\hat{\theta}_k = \bar{\theta}_{\hat{n}_k} + \underset{\delta\in\Theta_{\hat{n}_k}-\bar{\theta}_{\hat{n}_k}}{\arg\max} \frac{|\langle \Psi(\delta), \boldsymbol{U}^\top\hat{\boldsymbol{\beta}}_k\rangle|}{\|\Psi(\delta)\|_2} \tag{5.40}$$

Alternatively, the step (5.38) can be interpreted as creating a "forbidden zone" around selected grid points whose width is determined by the order of the interpolation solely. This results in a more aggressive removal of atoms' contribution to update the residue. The algorithm is then expected to provide better performance for measurement in which the K atoms corresponding to the ground truth parameters are *moderately* correlated. Indeed, suppose these atoms are close to orthogonal. In that case, there is no gain provided by the ability to tune the *off-the-grid* shift adaptively, and hence, no difference can be observed with the first interpolation-based method. If the atoms are very close, then the aggressive removal may erase the contribution of other atoms from the ground truth before being able actually to detect them. This effect is illustrated in Fig. 5.8, showing the correlation functions that appear in the two first iterations of the Standard-COMP in (a) and of the above-relaxed method in (b) when $\boldsymbol{y}$ is a noiseless measurement obtained from a combination of two close atoms in the parametric dictionary of Ex. 3.1. Nonetheless, our proposition appears most efficient as long as we assume a minimal separation between the true parameters from which $\boldsymbol{y}$ originates, as shown by the numerical results below.

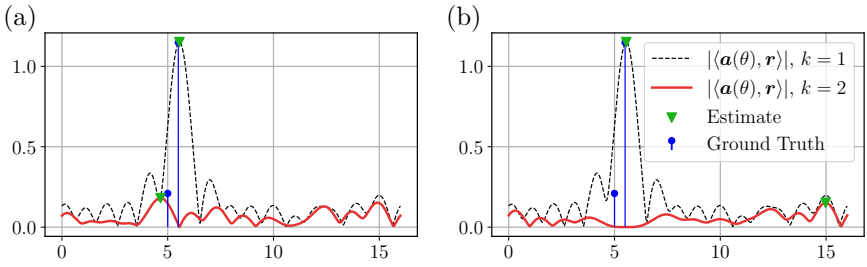

Figure 5.8: Representation of selections step's objective function of the Standard-COMP in (a) and of our relaxed proposition in (b), respectively performed during the first (dashed) and the second (solid) iterations of the algorithm.

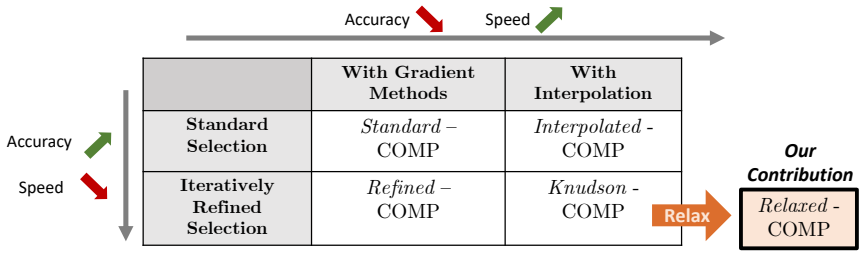

Table 5.1: Summary of the different implementations of a Continuous OMP algorithm explored in this document.

5.4.2 Numerical Results

In this section, we present the results of Monte Carlo simulations that compare the different implementations of COMP we explored in Chapter 3 and in this chapter. The five competing algorithms are summarized in Table 5.1.

We used a Fourier dictionary composed of atoms with $M = 256$ components for all the results shown in this section. We generated 50,000 random realizations of $K = 5$ true parameters $\{\theta_k^*\}_{k=1}^K$ from which we built *noiseless* measurements $\boldsymbol{y}$ according to the model (3.1) (excluding the noise term). We restricted the random selection of these parameters to a minimal separation condition given by

$$|\theta_k^* - \theta_{k'}^*| \geq e, \tag{5.41}$$

and studied the impact of the value of e on the competing algorithms.

The coefficients $\{\alpha_k^*\}_{k=1}^K$ are randomly selected from independent complex normal distributions. For each realization, each of the three algorithms provided estimates $\{\hat{\theta}_k\}_{k=1}^K$ that we sorted to minimize the average error, which reads

$$E := \frac{1}{K} \sum_{k=1}^K E_k, \tag{5.42}$$

where we define the error $E_k := |\hat{\theta}_k - \theta_k^*|$. We evaluated the performances of the competing algorithms using two metrics that convey more information than the average error given above, namely the Miss Rate (MR) and the Average Hit Error (AHE). In short, the MR counts the ratio of estimates such that $E_k > 1$ while the AHE averages the errors that were not counted as misses. In a nutshell, this enables us to distinguish the algorithms' detection and precision abilities. This aspect was stressed in Fig. 5.8. The sub-figure (a) shows a scenario where the lower spike is detected (possibly with low precision) in the second iteration of the Standard-COMP. In opposition, the sub-figure (b) shows a situation in which the second iteration of a greedy procedure fails to detect the second parameter, causing a "miss", which we formally define as an error larger than the width of the main lobe of the kernel of the parametric dictionary. This width is 1 for the Fourier dictionary of interest in that section.

We start our comparison by restricting the results to the algorithms that do not perform "refinement" of the parameter selections. More precisely, we compare first the Standard-COMP with the Interpolated-COMP. We also show the performance of the conventional discrete OMP algorithm, as a reference. We implemented the Interpolated-COMP with two interpolation patterns: a first-order polynomial ($I = 2$) and a polar interpolation ($I = 3$). The performance and the computation time characterizing the output of each of the four competing implementations are depicted by Fig. 5.9, with respect to the density of Θ_G and with a constant minimal spacing $e = 1$.

As expected we first observe that all *off-the-grid* techniques outperform the OMP, which is restricted to the grid. The performance of the Standard-COMP does not depend on the grid density since the simulations were performed only with densities sufficiently high for the grid search always to provide initialization, which fell in the basin of attraction of (3.14). Its computation time, however, decreases with the grid density since a more robust initialization leads the gradient descent procedure to require fewer iterations to reach its exit criterion, met when the Newton step is smaller than 10^{-6}. Eventually, this effect ceases, and the computation time grows due to the complexity of the grid search starting to dominate.

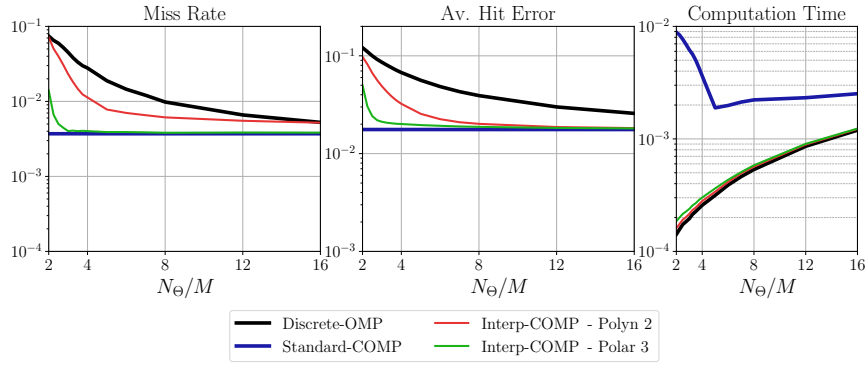

Figure 5.9: Miss Rate, Average Hit Error and computation times resulting from 10,000 realizations of noiseless signals that are exactly 5-sparse ($K = 5$) in a continuous Fourier dictionary with $M = 256$. The minimal separation e is set to 1. The results are shown with respect to the number of grid points $N_\Theta = |\Theta_G|$.

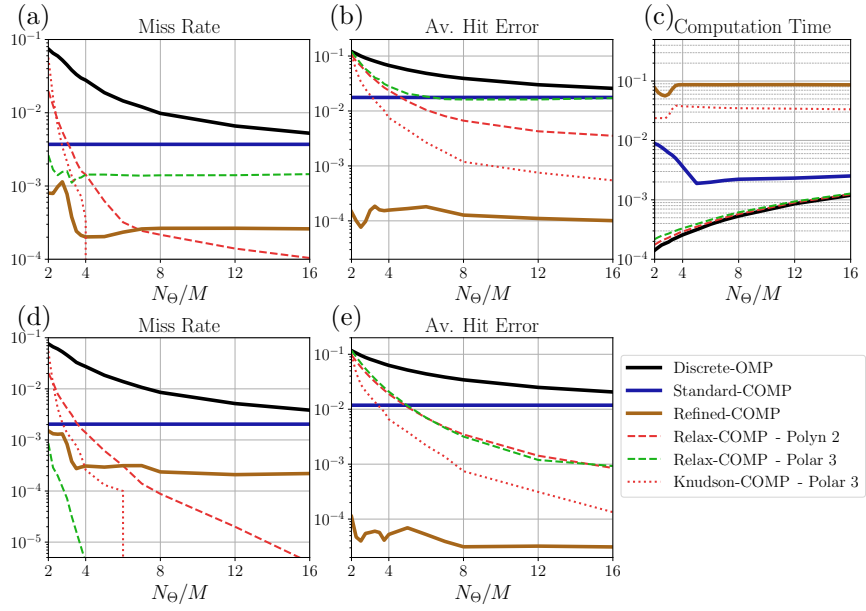

Figure 5.10: Performance of the Refined-COMP and its interpolation-based extension in the same conditions as Fig. 5.9. In (a) and (b), a minimal separation of 1 between elements of the ground truth is imposed. In (d) and (e), this separation is set to 2.

The Interpolated-COMP faithfully approximates the Standard-COMP and, thus, yields performance that asymptotically tends toward the one of this algorithm. Its computation time is close to one of the OMP because we focused on interpolation patterns endowed with the (quasi) closed-form expressions to recover off-grid shifts, as explained in Sec. 5.2 *i.e.,* the first order polynomial and the standard polar interpolations. The polar interpolation appears to be more efficient since a lower-density grid can be used to get the same performance as those obtained from the polynomial one.

We continue our analysis by adding the performance of algorithms based on the Refined-COMP. This includes the Knudson-COMP and our proposition: the Relaxed-COMP. Fig. 5.10(a)-(c) shows results obtained in similar conditions as Fig. 5.9 for the three algorithms of interest. Fig. 5.10(d)-(e) shows similar results obtained with $e = 2$ instead of one. We showed the same two interpolation patterns for the Relaxed-COMP. For the Knudson-COMP, we only displayed results for the polar interpolation for clarity purposes. The reader can refer to [KYHP15] to observe how other interpolation patterns behave with this algorithm.

As expected, these algorithms all offer better performance compared to the Standard-COMP approach. They thus also offer better performance than the Interpolated-COMP, which imitates the Standard-COMP. Increasing e from 1 to 2 naturally improves the performance of all algorithms since it provides them with an "easier" instance of the estimation problem. As already hinted, our proposition is more impacted by this parameter than the other algorithms. Indeed, our proposition provides significantly less robust results than the Knudson-COMP and the Refined-COMP for $e = 1$ but better results than these counterparts when $e = 2$. Moreover, when $e = 1$, the MR is capped only for the polar interpolation, while the lower-order polynomial pattern results in a lower error. When $e = 2$, the apparent lower cap disappears, and the MR from the polar Relaxed-COMP quickly drops near zero for a sufficiently dense grid.

The effect of e is fully explained by observing Fig. 5.11 where the same performance metrics are shown for a fixed density $N_\Theta/M = 4$ and with respect to the separation e. In this simulation, we forced each realization of ground truth to have one pair of parameters *exactly* separated by e. While applying Relaxed-COMP with the two patterns yields similar AHE and hence similar precision, the MR is strongly affected by the separation. With this figure, we can draw two conclusions. First, our method is specifically tailored to treat cases with minimal separation constraints on the

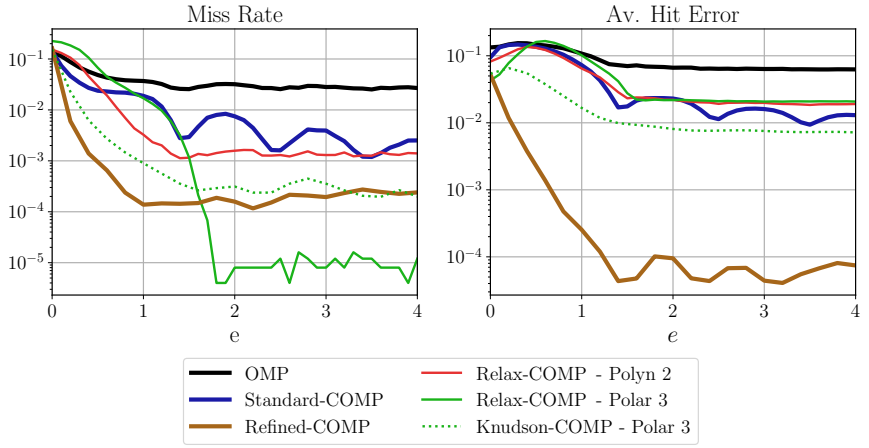

Figure 5.11: Results of a simulation similar to Fig. 5.10 presented for a fixed value $N_\Theta/M = 4$ and varying the minimal separation e.

ground truth. If one expects no such separation, then using one of the four other *off-the-grid* algorithms is preferred, with the exact choice depending on the available computational power. Second, when using our Relaxed-COMP, higher-order interpolation patterns constitute a better choice than lower-order ones only if we expect a sufficiently separated ground truth to estimate.

We end the section by comparing the application of our algorithm with all the patterns presented in Sec. 5.2. The results are shown in Fig. 5.12. We directly notice that the absence of close form estimation of the *off-the-grid* shift negatively impacts the computation times exhibited by the other patterns than the one used to obtain the previous results. Also, the patterns with order $I = 5$ have their MR being capped, indicating that they require a higher separation than the value $e = 2$ used to get those results. Finally, although the Relaxed-COMP with the LS-optimal interpolation yields the best performance regarding both the MR and the AHE, its high computation time makes it not suitable for the real-time applications which constitute the focus of this work.

In the next section, we will build upon the Relaxed-COMP to derive an extension that exploits factorization features from the measurement model. Such factorization notably appeared in the FMCW radar applications that we presented in Chapter 2.

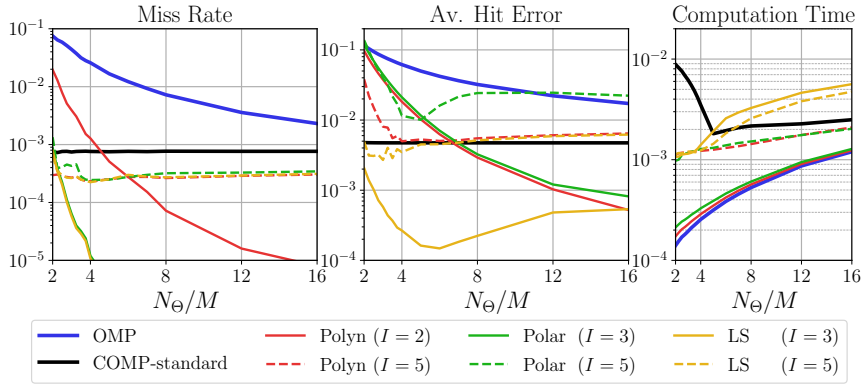

Figure 5.12: Results of a simulation similar to Fig. 5.10 focusing on more interpolation patterns used with the relaxed COMP.

5.5 Factorized COMP

The content of this section and the next one have been presented at iTwist'20 [dGVJ20] and ICASSP'21 [dGFVJ21].

We can now enter the final part of this chapter, focusing on factorizable dictionaries. In this case, greedy algorithms extending the OMP can be accelerated. For instance, in work [ZW13, FHW11], the authors consider an adaptation of OMP, that we coin Factorized OMP (F-OMP), which leverages the factorization to reduce the dimensionality of the recovery problem. Yet, F-OMP only enables the estimation of *on-the-grid* parameters from a finite discrete set of parameters. In this section, we will propose the Factorized COMP (F-COMP), which extends the Relaxed-COMP to the factorized framework used by F-OMP. Note that from now on, we drop the "Relaxed" prefix of the Relaxed-COMP and simply call it COMP.

Precisely, we consider in this section the particular case of multidimensional parameters $\theta = (\theta_1, \dots, \theta_L)^\top$ taken from the factorizable domain

$$\Theta = \Theta_1 \times \cdots \times \Theta_L, \tag{5.43}$$

and the parametric atom $\boldsymbol{a}(\theta)$ that factorizes in L sub-atoms. More precisely, introducing the tensor atom $\boldsymbol{B}(\theta) \in \mathbb{C}^{M_1 \times \cdots \times M_L}$ reshaping $\boldsymbol{a}(\theta) \in \mathbb{C}^M$ as

$$B_{m_1,\cdots,m_L}(\theta) := a_{\bar{m}}(\theta), \tag{5.44}$$

with $\bar{m} := m_L + \sum_{\ell=1}^{L-1}(m_\ell - 1)\prod_{i=\ell+1}^{L} M_i \in [M]$, $m_\ell \in [M_\ell]$ for all $\ell \in [L]$ and $M = M_1 M_2 \cdots M_L$, we assume that the atom $\boldsymbol{B}(\theta)$ decomposes in

$$\boldsymbol{B}(\theta) := \boldsymbol{b}_1(\theta_1) \otimes \cdots \otimes \boldsymbol{b}_L(\theta_L). \tag{5.45}$$

In (5.45), each $\boldsymbol{b}_\ell(\theta_\ell) \in \mathbb{C}^{M_\ell}$ is a sub-atom taken from the continuous sub-dictionary $\mathcal{B}_\ell := \{\boldsymbol{b}_\ell(\theta) : \theta \in \Theta_\ell\}$. In the tensor reshaped domain, the decomposition (3.1) becomes

$$\boldsymbol{Y} = \sum_{k=1}^{K} \alpha_k \boldsymbol{B}(\theta) + \boldsymbol{W}, \tag{5.46}$$

where $\boldsymbol{Y} \in \mathbb{C}^{M_1 \times \cdots \times M_L}$ is the tensor-shaped measurement, *i.e.,* $Y_{m_1,\cdots,m_L} := y_{\bar{m}}$.

We emphasize that the factorized model (5.46) particularizes to the simplified model of FMCW radar's signal (2.28), given $L = 2$ (one dimension for the inner signal connected to the range of the target, and the other for the outer signal connected to its speed). Still, as stated in Chapter 2, this factorization only applies to FMCW radars under a few simplifications on the model, which can cause mismatches that are studied in the next and final section of this chapter.

We now derive the Factorized COMP (F-COMP). To accelerate the Relaxed-COMP studied in the previous sections, we extend the interpolation of the atom presented in Sec. 5.1 to the above-factorized model. Practically, it directly results from (5.45) that the interpolant atoms used to derive $\tilde{\boldsymbol{a}}(\theta)$ are factorizable, regardless of the chosen interpolation pattern. We can thus apply the *factorization-over-interpolation* principle that we presented in [dGVJ20]. Partitioning Θ in a similar fashion as explained in Sec. 5.1 provides, for each n-th element of the partition, the tensor-shaped interpolation that reads

$$\tilde{\boldsymbol{B}}(\theta) := \sum_{i=1}^{I} c_i \boldsymbol{B}(\bar{\boldsymbol{\theta}}_{n,i}), \tag{5.47}$$

where the coefficients $\boldsymbol{c} := (c_1, \ldots, c_I)^\top$ are obtained from the interpolating function $\Phi_n(\theta)$.

We can now explain how the selection (5.37) and the projection (5.38) both take advantage of the factorization to reduce their respective required computation time. With $\boldsymbol{R}$ denoting the tensor-shaped residue being used in a certain iteration of the relaxed COMP, the correlation in the selection

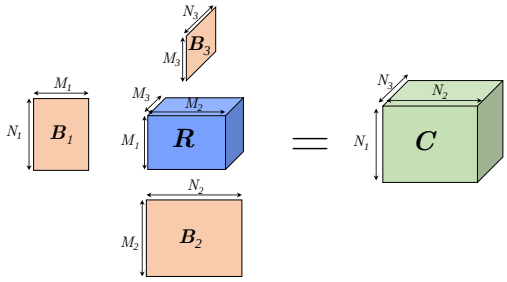

Figure 5.13: Representation of the successive mode products $\boldsymbol{R} \times_1 \boldsymbol{B}_1 \times_2 \boldsymbol{B}_2 \times_3 \boldsymbol{B}_3$

process (5.37) is equivalently computed with the sequence of tensor *mode* products

$$C(\theta) := \langle \boldsymbol{a}(\theta), \boldsymbol{r} \rangle = \boldsymbol{R} \times_1 \boldsymbol{b}_1(\theta_1) \times_2 \cdots \times_L \boldsymbol{b}_L(\theta_L). \tag{5.48}$$

Remark 5.6 (Reminder on mode products). Given the tensor $\boldsymbol{T} \in \mathbb{C}^{M_1 \times \cdots \times M_L}$ and the matrix $\boldsymbol{V} \in \mathbb{C}^{M_\ell \times N}$, the output of the ℓ-mode product $\boldsymbol{S} = \boldsymbol{T} \times_\ell \boldsymbol{V}$ is such that for all $n \in [N]$,

$$S_{m_1,\cdots,m_{\ell-1},n,m_{\ell+1},\cdots,m_L} = \sum_{m_\ell=1}^{M_\ell} T_{m_1,\cdots,m_\ell,\cdots m_L} V_{n,i}. \tag{5.49}$$

A schematic example of a series of $L = 3$ mode products with three matrices is shown in Fig. 5.13.

Assuming for simplicity $M_\star := M_1 = M_2 = \cdots = M_L$ and that the grid Θ_{G} results from the discretization of each sub-domain Θ_ℓ into $N_\star \geq M_\star$ grid bins, the computational complexity of the evaluation of (5.48) on the grid Θ_{G} is given by $\mathcal{O}(L\, M_\star\, N_\star^L)$ instead of the $(M_\star N_\star)^L$ products that would be required for the direct evaluations of (5.37) without leveraging the factorization of the atom function. When multidimensional Fourier atoms are considered, like in the simplified radar model (2.18), then the factorization above is intrinsically leveraged by the multidimensional FFT algorithm to attain the mentioned reduction of computational complexity.

Next, to leverage the factorization in the computation of $\hat{\boldsymbol{\beta}}$ performed by (5.38), we propose the following procedure, inspired by the implementation of the Factorized 2D-OMP (F-OMP) in [FHW11]. Given the in-

dex $\hat{n}_k$ resulting from the selection step in the k-th iteration of the relaxed COMP, the i-th interpolant grid component in Θ_{G} decomposes as $\bar{\theta}_{\hat{n}_k,i} = (\bar{\theta}_{\hat{n}_k,i,1}, \cdots, \bar{\theta}_{\hat{n}_k,i,L})$. With this notation, one takes advantage of the factorization to perform (5.38) faster as follows,

$$\hat{\boldsymbol{\beta}} = \boldsymbol{H}^{-1}\boldsymbol{f}, \tag{5.50}$$

where $\boldsymbol{H} = \boldsymbol{H}_1 \odot \cdots \odot \boldsymbol{H}_L \in \mathbb{C}^{Ik \times Ik}$ and $\boldsymbol{H}_\ell = \boldsymbol{\Xi}_\ell^H \boldsymbol{\Xi}_\ell$ with

$$\boldsymbol{\Xi}_\ell = \left[\boldsymbol{b}_\ell(\bar{\theta}_{\hat{n}_1,1,\ell}), \cdots, \boldsymbol{b}_\ell(\bar{\theta}_{\hat{n}_k,I,\ell})\right] \tag{5.51}$$

and $\boldsymbol{f} = (f_{1,1}, \cdots f_{1,I}, f_{2,1}, \cdots, f_{k,I}) \in \mathbb{C}^{Ik}$ where

$$f_{j,i} = C(\bar{\theta}_{\hat{n}_j,i}), \tag{5.52}$$

which is computed by (5.48).

Overall, the computational complexity of F-COMP is dominated by the selection step, to which we must add the computation of the *off-the-grid* shift. Motivated by the study of the previous section, we restrict the implementation of F-COMP to interpolation patterns provided with (quasi) closed form solution for the estimation of off-grid shifts. Although the proposition in [CH21] suggests that the polar interpolation straightforwardly extends, using their model, to the factorized framework of this section, we restrict to a first-order polynomial interpolation our numerical application of F-COMP to FMCW radars, which is shown in the next section. When the parameter domain factorizes by (5.43), then decomposing the *off-the-grid* shift as $\delta = (\delta_1, \cdots, \delta_L)^\top$, this pattern is defined by

$$\Psi(\delta) = \left[1, \delta_1, \cdots, \delta_L\right], \tag{5.53}$$

and hence $I = L + 1$. Each *off-the-grid* shift δ_ℓ can then be approximated by applying the expression (5.15) in which $\hat{\boldsymbol{\beta}}_2$ is replaced by $\hat{\boldsymbol{\beta}}_{\ell+1}$. This pattern, along with the factorization-based accelerations detailed in this section, is applied to numerical simulations of a FMCW monostatic radar in the next section.

5.6 Numerical Validation of F-COMP

To compare the effectiveness of both our relaxed version of COMP and its extension F-COMP explained above, in a practical scenario, we simulated

the signals received from an FMCW monostatic radar by generating signals given by the model (2.27) while ignoring the noise term. Formally, this model requires the simplifications explained in Sec. 2.3.1 to allow its factorization from which F-COMP is built. Therefore, we expect COMP to yield better estimation performance since we build it upon the exact model.

In practice, this section presents the results published in [dGFVJ21] in which we compared four algorithms,

1. OMP which uses the exact radar model while being restricted to the grid and acts as the reference algorithm,
2. F-OMP leveraging the factorization of the simplified radar model and thereby accelerating OMP at a cost in the resulting precision,
3. COMP which uses the exact radar model and the interpolation to leave the grid,
4. F-COMP which accelerates COMP with the factorization with a reduced performance due to the model mismatch.

The range-Doppler maps shown in Chapter 2 result from the 2D Fourier transform of the received signals and hence, amount to the correlation functions used by both F-OMP and F-COMP. By construction, we expect COMP to provide the best performance with a high computational time and F-OMP to be the fastest algorithm while yielding the most approximate results. Our approach, F-COMP, is a compromise between these two. The question addressed in this section is as follows. Under which conditions on the radar system can F-COMP significantly improve the performance of F-OMP while being significantly faster than COMP? This question originates from the observation that the precision gained by leaving the grid with the interpolation may appear as negligible with respect to the mismatch errors that we highlighted in Sec. 2.4.

The simulated radar was characterized by the quantities $B = 250\text{MHz}$, $f_0 = 24\text{GHz}$, $T_\text{s} = 5\mu\text{s}$ and $T_\text{c} = M_\text{s}T_\text{s}$, for multiple values for the respective numbers of chirps and samples per chirp, *i.e.,* M_c and M_s. For each configuration, we generated 10,000 realizations of the measurement $\boldsymbol{y}$ by picking, for each of them, $K = 5$ random values of ranges and speeds from uniform distribution in the range and speed domains, $\mathcal{R}$ and $\mathcal{U}$, respectively defined as in (2.21) and (2.22). The $K = 5$ scattering coefficients were independently drawn from complex normal distributions of unitary variance.

The performance is evaluated by means of the MR and the AHE, defined in Sec. 5.4 from the following expression for the error E_k. Given the ground

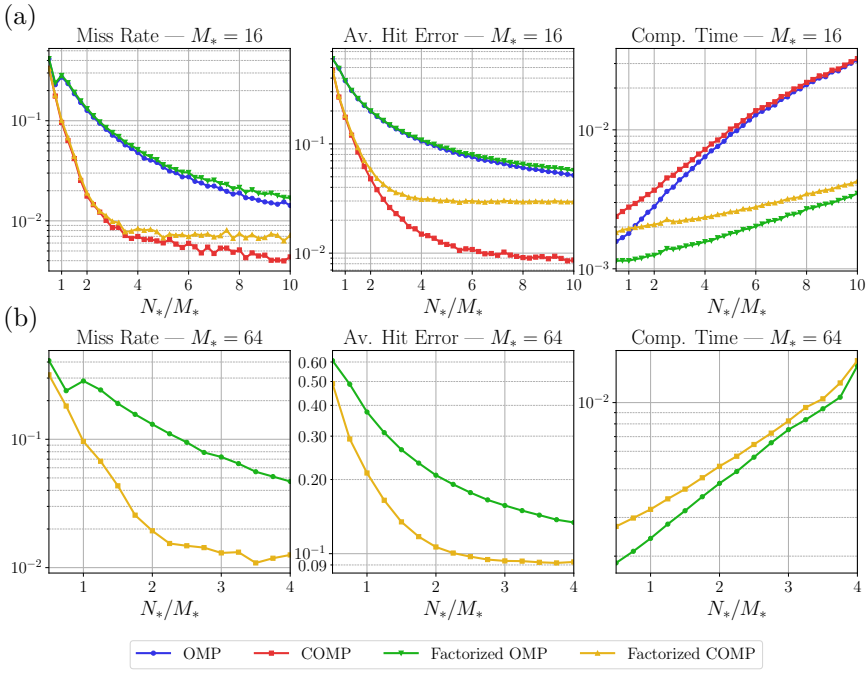

Figure 5.14: Results of the Monte-Carlo simulations of an FMCW radar with: **(a)** the comparison the four algorithms (F)-(C)OMP given $M_\star = M_\mathrm{s} = M_\mathrm{c} = 16$ and for any value of $N_\star = N_\mathrm{r} = N_\mathrm{u}$, **(b)** the comparison of the two factorized algorithms given $M_\star = 64$.

truth ranges and speeds respectively denoted by $\{r_k^*\}_{k=1}^K$ and $\{u_k^*\}_{k=1}^K$, and their respective estimates $\{\hat{r}_k\}_{k=1}^K$, $\{\hat{u}k\}_{k=1}^K$, we use for each $k \in [K]$

$$E_k = \|(\hat{r}_k, \hat{u}_k)^\top - (r_k^*, u_k^*)^\top\|_2, \tag{5.54}$$

The concatenation of the range and speed into a single vector done above makes sense since we consider the normalization explained in Def. 2.1.

In real radar applications, the “non-factorized algorithms" that exploit the exact signal model are often not practicable because of their excessively high memory and time requirements. Before comparing the two “factorized algorithms" with realistic radar parameter values, we first compared the results of the four algorithms with low numbers of acquired samples, *i.e.,* $M_\star := M_\mathrm{s} = M_\mathrm{c} = 16$. This enables separable observations of the effect of the factorization and the interpolation. From Fig. 5.14, it is clear that

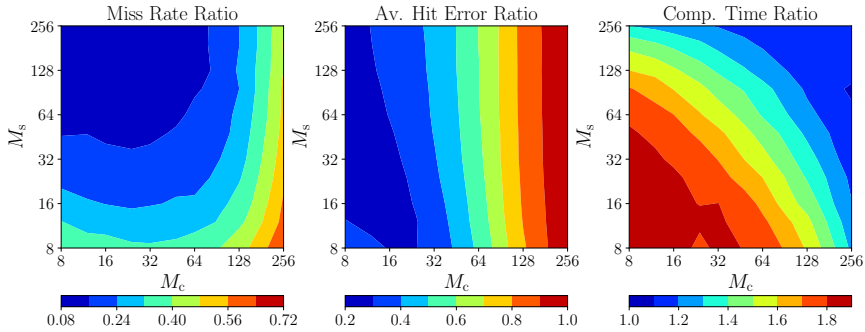

Figure 5.15: Improvement ratio when coming from F-OMP to F-COMP. For instance, the Miss Rate Ratio corresponds to $\frac{\mathrm{MR(F-COMP)}}{\mathrm{MR(F-OMP)}}$, showing to what extent F-COMP provides better performance with respect to F-OMP for many values of M_s and M_c, given $N_\mathrm{r} = 2M_\mathrm{s}$ and $N_\mathrm{u} = 2M_\mathrm{c}$ for all set.

the computation times of non-factorized algorithms grow faster with $N_\star$ than the factorized ones. The gaps in computation time between the *Continuous* algorithms and their respective *on-the-grid* counterpart correspond the computation time of the *off-the-grid* corrections. Fig. 5.14 also shows that the continuous algorithms provide better estimations than *on-the-grid* algorithms for all values of $N_\star$.

The performance loss caused by the approximation used by the factorized algorithms only appears with a dense grid (high value of $N_\star/M_\star$). Indeed, Fig. 5.14 shows that when the grid density is low, F-COMP and COMP almost have identical performance because the effect of this approximation is dominated by the grid error. Therefore, non-factorized algorithms are only advantageous when using a dense grid, in which case they are not practicable in terms of computation time and memory requirement. Regarding the factorized algorithms, F-COMP is only slower than F-OMP by a constant term (which depends on the interpolation scheme) while providing significant improvement by enabling *off-the-grid* target estimation. For this reason F-COMP appears as the best trade-off between performance and computation time for most values of MR and AHE it can reach.

Fig. 5.14 (b) presents the result from a simulated radar with the more realistic value $M_\star = 64$. It also reveals that the F-COMP curves saturate to greater (*i.e.,* worse) values of MR and AHE than the results from Fig. 5.14(a). This deterioration is caused by an increase of the model mismatch because $M_\mathrm{c}\frac{B}{f_0}$ is larger, which diminishes the validity of neglecting

the distortion to go from the exact to the factorizable radar model.

We finally analyzed the effect of M_s and M_c, through their impact on the model mismatch, on the performance gain between F-COMP and F-OMP. We varied these numbers from 8 to 256 while maintaining $N_\mathrm{r} = 2M_\mathrm{s}$ and $N_\mathrm{u} = 2M_\mathrm{c}$. Fig. 5.15 shows the ratios between the computation times and our performance metrics of these two algorithms. Since the targets are taken uniformly at random in the domain Θ whose size is $M_\mathrm{s} \times M_\mathrm{c}$, increasing either M_s of M_c increases the proportion of realizations in which the ground truth parameters are more separated with respect to the radar's resolution. Indeed, on the one hand, Θ gets larger, while on the other hand, both the speed and the range resolutions of the radar —corresponding to the width of the main lobe of the kernel associated with $\mathcal{A}$— are fixed to 1 by definition of their normalization. Therefore such an increase should lead to a particular improvement in the performance of the interpolated procedure, as dictated by the conclusion of Sec. 5.4. This improvement of MR is, however, observed only along the M_s direction in Fig. 5.15. Along the direction of M_c, this improvement is overwhelmed by the increasing model mismatch, deteriorating the MR and AHE ratios for large values of M_c. To sum up, Fig. 5.15 reveals that for all values M_s and M_c, F-COMP always exhibits better performance than F-OMP. Yet, the improvement tends to fade as $M_\mathrm{c}\frac{B}{f_0}$ increases.

5.7 Discussion

This chapter introduced the reader to the use of interpolation techniques in the context of the sparse decomposition of signals. This concept is of paramount interest for the subsequent chapters, in which we leverage the techniques presented in this chapter to accelerate so-called "single step" estimation methods for sensor network applications.

We started the chapter with a review of various popular interpolation patterns. We reformulated these interpolations with a unified notation enabling us to fully understand them by only providing their associated "interpolating function" (denoted by Ψ). This part of the chapter ended with the definition of the interpolation gap ξ, which is a useful tool to evaluate the precision of an interpolation pattern. We added our first contribution to this topic by exploiting the interpolation gap to provide a theoretical estimation bound for the interpolation-based selection step in OMP-based algorithms.

Next, the study from Sec. 5.4 compared distinct existing implementations of a Continuous OMP (COMP) exploiting the interpolation. Identifying the respective limitations of each implementation enabled us to derive our proposition coined the "Relaxed" COMP. Our proposition does not necessarily surpass existing greedy algorithms in terms of performance. However, its interest resides in its ability to reduce to near-zero the Miss Rate (MR) of the parameter estimation with low computation time when a constraint of sufficient separation in the ground truth is met. This makes it an excellent alternative for radar applications, as we emphasized in the last sections of this chapter using extensive Monte Carlo simulations.

For radar applications, we ended by developing the Factorized COMP (F-COMP), extending our "Relaxed" COMP to exploit the factorized nature of radar signals. We showed how some values of the radar system parameters affect, but do not remove, the performance gain of F-COMP over F-OMP. Future work on this topic includes the application of F-COMP to radars with strong interpolation patterns such as polar interpolation. Also, theoretical bounds on the continuous sparse decomposition with more than $K = 1$ targets could be investigated in the future, notably extending Tropp's result on OMP [Tro04].

5.8 Appendix: Proofs

5.8.1 Proof of Lemma 5.1

Since the interpolant points are distinct from one another, the rows of $\boldsymbol{U}^{-1}$ are linearly independent. It results for all $i \in [I]$ that $\boldsymbol{U}\Psi(\bar{\theta}_{n,i} - \bar{\theta}_n)$ yields a vector of zeros except in its i-th entry, which contains 1. This directly implies (5.5) and concludes the proof.

5.8.2 Proof of Lemma 5.2

Proof. This proof relies on the equivalence

$$\hat{\delta} := \underset{\delta\in[-\frac{\Delta}{2},\frac{\Delta}{2}]}{\arg\min}\ \tfrac{1}{2}|\langle\Psi(\delta),\boldsymbol{g}\rangle|^2 = \underset{\delta\in[-\frac{\Delta}{2},\frac{\Delta}{2}]}{\arg\min}\ \min_{\alpha\in\mathbb{C}}\ \tfrac{1}{2}\|\alpha\Psi(\delta)-\boldsymbol{g}\|_2^2. \tag{5.55}$$

This holds since the inner minimization of the right-hand side of the above has a closed-form expression for its minimizer given by $\alpha_\delta = \frac{\langle\Psi(\delta),\boldsymbol{g}\rangle}{\|\Psi(\delta)\|_2^2}$ whose denominator $\|\Psi(\delta)\|_2^2 = 1 + R^2$ is constant (see model (5.18)).

Next, we use the notation $\zeta := \alpha_{\hat{\delta}}$, *i.e.,* ζ is the minimizer of the inner minimization associated with the minimizer $\hat{\delta}$ of (5.55). By construction, we have

$$\hat{\delta} = \underset{\delta\in[-\frac{\Delta}{2},\frac{\Delta}{2}]}{\arg\min}\ \tfrac{1}{2}\|\zeta\Psi(\delta)-\boldsymbol{g}\|_2^2. \tag{5.56}$$

Using the polar interpolation function (5.18), it can be written as

$$\hat{\delta} = \underset{\delta\in[-\frac{\Delta}{2},\frac{\Delta}{2}]}{\arg\min}\ \tfrac{1}{2}R\big(|\zeta R\cos(\tfrac{\phi}{\Delta}\delta)-g_2|^2 + |\zeta R\sin(\tfrac{\phi}{\Delta}\delta)-g_3|^2\big) \tag{5.57}$$

$$= \underset{\delta\in[-\frac{\Delta}{2},\frac{\Delta}{2}]}{\arg\max}\ \Re\{g_2/\zeta\}\cos(\tfrac{\phi}{\Delta}\delta) + \Re\{g_3/\zeta\}\sin(\tfrac{\phi}{\Delta}\delta). \tag{5.58}$$

We can now use the lemma 1 from [CH19] on (5.58) to provide its optimizer. We thus have

$$\hat{\delta} = \Delta\ \mathrm{clip}_{[-\frac{1}{2},\frac{1}{2}]}\left(\phi^{-1}\arccos\left(\frac{\Re\{g_2/\zeta\}}{\sqrt{\Re\{g_2/\zeta\}^2+\Re\{g_3/\zeta\}^2}}\right)\right), \tag{5.59}$$

which proves the expression (5.24). Finally, we inject the above into $\zeta := \frac{\langle\Psi(\hat{\delta}),\boldsymbol{g}\rangle}{\|\Psi(\hat{\delta})\|_2^2}$. We drop the constant denominator in the expression of ζ since it

has no consequence on the expression of $\hat{\delta}$. It leads to (5.25) and concludes this proof. □

5.8.3 Proof of Theorem 5.3

Given the ground truth $\theta^* \in \Theta$, the noiseless selection step can be simplified by $\hat{\theta} = \theta^* + \hat{\delta}$ with

$$\hat{\delta} = \arg\max_{\delta \in \Theta} |\tilde{\kappa}(\delta)|^2, \tag{5.60}$$

where the kernel κ is approximated by $\tilde{\kappa}$ with the interpolation, *i.e.*,

$$\tilde{\kappa}(\delta) := \langle \tilde{\boldsymbol{a}}(\theta + \delta),\, \boldsymbol{a}(\theta) \rangle. \tag{5.61}$$

The proof of Thm. 5.3 requires the three following lemmas. The first one is a direct consequence of the definition of lobe-r functions and the two others enable us to connect the interpolation gap to the difference between the exact and the interpolation-based objective functions of the selection step. They are proven in the next subsections.

Lemma 5.4. *Given the lobe-r function f with bounding constants U and L, it holds*

$$\max_{\boldsymbol{z}} \{f(\boldsymbol{z}) : \boldsymbol{z} \in \mathcal{Z} \backslash \mathbb{B}_\infty(R)\} < 1 - Ur^2 \tag{5.62}$$

Proof. The inequality (5.29) yields $f(\boldsymbol{z}) \leq 1 - Ur^2$ for all $\boldsymbol{z} \in \mathbb{B}_2(r)$ and hence, the definition (5.28) directly provides (5.62). □

Lemma 5.5. *Given three vectors $\boldsymbol{v}_1, \boldsymbol{v}_2$ and $\boldsymbol{v}_3$ with unit ℓ_2 norm and a positive gap ε, if*

$$\|\boldsymbol{v}_1 - \boldsymbol{v}_2\|_2 \leq \varepsilon, \tag{5.63}$$

then

$$\big| |\langle \boldsymbol{v}_1,\, \boldsymbol{v}_3 \rangle|^2 - |\langle \boldsymbol{v}_2,\, \boldsymbol{v}_3 \rangle|^2 \big| \leq \sqrt{2}\varepsilon \tag{5.64}$$

Lemma 5.6. *Given $0 \leq \varepsilon \leq 1$, $0 \leq \eta \leq 1$ and three vectors $\boldsymbol{v}_1, \boldsymbol{v}_2$, and $\boldsymbol{v}_3$ with unit ℓ_2 norm and such that*

$$|\langle \boldsymbol{v}_1,\, \boldsymbol{v}_2 \rangle|^2 \geq 1 - \varepsilon^2, \tag{5.65}$$

$$|\langle \boldsymbol{v}_1, \boldsymbol{v}_3 \rangle|^2 \geq 1 - \eta^2, \tag{5.66}$$

then

$$\big| |\langle \boldsymbol{v}_1, \boldsymbol{v}_3 \rangle|^2 - |\langle \boldsymbol{v}_2, \boldsymbol{v}_3 \rangle|^2 \big| \leq 2\varepsilon\eta + \varepsilon^2 \tag{5.67}$$

We prove Thm. 5.3 in two steps. First we use Lem. 5.5 to show that the condition (5.31) is sufficient to guarantee that $\hat{\delta}$ is inside $\mathbb{B}_2(r)$, which allows us to use the lobe bounding property (5.29). Next, we use Lem. 5.6 to prove the bound (5.32). We start by noticing from Lem. 5.5 that

$$\big| |\tilde{\kappa}(\delta)|^2 - |\kappa(\delta)|^2 \big| \leq \sqrt{2}\xi. \tag{5.68}$$

Since $|\kappa(\cdot)|^2$ is a lobe-r function, combining the Lem. 5.4 with the above leads to

$$\max_{\delta} \{|\tilde{\kappa}(\delta)|^2 : \delta \in \Theta \backslash \mathbb{B}_\infty(R)\} < 1 - Ur^2 + \sqrt{2}\xi, \tag{5.69}$$

while we have $|\tilde{\kappa}(\hat{\delta})|^2 \geq 1 - \sqrt{2}\xi$. Therefore (5.31) is sufficient to have $|\tilde{\kappa}(\hat{\delta})|^2 > \max_\delta \{|\tilde{\kappa}(\delta)|^2 : \delta \in \Theta \backslash \mathbb{B}_\infty(R)\}$. This ends the first part of this proof.

In the second part, we use Lem. 5.6 with the three vectors $\boldsymbol{a}(\theta^* + \hat{\delta})$, $\tilde{\boldsymbol{a}}(\theta^* + \hat{\delta})$ and $\boldsymbol{a}(\theta^*)$. The definition of the interpolation gap implies for all $\theta \in \Theta$,

$$\xi^2 \geq \|\tilde{\boldsymbol{a}}(\theta) - \boldsymbol{a}(\theta)\|_2^2 = 2(1 - \Re\{\langle \tilde{\boldsymbol{a}}(\theta), \boldsymbol{a}(\theta) \rangle\}) \geq 2(1 - |\langle \tilde{\boldsymbol{a}}(\theta), \boldsymbol{a}(\theta) \rangle|), \tag{5.70}$$

which in turn leads to

$$|\tilde{\kappa}(0)|^2 = |\langle \boldsymbol{a}(\theta^* + \hat{\delta}), \tilde{\boldsymbol{a}}(\theta^* + \hat{\delta}) \rangle|^2 \geq 1 - \xi^2. \tag{5.71}$$

Additionally, the lobe bounding property of $|\kappa(\delta)|^2$ gives

$$|\kappa(\hat{\delta})|^2 = |\langle \boldsymbol{a}(\theta^* + \hat{\delta}), \boldsymbol{a}(\theta^*) \rangle|^2 \geq 1 - L\|\hat{\delta}\|_2^2, \tag{5.72}$$

and thus, Lem. 5.5 yields $\big| |\tilde{\kappa}(\hat{\delta})|^2 - |\kappa(\hat{\delta})|^2 \big| \leq 2\sqrt{L}\xi\|\hat{\delta}\|_2 + \xi^2$.

Finally, by definition of the maximizer $\hat{\delta}$ and the lobe-bounding prop-

erty (5.29), this inequality holds:

$$\begin{aligned} 1-\xi^2 &\leq |\tilde{\kappa}(0)|^2 \\ &\leq |\tilde{\kappa}(\hat{\delta})|^2 \\ &\leq |\kappa(\hat{\delta})|^2 + 2\sqrt{L}\xi\|\hat{\delta}\|_2 + \xi^2 \\ &\leq 1 - U\|\hat{\delta}\|_2^2 + 2\sqrt{L}\xi\|\hat{\delta}\|_2 + \xi^2. \end{aligned}$$

This can only be true if the bound (5.32) holds, and thereby, we conclude this proof.

5.8.4 Proof of Lem. 5.5

The left part of (5.64) factorizes in conjugate symmetry as

$$\left| |\langle \boldsymbol{v}_1, \boldsymbol{v}_3\rangle|^2 - |\langle \boldsymbol{v}_2, \boldsymbol{v}_3\rangle| \right|^2 \leq |\langle \boldsymbol{v}_1 - \boldsymbol{v}_2, \boldsymbol{v}_3\rangle||\langle \boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{v}_3\rangle|, \tag{5.73}$$

which directly yields (5.64).

5.8.5 Proof of Lem. 5.6

For this proof, we use the notation $\boldsymbol{V}_i = \boldsymbol{v}_i\boldsymbol{v}_i^H$ for $i = 1, 2, 3$. We can rewrite the left part of (5.67) and bound it with the Cauchy-Scharwz inequality as

$$\begin{aligned} \left| |\langle \boldsymbol{v}_1, \boldsymbol{v}_3\rangle|^2 - |\langle \boldsymbol{v}_2, \boldsymbol{v}_3\rangle|^2 \right| &= \left| \langle \boldsymbol{V}_1 - \boldsymbol{V}_2, \boldsymbol{V}_3\rangle_{\mathrm{F}} \right| && (5.74) \\ &= \left| \langle \boldsymbol{V}_1 - \boldsymbol{V}_2, \boldsymbol{V}_1 - \boldsymbol{V}_3\rangle_{\mathrm{F}} - \langle \boldsymbol{V}_1 - \boldsymbol{V}_2, \boldsymbol{V}_1\rangle_{\mathrm{F}} \right| && (5.75) \\ &\leq \|\boldsymbol{V}_1 - \boldsymbol{V}_2\|_{\mathrm{F}}\|\boldsymbol{V}_1 - \boldsymbol{V}_3\|_{\mathrm{F}} + |1 - \langle \boldsymbol{V}_1, \boldsymbol{V}_2\rangle_{\mathrm{F}}|. && (5.76) \end{aligned}$$

Since $\|\boldsymbol{V}_i - \boldsymbol{V}_j\|_{\mathrm{F}}^2 = 2 - 2|\langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle|^2$ and $\langle \boldsymbol{V}_1, \boldsymbol{V}_2\rangle_{\mathrm{F}} = |\langle \boldsymbol{v}_1, \boldsymbol{v}_2\rangle|^2$, injecting the conditions (5.65) and (5.66) in the above directly provides (5.67) and conclude the proof.

Part III

Accelerated Estimation Algorithms for Sensor Networks

Chapter 6

Two-step and Single-step Methods in Sensor Networks

MULTISTATIC radar systems, briefly introduced in Chapter 2, share multiple characteristics with other sensor networks. Notably, many algorithms developed independently by distinct research communities on specific sensor networks can be unified in a single framework and connected to the sparse decomposition of signals we studied in Part II. Part III thereby expands beyond the single sensor scenario —that was implicitly considered in Chapter 5— by studying a generic sensor network observing different transformations of a common scene. Essentially, the overall purpose of this chapter is to establish the mathematical framework that will be extensively employed in Chapter 7, where our key contribution to the topic is derived.

More specifically, this chapter first aims, in Sec. 6.1, to define a novel unified formalism that encompasses several applications, including Multistatic Radar (MSR), sound or RF source localization from Time Difference of Arrival (TDoA)s or Direction of Arrival (DoA)s. This enables us to generalize the conclusions we will draw in the whole Part III beyond the MSR context. In Sec. 6.2, we connect to our formalism a common classification of estimation algorithms designed for the multi-sensor context that emerged separately in multiple research communities. This classification gathers, on the one hand, two-step methods that are known for their simplicity of

implementation and their low complexity [Han87, FHW17, CT18, SNH09, SC19, CBH06, FRMTS17, KGP20]. On the other hand, we name single-step methods the strategies that jointly process all the received signals, hence providing more robust results [BM11, GN11, YGD13, MFJV19, YLGD21, BKY16, HPKS12, CML11, LMN+15, BD20a, MCLE13, COH18, BD22].

In Sec. 6.3, we mathematically exemplify the particularization of our general framework to a few typical applications. Notably, an emphasis on MSR systems is given in Sec. 6.3.3 where we highlight elements of a previous work [MFJV19] that concluded my Master thesis on factorization strategies for the single-step estimation in the specific context of FMCW MSR. That section extends it with a novel theoretical bound on the estimation error caused by the factorized estimation process. Finally, the single-step methods have multiple times been shown to provide more robust estimation results than their two-step counterpart when the signal transmission is submitted to high noise [BKY16] or is subjected to reverberation [DSB01, P.03, ZFZ08], for given application fields. With this chapter, we bring more evidence in this direction by showing how those conclusions still hold for FMCW MSR systems whose signal model follows the expressions from Chapter 2. This is shown through Monte-Carlo simulations in which noise is the only nuisance.

All the proofs of the theorems and lemmas given in this chapter are gathered in the last section of this chapter.

6.1 The Multi-sensor Sensing Model

This chapter generalizes the estimation problem originating from the multi-static radar application. In substance, we consider the abstract problem of estimating a *parameter of interest* denoted by $\boldsymbol{p}^*$ from the signals received by Q distinct sensors. We specifically consider the case of multi-sensor systems in which each sensor taken individually is only sensitive to a certain projection of $\boldsymbol{p}^*$ onto a low-dimensional observation space, from which $\boldsymbol{p}^*$ typically cannot be retrieved unambiguously. For instance, any sensor measuring transmission delays observes a projection of a 2D (or 3D) target location onto a 1D domain of distances between the target and the sensor. Practically, this framework encapsulates any application that could require triangulation, multilateration, or hyperbolas intersection for a given estimation context. Examples of applications in connection with existing estimation problems from the literature and for which our study applies are detailed in Sec. 6.3.

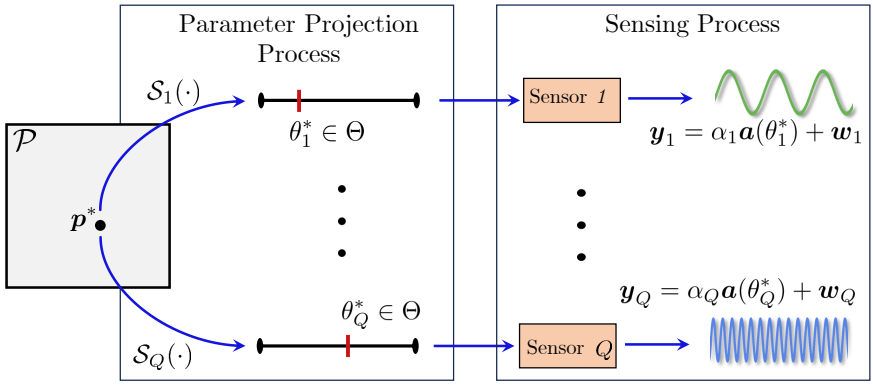

Figure 6.1: Schematic representation of the two processes that characterize the acquisition model of a sensor network. Each q-th sensor observes a signal $\boldsymbol{y}_q$ which is parameterized by the sensed parameter $\boldsymbol{\theta}_q$. The q-th sensed parameter is a projection parameter of interest $\boldsymbol{p}$ from the domain of interest $\mathcal{P}$ onto the sensing domain Θ.

Our generalized mathematical model for the targeted sensor networks interprets the measurement as a composition of two processes: *(i)* the parameter projection process, which projects the parameter of interest into the sensing space, and *(ii)* the sensing process that provides the received signals as functions of the sensed parameters. These processes are represented in Fig. 6.1, using the notations that are defined hereafter.

6.1.1 Parameter Projection Model

As depicted by Fig. 6.1, each sensor, indexed by $q \in [Q]$, may observe the parameter of interest $\boldsymbol{p}^*$ only through a specific parameter projection operator $\mathcal{S}_q$. Assuming that $\boldsymbol{p}^*$ belongs to a domain of interest $\mathcal{P} \subset \mathbb{R}^d$, this function maps from $\mathcal{P}$ onto the sensing domain $\Theta \subset \mathbb{R}^{\tilde{d}}$.

Formally, we denote by θ_q^* the q-th *sensed parameter* related to $\boldsymbol{p}^*$ through

$$\theta_q^* = \mathcal{S}_q(\boldsymbol{p}^*). \tag{6.1}$$

While $\boldsymbol{p}^*$ has a physical meaning of interest, θ_q^* may either have another physical meaning of lesser interest or none at all. Fig. 6.2 represents simple examples of Parameter Projection Function (PPF)s where $\boldsymbol{p}^*$ is a location and θ_q^* is a projection of this location on a reduced domain such as the 2D photography of a 3D space or a distance between $\boldsymbol{p}^*$ and a location of reference.

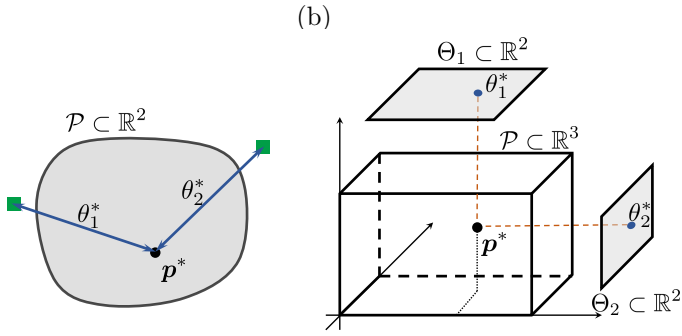

Figure 6.2: Representation of two examples of PPFs $\mathcal{S}(\cdot)$ where the parameter of interest is a location. (a) The sensed parameter is a distance between $\boldsymbol{p}^*$ and a reference location. (b) The sensed parameter is a 2D location in a plane corresponding to a photography of the 3D domain of interest.

The definition of the PPF comes with a few crucial assumptions.

Assumption 6.1 (Assumption on the Parameter Projection Function)**.**

1. Each *local* projection function $\mathcal{S}_q$ taken alone can be non-injective.
2. The *global* projection function

$$\mathcal{S}(\boldsymbol{p}) := \big(\mathcal{S}_1(\boldsymbol{p}), \cdots, \mathcal{S}_Q(\boldsymbol{p})\big)^\top. \tag{6.2}$$

   is injective over $\mathcal{P}$. Equivalently, $\mathcal{S}$ is a bijection between $\mathcal{P}$ and the *admissible sensing domain*

$$\mathcal{T} := \mathcal{S}(\mathcal{P}) := \{\boldsymbol{\theta} = \mathcal{S}(\boldsymbol{p}) : \boldsymbol{p} \in \mathcal{P}\}. \tag{6.3}$$

3. For all $q \in [Q]$, $\mathcal{S}_q$ is differentiable over $\mathcal{P}$.
4. The domain of interest $\mathcal{P}$ and the sensing domain Θ are bounded. Thus $\mathcal{T}$ is also bounded.

Additionally, for simplicity, we assume that the sensing domain Θ is identical for all index q. We also set $\tilde{d} = 1$ for the remaining of this dissertation. Still, the approaches studied in this chapter and the next one can be extended beyond these two restrictions.

The assumptions 6.1.1 and 6.1.2 translate the contrast between the potential inability of one sensor alone to recover $\boldsymbol{p}^*$ unambiguously but the ability of the whole network to do so. By studying the jacobian of $\mathcal{S}$ using the definition below, we can evaluate its ability to provide outputs that enable us to efficiently discriminate between input parameters of interest. This definition will be used for deriving theoretical error bounds later in this chapter and the next.

Definition 6.2 (Variations of parameter projection)**.** Given the global projection function $\mathcal{S} : \mathcal{P} \mapsto \mathcal{T}$, we respectively define the maximal and the minimal variation coefficients by

$$\lambda_{\min} := \min_{\boldsymbol{p}\in\mathcal{P}} \ \lambda_{\min}\big(\boldsymbol{J}(\boldsymbol{p})^\top \boldsymbol{J}(\boldsymbol{p})\big) \tag{6.4}$$

$$\lambda_{\max} := \max_{\boldsymbol{p}\in\mathcal{P}} \ \lambda_{\max}\big(\boldsymbol{J}(\boldsymbol{p})^\top \boldsymbol{J}(\boldsymbol{p})\big) \tag{6.5}$$

where $\boldsymbol{J}(\boldsymbol{p})$ denotes the jacobian matrix of $\mathcal{S}$ evaluated in $\boldsymbol{p}$.

The quantities $\lambda_{\min}$ and $\lambda_{\max}$ carry information about the ability of the global parameter projection function $\mathcal{S}$ to observe $\mathcal{P}$ unambiguously. Said differently, they relate the inherent *geometry* of the sensing problem for a given domain of interest $\mathcal{P}$ to the "injectivness" of the parameter projection. We emphasize this aspect with the following lemma, whose proof is given in Sec. 6.7.1.

Lemma 6.1. *Given a global PPF denoted $\mathcal{S}$ differentiable according to the assumption 6.1.3 and a* convex *domain $\mathcal{P}$ meeting As. 6.1.4. If $\mathcal{S}$ does not meet As. 6.1.2, meaning that is it non injective over $\mathcal{P}$, then*

$$\lambda_{\min} = 0,$$

where $\lambda_{\min}$ is defined according to Def. 6.2.

With this lemma, we notice that the ill-posed situation of a non-injective global projection function yields $\lambda_{\min} = 0$. Conversely, the ideal and usually unreachable geometry implies $\lambda_{\min} = \lambda_{\max}$.

6.1.2 Sensing Model

Our sensing scenario considers that each q-th sensor records a signal directly parameterized by the value of the sensed parameter θ_q^*. In all generality, the sensing process is directly explained by the atom function $\boldsymbol{a}$ and the parametric dictionary it defines, according to Def. 3.1.

If the network observes a scene from which a single parameter of interest needs to be estimated (*e.g.,* a multistatic radar observing a single moving target, or a single source of sound captured by a network of microphones), then up to an additive noise, the q-th recorded signal $\boldsymbol{y}_q$ is 1-sparse in the *continuous* dictionary $\mathcal{A}$. Mathematically,

$$\boldsymbol{y}_q = \alpha_q \boldsymbol{a}(\theta_q^*) + \boldsymbol{w}_q. \tag{6.6}$$

In (6.6), the complex coefficient $\alpha_q \in \mathbb{C}$ models, in all generality, an unknown amplitude and phase. The term $\boldsymbol{w}_q$ is an additive measurement noise assumed white and Gaussian.

It is important to note that our model assumes a sensor network observing the common scene in an *incoherent* manner. This concept is related to the prior knowledge on the coefficients $\{\alpha_q\}_{q=1}^Q$ only. When the phase shifts they cause are not only unknown but also independent from one another, the measurements are said to be incoherent. This has an impact on the design and the performance of the single-step method that we describe in Sec. 6.2. Indeed, in applications where coherent measurements are obtained (*e.g.,* in which all α_q are guaranteed to be a positive real number), we can leverage this additional information to develop more robust single-step algorithms. Nonetheless, any algorithm designed for incoherent measurement is suited (yet suboptimal) to process signals originating from a coherent sensing process. The opposite is not true. This distinction between coherent and incoherent single-step methods is notably examined in [BM11]. We will briefly revisit this discussion when presenting implementations of single-step methods in Sec. 6.2.

For simplicity of notations, we assume that the atom function $\boldsymbol{a}$ is identical for all index $q \in [Q]$. Practically, this implies that we consider a network composed of homogeneous sensors. For example, the multistatic radar presented in Sec. 6.3.3 must be composed of radars operating on identical modulation and acquisition settings. Similarly, the arrays of sub-receivers composing the DoA-based localization system described in Sec. 6.3.2 must all have identical spatial geometry of sub-receivers. While all the conclusions

drawn in this chapter and the next one also extend beyond this restriction, the performance bounds derived in those two chapters are built upon it. Considering the injectivity assumptions of the projection operator $\mathcal{S}$, the assumption 3.2 on the atom function guarantees the ability of the network to recover $\boldsymbol{p}^*$ unambiguously.

Finally, assuming a more general model where multiple parameters of interest are simultaneously observed, the sensor networks are assumed to record a linear combination of waveforms, each associated with one parameter of interest. This is often the case for localization systems when multiple targets have to be located simultaneously, as highlighted in Sec. 6.3. Given K parameters of interest denoted by $\boldsymbol{p}_1^*, \ldots, \boldsymbol{p}_K^*$ to be estimated, the multi-parameters model addressed by our work is

$$\boldsymbol{y}_q = \sum_{k=1}^{K} \alpha_{q,k} \boldsymbol{a}\big(\mathcal{S}_q(\boldsymbol{p}_k^*)\big) + \boldsymbol{w}_q, \tag{6.7}$$

where the coefficient $\alpha_{q,k}$ depends on both k and q which respectively index the observed parameter of interest and the sensor. Let us point out that in the context of the model (6.7), the concept of coherent or incoherent measurement refers to the dependence or not of the phase shifts in the coefficients $\alpha_{q,k}$ only along the index q. By definition of the sensing scenario, the independence along the index k is guaranteed since it corresponds to distinct elements of the observed scene.

6.2 Two-step and Single-step Methods

Starting from the above general signal model for sensor networks, we can now formalize the two conventional classes of techniques, namely the single-step and the two-step methods, to recover the parameters of interest from the measurements $\{\boldsymbol{y}_q\}_{q=1}^{Q} \subset \mathbb{C}^M$. In brief, on the one hand, the two-step methods separately estimate intermediate parameters in each sensor before combining them. On the other hand, the single-step methods jointly process all the received signals.

Two-step methods: The two-step framework is schematically represented by Fig. 6.3. In short, each signal $\boldsymbol{y}_q$ is first separately processed by working in the sensed domain Θ solely, thereby ignoring the multi-sensor aspect of the system. For each receiver, the algorithm behind this first step thus

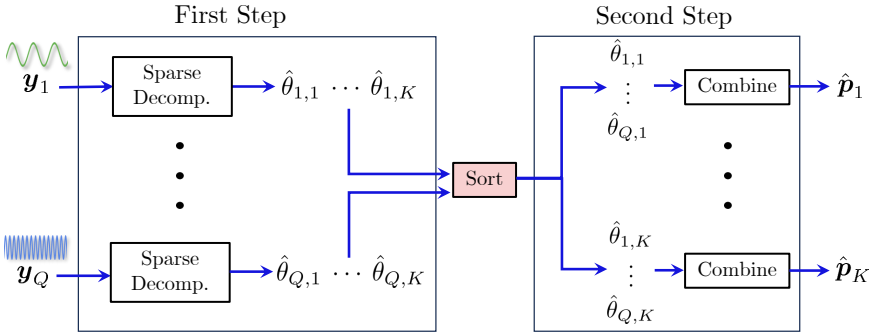

Figure 6.3: Schematic representation of the two-step procedure.

corresponds to the decomposition of $\boldsymbol{y}_q$ onto a sparse representation in the dictionary $\mathcal{A}$, parameterized over Θ. In this research, we consider for this task the application of OMP or any of its *off-the-grid* extensions that we explored in Chapter 5. In the second step, for each k-th parameter of interest, a group of Q intermediate estimates is combined to output the estimate $\hat{\boldsymbol{p}}_k$.

Let us emphasize that in the estimation context of multiple parameters of interest, any two-step method requires an intermediate "sorting" task, represented in Fig. 6.3. This challenging task is sometimes named the "data association problem" [BP07, ABM15, WL15]. In short, the sensed parameters provided by each sensor need to be correctly associated with one another for the second step. In [AM18], the authors propose to solve this with feature vectors based on the structure of speech signals. In other applications, the problem is addressed through tracking [GTK13, JFM16, LJT+20].

For the sake of simplicity, we mathematically provide the expression corresponding to both steps when restricted to the estimation of a single parameter $\boldsymbol{p}^*$. We remind that, in this simple situation, OMP is reduced to its atom selection step. Its expression in (3.14) is here re-expressed by

$$\hat{\theta}_q = \arg\max_{\theta \in \Theta} |C_q(\theta)|^2, \quad q \in [Q], \tag{6.8}$$

where

$$C_q(\theta) := \langle \boldsymbol{y}_q, \boldsymbol{a}(\theta) \rangle. \tag{6.9}$$

As stated earlier, the above can be solved quickly by reducing the problem to a grid search on Θ_{G} or by leveraging the interpolation to leave the grid.

The second step estimates $\hat{\boldsymbol{p}}$ from the non-linear least squares minimization:

$$\hat{\boldsymbol{p}} = \underset{\boldsymbol{p}\in\mathcal{P}}{\arg\min} \, \|\hat{\boldsymbol{\theta}} - \mathcal{S}(\boldsymbol{p})\|_2. \tag{6.10}$$

This second step is referred to as *multilateration* for ellipses or hyperbolas intersection [KGP20, FHW17], and *triangulation* for directions intersection [Han87].

Single-step methods: In the single-step algorithms, all the signals are processed together to directly output estimates of the parameters of interest. Given a single parameter $\boldsymbol{p}^*$ to estimate, the method is reduced to

$$\hat{\boldsymbol{p}} = \underset{\boldsymbol{p}\in\mathcal{P}}{\arg\max} \, \textstyle\sum_{q=1}^{Q} |C_q(\mathcal{S}_q(\boldsymbol{p}))|^2. \tag{6.11}$$

In this single step, one thus looks for a value of $\hat{\boldsymbol{p}} \in \mathcal{P}$ directly maximizing the non-coherent sum of all measurement correlations with $\mathcal{S}_q(\boldsymbol{p})$. Although the interpolation-based strategies studied in the last chapter can be extended to efficiently solve (6.11) in an *off-the-grid* fashion, we postpone this study to Chapter 7. We assume for this chapter that problem (6.11) is either approximately solved by a grid search over $\mathcal{P}_{\mathrm{G}}$, uniformly discretizing $\mathcal{P}$ into N_P grid bins, or refined by means of a gradient ascent initialized by the *on-the-grid* estimate.

Remark 6.1 (Connection with Maximum Likelihood Estimation). As derived in multiple work on single-step methods [P.03, YLGD21, NBVD12, BKY16], considering additive white Gaussian noise vectors independent from one sensor to another in the model (6.6), the single step above amounts to a maximum likelihood estimator. This results from the log-likelihood function, which reads up to an additive constant,

$$\log \mathrm{P}(\{\boldsymbol{y}_q\}_{q=1}^{Q} \,|\, \boldsymbol{p}, \boldsymbol{\alpha}) := -\sum_{q=1}^{Q} \|\boldsymbol{y}_q - \alpha_q \boldsymbol{a}(\mathcal{S}_q(\boldsymbol{p}))\|_2^2. \tag{6.12}$$

Since all α_q are independent from one another and from $\boldsymbol{p}$, and since the atom function $\boldsymbol{a}$ is normalized, maximizing (6.12) is equivalent to (6.11).

Algorithm 4: Block OMP in the Multi-sensor Context

Input : $\boldsymbol{y}_q$, K
Output: $\hat{\boldsymbol{p}}_1, \cdots, \hat{\boldsymbol{p}}_K$
begin
Initialization: $\boldsymbol{r}_q \longleftarrow \boldsymbol{y}_q$ for $q \in [Q]$;
for $k = 1$ **to** K **do**

1. Selection Step

$$\hat{\boldsymbol{p}}_k = \arg\max_{\boldsymbol{p}\in\mathcal{P}} \sum_{q=1}^{Q} \left|\left\langle \boldsymbol{a}\big(\mathcal{S}_q(\boldsymbol{p})\big),\, \boldsymbol{r}_q \right\rangle\right|^2 \tag{6.14}$$

2. Fitting Step

$$\hat{\boldsymbol{\alpha}}_q \longleftarrow \arg\min_{\boldsymbol{\alpha}\in\mathbb{C}^k} \tfrac{1}{2}\left\|\boldsymbol{y}_q - \boldsymbol{B}_q\,\boldsymbol{\alpha}\right\|_2^2, \quad \forall q \in [Q] \tag{6.15}$$

where $\boldsymbol{B}_q := \big[\boldsymbol{a}\big(\mathcal{S}_q(\hat{\boldsymbol{p}}_1)\big), \cdots, \boldsymbol{a}\big(\mathcal{S}_q(\hat{\boldsymbol{p}}_k)\big)\big]$

3. Residue Update

$$\boldsymbol{r}_q \longleftarrow \boldsymbol{y}_q - \boldsymbol{B}_q\,\hat{\boldsymbol{\alpha}}, \quad \forall q \in [Q] \tag{6.16}$$

Following up on the comments from the previous section, We also emphasize that given a coherent set of measurements, the single-step procedure given above can be replaced by

$$\hat{\boldsymbol{p}}_{\text{coherent}} = \arg\max_{\boldsymbol{p}\in\mathcal{P}} \left| \sum_{q=1}^{Q} \zeta_q C_q(\mathcal{S}_q(\boldsymbol{p})) \right|^2, \tag{6.13}$$

with $\zeta_q \in \mathbb{C}$ denoting known or estimated phase coefficients enabling the coherent summation. In a nutshell, the coherent single-step estimation concatenates all the signals $\{\boldsymbol{y}_q\}_{q=1}^{Q}$ into one vector and does the same for the atoms $\{\boldsymbol{a}(\mathcal{S}_q(\boldsymbol{p}))\}_{q=1}^{Q}$ and thus proceeds like a single sensor estimation in the domain of interest $\mathcal{P}$.

When multiple parameters of interest $\boldsymbol{p}_1^*, \ldots, \boldsymbol{p}_K^*$ must be jointly estimated from the Q measurements, the single step (6.11) straightforwardly provides the selection step of an extension of OMP, commonly referred to as Block Orthogonal Matching Pursuit (BOMP) [GN11, YGD13, SCWL15,

LP15]. This algorithm, detailed in Alg. 4, specifically corresponds to the application of a proposition of Y. C. Eldar [EKB10] to the sensor network framework of this chapter. In a nutshell, Block Orthogonal Matching Pursuit (BOMP) leverages the prior knowledge that all the signals $\{\boldsymbol{y}_q\}_{q=1}^{Q}$ exhibit sparse representations with identical support in the domain of interest $\mathcal{P}$. Thereby, it greedily searches for the best collection of atoms that jointly fits all the received signals while corresponding to a single value $\boldsymbol{p}$, viewed through the projection operators $\mathcal{S}_q$ associated with each received signal. Again, in the case of coherent measurements, BOMP can be adapted by replacing the selection step with a similar step as (6.13).

6.3 Examples of Applications

In this section, we overview a few *standard* examples of applications for which the model and the two classes of methods described above apply. For each application, we describe the contextual meaning of four quantities, *(i)* the number Q of acquired signals, *(ii)* the length M of the measurements, *(iii)* the sensed parameters $\{\theta_q^*\}_{q=1}^{Q}$, and *(iv)* the parameter of interest $\boldsymbol{p}^*$. Moreover, we provide the parametric expression of the atom $\boldsymbol{a}(\theta) = (a_1(\theta), \cdots, a_M(\theta))^\top$ and the expression of the PPF $\mathcal{S}_q(\boldsymbol{p})$ which links the parameter of interest with the q-th sensed parameter. For the sake of simplicity, we express the most basic form of those applications to show how they relate to our model. We also provide references for more details and advanced configurations.

We begin with the source localization by means of Time Difference of Arrival (TDoA) before showing the same kind of application based on Direction of Arrival (DoA). As a third example, we will show how the FMCW MSR, introduced in Chapter 2, fits our general model. The reader may notice that those applications are not mutually exclusive since, for example, there exist radars based on the estimation of TDoA [KGBR21, YGL+19]. Also, some source localization systems are able to perform both the TDoA and the DoA estimation. Additionally, radars with antenna arrays on each receiving node (such as our system presented in Sec. 2.5) are sensitive to DoA on top of the bistatic ranges and speeds related to each observed target. All these examples can be explained by appropriately adapting the projection function $\mathcal{S}$ and the atom function $\boldsymbol{a}$. The simplified standard applications reviewed in this section thus provide the essential building block for the design of these extensions.

Application	Q	M	$\boldsymbol{p}^*$
Localization with MSR	RX antennas × number of chirps	samples per chirp	target's location
Velocity with MSR	RX antennas	chirps	target's velocity
Source Loc. from TDoA	pairs of RXs	Fourier bins used in GCC	source location
Source Loc. from DoA	sensor arrays	sub-receivers per array	source location

Application	parameter projection $\mathcal{S}_q(\boldsymbol{p}) =$	atom $a_m(\theta) =$
Localization with MSR	$\frac{1}{2}(\|\boldsymbol{p}-\boldsymbol{x}^{\mathrm{TX}}\|_2 + \|\boldsymbol{p}-\boldsymbol{x}_q^{\mathrm{RX}}\|_2)$	$\exp(\mathrm{j}2\pi\theta m/M)$
Velocity with MSR	$\frac{1}{2}\langle \boldsymbol{h}^{\mathrm{TX}} + \boldsymbol{h}_q^{\mathrm{RX}}, \boldsymbol{p}\rangle$	$\exp(\mathrm{j}2\pi\theta m/M)$
Source Loc. from TDoA	$\frac{1}{cT_s}(\|\boldsymbol{p}-\boldsymbol{x}_q\|_2 - \|\boldsymbol{p}-\boldsymbol{x}'_q\|_2)$	$\exp(\mathrm{j}2\pi\theta m/M)$
Source Loc. from DoA	$\angle(\boldsymbol{p}-\boldsymbol{x}_q^{\mathrm{RX}})$	$\exp(\mathrm{j}2\pi\lambda^{-1}\,\boldsymbol{h}_m^\top \boldsymbol{e}_\theta)$

Table 6.1: Summary of the applications presented in Sec. 6.3. The meaning of each quantity in the model (6.6) is detailed for each application.

Furthermore, depending on the practical context of each application, the Q measurements can either be coherent or incoherent [BM11]. As explained in the previous sections, this aspect do not alter the general model (6.7). The single-step method must simply be designed in accordance with the nature of the measurement. Let us note that some applications perform hybrid coherent-incoherent measurements. For example, our MSR system with antenna arrays on each receiving node yields coherent measurement between receivers inside the same array but incoherent from one node array to another. This aspect will be taken into account in the next chapter, in which the measurement results are extensively exploited to demonstrate the effectiveness of our method "Grid Hopping".

The instantiation of (6.6) to the presented standard applications are summarized in Table 6.1

6.3.1 Source Localization with Time Difference of Arrival

The model provided in Sec. 6.1 first applies to systems focusing on the estimation of the location of the source of a signal, which thereby acts as the parameter of interest $\boldsymbol{p}^*$. In most applications, the receivers are not –or

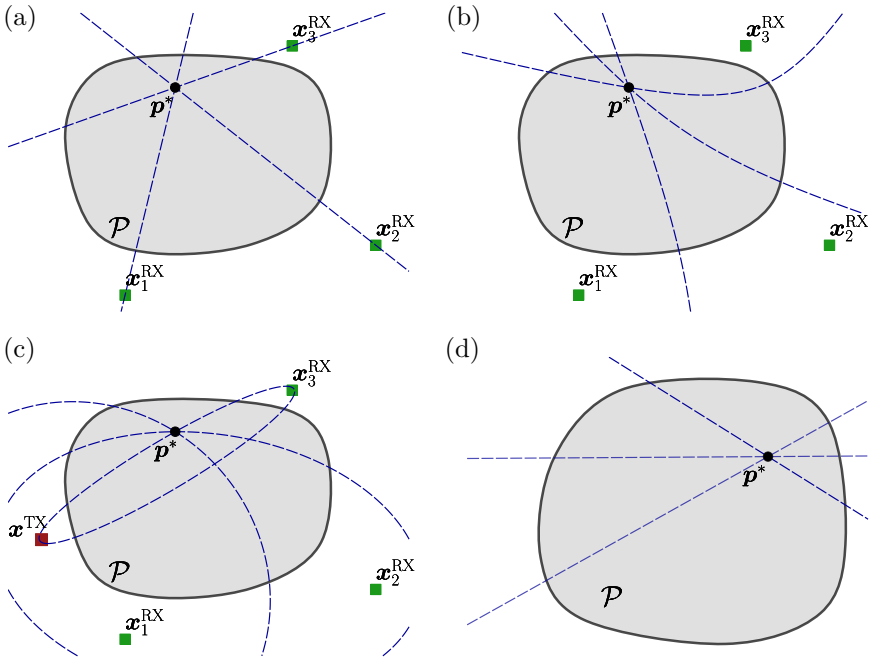

Figure 6.4: Representations of the three standard application that enters into the scope of this chapter. The dashed lines represent the locus of identical sensed parameters in the domain of interest. **(a)** A source localization system where $\boldsymbol{p}$ is the source's location to estimate and composed of four RX. The locus identical sensed parameters are hyperbolas. **(b)** A source localization system composed of four receiving nodes, each composed of an array of L receivers. **(c-d)** A multistatic radar composed of one TX and three RX. Each RX provides one measurement such that $Q = 3$. For the location estimation, the locus of identical sensed parameters is given by an ellipse whose focuses are the locations of the TX and the given RX. For the velocity, the locus are lines whose slope matches the tangent to the ellipses shown in (c).

imperfectly– synchronized with the transmitter, and hence, the delays cannot be used for the intersection of ellipses like in radar systems. Therefore, evaluating the TDoA between pairs of receivers is a commonly used strategy that leads to the intersection of hyperbolas, as shown in Fig. 6.4(a). This method is typically used in sound source localization [P.03] or in Ultra-Wide Bandwidth (UWB) positioning systems [ADGD17, DDG+20]. The reader can refer to review articles for a detailed description of those applications [CAA+17, DM22].

Connecting this kind of system to our general sensing model is not straightforward because the signals that we denote by $\boldsymbol{y}_q$ refer here to an

intermediate product of the conventional processing applied to the raw signals. This "pre-processing" step is briefly explained hereafter. Let us denote by L the number of receivers (such as microphones for sound source localization) located around $\boldsymbol{p}^*$ as depicted in Fig. 6.4(b). There are thus $Q = L(L-1)/2$ pairs of receivers that we associate to record the TDoA. Each of the Q pairs ends up providing one pre-processed measurement $\boldsymbol{y}_q$. In substance, each q-th pair yields $\boldsymbol{y}_q$ at the output of step called "Generalized Cross-Correlation (GCC)" using a *phase-transform weighting function*. This amounts to computing the Discrete Fourier Transform (DFT) of each of the L received signals, and, given a pair of receivers, multiplying their two spectra and keeping only their phase by modulus division.

Mathematically, for a single source, all receivers record the same signal with different unknown delays τ_ℓ. Therefore, given $F(\omega)$ to denote the output of the DTFT of the signal transmitted by the source, the spectra of the ℓ-th received signal reads

$$F(\omega)\exp(-\mathrm{j}\omega\tau_\ell), \tag{6.17}$$

up to an amplitude and phase coefficient and some additive noise. Multiplying the conjugate ℓ-th spectra with the ℓ'-th one yields $|F(\omega)|^2\exp(-\mathrm{j}\omega(\tau_\ell - \tau_{\ell'}))$. The commonly used "phase transform" [CAA+17] keeps only the phase of this expression. Finally, taking into account the amplitude and phase coefficient and the additive noise, we obtain an expression that fits our signal model. Indeed, since the DFT is tantamount to discretizing the ouput of the DTFT, setting the sensed parameter $\theta_q^* = \tau_\ell - \tau_{\ell'}$ results in the expression (6.6) in which

$$a_m(\theta) = \exp(\mathrm{j}2\pi\theta m/M), \quad m \in [M]. \tag{6.18}$$

The PPF, which maps the location $\boldsymbol{p}^*$ to the TDoA (normalized with respect to the sampling time T_s for simplicity) has thus the following expression:

$$\mathcal{S}_q^{\mathrm{tdoa}}(\boldsymbol{p}) = \tfrac{1}{cT_\mathrm{s}}(\|\boldsymbol{p}-\boldsymbol{x}_q\|_2 - \|\boldsymbol{p}-\boldsymbol{x}'_q\|_2), \tag{6.19}$$

where $\boldsymbol{x}_q$ and $\boldsymbol{x}'_q$ denote the locations of the two receivers of the q-th pair.

In the case of multiple sources, (6.7) can hold under additional assumptions. For cooperative transmitters, as in UWB applications, the transmitted waveforms are usually adjusted to have non-overlapping spectra (or, equivalently, to have good cross-correlation properties), which directly leads to separable measurements $\boldsymbol{y}_q$ as expressed in (6.7). In sound source local-

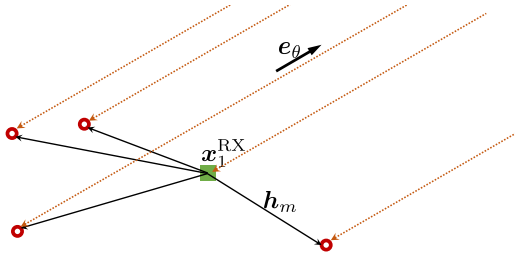

Figure 6.5: Representation of one node of a DoA-based application. A node is composed of M nodes of co-located sub-receivers such that the source is located in the far field, which implies that the angle of arrival θ (which acts as the sensed parameter) is identical for all receivers of a given array.

ization for speech processing, speakers are typically non-cooperative. However, under the Window-Disjoint Orthogonality (WDO) assumption [SH08], as one source is dominant in each time-frequency bin of the received signal, the spectra of two sources are non-overlapping at each instant. In this case, the multiple targets model (6.7) still holds for speech signal processing applications.

6.3.2 Source Localization from Directions of Arrival

The model (6.6) is also applicable in the context of the estimation of the location of a source (or a reflector) from a collection of Q sensor arrays, such that each array senses a Direction of Arrival (DoA) of the wave emitted by the source (or reflector) [Dog05, DOR08, WYL20, Wei04, WLW22]. In these applications, $\boldsymbol{p}^*$ is the location to estimate from the intersection of lines given by the DoAs, such as depicted in Fig. 6.4(b) Practically, DoAs are often observed from arrays of microphones [WYL20, AAM21, PGPM13], or radio frequency receivers [Wei04, WLW22].

In those applications, each sensor array is composed of M sub-receivers as depicted in Fig. 6.5. The atom function in model (6.6) corresponds to the *steering vector* [CKS89, KSF14] which can, for a generic array of receivers, be written as in [DOR08, C.22]. With our notations, this writes

$$a_m(\theta) = \exp(\mathrm{j}2\pi\lambda^{-1}\, \boldsymbol{h}_m^\top \boldsymbol{e}_\theta), \quad m \in [M]. \tag{6.20}$$

In the above, λ is the wavelength of the received signal, $\boldsymbol{h}_m$ is the vector pointing from a chosen reference sub-receiver's location toward the m-th sub-receiver, and $\boldsymbol{e}_\theta$ is the unit vector pointing in the direction given by

the sensed parameter θ. If the location to estimate $\boldsymbol{p}^*$ lies in 2D, then θ is an angle of arrival defined with respect to an arbitrary reference direction. If, however, $\boldsymbol{p}^*$ is a 3D position, then θ is a pair of azimuth and elevation angles. In both cases the PPF reads

$$\mathcal{S}_q^{\text{doa}}(\theta) = \angle(\boldsymbol{p} - \boldsymbol{x}_q^{\text{RX}}). \tag{6.21}$$

The extension for the localization of multiple sources is valid in DoA-based applications for the same reason as provided in Sec. 6.3.1 for TDoA.

6.3.3 Multistatic FMCW radar

Last but not least, we particularize our model to the context of a multistatic radar composed of one TX and Q RXs. Formally, the parameter of interest $\boldsymbol{p}^* = [\boldsymbol{x}^*, \boldsymbol{v}^*]^\top$ represents the couple of normalized location-velocity of a target, according to Def. 2.1. The sensed parameter θ_q corresponds to the couple of bistatic range and speed, as observed by the q-th receiver. Extending the model defined in Sec. 2 to the multistatic context, we obtain expressions that fit the model (6.7) with the atom function being the function $\boldsymbol{a}^{\text{fmcw}}(\cdot)$ —given in (2.16)— and the PPF defined, with $\boldsymbol{p} = (\boldsymbol{x}, \boldsymbol{v})$, by

$$\mathcal{S}_q^{\text{fmcw}}(\boldsymbol{p}) = \tfrac{1}{2} \begin{bmatrix} \mathcal{S}_q^{\text{range}}(\boldsymbol{x}) + \sigma_{\text{u}} \mathcal{S}_q^{\text{speed}}(\boldsymbol{x}, \boldsymbol{v}) \\ \mathcal{S}_q^{\text{speed}}(\boldsymbol{x}, \boldsymbol{v}), \end{bmatrix} \tag{6.22}$$

where, following the conclusions from Sec. 2.3.1 in the equations (2.10) and (2.11), and ignoring for simplicity the range calibration due to the length of the RF cables, we have

$$\begin{aligned} \mathcal{S}_q^{\text{range}}(\boldsymbol{x}) &:= \|\boldsymbol{x} - \boldsymbol{x}^{\text{TX}}\|_2 + \|\boldsymbol{x} - \boldsymbol{x}_q^{\text{RX}}\|_2, \\ \mathcal{S}_q^{\text{speed}}(\boldsymbol{x}, \boldsymbol{v}) &:= \langle \boldsymbol{h}^{\text{TX}} + \boldsymbol{h}_q^{\text{RX}}, \boldsymbol{v} \rangle. \end{aligned} \tag{6.23}$$

We remind the reader that the directional vectors $\boldsymbol{h}^{\text{TX}}$ and $\boldsymbol{h}_q^{\text{RX}}$ are represented in Fig. 2.3(b) and depend on the value of $\boldsymbol{x}$, following the expression (2.12).

This joint location-velocity model has been used multiple times to derive single-step methods for target detection using FMCW MSR [GN11, YGD13, SCWL15, BM11]. Yet, the required discretization of the 4D location-velocity space often yields algorithms with unscalable computational complexity regarding the real-time requirement of a multistatic radar, when

working with realistic dimensions of the radar signals. Inspired by works performed on pulse-Doppler radars [LP15], I proposed during my Master thesis a factorization-based methodology that enabled us to circumvent this limitation. While in this work, we shall directly consider the use of this factorization, the reader may refer to [MFJV19, dG19] for more details on this topic and simulation results showing the effectiveness of the technique with respect to its non-factorized counterpart.

When applying the factorization, the BOMP's selection step (6.14) is split into two parts with the following order: first, the estimation of location, performed as though the M_c chirps were distinct and non-coherent signals, and next, the estimation of the velocity. The estimation of the location, performed from virtually QM_c signals, and the subsequent estimation of velocity, which assumes a known target's location, can thus be interpreted as separated applications of our general framework.

Let us stress the approximations required for this separation by rewriting, in the multistatic context, the factorized expression of the FMCW radar signal that we formulated in Chapter 2 by (2.18). Neglecting the distortion term in (2.16) and using the PPFs given in (6.23), the matrix-reshaped q-th RX's measured signal, acquired in the presence of a single target, reads up to some additive noise,

$$\boldsymbol{Y}_q = \alpha_q \boldsymbol{a}^{\mathrm{in}}\big(\,\mathcal{S}_q^{\mathrm{range}}(\boldsymbol{x}^*) + \sigma_\mathrm{u}\mathcal{S}_q^{\mathrm{speed}}(\boldsymbol{x}^*,\boldsymbol{v}^*)\,\big) \otimes \boldsymbol{a}^{\mathrm{out}}(\mathcal{S}_q^{\mathrm{speed}}(\boldsymbol{x}^*,\boldsymbol{v}^*)). \quad (6.24)$$

Unlike the range and the speed, the location $\boldsymbol{x}^*$ and the velocity $\boldsymbol{v}^*$ are non-linearly coupled in the input of both $\boldsymbol{a}^{\mathrm{in}}$ and $\boldsymbol{a}^{\mathrm{out}}$ in the above, hence preventing the desired separation of the BOMP's selection step. The method that we proposed in [MFJV19] is built upon the following approximation, which reads, given the shifted location $\boldsymbol{x}_\mathrm{S}^* := \boldsymbol{x}^* + \sigma_\mathrm{u}\boldsymbol{v}^*$,

$$\begin{aligned}\boldsymbol{Y}_q &\simeq \alpha_q \boldsymbol{a}^{\mathrm{in}}\big(\,\mathcal{S}_q^{\mathrm{range}}(\boldsymbol{x}^* + \sigma_\mathrm{u}\boldsymbol{v}^*)\,\big) \otimes \boldsymbol{a}^{\mathrm{out}}(\mathcal{S}_q^{\mathrm{speed}}(\boldsymbol{x}^*,\boldsymbol{v}^*)).\\ &\simeq \alpha_q \boldsymbol{a}^{\mathrm{in}}\big(\,\mathcal{S}_q^{\mathrm{range}}(\boldsymbol{x}_\mathrm{S}^*)\,\big) \otimes \boldsymbol{a}^{\mathrm{out}}(\mathcal{S}_q^{\mathrm{speed}}(\boldsymbol{x}_\mathrm{S}^* - \sigma_\mathrm{u}\boldsymbol{v}^*,\boldsymbol{v}^*)). \quad (6.25)\end{aligned}$$

Taking advantage of the definition of the outer product $\otimes$, each column of $\boldsymbol{Y}_q$ can now be interpreted as one signal, distinct from the other columns, thereby yielding a total of QM_c signals to which BOMP can be applied. Each resulting signal follows our model (6.6) with the shifted location as the parameter of interest, $\boldsymbol{a}^{\mathrm{in}}(\cdot)$ as the atom function (see (2.19)) and $\mathcal{S}_q^{\mathrm{range}}(\cdot)$ for the projection function. The resulting problem's geometry is illustrated in Fig. 6.4(c), showing the elliptic loci that originate from the definition of

$\mathcal{S}_q^{\text{range}}(\cdot)$.

Next, assuming that the shifted location $\hat{\boldsymbol{x}}_{\text{S}}$ has been estimated from the above, the velocity estimation is linked to our formalism by setting the target's velocity for the parameter of interest. We then set $\boldsymbol{a}^{\text{out}}(\cdot)$ as the atom function (see (2.20)) and $\mathcal{S}_q^{\text{speed}}(\hat{\boldsymbol{x}}_{\text{S}} - \sigma_{\text{u}}\boldsymbol{v}, \boldsymbol{v})$ for the projection function, in which $\boldsymbol{v}$ is the only variable since $\hat{\boldsymbol{x}}$ is fixed by the previous step.

Putting all together, using $\boldsymbol{Y}^q_{m_{\text{c}}}$ to denote the m_{c}-th column of the $M_{\text{s}} \times M_{\text{c}}$ matrix-reshaped q-th measurement, the selection step (6.14) of the first iteration of BOMP is thus replaced by the factorized selection:

$$\hat{\boldsymbol{x}}_{\text{S}} := \underset{\boldsymbol{x}\in\mathcal{X}}{\arg\max} \sum_{q=1}^{Q} \sum_{m_{\text{c}}=1}^{M_{\text{c}}} \left|\langle \boldsymbol{a}^{\text{in}}\big(\mathcal{S}_q^{\text{range}}(\boldsymbol{x})\big),\, \boldsymbol{Y}^q_{m_{\text{c}}}\rangle\right|^2, \tag{6.26}$$

$$\hat{\boldsymbol{v}} := \underset{\boldsymbol{v}\in\mathcal{V}}{\arg\max} \sum_{q=1}^{Q} \left|\langle \boldsymbol{a}^{\text{out}}\big(\mathcal{S}_q^{\text{speed}}(\hat{\boldsymbol{x}}_k - \sigma_{\text{u}}\boldsymbol{v}, \boldsymbol{v})\big),\, \boldsymbol{g}_q\rangle\right|^2, \tag{6.27}$$

where the m_{c}-th component of $\boldsymbol{g}_q$ is $\langle \boldsymbol{a}^{\text{in}}\big(\mathcal{S}_q^{\text{range}}(\hat{\boldsymbol{x}}_{\text{S}})\big),\, \boldsymbol{Y}^q_{m_{\text{c}}}\rangle$. The next iterations of a BOMP procedure are similar to the above, using the residue which is progressively updated by the greedy method instead of $\boldsymbol{Y}^q_{m_{\text{c}}}$.

While the impact of this simplification has been empirically studied in our previous work [MFJV19], we deepen our understanding of this effect by means of the theoretical bound on the selection step's error, derived in the next section.

6.3.4 Illustrations of the Single-step Selection

We finalize this section with a few illustrations of correlation functions, maximized in the selection of the first iteration of BOMP. For each of the four contexts appearing in Table 6.1, we considered a system whose geometry is described by Fig. 6.6 and in which the TX's location $\boldsymbol{x}^{\text{TX}}$ is only to be taken into account for the MSR context. For the estimation of source location from DoA, each of the five nodes is zoomed-in by Fig. 6.6(b), showing a circular array of M sub-receivers. In all applications, we set $M = 64$ and $Q = 5$, except for the intersection of TDoA in which 5 receiving nodes yield $Q = 10$ measurements. We build measurements $\boldsymbol{y}_q$ by considering a single target located in the normalized location [-16, 32] and, for the estimation of velocity with MSR, a normalized velocity [12, -5]. We added some noise to the measurement given a 20dB SNR.

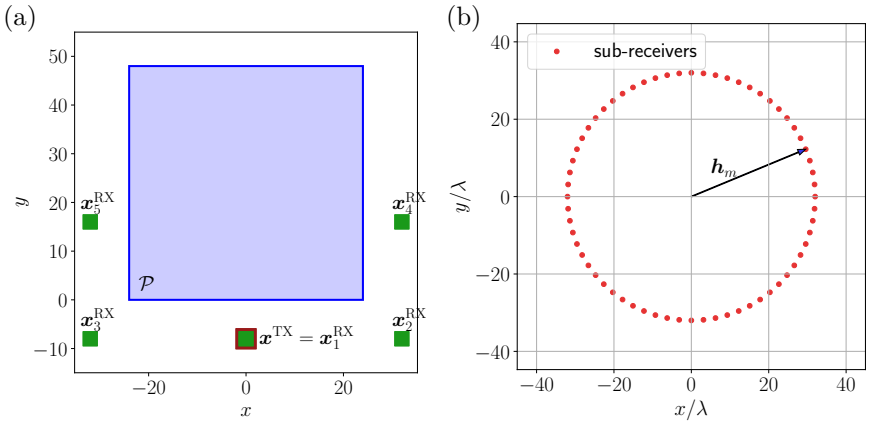

Figure 6.6: Schematic representation of the geometry of the system used to illustrate the single-step and the two-step principles. **(a)** Representation of the locations of the five nodes, hence yielding $Q = 5$, except for TDoA-based localization where we get $L = 5$ and $Q = 10$. **(b)** Zoom on each node for the DoA-based localization, where a circular array is considered.

The resulting global correlation functions computed by single-step methods are depicted in Fig. 6.7, following the same order as the schematic previously shown in Fig. 6.4. We naturally observe that the mapping intrinsically performed by the composition of the functions C and $\mathcal{S}_q$ let appear in Fig. 6.7 the loci represented in Fig. 6.4. The resulting intersection of these loci somehow enables the single-step method to counter the possibility of one (or several) sensors taken alone to fail in the detection of the target. Indeed, the single-step decision is taken from a consensus on the value of $\boldsymbol{p}$ that delivered the most overall power across all the received signals. Said differently, it directly gathers all available information and promotes, via the mappings $\mathcal{S}_q(\boldsymbol{p})$, consistent models across all sensors, thereby potentially offering a more robust estimation of $\boldsymbol{p}^*$.

In comparison, Fig. 6.8 shows the objective functions from (6.8) for the first received signal (*i.e.,* for $q = 1$). Since the sensing domain Θ is a 1D domain of distances, defining a grid Θ_{G} with sufficient density to find the main lobe of the objective function of the two-step method appears to require the use of significantly fewer bins than required for a similar sampling of the 2D correlations used by single-step methods. Given $N_\Theta := |\Theta_{\mathrm{G}}|$ and $N_P := |\mathcal{P}_{\mathrm{G}}|$ denoting the respective cardinality of the sensed grid and of the grid of interest, it yields the relation $N_P \gg N_\Theta$ which intensifies as a

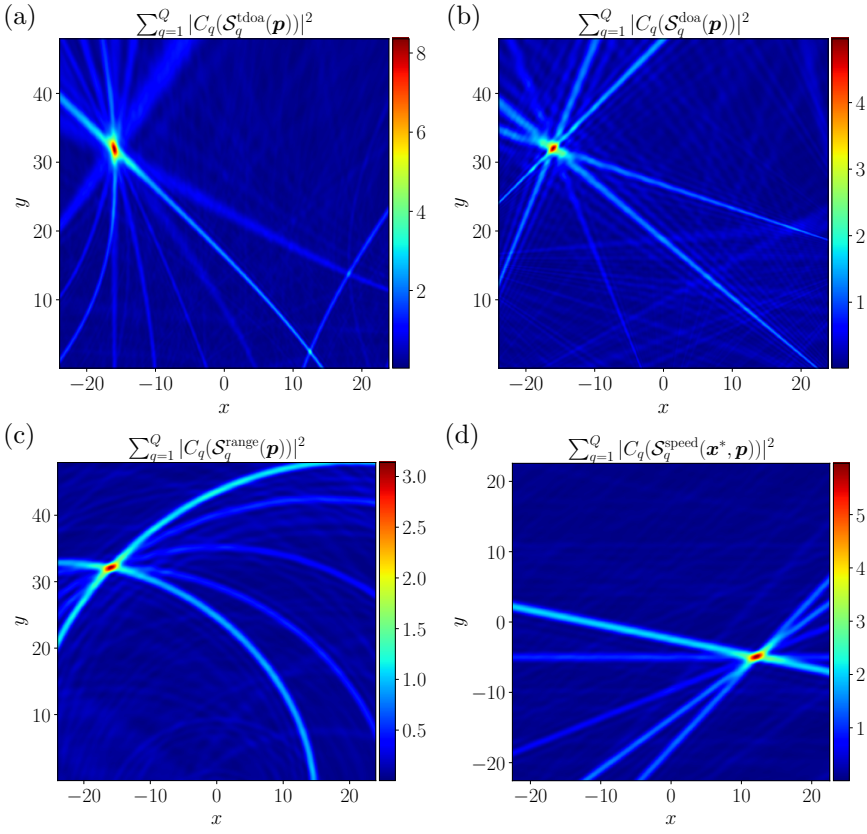

Figure 6.7: Single-step correlation functions obtained for our four standard examples of applications in the first selection step of BOMP given the setup shown in Fig. 6.6. **(a)** Source Localization (SL) from TDoA, **(b)** SL from DoA, **(c)** target localization from MSR, and **(d)** velocity estimation from MSR.

higher precision is targeted by the use of denser grids. This is why two-step methods are known to be faster than their single-step counterparts.

The upcoming sections of this chapter highlight the superiority of the estimation performance of the single-step parameter selection compared to its two-step counterpart, with a specific focus on the MSR application. However, in the following chapter, our focus will shift towards addressing the question of preserving the strengths of both the two-step and single-step approaches through a hybrid strategy coined Grid Hopping method.

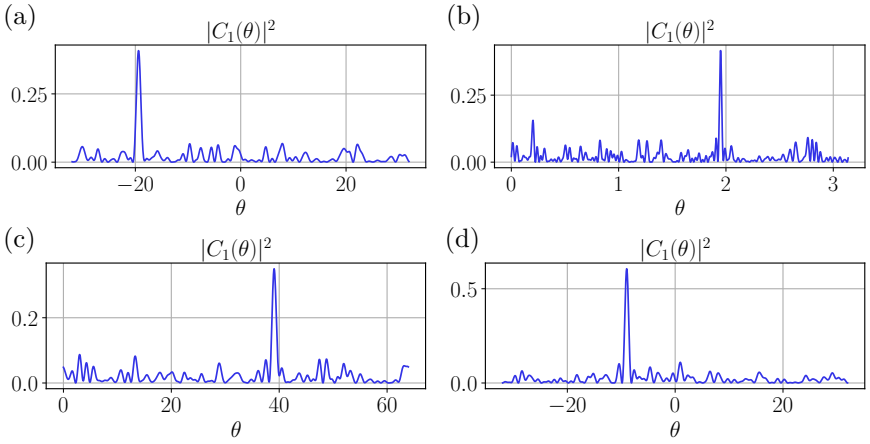

Figure 6.8: First (for $q = 1$) correlation function obtained in the two-step method applied to our four standard examples of applications given the setup shown in Fig. 6.6. **(a)** SL from TDoA, **(b)** SL from DoA, **(c)** target localization from MSR, and **(d)** velocity estimation from MSR.

6.4 Bound on the Factorized One-sep Selection in Radar Applications

Applying the approximation (6.25) has consequences on the estimation performance at the output of the factorized selection (6.26) when compared to the output of (6.11) used with the complete model (6.22). First, as studied in [BM11], it reduces the robustness against noise since (6.26) adds up the M_c correlations in a non-coherent manner. Moreover, the approximation which removes the function $\mathcal{S}_q^{\text{speed}}$ in the input of $\boldsymbol{a}^{\text{in}}$ generates a model mismatch on the estimation performance, thereby preventing the ellipses that we observe in Fig. 6.4(c) to exactly meet in one point. To evaluate the impact of this mismatch, let us study a simple case in which a multistatic radar observes a single target in noiseless conditions. In that situation, the estimation error from the exact method, with no approximation on the model, and not restricted to a grid search, is always zero. Since this is no longer the case with the approximation and the use of the factorized selection step (6.26), we hence provide an upper bound on the resulting error.

The bounding theorem presented in this section relies on the definition of "lobe-r" functions, introduced in Def. 5.3. We remind the reader that the atom function $\boldsymbol{a}^{\text{in}}$ is nothing more than a Fourier atom and thereby, its associated correlation kernel κ is translation invariant and such that $|\kappa(\cdot)|^2$ is a lobe-r function with lower and upper bounding constants L and U, as notably shown in Ex. 5.5.

We now define the notations used in the writing of this theorem. To this end, we first remind that the location domain $\mathcal{X}$ is assumed to meet the condition (2.23) in which $r_{\min}$ is a minimal normalized (and hence dimensionless, see Def. 2.1) distance between any location from $\mathcal{X}$ and any transmitting or receiving node composing the MSR system. Additionally, given the global PPF defined for all $\boldsymbol{x} \in \mathcal{X}$ by $\mathcal{S}^{\text{range}}(\boldsymbol{x}) := (\mathcal{S}_1^{\text{range}}(\boldsymbol{x}), \cdots, \mathcal{S}_Q^{\text{range}}(\boldsymbol{x}))^\top$, the associated coefficient of variation defined by Def. 6.2 are denoted by $\lambda_{\min}$ and $\lambda_{\max}$. Finally, we define the square-root power ratio by

$$R_{\text{P}} := \sqrt{\frac{|\alpha_{\max}|^2 L}{|\alpha_{\min}|^2 U}}, \tag{6.28}$$

where $\alpha_{\max} := \max_{q\in[Q]} |\alpha_q|$ and $\alpha_{\min} := \min_{q\in[Q]} |\alpha_q|$.

Theorem 6.2 (Factorized-MSR Selection Error Bound)**.** *Let us consider a linear-chirp FMCW MSR observing (i) a* single *target, (ii) in noiseless condition, and (iii) such that the expression of each q-th receiver's signal follows* (6.24).
Given a location $\boldsymbol{x}^*$ and a known *velocity $\boldsymbol{v}^*$, respectively taken from the location domain $\mathcal{X}$ and the velocity domain $\mathcal{V}$ meeting the conditions* (2.23) *and* (2.24)*, with $\hat{\boldsymbol{x}}_{\text{S}}$ denoting the output of the location selection step* (6.26) *and with*

$$\hat{\boldsymbol{x}} = \hat{\boldsymbol{x}}_{\text{S}} - \sigma_{\text{u}} \boldsymbol{v}^* \tag{6.29}$$

denoting the corrected location estimate. If

$$\frac{Q}{4} \frac{1}{2r_{\min} - 1} \leq R_{\text{P}}^{-1} r, \tag{6.30}$$

the following bound on the location estimation error holds

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2 \leq \lambda_{\min}^{-1} \big(R_{\text{P}} + 1\big) \tfrac{Q}{4} \frac{1}{2r_{\min} - 1}. \tag{6.31}$$

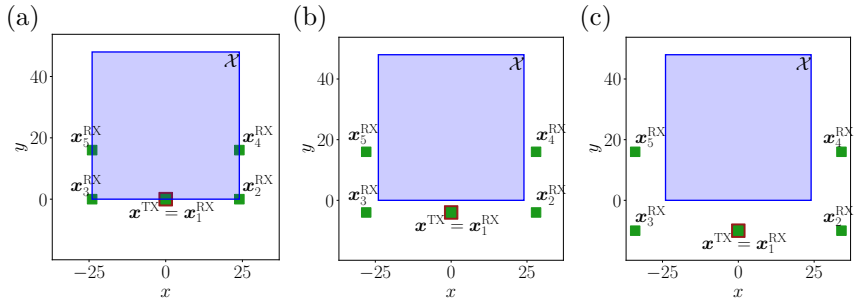

Figure 6.9: Examples of the MSR geometry with **(a)** $\varepsilon = 0$, **(b)** $\varepsilon = 4$ and **(c)** $\varepsilon = 10$. When $\varepsilon = 0$, the receivers are located on the edge of the location domain. Else they are located at a distance of either ε of $\sqrt{2}\varepsilon$ away from $\mathcal{X}$.

In short, this theorem states that the effects of the approximation become negligible as the value of $r_{\min}$ increases. Indeed, in the ideal conditions of this theorem, the estimate $\hat{\boldsymbol{x}}$ only differs from the ground truth $\boldsymbol{x}^*$ because of the model approximation. Therefore, the bound (6.31) states that its effect vanishes for a sufficiently high value of $r_{\min}$. This could have been anticipated since the approximation going from (6.24) to (6.25) can be interpreted as a linearization of the function $\mathcal{S}_q^{\mathrm{range}}(\boldsymbol{x} + \sigma_{\mathrm{u}}\boldsymbol{v})$ into $\mathcal{S}_q^{\mathrm{range}}(\boldsymbol{x}) + \sigma_{\mathrm{u}}\mathcal{S}_q^{\mathrm{speed}}(\boldsymbol{x}, \boldsymbol{v})$. Indeed, $\mathcal{S}_q^{\mathrm{speed}}(\boldsymbol{x}, \boldsymbol{v})$ amounts to the first order derivative with respect to $\boldsymbol{v}$ of $\mathcal{S}_q^{\mathrm{range}}(\boldsymbol{x}+\sigma_{\mathrm{u}}\boldsymbol{v})$. As the ellipses are more keen to be approximated by lines when their foci are further apart, increasing $r_{\min}$ leads to a smoother linearization.

Note that the bound hints that better results are expected when Q is small. Decreasing Q is, however, likely to significantly reduce $\lambda_{\min}$ and possibly make it tend toward zero if the bijectivity of $\mathcal{S}^{\mathrm{range}}$ is no longer guaranteed, pushing the bound toward positive infinity. Furthermore, a tighter bound, taking into account the fact that a distance equal $r_{\min}$ is less likely to be met at the same time with respect to all RXs as Q increases may be developed in the future to remove the presumably false intuition given by the above that one must decrease Q to obtain a smaller error.

6.5 Simulations of MSR Applications

In this section, we present the results of Monte-Carlo simulations of a multistatic radar composed of one transmitter and $Q = 5$ receivers and observing a single target located in $\boldsymbol{x}^*$ and moving with a velocity $\boldsymbol{v}^*$. The goal of this

section is two-fold: (i) compare the performance two-step and single-step methods when applied to MSR and (ii) observe the impact of the antennas' location on the estimation performance that we hinted by means of Thm. 6.2.

6.5.1 Simulated System

We simulated multiple geometries for our MSR system characterized by different locations of their constituting antennas. These locations are fully described by a distance parameter ε modeling the distance between the scene (*i.e.,* the location domain $\mathcal{X}$) and the antennas. More precisely, Fig. 6.9 shows examples of geometry obtained for different values of ε, which is actually equivalent to the quantity $r_{\min}$ used in Thm. 6.2. We then considered two measurement scenarios:

- **Scenario 1**: The estimation of the location $\boldsymbol{x}^*$ only, using a single chirp ($M_{\mathrm{c}} = 1$) of $M_{\mathrm{s}} = 64$ samples.
- **Scenario 2**: The estimation of both the location $\boldsymbol{x}^*$ and the velocity $\boldsymbol{v}^*$, using $M_{\mathrm{c}} = 64$ chirps, each of $M_{\mathrm{s}} = 64$ samples. In that scenario, the distortion-free model (6.24) was used to generate the signals.

The location domain is always $\mathcal{X} = [-24, 24] \times [0, 48]$ and the velocity domain is $\mathcal{V} := [-32/\sqrt{2}, 32/\sqrt{2}]^2$. For each geometry, we generated 10,000 realizations of radar measurement from the true location $\boldsymbol{x}^*$ and the true velocity $\boldsymbol{v}^*$ taken randomly following a uniform distribution on respectively $\mathcal{X}$ and $\mathcal{V}$.

6.5.2 Competing Estimation Methods

Since there is only one target to locate, the estimation methods are reduced to their selection step, introduced in Sec. 6.2.

The two-step estimation is thus fully explained by the equations (6.8) and (6.10). We solved the maximization (6.8) in two ways either by restricting it to an exhaustive grid search on the grid Θ_{G} or by refining this grid search with a gradient ascent. The first case is labeled "ONG" and the second case "OFG", respectively standing for *on-the-grid* and *off-the-grid*. In both cases, the used grid Θ_{G} was uniformly discretizing Θ with a constant bin spacing 1/4. In the scenario 1, Θ is the range domain $\mathcal{R} = [0, M_{\mathrm{s}}]$, leading to 256 bins. In the scenario 2, Θ is the range-speed domain $\mathcal{R} \times \mathcal{U} = [0, M_{\mathrm{s}}] \times [-M_{\mathrm{c}}/2, M_{\mathrm{c}}/2]$, leading to 65536 bins. Note that

	S-S. (ONG)	S-S. (OFG)	T-S. (ONG)	T-S. (OFG)
Scenario 1	15.22ms	30.79ms	0.64ms	3.70ms
Scenario 2	812ms	1010ms	38ms	62ms

Table 6.2: Average computation times of obtained for two-step (T-S) and single-step (S-S) selections in both simulated scenarios.

we took advantage of the Fourier nature of the used dictionary and the use of a uniform grid to use either a 1D FFT or a 2D FFT to perform the grid search in scenarios 1 and 2, respectively.

The single-step estimation corresponds to solving either (6.26) only (with $M_c = 1$) for the scenario 1 or both this equation and (6.27) for the scenario 2. We also performed either *on-the-grid* estimations using an exhaustive grid search over uniform grids with a 1/4 bin spacing or *off-the-grid* estimations by refining the grid estimate with a gradient ascent. Note that no FFT can be used in that situation.

Each method outputs either the estimate $\hat{\boldsymbol{x}}$ for scenario 1 or both $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{v}}$ for scenario 2. We thus defined the error E term which is contextually defined by $\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2$ or $\|(\hat{\boldsymbol{x}}, \hat{\boldsymbol{v}}) - (\boldsymbol{x}^*, \boldsymbol{v}^*)\|_{\mathrm{F}}$. The performance metrics are similar to those used in Sec. 5.6. That is, we use (i) the Average Error (AE) over the 10,000 realizations, (ii) the Miss Rate (MR) that counts the ratio of estimates such that $E \geq 1$, and (iii) the Average Hit Error (AHE) that averages only the errors whose value is below 1.

6.5.3 Comparing Two-step and Single-step Methods

We begin with scenario 1 in Fig. 6.10, where the performance is plotted with respect to the SNR that characterizes the measurement. Overall, we observe that the single-step method provides a stronger robustness against noise that is translated in the three metrics. We observe that the *on-the-grid* version of the methods does not alter when compared to their *off-the-grid* counterparts, the detection ability of the algorithms since the MR is identical in both cases. The only consequence of the *on-the-grid* restriction is the precision of the estimation, which is limited by the density of the used grid. The advantage of using a single-step method in the *on-the-grid* context is only appearing when the nuisance (the noise here) is sufficiently strong to overcome this grid limitation.

When inspecting the *off-the-grid* versions of the algorithms, we observe two regimes in the Average Error (AE) curves, characterized by a sudden change of slope. While the second and final slope of the AE curve is identical

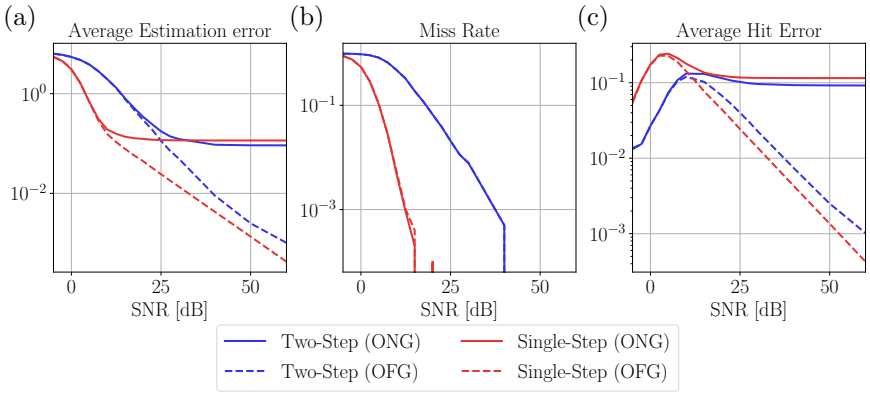

Figure 6.10: Comparative performance results of the application of the four competing algorithms in scenario 1.

for both methods, the single-step process starts its evolution with respect to the SNR with a more negative slope. This can be explained since the consensus taken across all the receivers leads to a stronger detection ability (smaller probability of completely missing the target), highlighted by the MR curves that also show distinct slopes between the two methods. The precision, however, highlighted by the AHE exhibits a similar slope for both methods.

Table 6.2 gathers the average computation of each algorithm. This confirms the intuition that despite single-step methods being more robust against noise, they require a drastically higher computation time than their two-step counterparts.

Let us continue with scenario 2, for which the results are shown in Fig. 6.11. While the *on-the-grid* versions of the algorithm tell a similar story as for scenario 1, this is not the case for the *off-the-grid* versions. The additional saturation is caused by the mismatch we pointed out and studied in Thm. 6.2. Again, Table 6.2 shows the average computation and leads to the same conclusion as for scenario 1. Note that the computation time for scenario 2 is significantly higher since we have $M_c = 64$ instead of 1.

6.5.4 Impact of the MSR's Geometry

We end this section with the evaluation of the impact of the spacing ε of the antennas with respect to the scene $\mathcal{X}$ on the estimation performance. The

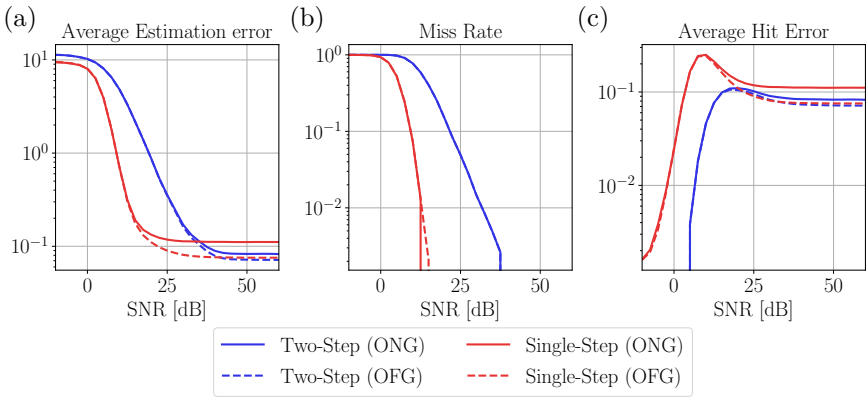

Figure 6.11: Comparative performance results of the application of the four competing algorithms in scenario 2.

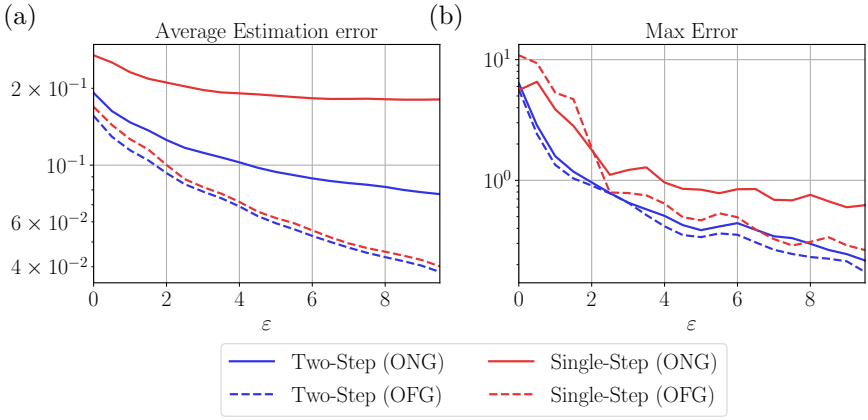

Figure 6.12: Estimation error as a function of the spacing between the transmitting and receiving nodes with respect to the observed scene, in an MSR context.

resulting average estimation error is shown in Fig. 6.12(a), and the maximal error obtained among all the realizations is plotted in Fig. 6.12(b). Since the simulations were performed in noiseless conditions, the fact that *off-the-grid* algorithms provide non-zero estimation errors can only be explained by the model mismatch studied in Thm. 6.2. We observe a neat decrease of this error in both metrics when the value of ε increases. Interestingly, the conclusions from Thm. 6.2 also apply to the two-step parameter selection. We note that the *on-the-grid* curves saturate because of another limitation: the discretization error. This aspect will be more deeply studied in the next chapter.

To summarize, we confirm with this simulation that the model mismatch induced by the factorization of the MSR model (6.25) can be neglected given transmitting receiving nodes that are sufficiently far away from the observed scene. Therefore, we can safely neglect this aspect in the subsequent approaches we will propose in the next chapter to improve the single-step methods.

6.6 Discussion

In this chapter, we presented and illustrated the distinction between two-step and single-step estimation methods in the sensor network context. These two classes of methods will constitute the starting point of the derivation of our "Grid Hopping" framework that we describe in the next chapter. We dedicated careful efforts to unifying multiple applications (such as sound source localization or radars) under a single mathematical formalism. Unlike literature on similar topics, this enables us, in the next chapter, to provide accelerations that remain valid independently of the targetted application.

In the second part of this chapter, we focused on Multistatic Radar (MSR) systems. We theoretically showed under which conditions on the system's geometry the factorization that we intuitively used in previous research [MFJV19] yields negligible model mismatch. We also confirmed with Monte Carlo simulations the superiority of single-step methods over their two-step counterparts in terms of estimation performance.

Although this chapter brings only minor contributions to the sensor network community, the mathematical formalism it introduces paves the way for the next chapter, which contains our main contribution to the topic.

6.7 Appendix: Proofs

6.7.1 Proof of Lemma 6.1

Proof. If $\mathcal{S}$ is non-injective over $\mathcal{P}$, then there exist two distinct parameters $\boldsymbol{p}_1, \boldsymbol{p}_2 \in \mathcal{P}$ such that $\mathcal{S}(\boldsymbol{p}_1) = \mathcal{S}(\boldsymbol{p}_2)$. From the Taylor's theorem, there is at least one scalar $\varepsilon \in [0, 1]$ such that

$$\mathcal{S}(\boldsymbol{p}_2) = \boldsymbol{J}\big(\boldsymbol{p}_1 + \varepsilon(\boldsymbol{p}_2 - \boldsymbol{p}_1)\big)(\boldsymbol{p}_2 - \boldsymbol{p}_1), \tag{6.32}$$

Denoting $\bar{\boldsymbol{p}} := \boldsymbol{p}_1 + \varepsilon(\boldsymbol{p}_2 - \boldsymbol{p}_1)$, the convexity of the set $\mathcal{P}$ imposes that $\bar{\boldsymbol{p}} \in \mathcal{P}$. Moreover, (6.32) associated to the equality $\mathcal{S}(\boldsymbol{p}_1) = \mathcal{S}(\boldsymbol{p}_2)$ implies that $\boldsymbol{J}(\bar{\boldsymbol{p}})(\boldsymbol{p}_2 - \boldsymbol{p}_1) = 0$. Therefore, $(\boldsymbol{p}_2 - \boldsymbol{p}_1)$ is an eigenvector of $\boldsymbol{J}(\bar{\boldsymbol{p}})^\top \boldsymbol{J}(\bar{\boldsymbol{p}})$ associated with the eigenvalue 0. Finally, since $\boldsymbol{J}(\bar{\boldsymbol{p}})^\top \boldsymbol{J}(\bar{\boldsymbol{p}})$ is positive semidefinite by construction, we conclude that $\lambda_{\min} = 0$ and this proof as well. □

6.7.2 Proof of Theorem 6.2

Thm. 6.2 is a particular application of the following theorem, proven in the next section.

Theorem 6.3 (Lobe Function Maximization Bound)**.** *Given the lobe-r function $f : \mathcal{Z} \mapsto \mathbb{R}$ with lower and upper bounding constants L and U, and given a set $\mathcal{Z}' \subset \mathcal{Z}$, if there is*

$$\boldsymbol{z}_{\text{ref}} \in \mathcal{Z}' \cap \mathbb{B}_2(r), \tag{6.33}$$

then the estimator $\hat{\boldsymbol{z}}$ defined by

$$\hat{\boldsymbol{z}} := \arg\max_{\boldsymbol{z} \in \mathcal{Z}'} \tilde{f}(\boldsymbol{z}) \tag{6.34}$$

is such that $\|\hat{\boldsymbol{z}}\|_2 \leq \sqrt{\frac{L}{U}} \|\boldsymbol{z}_{\text{ref}}\|_2$.

In short, Thm. 6.3 provides a bound on the maximization of a lobe-r function restricted to a domain $\mathcal{Z}'$. if $\mathcal{Z}'$ contains 0, then the error is naturally 0 by definition of a lobe-r function. Else, given a candidate $\boldsymbol{z}_{\text{ref}}$ in the inner lobe of f, the decreasing properties of f guarantee given that our candidate is not the maximizer, then the maximizer $\hat{\boldsymbol{z}}$ can not be further away from 0 than a value proportional to the norm of our candidate.

The application of the above theorem to obtain Thm. 6.2 relies on three lemmas, given hereafter and whose proofs are given at the end of this appendix.

- The first one (Lem. 6.4) uses the concepts from Def. 6.2 to enable us to navigate the bounds from the domain of interest $\mathcal{P}$ (which is $\mathcal{X}$ in the MSR context) to the admissible sensed domain $\mathcal{T}$ that we defined in the general context in (6.3).

- The second lemma (lem. 6.5) explains how we can translate any outer summation of similar lobe-r functions into a novel higher dimensional lobe-$\tilde{r}$ function, with a new lobe radius $\tilde{r}$ and new bounding constants, expressed from the constants defining the initial lobe-r function.

- Unlike the two first lemmas, the third one (Lem. 6.6) is specific to radar applications and directly bounds the approximation error on the PPF modification that was performed when going from (6.24) to (6.25).

Lemma 6.4. *Given a global PPF following assumption 6.1 and denoted by $\mathcal{S}$, for all pairs $\boldsymbol{p}_1, \boldsymbol{p}_2 \in \mathcal{P}$ and $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \mathcal{T}$ such that $\boldsymbol{\theta}_1 = \mathcal{S}(\boldsymbol{p}_1)$ and $\boldsymbol{\theta}_2 = \mathcal{S}(\boldsymbol{p}_2)$, the following holds*

$$\lambda_{\min}\|\boldsymbol{p}_2 - \boldsymbol{p}_1\|_2^2 \leq \|\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1\|_2^2 \leq \lambda_{\max}\|\boldsymbol{p}_2 - \boldsymbol{p}_1\|_2^2. \tag{6.35}$$

Lemma 6.5 (Summation of lobe functions)**.** *Given positive scalars $b_1, \cdots, b_Q$ such that $\sum_{q=1}^{Q} b_q = 1$ and the separable function*

$$f_\Sigma \,:\, \mathcal{Z}^Q \mapsto \mathbb{R} \,:\, f_\Sigma(\boldsymbol{z}) = \textstyle\sum_{q=1}^{Q} b_q f(\boldsymbol{z}_q) \tag{6.36}$$

with $\boldsymbol{z} = [\boldsymbol{z}_1^\top, \cdots, \boldsymbol{z}_Q^\top]^\top$, f_Σ is lobe-$\tilde{r}$ whose lobe is bounded by $\tilde{L}$ and $\tilde{U}$ obtained from

$$\tilde{r}^2 = \tfrac{U}{L}\tfrac{b_{\min}}{b_{\max}} r^2, \tag{6.37}$$

$$\tilde{L} = L b_{\max}, \tag{6.38}$$

$$\tilde{U} = U b_{\min}, \tag{6.39}$$

and where $b_{\max} := \max_{q\in[Q]} b_q$ *and* $b_{\min} := \min_{q\in[Q]} b_q$.

Lemma 6.6 (Range Approximation Error)**.** *Let* $\mathcal{S}_q^{\text{range}}(\cdot)$ *and* $\mathcal{S}_q^{\text{speed}}(\cdot)$ *be defined as in* (6.23)*. Given a normalized location* $\boldsymbol{x} \in \mathcal{X}$ *and a normalized velocity* $\boldsymbol{v} \in \mathcal{V}$ *with* $\mathcal{X}$ *and* $\mathcal{V}$ *respectively defined following* (2.23) *and* (2.24)*, then*

$$\begin{aligned}\left|\mathcal{S}_q^{\text{range}}(\boldsymbol{x}+\sigma_{\text{u}}\boldsymbol{v}) - \left(\mathcal{S}_q^{\text{range}}(\boldsymbol{x}) + \mathcal{S}_q^{\text{speed}}(\boldsymbol{x},\boldsymbol{v})\right)\right| \\ \leq \tfrac{1}{8}\left(\frac{1}{2\|\boldsymbol{x}-\boldsymbol{x}^{\text{TX}}\|_2 - 1} + \frac{1}{2\|\boldsymbol{x}-\boldsymbol{x}_q^{\text{RX}}\|_2 - 1}\right)\end{aligned} \tag{6.40}$$

Corollary 6.7 (Loose Range Approximation Error)**.** *Under the same conditions of Lem. 6.6, the bound it provides can be simplified using the loose upper bound*

$$\left|\mathcal{S}_q^{\text{range}}(\boldsymbol{x}+\sigma_{\text{u}}\boldsymbol{v}) - \left(\mathcal{S}_q^{\text{range}}(\boldsymbol{x}) + \mathcal{S}_q^{\text{speed}}(\boldsymbol{x},\boldsymbol{v})\right)\right| \leq \tfrac{1}{4}\frac{1}{2r_{\min}-1}, \tag{6.41}$$

where $r_{\min}$ *is the minimal distance between* $\boldsymbol{x}$ *and any transmitting or receiving node.*

Proof. This is a direct consequence of the application of the condition (2.23) on $\mathcal{X}$ to (6.40). □

We can now start the demonstration whose main principle is to make appear a formulation that meets the conditions of Thm. 6.3. To this end, and to improve the reading of this demonstration, let us define the following notations.

- First, returning to the physical meaning of the output of the PPFs $\mathcal{S}^{\text{range}}$ and $\mathcal{S}^{\text{speed}}$, we use the notations from Chapter 2 to denote the true ranges and speeds associated to the observed target, *i.e.,* for all $q \in [Q]$,

$$r_q^* := \mathcal{S}_q^{\text{range}}(\boldsymbol{x}^*) \quad \text{and} \quad u_q^* := \mathcal{S}_q^{\text{speed}}(\boldsymbol{x}^*, \boldsymbol{v}^*). \tag{6.42}$$

- To smoothly meet the conditions of Lem. 6.5, we define the normalized scattering coefficients

$$\gamma_q = |\alpha_q|^2 \,/\, \textstyle\sum_{q'=0}^{Q} |\alpha_{q'}|^2, \tag{6.43}$$

  and more specifically $\gamma_{\max} := \max_{q\in[Q]} \gamma_q$ and $\gamma_{\min} := \min_{q\in[Q]} \gamma_q$.

- Using the above, we define the total squared kernel function, resulting from the outer summation of the squared kernels associated with $\boldsymbol{a}^{\text{in}}$ by

$$\kappa_\Sigma(\boldsymbol{\delta}) := \textstyle\sum_{q=1}^{Q} \gamma_q \big|\big\langle \boldsymbol{a}(\theta_q),\, \boldsymbol{a}(\theta_q + \delta_q)\big\rangle\big|^2, \tag{6.44}$$

  for all $\boldsymbol{\theta} := (\theta_1, \cdots, \theta_Q)^\top \in \Theta^Q$ and for all $\boldsymbol{\delta} := (\delta_1, \cdots, \delta_Q)^\top \in \Theta^Q$.

- Finally, while the parameters of interest in this context are the shifted true and approximated locations, ($\hat{\boldsymbol{x}}_{\text{S}}$ and $\boldsymbol{x}^*_{\text{S}}$, respectively, the quantity acting as the sensed parameters are the projection of these shifted locations into the range domain. In this MSR context, we thus define them as

$$\begin{aligned} \hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \cdots, \hat{\theta}_Q)^\top &:= \mathcal{S}^{\text{range}}(\hat{\boldsymbol{x}}_{\text{S}}), \\ \boldsymbol{\theta}^* = (\theta^*_1, \cdots, \theta^*_Q)^\top &:= \mathcal{S}^{\text{range}}(\boldsymbol{x}^*_{\text{S}}). \end{aligned} \tag{6.45}$$

  We emphasize that given that $\mathcal{X}$ acts here as the domain of interest, and since from assumption 6.1, the global PPF, denoted by $\mathcal{S}$, is a bijection from $\mathcal{X}$ to the admissible sensed domain $\mathcal{T}$. This latest corresponds, in this context, to the set of possible collections of range that matches a location in $\mathcal{X}$.

The demonstration now begins by using Lem. 6.4 to obtain the inequality

$$\begin{aligned} \|\boldsymbol{x}^* - \hat{\boldsymbol{x}}\|_2 &= \|\boldsymbol{x}^* + \sigma_{\text{u}}\boldsymbol{v}^* - (\hat{\boldsymbol{x}} + \sigma_{\text{u}}\boldsymbol{v}^*)\|_2 \hat{\boldsymbol{x}}_{\text{S}}\|_2 \\ &\leq \lambda_{\min}^{-1} \|\boldsymbol{\theta}^* - \hat{\boldsymbol{\theta}}\|_2 \end{aligned} \tag{6.46}$$

With $\hat{\boldsymbol{\delta}} := \boldsymbol{\theta}^* - \hat{\boldsymbol{\theta}}$, the demonstration now amounts to derive an upper bound on $\|\hat{\boldsymbol{\delta}}\|_2$. With the notations presented here above, we rewrite the

selection step (6.26) as

$$\hat{\boldsymbol{x}}_{\mathrm{S}} = \operatorname*{arg\,max}_{\boldsymbol{x}\in\mathcal{X}} \sum_{q=1}^{Q} \gamma_q \left|\left\langle \boldsymbol{a}\big(r_q^* + \sigma_{\mathrm{u}} u_q^*\big),\, \boldsymbol{a}\big(\mathcal{S}_q^{\mathrm{range}}(\boldsymbol{x})\big)\right\rangle\right|^2 \tag{6.47}$$

$$= \operatorname*{arg\,max}_{\boldsymbol{x}\in\mathcal{X}} \kappa_\Sigma\big(r_q^* + \sigma_{\mathrm{u}} u_q^* - \mathcal{S}_q^{\mathrm{range}}(\boldsymbol{x})\big) \tag{6.48}$$

Now defining the error term $\boldsymbol{e} := (e_1, \cdots, e_Q)^\top$ such that

$$e_q := r_q^* + \sigma_{\mathrm{u}} u_q^* - \mathcal{S}_q^{\mathrm{range}}(\boldsymbol{x}^* + \sigma_{\mathrm{u}}\boldsymbol{v}^*), \tag{6.49}$$

that we shall later bound by means of Lem. 6.6, it follows

$$\hat{\boldsymbol{x}}_{\mathrm{S}} = \operatorname*{arg\,max}_{\boldsymbol{x}\in\mathcal{X}} \kappa_\Sigma\big(\mathcal{S}^{\mathrm{range}}(\boldsymbol{x}^* + \sigma_{\mathrm{u}}\boldsymbol{v}^*) + \boldsymbol{e} - \mathcal{S}^{\mathrm{range}}(\boldsymbol{x})\big) \tag{6.50}$$

$$= \operatorname*{arg\,max}_{\boldsymbol{x}\in\mathcal{X}} \kappa_\Sigma\big(\boldsymbol{\theta}^* + \boldsymbol{e} - \mathcal{S}^{\mathrm{range}}(\boldsymbol{x})\big). \tag{6.51}$$

Given that $\mathcal{S}$ is a bijection from $\mathcal{X}$ to $\mathcal{T}$, and using the notations from (6.45), calculating $\hat{\boldsymbol{x}}_{\mathrm{S}}$ from the above is tantamount to finding $\hat{\boldsymbol{\theta}}$ from

$$\hat{\boldsymbol{\theta}} := \operatorname*{arg\,max}_{\boldsymbol{\theta}\in\mathcal{T}} \kappa_\Sigma\big(\boldsymbol{\theta}^* + \boldsymbol{e} - \boldsymbol{\theta}\big). \tag{6.52}$$

Ultimately, we can write

$$\hat{\boldsymbol{\delta}} := \operatorname*{arg\,max}_{\boldsymbol{\delta}\in\mathcal{T}-\boldsymbol{\theta}^*} \kappa_\Sigma\big(\boldsymbol{\delta} + \boldsymbol{e}\big), \tag{6.53}$$

and, with $\hat{\boldsymbol{\delta}}' := \hat{\boldsymbol{\delta}} + \boldsymbol{e}$,

$$\hat{\boldsymbol{\delta}}' := \operatorname*{arg\,max}_{\boldsymbol{\delta}\in\mathcal{T}-(\boldsymbol{\theta}^*-\boldsymbol{e})} \kappa_\Sigma\big(\boldsymbol{\delta}\big), \tag{6.54}$$

which is *uniquely* connected to our initial problem through $\hat{\boldsymbol{x}}_{\mathrm{S}} = \mathcal{S}^{-1}\big(\boldsymbol{\theta}^* + \boldsymbol{e} + \hat{\boldsymbol{\delta}}'\big)$. While $\mathcal{T} - \boldsymbol{\theta}^*$ contains 0 —since $\boldsymbol{\theta}^* \in \mathcal{T}$ by definition—, $\mathcal{T} - (\boldsymbol{\theta}^* - \boldsymbol{e})$ is unlikely to contain it. We know however from this reasoning that $\boldsymbol{e} \in \mathcal{T} - (\boldsymbol{\theta}^* - \boldsymbol{e})$.

Since there exist positive constants r, L, U such that the squared kernel $|\kappa(\cdot)|^2$ associated to the Fourier atom $\boldsymbol{a}^{\mathrm{in}}$ is a lobe-r function with bounding constants L and U (see Ex. 5.5), Lem. 6.5 guarantees that κ_Σ is a lobe-$\tilde{r}$

with the constants

$$\tilde{L} = \gamma_{\max} L, \quad \tilde{U} = \gamma_{\min} U, \qquad \tilde{r} = \sqrt{\frac{\gamma_{\min} U}{\gamma_{\max} L}} r \tag{6.55}$$

We can now use Thm. 6.3 on the problem (6.54). Given that $\boldsymbol{e} \in \mathcal{T} - (\boldsymbol{\theta}^* - \boldsymbol{e})$, it yields

$$\|\hat{\boldsymbol{\delta}}'\|_2 \leq \sqrt{\frac{\gamma_{\max} L}{\gamma_{\min} U}} \|\boldsymbol{e}\|_2, \tag{6.56}$$

and hence with the triangular inequality, we have

$$\|\hat{\boldsymbol{\delta}}\|_2 \leq \left(\sqrt{\frac{\gamma_{\max} L}{\gamma_{\min} U}} + 1\right) \|\boldsymbol{e}\|_2, \tag{6.57}$$

We end this demonstration by using the Corollary. 6.7 to obtain the upper bound

$$\|\boldsymbol{e}\|_2 \leq \|\boldsymbol{e}\|_1 \leq \frac{Q}{4} \frac{1}{2r_{\min} - 1}. \tag{6.58}$$

Injecting the above in (6.57) and then in (6.46) yields (6.31) and concludes this proof.

6.7.3 Proof of Theorem 6.3

The condition (6.33) guarantees that $\boldsymbol{z}_{\text{ref}}$ lies inside the inner lobe of f. Since by definition, $f(\hat{\boldsymbol{z}}) \geq f(\boldsymbol{z}_{\text{ref}})$, the condition (5.28) implies that $\hat{\boldsymbol{z}}$ is inside the outer lobe of f and hence, the lobe bounding properties (5.29) applies. We can now write

$$1 - L\|\boldsymbol{z}_{\text{ref}}\|_2^2 \leq f(\boldsymbol{z}_{\text{ref}}) \leq f(\hat{f}) \leq 1 - U\|\hat{\boldsymbol{z}}\|_2^2, \tag{6.59}$$

which directly leads to the conclusion of Thm. 6.3 and concludes this proof.

6.7.4 Proof of Lemma 6.4

From a first-order Taylor series error bound, we have

$$\|\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1\|_2 \leq \max_{\boldsymbol{p} \in \mathcal{P}} \frac{(\mathcal{S}(\boldsymbol{p}) - \mathcal{S}(\boldsymbol{p}_1))^\top}{\|\mathcal{S}(\boldsymbol{p}) - \mathcal{S}(\boldsymbol{p}_1)\|_2} \boldsymbol{J}(\boldsymbol{p})(\boldsymbol{p}_2 - \boldsymbol{p}_1) \tag{6.60}$$

$$\leq \max_{\boldsymbol{p} \in \mathcal{P}} \|\boldsymbol{J}(\boldsymbol{p})\|_2 \|\boldsymbol{p}_2 - \boldsymbol{p}_1\|_2. \tag{6.61}$$

As $\|\boldsymbol{J}(\boldsymbol{p})\|_2 = \lambda_{\max}^{1/2}\big(\boldsymbol{J}(\boldsymbol{p})^\top \boldsymbol{J}(\boldsymbol{p})\big)$, the right part of (6.35) is proven. The inverse $\mathcal{S}^{-1} : \mathcal{T} \mapsto \mathcal{P}$ has its Jacobian evaluated in $\boldsymbol{\theta}$ denoted by $\tilde{\boldsymbol{J}}(\boldsymbol{\theta})$. A

similar reasoning leads to

$$\|\boldsymbol{p}_2 - \boldsymbol{p}_1\|_2 \leq \max_{\boldsymbol{\theta}\in\mathcal{T}} \|\tilde{\boldsymbol{J}}(\boldsymbol{\theta})\|_2 \|\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1\|_2 \tag{6.62}$$

The function $\mathcal{S}$ is a bijection and hence $\left(\tilde{\boldsymbol{J}}(\boldsymbol{\theta})^\top \tilde{\boldsymbol{J}}(\boldsymbol{\theta})\right)^{-1} = \boldsymbol{J}(\boldsymbol{p})^\top \boldsymbol{J}(\boldsymbol{p})$. Therefore, we have $\lambda_{\min}^{-1} = \max_{\boldsymbol{\theta}\in\mathcal{T}} \|\tilde{\boldsymbol{J}}(\boldsymbol{\theta})\|_2$, which concludes the proof.

6.7.5 Proof of Lemma 6.5

We first prove that property (5.28) holds for f_Σ with $\tilde{r}$ defined by (6.37). We define

$$\boldsymbol{z}^{\text{in}} := [(\boldsymbol{z}_1^{\text{in}})^\top, \cdots, (\boldsymbol{z}_Q^{\text{in}})^\top]^\top, \tag{6.63}$$

$$\boldsymbol{z}^{\text{out}} := [(\boldsymbol{z}_1^{\text{out}})^\top, \cdots, (\boldsymbol{z}_Q^{\text{out}})^\top]^\top. \tag{6.64}$$

such that $\boldsymbol{z}^{\text{out}} \in \mathcal{Z}^Q \backslash \mathbb{B}_\infty(R)$ and that $\boldsymbol{z}_q^{\text{in}} \in \mathbb{B}_2(r) \subset \mathcal{Z}$ for all $q \in [Q]$. By construction, there is at least one index $q' \in [Q]$ such that $\boldsymbol{z}_{q'}^{\text{out}} \in \mathcal{Z} \backslash \mathbb{B}_\infty(R)$. For such index q', the properties of a lobe-r function guarantee that

$$f(\boldsymbol{z}_{q'}^{\text{out}}) < \min_{\boldsymbol{z}\in\mathbb{B}_1(r)} f(\boldsymbol{z}) \tag{6.65}$$

$$\leq 1 - U \max_{\boldsymbol{z}\in\mathbb{B}_1(r)} \|\boldsymbol{z}\|_2^2 \tag{6.66}$$

$$= 1 - Ur^2. \tag{6.67}$$

For other indexes $q \neq q'$, we have $f(\boldsymbol{z}_q^{\text{out}}) \leq 1$ since 1 is the maximum value of f by definition. Using the above, we can write

$$f_\Sigma(\boldsymbol{z}^{\text{out}}) < 1 - b_{q'} Ur^2. \tag{6.68}$$

$$\leq 1 - b_{\min} Ur^2. \tag{6.69}$$

Moreover, using property (5.29), we have

$$f_\Sigma(\boldsymbol{z}^{\text{in}}) \geq \textstyle\sum_{q=1}^Q b_q (1 - L\|\boldsymbol{z}_q^{\text{in}}\|_2^2) \tag{6.70}$$

$$\geq 1 - \textstyle\sum_{q=1}^Q b_q L \|\boldsymbol{z}_q^{\text{in}}\|_2^2 \tag{6.71}$$

$$\geq 1 - b_{\max} L \|\boldsymbol{z}^{\text{in}}\|_2^2 \tag{6.72}$$

From (6.68) and (6.72), we deduce that any couple $(\boldsymbol{z}^{\text{in}}, \boldsymbol{z}^{\text{out}})$ defined as above exhibits the inequality

$$f_\Sigma(\boldsymbol{z}^{\text{in}}) - f_\Sigma(\boldsymbol{z}^{\text{out}}) > b_{\min} U r^2 - b_{\max} L \|\boldsymbol{z}^{\text{in}}\|_2^2. \tag{6.73}$$

A sufficient condition for $f_\Sigma(\boldsymbol{z}^{\text{in}}) - f_\Sigma(\boldsymbol{z}^{\text{out}})$ to be strictly positive is that $\|\boldsymbol{z}^{\text{in}}\|_2^2 \leq \frac{U}{L}\frac{b_{\min}}{b_{\max}} r^2 := \tilde{r}^2$, which proves (6.37).

Now we prove that the bounding property (5.29) holds for f_Σ with $\tilde{L}$ and $\tilde{U}$. Given $\boldsymbol{z} = [\boldsymbol{z}_1^\top, \cdots, \boldsymbol{z}_Q^\top]^\top \in \mathbb{B}_\infty(R) \subset \mathcal{Z}^Q$, we have that for all $q \in [Q]$, $\boldsymbol{z}_q \in \mathbb{B}_\infty(R) \subset \mathcal{Z}$. Therefore, lower bound in (5.29) applies and we have

$$\begin{aligned} f_\Sigma(\boldsymbol{z}) &\geq 1 - \textstyle\sum_{q=1}^Q b_q L \|\boldsymbol{z}_q\|_2^2 && (6.74)\\ &\geq 1 - b_{\max} L \|\boldsymbol{z}\|_2^2 && (6.75) \end{aligned}$$

which proves (6.38). Similarly, the upper bound in (5.29) applies and we have

$$\begin{aligned} f_\Sigma(\boldsymbol{z}) &\leq 1 - \textstyle\sum_{q=1}^Q b_q U \|\boldsymbol{z}_q\|_2^2 && (6.76)\\ &\leq 1 - b_{\min} U \|\boldsymbol{z}\|_2^2 && (6.77) \end{aligned}$$

which proves (6.39) and concludes this proof.

6.7.6 Proof of Lemma 6.6

This lemma is a consequence of the error bound obtained from the following second-order Taylor series, holding for all $\boldsymbol{x}, \boldsymbol{b} \in \mathbb{R}^n$,

$$\left| \|\boldsymbol{x} + \boldsymbol{b}\|_2 - \left(\|\boldsymbol{x}\|_2 + \left\langle \tfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}, \boldsymbol{b} \right\rangle\right) \right| \leq \tfrac{1}{2} \max_{0 \leq \varepsilon \leq 1} \left| \boldsymbol{b}^\top \boldsymbol{H}(\boldsymbol{x} + \varepsilon \boldsymbol{b}) \boldsymbol{b} \right|, \tag{6.78}$$

where given $\boldsymbol{I}$ to denote the identity $n \times n$ matrix, the Hessian reads $\boldsymbol{H}(\boldsymbol{x}) := \frac{1}{\|\boldsymbol{x}\|_2} \boldsymbol{I} - \frac{1}{\|\boldsymbol{x}\|_2^3} \boldsymbol{x}\boldsymbol{x}^\top$. It directly follows

$$\left| \|\boldsymbol{x} + \boldsymbol{b}\|_2 - \left(\|\boldsymbol{x}\|_2 + \left\langle \tfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}, \boldsymbol{b} \right\rangle\right) \right| \leq \tfrac{1}{2} \frac{\|\boldsymbol{b}\|_2^2}{\|\boldsymbol{x} + \boldsymbol{b}\|_2}. \tag{6.79}$$

By appropriately substituting $\boldsymbol{x}$ by $\boldsymbol{x}-\boldsymbol{x}^{\mathrm{TX}}$ or $\boldsymbol{x}-\boldsymbol{x}_q^{\mathrm{RX}}$ and $\boldsymbol{b}$ by $\sigma_{\mathrm{u}}\boldsymbol{v}$ above to fit the definitions (6.23), we directly get

$$\begin{aligned}\left|\mathcal{S}_q^{\mathrm{range}}(\boldsymbol{x}+\sigma_{\mathrm{u}}\boldsymbol{v})-\left(\mathcal{S}_q^{\mathrm{range}}(\boldsymbol{x})+\mathcal{S}_q^{\mathrm{speed}}(\boldsymbol{x},\boldsymbol{v})\right)\right| \\ \leq \tfrac{1}{4}\frac{\sigma_{\mathrm{u}}^2\|\boldsymbol{v}\|_2^2}{\|\boldsymbol{x}-\boldsymbol{x}^{\mathrm{TX}}+\sigma_{\mathrm{u}}\boldsymbol{v}\|_2}+\tfrac{1}{4}\frac{\sigma_{\mathrm{u}}^2\|\boldsymbol{v}\|_2^2}{\|\boldsymbol{x}-\boldsymbol{x}_q^{\mathrm{RX}}+\sigma_{\mathrm{u}}\boldsymbol{v}\|_2}. \quad (6.80)\end{aligned}$$

Now since $\sigma_{\mathrm{u}}\|\boldsymbol{v}\|_2 \leq \frac{1}{2}$ by the definition (2.24) of $\mathcal{V}$, the above leads to (6.40), which concludes the proof.

Chapter 7

The Grid Hopping Framework

The single-step and the two-step approaches presented in Chapter 6 both come with their advantages and disadvantages. We concluded that although the superior robustness of the single-step strategy provides significantly higher estimation performance when applied to most systems, including Multistatic Radar (MSR), the computational cost it implies makes it seldom used in practice, where real-time and low-cost applications are often targeted. This chapter builds up from the concepts introduced in both Chapters 5 and 6 to derive our main contribution in sensor networks: "Grid Hopping (GH)". The proposed framework encapsulates and extends existing acceleration approaches from the literature related to different applicative fields, including the Source Localization (SL) of sound or from Ultra-Wide Bandwidth (UWB) systems, as well as multistatic radars. The connection of GH with these contributions is reviewed in Sec. 7.1. Practically, GH is detailed in Sec. 7.2 and relies on the interpolation from one grid, which discretizes the sensed parameter's domain, to another grid which corresponds to a set of candidates in the domain of the parameters of interest.

The performance of GH is studied in three parts. First in Sec. 7.3 with the derivation of a novel theoretical bound quantifying the impact of GH on the performance of the single-step methods. In short, this result is a combined extension of both the bounds proposed in Chapter 5 for the selection step using atom interpolation and in Chapter 6 for the single-step

selection restricted to a subset of the domain of interest. Our GH bound enables us to see the separate impacts of (i) the properties of the projection process between the parameter of interest and the sensed parameters (*e.g.,* from a position to a collection of distances), (ii) the intrinsic precision of the single-step methodology and (iii) the precision loss due to GH. These three aspects are respectively connected to the nature and the locations of the sensors, the density of the grid of parameters of interest, and the interpolation strategy from the grid of sensed parameters.

Secondly, we evaluate in Sec. 7.4 the effectiveness of our method by means of extensive Monte-Carlo simulations of multiple systems to which the GH framework applies. Finally, we particularize our method to the factorized BOMP algorithm designed for MSR application, introduced in Sec. 6.3.3 and next, we validate our method by applying it on data resulting from measurement performed using the active MSR system that we designed and presented in Chapter 2.

All the proofs of the theorems and lemmas given in this chapter are gathered in Sec. 7.8

7.1 Related Work on Accelerated One-step Methods

Enhanced single-step methods have been studied independently in different signal-sensing communities, each proposing specific acceleration strategies.

In the MSR field, a single-step method, often referred to as *joint* [BM11, GN11, YGD13] or *direct* [BKY16, HPKS12, YLGD21] arises in connection with the sparse signal recovery literature as an application of the BOMP algorithm proposed by Y. C. Eldar in [EKB10]. As previously highlighted in Sec. 6.3.3, the discretization of the 4D (or 6D) location-velocity space is a limitation that makes such single-step method unsuitable for real-time applications. Factorization approaches such as those suggested in [LP15, MFJV19] and explored in the previous chapter enable the sequential estimation of locations and velocities which significantly reduces the computation time. They remain however not sufficient for practical applications to consider the use of single-step method in real-time applications targeted for multistatic radars.

When applied to the SL field, the single-step method corresponds to the Steered Response Power (SRP) algorithm. Many low-complexity strategies have been proposed to decrease the computational complexity of SRP.

In [DBA07] the authors propose different techniques to restrict the research to a subset of the domain of interest. In [BD20b], a hybrid approach between the single-step and two-step paradigms is developed by leveraging the intrinsic properties of speech signals to first perform a two-step estimation and then correct it with a single-step-like approach to avoid data-association issues. Other techniques such as the stochastic region contraction proposed in [DSY07] or grid rescaling methods from [MCLE13, COH18, BD22] propose to iteratively refine the search space, starting from a coarse search of candidates. In [CML11, LMN+15, BD20a], the authors propose a volumetric technique that enables to represent multiple grid bins into a single one, enabling the search on a coarser grid. In the contributions above, the SRP map is computed from the computation of the Generalized Cross-Correlation (GCC) [KC76]. As shown in ..., this mapping is an approximation which can be assimilated as the most simple particularization of GH.

In [DDSvW21], the authors propose a low-complexity single-step algorithm that exploits the Shannon-Nyquist sampling properties of the speech signals in order to build an interpolation-based mapping from the delay grid bins to the location grid. Grid Hopping appears as a generalization of this work for other applications as well as other interpolation strategies with different properties. The bound derived in Sec. 7.3 enables us to quantify the performance loss caused by the use of GH with different interpolations, including the one presented in [DDSvW21] and the polar interpolation [ETS11] which is popular in sparse continuous signal recovery community [FDJ15, KYHP15].

7.2 Enhancing the One-Step Approaches with Grid Hopping

The presentation of the GH approach is done in three parts. We first provide the reader with the intuition behind the technique by studying a simple example. Next, we formalize the mathematics of GH, enabling to "hop" from the grid Θ_{G} to the grid $\mathcal{P}_{\mathrm{G}}$. This includes a study of the computational complexity of the different algorithms for parameter estimation from sensor networks. Finally, we extend the technique by adding an *off-the-grid* step, built upon the COMP algorithm described in Chapter 5.

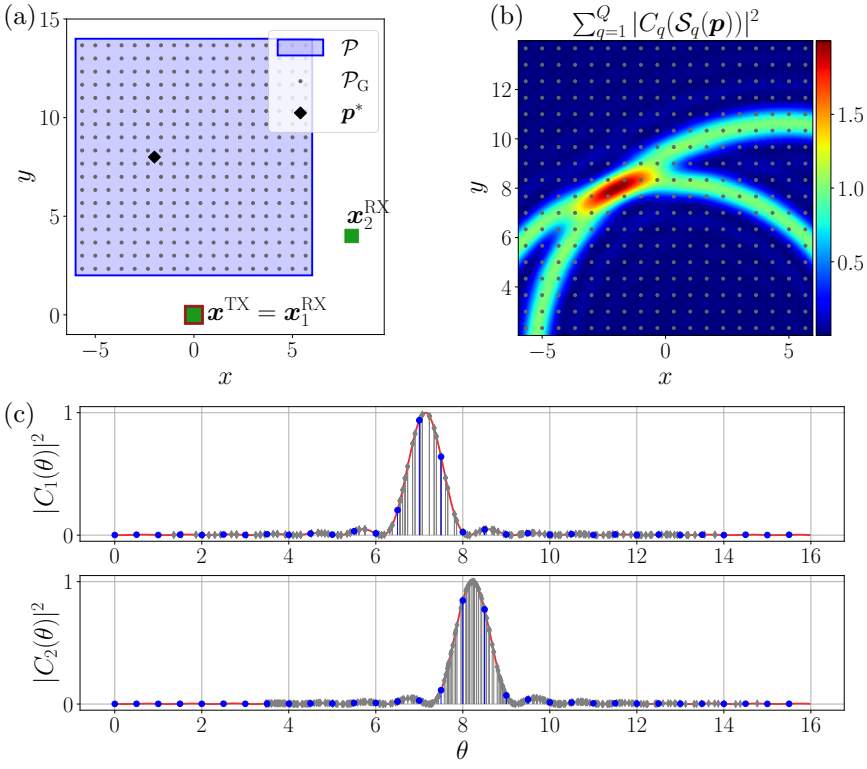

Figure 7.1: **(a)** Representation of a multistatic radar composed of one TX and three RXs. The grid of dots represents the elements of $\mathcal{P}_{\mathrm{G}}$. **(b)** The joint correlation function used in (6.11). The grid $\mathcal{P}_{\mathrm{G}}$ is also shown in grey. **(c)** The correlation function of the first step (6.8) for $q = 1$ (top) and $q = 2$ (bottom) corresponding to two of the three RXs in the situation described in (a). The big blue dots represent the samples taken on the grid Θ_{G}. The small grey dots are the samples taken on the grids $\mathcal{S}_q(\mathcal{P}_{\mathrm{G}})$.

7.2.1 Motivation for Grid Hopping

Let us first motivate and hint at the idea that resides behind GH by observing an example of the application of the single-step and two-step methods. We consider the case of a simple multistatic radar composed of one TX and two RXs, located according to Fig. 7.1(a). The system acquires a single chirp in which $M = 16$ samples are taken in order to determine the location $\boldsymbol{p}^*$ of a single target.

Fig. 7.1(b) shows an example of a resulting correlation function used by the single-step's selection. Due to the potentially high non-convexity of this

function, the problem of finding its maximum usually requires evaluating it on the grid $\mathcal{P}_{\mathrm{G}}$, represented by the gray dots in this figure. From this on-grid estimate, one is either satisfied or chooses to refine it off-the-grid. This is typically done for example by means of a gradient ascent or with an interpolation-based method following the ideas presented in Chapter 5. In each of these cases, evaluating the global correlation on the grid $\mathcal{P}_{\mathrm{G}}$ is a required step.

In Fig. 7.1(c), we observe the correlations that one maximizes when using a two-step approach. The blue grid represents the sensing grid Θ_{G} which discretizes the sensing domain Θ with an equivalent spacing as the grid $\mathcal{P}_{\mathrm{G}}$ discretizes $\mathcal{P}$. Formally, this grid density aspect is quantified using the *covering radius* defined by the following.

Definition 7.1 (Covering radius)**.** Given a subspace $\mathcal{Z}$, the grid $\mathcal{Z}_{\mathrm{G}}$ is said to discretize $\mathcal{Z}$ with a *covering radius* defined by

$$\max_{\boldsymbol{z}\in\mathcal{Z}} \min_{\boldsymbol{z}'\in\mathcal{Z}_{\mathrm{G}}} \|\boldsymbol{z}-\boldsymbol{z}'\|_2. \tag{7.1}$$

To motivate the principle of GH, we added in Fig. 7.1(c) the transformation of $\mathcal{P}_{\mathrm{G}}$ through the projection functions $\mathcal{S}_q$. Computationally speaking, the full grid search used to *exactly* solve the single-step problem (6.11) restricted to the grid $\mathcal{P}_{\mathrm{G}}$ requires, for each index $q \in [Q]$, to evaluate $C_q(\theta)$ for all θ taken from the grid $\mathcal{S}_q(\mathcal{P}_{\mathrm{G}})$ of cardinality N_P. With GH, we thus propose to use atom interpolation to "hop" from the grid Θ_{G} onto the grid $\mathcal{S}_q(\mathcal{P}_{\mathrm{G}})$ for the evaluations of C_q. This operation is expected to provide a good approximation of the exact single-step grid search with a computational complexity with the same order of magnitude as exhibited by two-step methods, which are commonly faster. Fig. 7.2 illustrates this idea, as applied to our simple example. The yellow curves correspond to the approximations obtained using the interpolated atoms instead of the actual ones in the correlation processes. Since any value in these approximated curves can be computed with no significant computation time from the blue grid, we expect this process to be faster than the computation of all the grey points.

Intuitively, the performance of GH can be motivated by the following reasoning. Given a ground truth $\boldsymbol{p}^*$, the precision of the estimate $\hat{\boldsymbol{p}}$ resulting from (6.11) *restricted to* $\mathcal{P}_{\mathrm{G}}$ is intrinsically limited by the density of this grid. Yet, as we observe respectively in Fig.7.1(b) and (c), the grids $\mathcal{S}_q(\mathcal{P}_{\mathrm{G}})$

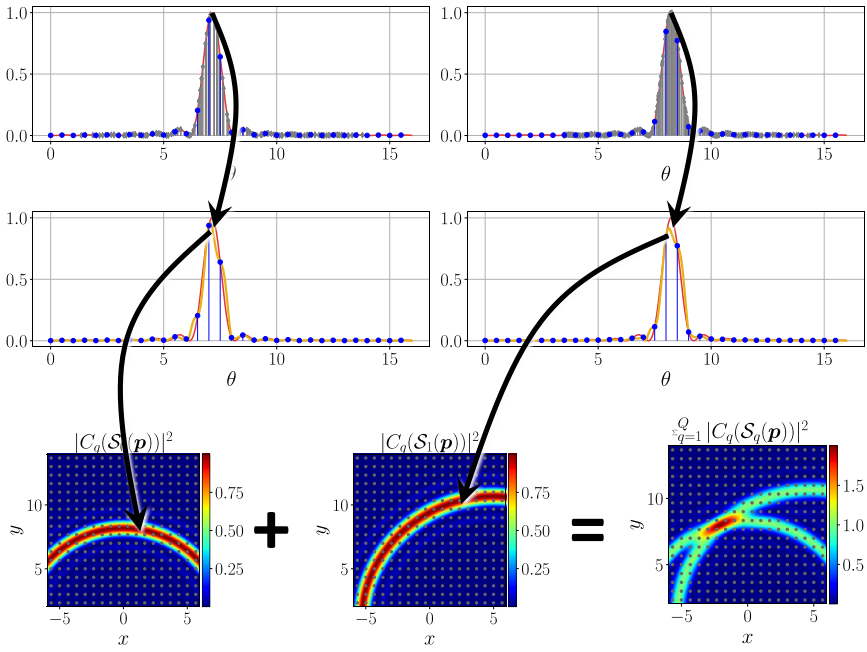

Figure 7.2: Intuitive representation of the grid hopping principle applied to the simple example from Fig. 7.1.

discretize Θ more densely than $\mathcal{P}_{\mathrm{G}}$ discretizes $\mathcal{P}$. Therefore, we expect —and later quantify with the bound derived in Sec. 7.3— that the error caused by an interpolation from the lower density grid Θ_{G} represented in Fig.7.1(b) to $\mathcal{S}_q(\mathcal{P}_{\mathrm{G}})$ is often negligible with respect to the limitation of precision intrinsically caused by the gridding $\mathcal{P}_{\mathrm{G}}$ of $\mathcal{P}$.

7.2.2 Hopping from One Grid to Another

We continue our exploration of GH by keeping the focus on the first iteration of BOMP. We leave for the next section the full description of an adaptation of BOMP with GH.

Implementing the principle exposed in the previous section, GH starts with the evaluation of the set $C_q(\Theta_{\mathrm{G}})$ for each $q \in [Q]$, like in the first step of the two-step method (6.8). Next, the set $C_q(\mathcal{S}_q(\mathcal{P}_{\mathrm{G}}))$ is approximated using the interpolation principle from Chapter 5. In practice, for all sensor indexed by $q \in [Q]$, each grid bin $\bar{\boldsymbol{p}}_n \in \mathcal{P}_{\mathrm{G}}$ with $n \in [N_P]$ is mapped to

one bin taken from Θ_{G}. This mapping operation amounts to the *offline* computation of Q lookup tables, each providing a connection between $\mathcal{P}_{\mathrm{G}}$ and Θ_{G} by minimizing

$$\min_{\ell \in [N_\Theta]} |\bar{\theta}_\ell - \mathcal{S}_q(\bar{\boldsymbol{p}}_n)|, \tag{7.2}$$

where we remind that N_Θ denotes the cardinality of Θ_{G}.

Next, for all $q \in [Q]$ and for all $n \in [N_P]$, a set of interpolating coefficients $\boldsymbol{c}$ are computed offline according to a chosen interpolation pattern. With ℓ denoting the index in Θ_{G} that is associated to n by (7.2), and using the notation of Def. 5.1, those coefficients are such that

$$\boldsymbol{a}(\mathcal{S}_q(\bar{\boldsymbol{p}}_n)) \simeq \tilde{\boldsymbol{a}}(\mathcal{S}_q(\bar{\boldsymbol{p}}_n)) = \boldsymbol{A}_\ell \boldsymbol{c}. \tag{7.3}$$

Once the mapping and the set of interpolating coefficients are determined offline, GH amounts to using $\tilde{\boldsymbol{a}}$ instead of $\boldsymbol{a}$ in the evaluation of the correlation functions. Thus, the following interpolated correlation function is computed online,

$$\tilde{C}_q(\mathcal{S}_q(\bar{\boldsymbol{p}}_n)) := \langle \boldsymbol{y}_q, \tilde{\boldsymbol{a}}(\mathcal{S}_q(\bar{\boldsymbol{p}}_n)) \rangle = \sum_{i=1}^{I} c_i \, C_q(\bar{\theta}_{\ell,i}). \tag{7.4}$$

This ultimately leads to the grid-hopping estimate

$$\hat{\boldsymbol{p}} = \underset{\boldsymbol{p} \in \mathcal{P}_{\mathrm{G}}}{\arg\min} \sum_{q=1}^{Q} |\tilde{C}_q(\mathcal{S}_q(\boldsymbol{p}))|^2. \tag{7.5}$$

In practice, for each index q, we store the interpolating coefficients (denoted by c_i in (7.4)) in a large sparse matrix $\boldsymbol{C}_q \in \mathbb{C}^{N_P \times N_\Theta}$. This matrix thus contains I non-zero component in each row, at locations driven by the lookup table computed according to (7.2).

Overall, the QMN_P products required for the grid search of the one-step estimation (6.11) is replaced by the computation of QMN_Θ products for the calculation of the Q sets of correlations $C_q(\Theta_{\mathrm{G}})$ plus additional QIN_P products for the interpolation (7.4). Because in non-degenerate conditions $I \ll M$ and $N_\Theta \ll N_P$, the GH-based grid search is expected to be computed faster than the standard single-step on-grid selection step.

We illustrate this effect with Fig. 7.3. We obtained those heatmaps using a simulated measurement based on the MSR system whose geometry is defined in Fig. 6.6. We used GH with a polar interpolation from a grid Θ_{G} with a spacing 1/8 onto a grid of interest $\mathcal{P}_{\mathrm{G}}$ with a regular spacing

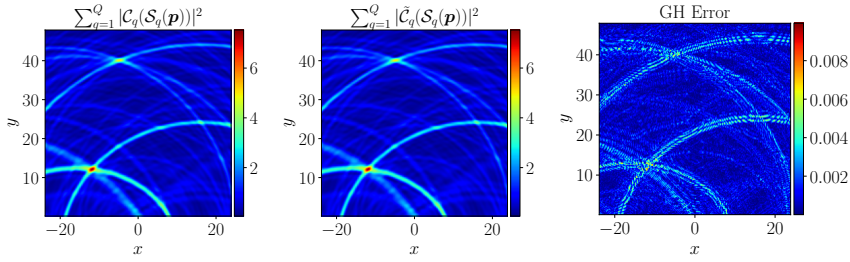

Figure 7.3: Example of the application of GH in the selection step of the first iteration of BOMP. The GH Error corresponds to the absolute difference between the exact (left) and the approximated (center) correlation.

1/4. Computing the exact single-step correlation took 12ms while the GH approximation only required 0.8ms to calculate. While this improvement is significant, it only accelerates the estimation of $\boldsymbol{p}^*$ *restricted to the grid* $\mathcal{P}_\mathrm{G}$. In the next section, we propose to use the interpolation a second time to efficiently leave this grid without dominating the computation time of the accelerated grid search that we just presented.

7.2.3 Hopping Off-the-Grid

Following the principle of our proposition of interpolation-based Continuous OMP, detailed in Alg. 3, we interpret the above *on-the-grid* estimation as a first step that can later be refined without additional computation of correlations. When GH is used, we dispose of the values of C_q evaluated on Θ_G. Therefore, in this section, we suggest applying an *off-the-grid* procedure that leaves the grid $\mathcal{P}_\mathrm{G}$ for the continuous space $\mathcal{P}$ while working with interpolations defined in the sensing domains Θ of each signal. Not only does this choice enable the use of simple interpolation patterns, already studied in Chapter 5, but it also leads to a fast way to interpolate the global correlation function outside $\mathcal{P}_\mathrm{G}$ by only using values already computed.

Practically, given that the $\hat{n}$-th bin of $\mathcal{P}_\mathrm{G}$ has been selected by the grid search (7.5), we refer to the lookup table (computed offline following (7.2)) and collect index ℓ_q that connects $\hat{n}$ to the sensing grid, for all $q \in [Q]$. Next, defining

$$\boldsymbol{g}_q := \boldsymbol{U}^\top \boldsymbol{A}_{\ell_q}^\mathrm{H} \boldsymbol{y}_q, \tag{7.6}$$

we interpolate $\hat{\boldsymbol{p}}$ off-the-grid with

$$\hat{\boldsymbol{p}} := \max_{\boldsymbol{p}\in\mathcal{P}} \sum_{q=1}^{Q} \left|\left\langle \boldsymbol{g}_q,\, \Psi\big(\mathcal{S}_q(\boldsymbol{p}) - \bar{\theta}_{\ell_q}\big)\right\rangle\right|. \tag{7.7}$$

Since $\boldsymbol{U}$ is computed offline and $\boldsymbol{A}_{\ell_q}^{\mathrm{H}}\boldsymbol{y}_q$ amounts to take I samples from the set $C_q(\Theta_{\mathrm{G}})$, retrieving the values $\boldsymbol{g}_q$ requires no computation time. We propose to solve (7.15) using a gradient ascent initialized in $\bar{\boldsymbol{p}}_{\hat{n}}$, denoting the on-grid estimate given by (7.5). Since the applications detailed in Sec. 6.3 yield PPFs with simple closed forms for both their expression and their derivatives, we use a Newton step in this ascent. Note that given a sufficiently dense grid, the initialization is commonly such that the procedure converges in a few iterations, that we limit arbitrarily to 10. We emphasize that by construction, its complexity is insensitive to the dimension of the problem, *i.e.,* M, N_Θ and N_P. That is, the complexity of evaluating the objective —and its derivatives— of (7.7) only depends on the number of processed signal Q and the order of interpolation I. As the limited number of evaluations is firmly limited, we conclude that the additional computation time of this step becomes negligible when compared to the grid search for a sufficiently large scale of the estimation problem.

Let us finally fully express the resulting extension of BOMP, endowed with Grid Hopping on and off the grid. The resulting algorithm is detailed in Alg. 5, given measurements that follow model (6.7).

7.3 Grid Hopping Estimation Error Bound

Now that we presented the GH framework, we evaluate its performance by first derivating a theoretical bound on the estimation error resulting from its application in a simple scenario. Note that even in the *off-the-grid* extension detailed in Alg. 5, the grid selection step (7.9) remains the critical aspect since the off-grid shift is computed around this estimate. Therefore, combining the messages that we previously conveyed in both Thm. 5.3 and Thm. 6.3, we study an upper bound on the *on-the-grid* GH estimation error. Precisely, we focus on the *noiseless* observations of *one* parameter $\boldsymbol{p}^* \in \mathcal{P}$ (*e.g.,* a single source to localize). Given the estimate $\hat{\boldsymbol{p}}$ provided by (7.5), we derive an upper bound on the value of $\|\hat{\boldsymbol{p}} - \boldsymbol{p}^*\|_2$. This bound is given with respect to the properties of the grids Θ_{G} and $\mathcal{P}_{\mathrm{G}}$ as well as the interpolator used for Grid Hopping (GH), showing the effect of each processing step.

Algorithm 5: Grid-Hopping Continuous BOMP

Input : K, $\boldsymbol{y}_q$, $\boldsymbol{A}$,lookup tables $\mathcal{L}_q$ and lookup sparse matrices $\boldsymbol{C}_q$

Output: $\hat{\boldsymbol{p}}_1, \cdots, \hat{\boldsymbol{p}}_K$

begin

Initialization: $\boldsymbol{r}_q \longleftarrow \boldsymbol{y}_q$ for $q \in [Q]$;

for $k = 1$ **to** K **do**

1. Selection Step

$$\boldsymbol{z}_q = \boldsymbol{A}^{\mathrm{H}} \boldsymbol{r}_q, \quad \forall q \in [Q] \tag{7.8}$$

$$\hat{n}_k = \arg\max_{n \in [N_P]} \sum_{q=1}^{Q} \left| \boldsymbol{C}_q \boldsymbol{z}_q \right|^2 \tag{7.9}$$

2. Fitting Step

$$\hat{\boldsymbol{\beta}}_q \longleftarrow \arg\min\nolimits_{\boldsymbol{\beta} \in \mathbb{C}^{kI}} \tfrac{1}{2} \left\| \boldsymbol{y} - \left[\boldsymbol{A}_{\ell_{q,1}}, \cdots, \boldsymbol{A}_{\ell_{q,k}} \right] \boldsymbol{\beta} \right\|_2^2, \quad \forall q \in [Q]$$

where $\ell_{q,1}, \cdots, \ell_{q,k} = \mathcal{L}_q(\hat{n}_k)$.

3. Residue Update

$$\boldsymbol{r}_q \longleftarrow \boldsymbol{y}_q - \left[\boldsymbol{A}_{\ell_{q,1}}, \cdots, \boldsymbol{A}_{\ell_{q,k}} \right] \hat{\boldsymbol{\beta}}_q, \quad \forall q \in [Q] \tag{7.10}$$

4. Off-grid Shift

$$\hat{\boldsymbol{p}} := \max_{\boldsymbol{p} \in \mathcal{P}} \sum_{q=1}^{Q} \left| \left\langle \boldsymbol{g}_q, \Psi\left(\mathcal{S}_q(\boldsymbol{p}) - \bar{\theta}_{\hat{\ell}_{q,k}} \right) \right\rangle \right|, \quad \forall k \in [K] \tag{7.11}$$

where $\boldsymbol{g}_q$ are obtained by (7.6).

In the following theorem, we use a notation for the normalized amplitude coefficients, defined as

$$\gamma_q = |\alpha_q|^2 \,/\, \textstyle\sum_{q'=0}^{Q} |\alpha_{q'}|^2, \tag{7.12}$$

and we set $\gamma_{\max} := \max_{q \in [Q]} \gamma_q$, as well as $\gamma_{\min} := \min_{q \in [Q]} \gamma_q$.

Theorem 7.1 (Grid Hopping Bound)**.** *Let us assume that*

(A.1) the atom function $\boldsymbol{a}$ *is translation invariant and such that* $|\kappa(\cdot)|^2$*, where* κ *is defined as in* (7.1)*, is a lobe-r function with* L*,* U *as lower and upper bound constants respectively,*

(A.2) the grid $\mathcal{P}_{\mathrm{G}}$ *discretizes* $\mathcal{P}$ *with a covering radius* ρ*, and*

(A.3) We use an interpolation scheme with an interpolation gap ξ*.*

Given measurements provided by (6.6) *with* no noise*, the following bound on the estimate* $\hat{\boldsymbol{p}}$ *resulting from the estimation* (7.5) *holds. Given the constants*

$$R_{\mathrm{P}} := \sqrt{\tfrac{L\gamma_{\max}}{U\gamma_{\min}}}, \qquad \tilde{U} := U\gamma_{\min}, \qquad \tilde{R}_{\mathrm{P}} := R_{\mathrm{P}}/\sqrt{\tilde{U}}, \tag{7.13}$$

if

$$(\lambda_{\max} R_{\mathrm{P}}^2)\,\rho^2 \leq r^2 - \tfrac{2\sqrt{2}}{\tilde{U}}\xi, \tag{7.14}$$

then

$$\|\hat{\boldsymbol{p}} - \boldsymbol{p}^*\|_2 \leq \lambda_{\min}^{-\frac{1}{2}} \left(2\tilde{R}_{\mathrm{P}}\,\xi + \sqrt{\tfrac{2}{\tilde{U}}(2R_{\mathrm{P}}^2 + 1)\,\xi^2 + (\lambda_{\max} R_{\mathrm{P}}^2)\,\rho^2} \right). \tag{7.15}$$

The proof of Thm. 7.1 is provided in Sec. 7.8.1. In words, it tells us that given a grid with a coverage radius ρ sufficiently smaller than the inner lobe r, the estimation error between $\hat{\boldsymbol{p}}$ and $\boldsymbol{p}^*$ is upper bounded by a weighted sum of the interpolation gap ξ and an ℓ_2 combination of ξ and the grid gap ρ. We remind for interpretation purposes that the radius r is here associated with the squared kernel function $|\kappa(\cdot)|^2$ of one sensor taken alone and hence quantifies the intrinsic resolution of each sensor.

If GH is not used, hence leading to the classic single-step method, then the only remaining error is the *gridding error*, proportional to ρ. If, on the other hand, the GH is used to estimate $\boldsymbol{p}$ in a continuous space — *e.g.,* with an infinitely dense grid $\mathcal{P}_{\mathrm{G}}$— then the gridding error is zero and only interpolation error remains, proportional to ξ. Thm. 7.1 confirms the intuition that since GH is a combination of these two scenarios, it results in an error that asymptotically linearly combines their respective bounding errors.

The factor $\lambda_{\min}^{-\frac{1}{2}}$ translates the impact on the geometry of the parameter projection function $\mathcal{S}$, which is often related to the sensor network's ability to observe the parameter of interest diversely. Notably, a poorly conditioned

network implies that $\lambda_{\min}$ is close to 0. The value of $\lambda_{\min}$ reaches 0 when one or more values of $\boldsymbol{p} \in \mathcal{P}$ cannot be estimated unambiguously by the system. This happens, for example, if one places all the receivers of any localization system in the same location.

It is also noticeable from this bound that the imbalance between the power levels of the received signals (*i.e.,* $\gamma_{\max} := \max_{q\in[Q]} \gamma_q \gg \gamma_{\min} := \min_{q\in[Q]} \gamma_q$) negatively impacts the estimation performance in single-step methods. Indeed, this aspect is highlighted in the definition of the term R_{P}, which is minimized when all the received signals exhibit similar power (with ideally $\gamma_{\max} := \max_{q\in[Q]} \gamma_q = \gamma_{\min} := \min_{q\in[Q]} \gamma_q$). We discuss in Chapter 8 hints of strategies to approximately reach this optimal aspect, opening perspectives of future work on the topic.

The bound also enables us to study the compromise between using dense grids or higher-order interpolation schemes. For the following reasoning, we remind the reader that the computation of $\hat{\boldsymbol{p}}$ with GH has a computational complexity proportional to $N_\Theta M + N_P I$. If the term is ρ dominates the error, then we can only decrease ρ to significantly improve the performance, increasing N_P. If, however, the term in ξ dominates the error, then we can decrease this term either by using a more precise interpolation scheme (*e.g.,* using a polar interpolation instead of a nearest-neighbourg strategy) or by increasing the density of the observation grid Θ_{G}. Computationally speaking, in the first case, we increase I; in the other case, we increase N_Θ. Therefore, for a given signal length M, determining the best choice between these two ways of decreasing ξ depends on the value of N_P. If $N_\Theta M$ dominates the computation time, then a higher-order interpolation scheme will be preferred. If $N_P I$ is high, then one should use a low-order interpolation scheme with a denser grid Θ_{G}. This effect will be observed in the simulation results shown in the next section.

7.4 Numerical Results

This section presents numerical results from extensive Monte-Carlo simulations. First, we compare the bound derived in Sec. 7.3 with results obtained from the simulation of the simple MSR application presented to motivate GH in Sec. 7.2, given a simplified noiseless scenario which fits the conditions for which Thm. 7.1 applies. Next, we assess the effectiveness of GH by means of simulations performed for two applications presented in Sec. 6.3 and using several different interpolation schemes described in Sec. 7.2.

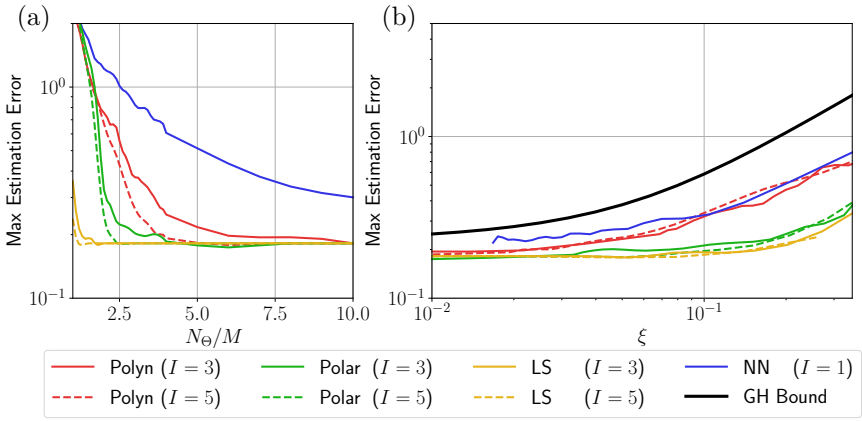

Figure 7.4: **(a)** Comparison of the maximal squared error $\|\hat{\boldsymbol{p}}-\boldsymbol{p}^*\|_2^2$ observed over 10,000 realizations of $\boldsymbol{p}^*$ for multiple inteprolation schemes used in GH. **(b)** The same results as in (a) plotted with respect to the interpolation gap ξ given in Fig. 5.4.

7.4.1 Evaluation of the GH Bound

For this simulation, we randomly selected 10,000 realizations of the true parameter of interest $\boldsymbol{p}^*$, each time taken following a uniform distribution in the domain $\mathcal{P}$ depicted in Fig. 7.1(a). For each selected value of $\boldsymbol{p}^*$, the collection of measurements $\{\boldsymbol{y}_q\}_{q=1}^Q$ were computed according to a noiseless scenario with equal power in each signal. This latest simplification leads to $\gamma_{\min} = \gamma_{\max} = 1/Q$. For each realization, the GH method was used with all the interpolation schemes that we compared in Fig. 5.4, outputting an estimate $\hat{\boldsymbol{p}}$ of $\boldsymbol{p}^*$. These estimations were performed for several *uniform* griddings Θ_{G} of Θ and griddings $\mathcal{P}_{\mathrm{G}}$ of $\mathcal{P}$. The grids $\mathcal{P}_{\mathrm{G}}$ all have a structure similar to the grid from Fig. 7.1(a) and are defined with different spacing between adjacent bins, directly proportional to the covering radius ρ.

For each grid and each interpolation scheme, we saved the realization which provided the worst (higher) estimation squared error $\|\hat{\boldsymbol{p}} - \boldsymbol{p}^*\|_2^2$. Fig. 7.4(a) depicts these worst-case errors with respect to the density of Θ_{G} expressed by N_Θ/M, while the density of $\mathcal{P}_{\mathrm{G}}$ is fixed with $\rho = 1/8$. In comparison with Fig. 5.4, we observe that, as expected, the interpolation schemes that exhibit lower interpolation gaps ξ lead to a lower worst-case error. As N_Θ increases, the error decreases according to the value of ξ shown in Fig. 5.4 before reaching the grid error, which results from the limited density of $\mathcal{P}_{\mathrm{G}}$.

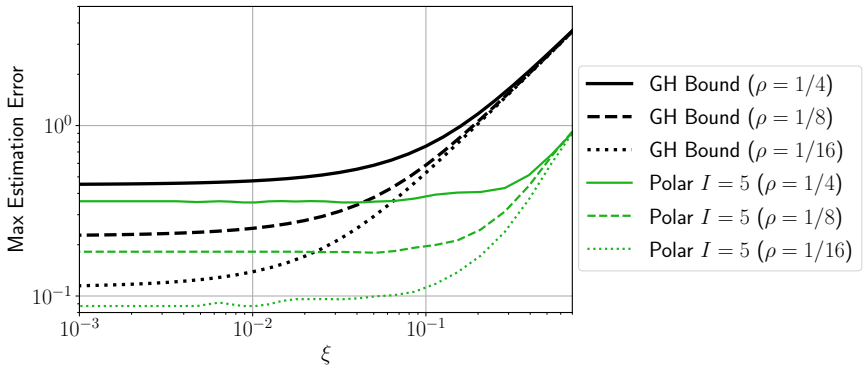

Figure 7.5: Comparison of the GH bound with the maximal squared error $\|\hat{\boldsymbol{p}} - \boldsymbol{p}^*\|_2^2$ obtained using GH with the order-5 polar interpolation [CH19]. Multiple uniform griddings $\mathcal{P}_{\mathrm{G}}$, with three different covering radius ρ are compared.

The link between the worst-case squared error and the interpolation gap ξ is better represented by Fig. 7.4(b) where the same data as in Fig.7.4(a) are plotted with respect to the values of ξ taken from Fig. 5.4. In this figure, we added the GH bound which has been computed given the constant covering radius $\rho = 1/8$. Similar results, restricted to the polar interpolation, are depicted in Fig. 7.5 for multiple values of ρ. Each curve exhibits two regimes. The left flat part corresponds to values of ξ such that the interpolation term dominates the error while the right increasing part corresponds to the regime where the term in ρ dominates the bound's expression. The same two regimes with equivalent slopes appear with the simulated results. The vertical shift between the theoretical bound and the results is caused by a constant factor while the horizontal shift translates a mismatch between the weightings of ρ and ξ in the expression of the bound with respect to the actual effective weightings given an interpolation scheme. Those are due to the unlikeliness of all inequalities of the bound to be tight simultaneously in a single value of $\boldsymbol{p}^*$. Nonetheless, the weighted-sum property of ρ^2 and ξ^2 is preserved when moving from the theory to the simulations.

7.4.2 Grid-Hopping Performance

We validate now the ability of GH to provide a good compromise between existing methods by comparing it to the two-step and the single-step methods proposed in Sec. 6.2. To this end, we simulated two different systems to which GH can apply under noisy conditions. While multiple contributions in the literature of each field have already shown the superiority of

single-step methods with respect to the two-step methods under sophisticated nuisance, including reverberation [DSB01, P.03, ZFZ08, BKY16], we restrict here our study to the robustness of the estimation against white noise. This simple case enables us to compare the effectiveness of GH, when used with different interpolations, to attain performance comparable to those of single-step methods with a computational complexity closer to two-step methods. The evaluation of GH in the presence of other nuisances is kept for future work.

The first application is a MSR with the same characteristic as the one described in Fig. 6.6(a). The other application is a network of *circular* sensor arrays such as depicted by Fig. 6.6(b). In both cases, we have measurements $\boldsymbol{y}_q$ each with $M = 64$ components.

We start with the results from the MSR's simulation. For multiple values of SNR, we averaged the estimation error $\|\hat{\boldsymbol{p}} - \boldsymbol{p}^*\|_2$ over 10,000 realizations of noisy signals (see model 6.6) in which $\boldsymbol{p}^*$ have been taken uniformly at random in $\mathcal{P}$ and the coefficients α_q have been taken independently for each index q from a complex normal distribution of variance equal to 1.

Fig. 7.6(a) shows the results of this simulation when the covering radius ρ of $\mathcal{P}_{\mathrm{G}}$ is equal to $1/4$ and when $N_\Theta/M = 4$. We observe first that the single-step method provides stronger robustness with respect to noise than its two-step counterpart. When the SNR is high, the performance saturates as it reaches the limitation due to the gridding $\mathcal{P}_{\mathrm{G}}$. The GH curves follow the one of the single-step but saturate sooner to reach, at high SNR, performances which depend on the choice of interpolation. The nearest neighbor strategy logically leads to the worst GH performance while higher order interpolations exhibit performance so close to the single-step algorithm that their curves cannot be distinguished in this figure.

Fig. 7.6(b) and (c) respectively show the average estimation error and the computation time with respect to the number N_Θ of grid bins of the sensing domain with a constant SNR set to 20dB. Increasing N_Θ leads to a smaller value of ξ for a given interpolation pattern of order I, but increases the computation time required for GH. Since in this situation, the term in $N_\Theta M$ is significant in the computation time of GH methods, increasing N_Θ deteriorates the speed of the algorithm. It results that the polar interpolation with $N_\Theta/M \simeq 4$ appears as potentially the best compromise between the performance and the computation time. Note that the LS interpolator requires a higher computation time than the polar one because it requires complex interpolating coefficients and hence, is virtually of order $I = 6$ instead of 3 as far as the computation time is concerned.

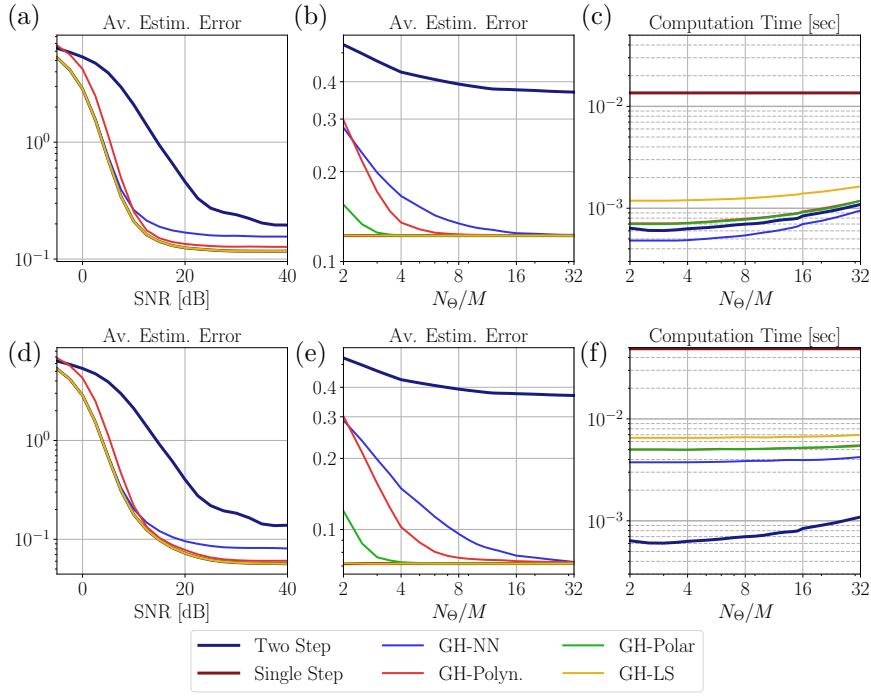

Figure 7.6: Simulation results from the MSR system showing the average of the estimation error $\|\hat{\boldsymbol{p}} - \boldsymbol{p}^*\|_2$ over 10,000 realizations of noisy signals following (6.6). The polynomial, the polar, and the LS interpolators used in GH for this simulation are all of order $I = 3$. In (a), the performance of the polar and LS interpolation-based GH are merged with the performance of the single-step method. In (f), the performance of the polynomial and LS interpolation-based GH are merged with the performance of the single-step method. The computation time of the polynomial is the same as the computation time for the polar method. **(a)-(c)** for $\rho = 1/4$. **(d)-(f)** for $\rho = 1/8$.

Let us remark that NN-based grid hopping method has been computed faster than the two-step method in this case. Although this may seem surprising, this happens because we limited the scale of the problem to $M = 64$, resulting in a computation for the multilateration step (6.10) that potentially surpasses the one of GH. If one increases the dimensionality of the problem, we expect to observe the computation time of the two-step method to be affected by a smaller increase than the GH-based algorithms, thereby leading to the conclusion that the two-step method is the fastest. In this higher dimension scenario, we also expect the single-step method's computation time to skyrocket. In fact, we set the limitation $M = 64$ in

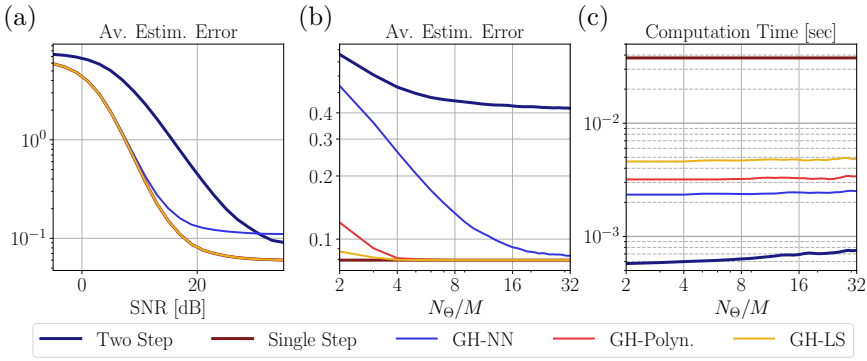

Figure 7.7: Simulation results from the DoA-based localization system following the same principle as Fig. 7.6, for $\rho = 1/8$.

order to maintain a scalable memory requirement for the exact single-step algorithm. Indeed, in order to fit real applications, we scale the location domain $\mathcal{P}$ with the radar's maximal non-ambiguous range (see the rule given by (2.23)). It results that in 2D, N_P is proportional to M^2 ans hence the single-step dictionaries $\{\boldsymbol{a}(\theta) : \theta \in \mathcal{S}_q(\mathcal{P}_\mathrm{G})\}$ are matrices that scale in $\mathcal{O}(M^3)$. Therefore, storing them appears as another limitation of the exact single-step method which is solved with grid hopping.

In Fig. 7.6(d) to (f), the same results are plotted for a covering radius $\rho = 1/8$ and with $N_\Theta/M = 8$ in (d). The same conclusions as before apply but the computation times tell another story. The increase in the number of grid bins N_P in the grid of interest has multiple consequences. First, the two-step method is now clearly the fastest method, confirming its lesser dependence on the scale of the problem. Moreover the term in $N_P I$ now dominates the computation time of GH. Therefore, seeking to improve the performance of the GH algorithms, it is now more efficient to reduce ξ by increasing N_Θ instead of I. Practically, this translates into the conclusion that the nearest neighbor strategy with a dense sensing grid Θ_G is the best choice in terms of compromise between performance and computation time. This observation confirms the intuition described in Sec. 7.3 from the inspection of the GH bound. In short, when N_P is sufficiently small, then a higher order interpolation is preferred to reach the single-step method's performance with GH. As N_P becomes larger, using a lower order interpolation with a dense sensing grid Θ_G becomes computationally more efficient.

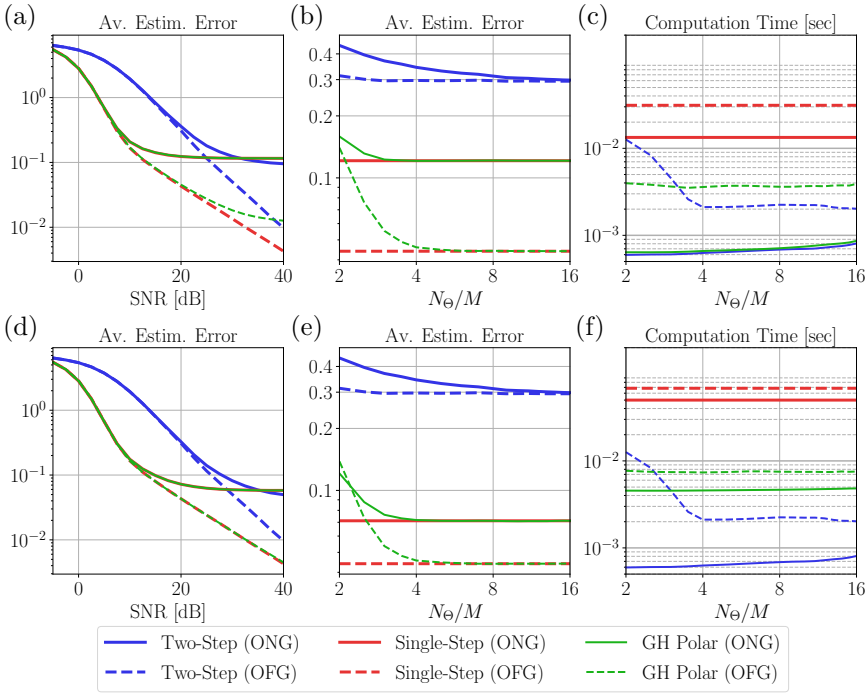

Figure 7.8: Results similar simations as in Fig. 7.6, comparing *on-the-grid* (ONG) methods and *off-the-grid* (OFG) methods. **(a)-(c)** for $\rho = 1/4$. **(d)-(f)** for $\rho = 1/8$.

We can now move on to the second application, for which results obtained in similar conditions are shown in Fig. 7.7. Since the atom function described by (6.20) is not translation invariant, we can no longer use the polar interpolation for which this property is a requirement. Also, this implies that our theoretical bound is no longer guaranteed. Still, we observe similar results as those from Fig. 7.6 in the MSR context and thus we conclude that the main intuition provided by the bound can be extended beyond the assumption it has been built upon.

Next, we decided to concentrate on the use of polar interpolation and to apply it to the *off-the-grid* GH technique that we proposed in Alg. 5. With the same conditions as Fig. 7.6, we compare in Fig. 7.8 the performance of six algorithms. The first three, labeled as "ONG", are restricted to the use of the grids $\mathcal{P}_\mathrm{G}$ and/or Θ_G that were already used for the results in Fig. 7.6. For each of them, we added an *off-the-grid* version. The off-grid single-step algorithm solves (6.11) by refining the grid search with a gradient ascent on

the global correlation function, initialized in the *on-the-grid* estimate. In the off-grid two-step algorithm, we similarly performed a gradient ascent to refine the *on-the-grid* estimate in the first step (6.8). Finally, the GH-based off-grid method amounts to Alg. 5 with a single iteration.

We observe in Fig. 7.8(a) that the *off-the-grid* GH yields results that exhibit robustness with respect to noise that is similar to the one of the exact single-step method. At high SNR, its performance saturates toward an error value which corresponds to the effect of the atom approximation intrinsically used in GH. We observe however in Fig. 7.8, in which a SNR of 20dB is constant, that this saturation can be pushed to a lower value by increasing the density of Θ_G (*i.e.*,, by increasing N_Θ). This is confirmed in Fig. 7.8(d) where $\rho = 1/8$ and $N_\Theta/M = 8$ are used. The *off-the-grid* GH seems to exactly fit the performance of the exact single-step algorithm in that case. By inspecting Fig. 7.8(c), we observe that the *off-the-grid* shift computation described in (7.7) leads to a significant increase in the computation time compared to the *on-the-grid* version of GH. Again, since this step's computation time is independent of the problem's dimension, we expect an attenuation of this effect in problems of higher scale, typically with higher values for M and/or N_P. Fig. 7.8(f) shows the results for a higher value of N_P and thereby confirms our intuition since we observe a reduced gap in the computation time between *on-the-grid* algorithms and their *off-the-grid* counterparts. Overall, the *off-the-grid* GH selection algorithm meets its promise by proposing an alternative that reaches the performance of the single-step method with a computational burden closer to the one of the two-step method.

7.5 Specific Application of GH to Multistatic Radars

The content of this section and the next one have been presented at the ISCS'23 conference [MFD+23].

We enter the final part of this chapter: the specific application of GH to the factorized single-step signal processing that we introduced in Sec. 6.3.3. Since the definition of GH is given in this chapter for the general sensor network context, it straightforwardly applies to the location selection step (6.26). For the velocity selection described by (6.27), however, the ap-

plication of GH comes with one challenge that translates into a restriction on the choice of speed grid and interpolation pattern. Indeed, let us point out that the definition of the effective PPF used in (6.27) to connect any velocity $\boldsymbol{v}$ from $\mathcal{V}$ to a set of speeds depends on the knowledge of the target's location. Therefore, the lookup table and the interpolating coefficients can no longer be computed offline. We instead need to compute them online from the value of the location estimate $\hat{\boldsymbol{x}}$. Fortunately, this task can be performed fast when working with a uniform grid in the speed domain and a nearest neighbor strategy as "interpolation". In that situation, the computation of the lookup table is reduced to a thresholding operation on the result of a small-scale matrix product. More specifically, from the estimate $\hat{\boldsymbol{x}}$, one can compute for all $q \in [Q]$, the directional vector

$$\hat{\boldsymbol{h}}_q = \tfrac{1}{2}\left(\frac{\hat{\boldsymbol{x}} - \boldsymbol{x}^{\mathrm{TX}}}{\|\hat{\boldsymbol{x}} - \boldsymbol{x}^{\mathrm{TX}}\|_2} + \frac{\hat{\boldsymbol{x}} - \boldsymbol{x}_q^{\mathrm{RX}}}{\|\hat{\boldsymbol{x}} - \boldsymbol{x}_q^{\mathrm{RX}}\|_2} \right). \tag{7.16}$$

We can then rewrite $\mathcal{S}_q^{\mathrm{speed}}(\boldsymbol{v}) = \hat{\boldsymbol{h}}_q^{\top} \boldsymbol{v}$. Therefore, given $\boldsymbol{V}_{\mathrm{G}}$ to denote the $2 \times N$ matrix containing the N velocities contained in the velocity grid $\mathcal{V}_{\mathrm{G}}$, Δ to denote the spacing between adjacent bins in the speed grid, the q-th lookup table is directly obtained from $\lfloor \hat{\boldsymbol{h}}_q^{\top} \boldsymbol{V}_{\mathrm{G}} / \Delta \rceil$, where $\lfloor \cdot \rceil$ is the rounding operator to the nearest integer. This is computed in an insignificant computation time. Moreover, no interpolating coefficients need to be computed since we consider the nearest-neighbor strategy.

7.6 Experimental Results

We finally validate the GH strategy by applying it to a sample of the data obtained during the measurement campaign described in Chapter 2. Although we recorded data from many different scenarios, potentially involving the simultaneous observation of multiple cars passing by the receiving antennas (see Fig. 2.11), this section is now restricted to the simplest scenario in which a single car is observed. This scenario was schematically represented in Fig. 2.11(a) redrawn in Fig. 7.10(a) with more precision. Moreover, to simplify the connection with our generalized sensing model, only one RX antenna in each station was used for the processing presented in this section. The processing of more sophisticated scenarios and using all height antennas present in our system —for which we already dispose of the recorded data— is left as a perspective of future work.

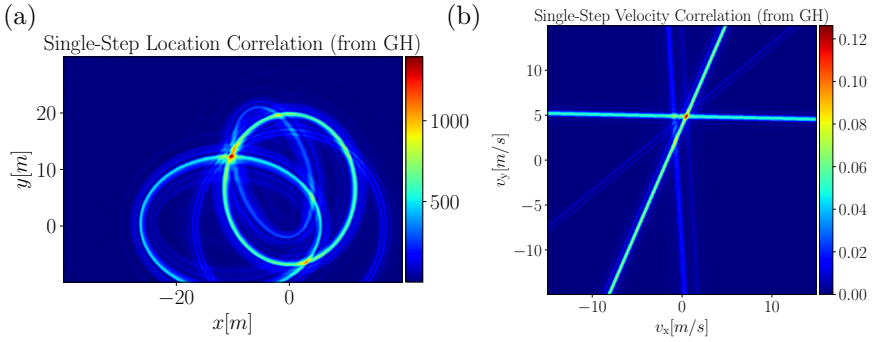

Figure 7.9: Example of the single-step correlation functions computed from one of the processing frames acquired during the measurement campaign.

For this scenario, we remind that we acquired signals for 60 seconds at a rate of one processed frame every 60ms on average. A processing frame is a block of $M_c = 128$ subsequent chirps, each composed of $M_s = 128$ samples. We remind the reader that the radar's ranging resolution is 60cm since it works on a 250MHz bandwidth. We first applied three estimation methods to each frame to retrieve the location and the speed of the single target. These are (i) the two-step estimation, (ii) the single-step estimation, and (iii) the single-step estimation enhanced with GH.

First, Let us illustrate with Fig. 7.9 the correlation functions used by the single-step method, approximated with GH, from one representative data frame. In the location correlation displayed in Fig. 7.9(a), we neatly observe the intersection of ellipses that we encountered many times from the simulations throughout Chapters 6 and 7. Similarly, the linear intersection corresponding to the velocity estimation appears when observing the corresponding correlation shown by Fig. 7.9(b). At least three comments can be made on this result. First, observing comparable figures from real-world experimental results in contrast to simulated data is a positive indication of the reliability and validity of the derivation proposed in those chapters. Second, since we used conventional patch antennas, there is no compensation for the path loss in the acquired data, leading to significantly distinct power levels in the signals received from different receivers. It results in ellipses with different intensities in Fig. 7.9(a) and ultimately in a potentially less precise estimation according to the conclusions drawn from Thm. 7.1. Third, both the ellipses in Fig. 7.9(a) and the lines in Fig. 7.9(b) do not all exactly meet in the same point. This difference with the theory mainly

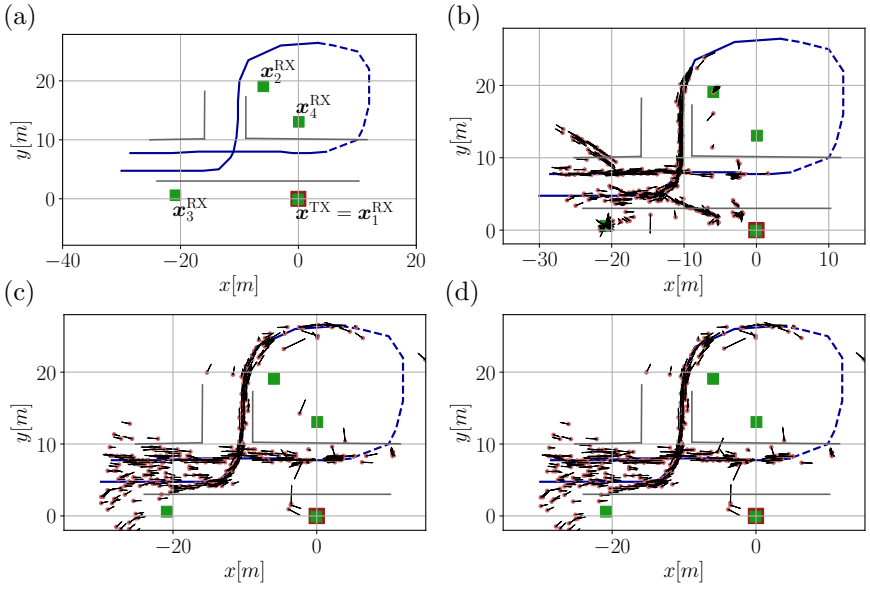

Figure 7.10: **(a)** representation of the antennas' locations with respect to the observed scene. The car moved along the blue curve. The dashed part corresponds to the locations of the cars that we discarded since they are outside the sight of the receivers. **(b)** Estimates obtained by processing the 433 frames of interest with the two-step method restricted to range and speed grids with spacing 1/8 times the radar's resolution. **(c)** Similar estimates from single-step method restricted to location and velocity grids with spacing 1/2 times the radar's resolution. **(d)** Similar estimates obtained by hopping from the grids used in (b) onto the grids used in (c) with polar interpolation.

comes from the non-point nature of the target. Since the observed car is significantly larger than the radar's resolution, this aspect that is not taken into account by our model negatively affects the results. This aspect is discussed in Chapter 8 in which we briefly investigate solutions left for future work to overcome the limitations of the current system.

Applying the three approaches to all processing frames yields location and velocity estimates plotted in Fig. 7.10. Let us point out that since the true path followed by the car is known, we discarded a set of frames that we know to correspond to the instant during which the car was located outside the sight of the receivers, hence leaving 433 relevant frames to process. We conclude from Fig. 7.10 that (i) the single-step method indeed produces more estimates that correctly follow the true path and (ii) the GH method produces estimates very close to those obtained from the single-step method.

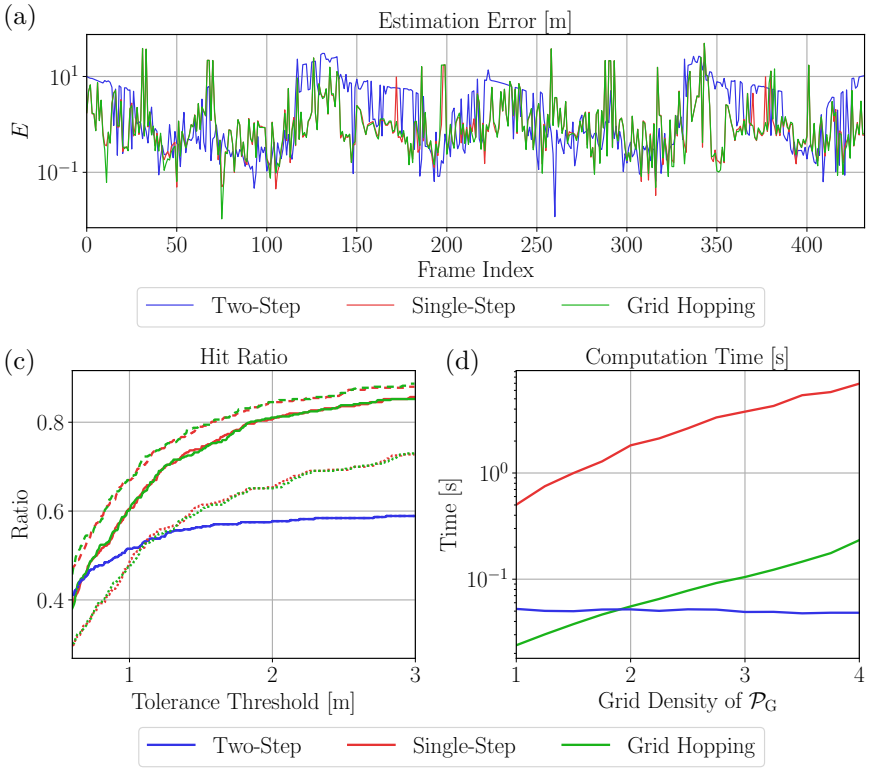

Figure 7.11: **(a)** Estimation error of all the 433 processing frames of interest, as obtained by the three estimation methods. **(b)** Hit Ratio as a function of the tolerance ratio (in meters). Three distinct densities of the grid $\mathcal{X}_G$ are respectively shown in solid line for a spacing of 1/2 the resolution, in dashed for a spacing 1/4 the resolution, and in dotted line for a spacing 1 the resolution. **(c)** Computation time of the methods with respect to the grid density of $\mathcal{X}_G$.

Finally, the three methods' performance and computation time are compared in Fig. 7.11. Each location estimate resulting from each algorithm was compared to the approximated ground truth drawn in Fig. 7.10(a). As the car is not a single point in space, such a comparison can only make sense with a tolerance threshold. When the error between an estimate and the reference is smaller than the "hit threshold", we counted it as a hit (see Fig. 7.11). We thus defined the hit ratio as the amount of hits divided by the total number of frames. The performance was tested for multiple densities of the location grid $\mathcal{X}_G$, defined as the inverse of the spacing between adja-

cent bins (in the normalized domain, *i.e.,* expressed as a number of times the resolution). We observe as expected that the single-step method performs better than the two-step one with a higher computational time. The hopping strategy provides an alternative with a computation time closer to the two-step method while exhibiting performance similar to the single-step method. This final result confirms the effectiveness of GH in a practical MSR application and concludes this chapter.

7.7 Discussion

With this chapter, we leveraged the concept provided by Chapters 5 and 6 to introduce and evaluate a novel technique that we coined Grid Hopping. This framework generalizes multiple acceleration strategies used for single-step estimation methods for networks of sensors. With Alg. 4, we also extended the BOMP algorithm to both apply the GH acceleration and perform *off-the-grid* estimation in the sensor network context. We validated, with Monte Carlo simulations, the ability of GH to perform similarly to standard single-step methods while exhibiting a computation time closer to the one of two-step methods.

By means of a few assumptions on the system and the sensing scenario, we provided a theoretical bound on the estimation error between the GH estimate and the ground truth. This bound leads to a rule that indicates under which conditions using a lower order interpolation with a dense grid is computationally more efficient than using a high order interpolating with coarse grids for similar performance. With the Monte-Carlo simulations, the rule was confirmed to apply beyond the restrictive assumption derivate of the bound. Our bound suggests that working with signals of equivalent power yields more robust results. This observation can open the way for future research on signal conditioning and pre-processing at the receiver level to meet this criterion efficiently.

Finally, we evaluated the efficiency of GH in real conditions by applying it to the measurement we performed with our multistatic radar. The current comparison is restricted to a single interpolation scheme and our most simple scenario with a single car moving. Still, the data bank we collected during our measurement contains many more sophisticated scenarios on which GH could be tested in the future. Those data can be found in open access at `https://gitlab.com/gilles.monnoyer.dg/msr-dataset-22-08-2022`.

7.8 Appendix: Proofs

7.8.1 Proof of Theorem 7.1

Theorem 7.1 results from the application of a more general theorem: Thm. 7.2, proven in the next section. We provide first this general theorem which is proven in the appendix of this paper. Note that in this whole section, $f : \mathcal{Z} \mapsto \mathbb{R}$ is a lobe-r function with lobe-bounding constants $L > 0$ and $U > 0$.

Theorem 7.2 (Lobe Function Maximization Bound)**.** *Let $\tilde{f} : \mathcal{Z} \mapsto \mathbb{R}$ be an approximation of the lobe-r function f such that for all $\boldsymbol{z} \in \mathcal{Z}$,*

$$|\tilde{f}(\boldsymbol{z}) - f(\boldsymbol{z})| \leq \eta. \tag{7.17}$$

Given a set $\mathcal{Z}' \subset \mathcal{Z}$, if there is $\boldsymbol{z}_{\text{ref}} \in \mathcal{Z}'$ such that

$$\|\boldsymbol{z}_{\text{ref}}\|_2^2 < \tfrac{Ur^2 - 2\eta}{L}, \tag{7.18}$$

then the estimator $\hat{\boldsymbol{z}}$ defined by

$$\hat{\boldsymbol{z}} := \underset{\boldsymbol{z} \in \mathcal{Z}'}{\arg\max}\ \tilde{f}(\boldsymbol{z}) \tag{7.19}$$

is such that $\|\hat{\boldsymbol{z}}\|_2^2 \leq \frac{L}{U}\|\boldsymbol{z}_{\text{ref}}\|_2^2 + \frac{2}{U}\,\eta$.

In words, Thm. 7.2 provides a bound on the distance between (i) the maximizer, restricted to $\mathcal{Z}'$, of $\tilde{f}$ which approximates the lobe-r function f and (ii) the exact maximizer of the f with no restriction, which is $\mathbf{0}$ by definition of a lobe function.

We provide some intuition on the idea behind this theorem. Given $\mathcal{Z}'$, we select the $\boldsymbol{z}_{\text{ref}} \in \mathcal{Z}'$ with the lowest ℓ_2 norm to obtain the best bound. Therefore, while the optimization (7.19) will not necessarily return $\boldsymbol{z}_{\text{ref}}$, this value is the best estimate we could hope for since (7.19) is restricted to $\mathcal{Z}'$. When the constraint (7.18) applies, we impose $\boldsymbol{z}_{\text{ref}}$ to belong inside the inner lobe of f, with a "security" coefficient $\frac{U-2\eta}{L}$ to account for the error between $\tilde{f}$ and f. Therefore, Thm. 7.2 tells us that if $\boldsymbol{z}_{\text{ref}}$ is sufficiently inside the inner lobe, not only we are guaranteed to have $\hat{\boldsymbol{z}}$ inside the outer lobe of f (which is a consequence of (5.28)), but also the decreasing properties of f

enable us to bound the ℓ_2 norm of $\hat{\boldsymbol{z}}$.

The bound is reached with equality in the unlikely event of $f(\boldsymbol{z}_{\mathrm{ref}})$ hitting the lobe's lower bound with a maximally unfavorable approximation error between $f(\boldsymbol{z}_{\mathrm{ref}})$ and $\tilde{f}(\boldsymbol{z}_{\mathrm{ref}})$. In contrast, another point of $\mathcal{Z}'$, whose ℓ_2 norm reaches the bound of Thm.7.2, exhibits a maximally favorable approximation error while hitting the lobe's upper bound (5.29).

To prove Thm. 7.1, we apply Thm. 7.2 on the functions we define hereafter. Given the ground truth $\boldsymbol{p}^*$ which is sensed through its associated sensed parameters $\boldsymbol{\theta}^* = (\theta_1^*, \cdots, \theta_Q^*)^\top := \mathcal{S}(\boldsymbol{p}^*)$, the approximated kernel, which results from the use of $\tilde{\boldsymbol{a}}$ instead of $\boldsymbol{a}$ in the correlation process, reads for all $q \in [Q]$,

$$\tilde{\kappa}_q(\delta) := \langle \tilde{\boldsymbol{a}}(\theta_q^* + \delta), \boldsymbol{a}(\theta_q^*) \rangle \tag{7.20}$$

Next, we consider the summation performed in the single-step method with and without GH respectively in (6.11) and (7.5). When particularized to the noiseless context for which this bound applies, and given $\boldsymbol{\delta} := (\delta_1, \cdots, \delta_Q)^\top$, this yields

$$\kappa_\Sigma(\boldsymbol{\delta}) := \textstyle\sum_{q=1}^{Q} \gamma_q \, |\kappa(\delta_q)|^2, \tag{7.21}$$

$$\tilde{\kappa}_\Sigma(\boldsymbol{\delta}) := \textstyle\sum_{q=1}^{Q} \gamma_q \, |\hat{\kappa}(\delta_q)|^2 \tag{7.22}$$

With these new functions, we can rewrite (7.5). Mathematically, with

$$\hat{\boldsymbol{\delta}} = \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*, \tag{7.23}$$

the optimization in (7.5) is equivalent to first solve

$$\hat{\boldsymbol{\delta}} = \underset{\boldsymbol{\delta} \in \mathcal{T}_{\mathrm{G}}}{\arg\max} \, \tilde{\kappa}(\boldsymbol{\delta}), \tag{7.24}$$

where $\mathcal{T}_{\mathrm{G}} := \{\mathcal{S}(\mathcal{P}_{\mathrm{G}}) - \boldsymbol{\theta}^*\} \subset \{\mathcal{T} - \boldsymbol{\theta}^*\}$, and then apply the bijection $\hat{\boldsymbol{p}} = \mathcal{S}^{-1}(\boldsymbol{\theta}^* + \hat{\boldsymbol{\delta}})$.

We demonstrate now in three steps that (7.24) meets the conditions for Thm. 7.2. Respectively those steps show that (i) $|\kappa_\Sigma(\delta)|^2$ is a lobe function, (ii) $|\tilde{\kappa}_\Sigma(\boldsymbol{\delta}) - \kappa_\Sigma(\boldsymbol{\delta})|$ is bounded, and (iii) the condition (7.14) is sufficient to guarantee the existence of a bin in $\mathcal{T}_{\mathrm{G}}$ which exhibits a bound that is assimilated to (7.18).

Step 1. Since A.1 imposes that $|\kappa(\delta)|^2$ is a lobe-r function and by construction $\sum_{q=1}^{Q} \gamma_q = 1$, we can use Lem. 6.5 to derive the properties of κ_Σ.

We conclude that κ_Σ is a lobe-$\tilde{r}$ with lobe-bounding constants $\tilde{L}$ and $\tilde{U}$ defined by

$$\tilde{L} = \gamma_{\max} L, \quad \tilde{U} = \gamma_{\min} U, \qquad \tilde{r} = \sqrt{\tfrac{\gamma_{\min} U}{\gamma_{\max} L}}\, r. \tag{7.25}$$

Setp 2. The definition of the interpolation gap implies for all $\theta \in \Theta$,

$$\xi^2 \geq \|\tilde{\boldsymbol{a}}(\theta) - \boldsymbol{a}(\theta)\|_2^2 = 2(1 - \Re\{\langle \tilde{\boldsymbol{a}}(\theta),\, \boldsymbol{a}(\theta)\rangle\}) \geq 2(1 - |\langle \tilde{\boldsymbol{a}}(\theta),\, \boldsymbol{a}(\theta)\rangle|), \tag{7.26}$$

, which in turn leads to

$$|\tilde{\kappa}(0)|^2 = |\langle \boldsymbol{a}(\theta^* + \hat{\delta}),\, \tilde{\boldsymbol{a}}(\theta^* + \hat{\delta})\rangle|^2 \geq 1 - \xi^2. \tag{7.27}$$

Additionally, the lobe bounding property of $|\kappa(\delta)|^2$ gives

$$|\kappa(\hat{\delta})|^2 = |\langle \boldsymbol{a}(\theta^* + \hat{\delta}),\, \boldsymbol{a}(\theta^*)\rangle|^2 \geq 1 - L\|\hat{\delta}\|_2^2. \tag{7.28}$$

Thus, applying both Lem. 5.5 and Lem. 5.6 to the functions $|\kappa(\cdot)|^2$ and $|\tilde{\kappa}_q(\cdot)|^2$ yields, for all $q \in [Q]$ and for all $\delta \in \{\Theta - \theta_q^*\}$,

$$\big||\tilde{\kappa}_q(\delta)|^2 - |\kappa(\delta)|^2\big| \leq \min\big\{\sqrt{2}\xi, 2\sqrt{L}\xi\|\delta\|_2 + \xi^2\big\}. \tag{7.29}$$

Next, we extend this result to $\tilde{\kappa}_\Sigma$. More precisely, for all $\boldsymbol{\delta} \in \{\mathcal{T} - \boldsymbol{\theta}^*\}$,

$$\begin{aligned} |\tilde{\kappa}_\Sigma(\boldsymbol{\delta}) - \kappa_\Sigma(\boldsymbol{\delta})| &= \big|\textstyle\sum_{q=1}^Q \gamma_q(|\tilde{\kappa}_q(\delta_q)|^2 - |\kappa(\delta_q)|^2)\big| && (7.30)\\ &= \textstyle\sum_{q=1}^Q \gamma_q\big||\tilde{\kappa}_q(\delta_q)|^2 - |\kappa(\delta_q)|^2\big| && (7.31)\\ &\leq \min\big\{\sqrt{2}\xi, 2\sqrt{L}\xi\|\hat{\delta}\|_2 + \xi^2\big\}. && (7.32) \end{aligned}$$

Step 3. The assumption A.2 implies that there is $\boldsymbol{p}_{\text{ref}} \in \mathcal{P}_{\text{G}}$ such that $\|\boldsymbol{p}_{\text{ref}} - \boldsymbol{p}^*\|_2^2 \leq \rho^2$. Applying Lem. 6.4 with $\boldsymbol{\theta}_{\text{ref}} = \mathcal{S}(\boldsymbol{p}_{\text{ref}})$ results in $\|\boldsymbol{\theta}_{\text{ref}} - \boldsymbol{\theta}^*\|_2^2 \leq \lambda_{\max}\rho^2$. By construction, we have $\boldsymbol{\delta}_{\text{ref}} := \boldsymbol{\theta}_{\text{ref}} - \boldsymbol{\theta}^* \in \mathcal{T}_{\text{G}}$. We can thus apply Thm. 7.2 to the maximization (7.24). Putting all together, if ρ is such that

$$\begin{aligned} \lambda_{\max}\rho^2 &\leq \tfrac{\tilde{U}\tilde{r}^2 - 4\xi^2}{\tilde{L}} && (7.33)\\ &= \tfrac{U\gamma_{\min}}{L\gamma_{\max}}(r^2 - \tfrac{2\sqrt{2}\xi}{U\gamma_{\min}}), && (7.34) \end{aligned}$$

then

$$\|\hat{\boldsymbol{\delta}}\|_2^2 \leq \tfrac{\tilde{L}}{\tilde{U}}\,\lambda_{\max}\rho^2 + \tfrac{4\sqrt{L}}{\tilde{U}}\,\xi\,\|\hat{\boldsymbol{\delta}}\|_2 + 2\xi^2. \tag{7.35}$$

$$\leq \tfrac{\tilde{L}}{\tilde{U}}\,\lambda_{\max}\rho^2 + \tfrac{4\sqrt{\tilde{L}}}{\tilde{U}}\,\xi\,\|\hat{\boldsymbol{\delta}}\|_2 + 2\xi^2. \tag{7.36}$$

Finding the only positive root of the hence formed order-2 polynomial in $\|\hat{\boldsymbol{\delta}}\|_2$ yields,

$$\sqrt{\tilde{U}}\,\|\hat{\boldsymbol{\delta}}\|_2 \leq 2\sqrt{\tfrac{\tilde{L}}{\tilde{U}}}\xi + \sqrt{(4\tfrac{\tilde{L}}{\tilde{U}} + 2)\xi^2 + \tilde{L}\|\boldsymbol{\delta}_{\text{ref}}\|_2^2}. \tag{7.37}$$

Since $\|\boldsymbol{\delta}_{\text{ref}}\|_2^2 \leq \lambda_{\max}\rho^2$, then using the correspondence from (7.25) and with R_{P} defined by (6.28), we obtain

$$\sqrt{U\gamma_{\min}}\,\|\hat{\boldsymbol{\delta}}\|_2 \leq 2R_{\text{P}}\xi + \sqrt{(4R_{\text{P}}^2 + 2)\xi^2 + L\gamma_{\max}\lambda_{\max}\rho^2}. \tag{7.38}$$

7.8.2 Proof of Theorem 7.2

The proof Thm. 7.2 exploits the decreasing properties (5.29) of the lobe-r function f. To use it, we must first guarantee that $\hat{\boldsymbol{z}}$ belongs inside the outer lobe, *i.e.,* $\hat{\boldsymbol{z}} \in \mathbb{B}_\infty(R)$. To this end, we show that $\boldsymbol{z}_{\text{ref}}$ is such that $\tilde{f}(\boldsymbol{z}_{\text{ref}}) > \tilde{f}(\boldsymbol{z})$ for all $\boldsymbol{z} \in \mathcal{Z}\backslash\mathbb{B}_\infty(R)$. If this holds, then since $\tilde{f}(\hat{\boldsymbol{z}}) \geq \tilde{f}(\boldsymbol{z}_{\text{ref}})$ by definiton of $\hat{\boldsymbol{z}}$, we conclude $\hat{\boldsymbol{z}} \in \mathbb{B}_\infty(R)$.

The constraint (7.18) on the ℓ_1 norm of $\boldsymbol{z}_{\text{ref}}$ is tantamount to

$$1 - L\|\boldsymbol{z}_{\text{ref}}\|_2^2 - \eta > 1 - Ur^2 + \eta \tag{7.39}$$

Since we remind that $L \geq U$ and $r \leq R$, the constraint (7.18) also guarantees that $\boldsymbol{z}_{\text{ref}} \in \mathbb{B}_2(R) \subset \mathbb{B}_\infty(R)$ and hence the lobe decreasing properties (5.29) apply to the above. This yields, for all $\boldsymbol{z} \in \mathcal{Z}\backslash\mathbb{B}_\infty(R)$,

$$f(\boldsymbol{z}_{\text{ref}}) - \eta > f(\boldsymbol{z}) + \eta, \tag{7.40}$$

and thus $\tilde{f}(\boldsymbol{z}_{\text{ref}}) > \tilde{f}(\boldsymbol{z})$ since (7.17) holds. As stated before, this proves $\hat{\boldsymbol{z}} \in \mathbb{B}_\infty(R)$.

We can now apply (5.29) on $f(\hat{\boldsymbol{z}})$. This leads to $1 - U\|\hat{\boldsymbol{z}}\|_2^2 + \eta \geq f(\hat{\boldsymbol{z}}) + \eta \geq \tilde{f}(\hat{\boldsymbol{z}})$, and $\tilde{f}(\boldsymbol{z}_{\text{ref}}) \geq f(\boldsymbol{z}_{\text{ref}}) - \eta \geq 1 - L\|\boldsymbol{z}_{\text{ref}}\|_2^2 - \eta$. Since $\tilde{f}(\hat{\boldsymbol{z}}) \geq \tilde{f}(\boldsymbol{z}_{\text{ref}})$, this implies $\|\hat{\boldsymbol{z}}\|_2^2 \leq \frac{L}{U}\|\boldsymbol{z}_{\text{ref}}\|_2^2 + \frac{2}{U}\eta$ and concludes the proof.

Part IV

Conclusions

Chapter 8

Conclusions and Perspectives

This thesis intended to contribute to setting one step forward in accelerating computationally heavy procedures to make them meet the specific real-time and low-cost constraints applied in many estimation applications with sensor networks. Motivated by the continuous advances in hardware performances, this research aimed to leverage the resulting procedures to enhance the performance of multiple sensor network applications, with specific attention drawn to Multistatic Radar (MSR) systems. To perform this task, we sequentially focused on three connected directions: (i) strengthening the box constraints in Branch-and-Bound (BnB) procedures to solve the ℓ_0 regularized sparse decomposition problem, (ii) leveraging interpolation to perform fast *off-the-grid* estimation and (iii) finding a computationally efficient trade-off between two-step and single-step methods in the estimation context with sensor networks.

The conclusions we draw from the outcome of this research are twofold. First, our interpolation-based extensions of existing *off-the-grid* and single-step approaches do not claim to offer the best possible solution that systematically surpasses existing approaches from the literature. In fact, the optimal best choice of algorithm strongly depends on the specific application context. Moreover, any practical implementation of our proposed methods requires application-specific adaptations to obtain optimal performances. This necessity became apparent when we tailored our techniques to the context of MSRs, and it also applies to other types of sensor networks.

Nonetheless, with our contributions providing both theoretical and experimental validation of the effectiveness of novel methods, we hope to have brought new perspectives that emphasize the credibility of considering *off-the-grid* and single-step approaches in localization systems, including MSRs and other systems discussed in Chapter 6.

Second, our numerical results demonstrated the effectiveness of our "peeling" technique to accelerate the sparse decomposition of signals by the ℓ_0 regularized problem in a simple scenario. Still, the method remains in its early developing stage. Notably, many proofs of effectiveness are still required if one wishes to see it as a novel state-of-the-art method for solving large-scale problems in statistics or machine learning. Additionally, our initial intention to connect this technique with real-time estimation applications still requires developing greedy algorithms improved with adapted peeling and screening principles. While this thesis did not achieve this task, we hope future work emerges on that topic.

In the following sections, we give an overview of this thesis's specific contributions and limitations, as well as some insights on possible further research.

8.1 Achievements and Limitations

The outcome of this thesis can be classified in three distinct yet connected levels of abstraction of the problems they address: (i) the sparse decomposition level, (ii) the sensor network level, and (iii) the specific MSR application level. The achievements and the limitations that lead to perspectives of future work are sorted and summarized in Table 8.1.

In **Chapter** 2, we described a small side contribution: the conception of an active Frequency-Modulated Continuous Wave (FMCW) MSR system and the associated measurement campaign we performed. Although its design is not assimilated as a full contribution, the measurement we acquired provided a large data bank. These contain the observation of various civil scenarios involving multiple cars passing by the antennas composing the system. We used a small subset of these measurements to validate the methods proposed in Chapter 7. However, all the advanced setups with the simultaneous observation of multiple targets have not been exploited yet. The whole recorded data samples remain available for future work either on Grid Hopping (GH) or any other research project on active FMCW MSR.

To the other side of the abstraction levels, **Chapter** 4 presented a

	Sparse Representation of Signals	**Sensor Networks**	**Specific MSR Considerations**
Chapter 2			• MSR design and large bank of data
			• Many scenarios yet to process
Chapter 4	• Strengthen convex relaxation in BnB → Reduced comp. Time	• Extension of BnB and Peeling to Multiple Measurement Vector Context	
	• Implications in Greedy methods ?		
Chapter 5	• Unified formulations of atom interpolation techniques	• Continuous OMP that spcifically focuses for localization systems' features	• Factorized COMP validated with Simulations
			• Experimental Validation ?
Chapter 6		• Unified model	• Factorization MSR aspects in Single-step
Chapter 7		• Grid Hopping • Performance Bound • Numerical validation	• MSR Factorization aspects in GH • Experimental validation
		• Optimizing Geometry for GH • Connection with CRLB	• Advanced Scenarios (K > 1) • Antenna Arrays

Table 8.1: Summary of the contributions and accomplishments of this thesis (in blue) as well as the current limitations, opening the way for future research projects (in orange).

tractable strategy, named "peeling", to tighten the box constraints used in a BnB procedure tailored to ℓ_0-regularized least-squares problems. Unlike the standard approach, which imposes one global constraint on the problem, our strategy aims to refine the box constraints at each decision tree node locally. This refinement strengthens the convex relaxations used in the pruning decisions made by the BnB procedure. As emphasized by our simulation results, it can lead to significant improvements in solving time. Following the above comment on this part of this thesis, we note that the conclusions from this chapter potentially open up new ways for future research, including the extension of ℓ_0 peeling and screening techniques to the multiple measurement vector context or the enhanced greedy algorithms.

Chapter 5 continued the journey on the acceleration of sparse decomposition by introducing the *off-the-grid* estimation problem. It started with the review of multiple commonly used atom interpolation patterns to ultimately compare their effectiveness within different variants of Continuous OMP (COMP). More specifically, we proposed a new version of a contin-

uous OMP acting as a compromise between existing approaches, intended to meet the desired computational time requirement for real-time sensor contexts. Although our method does not surpass the performances of the competing approaches for all sensing scenarios, it finds specific interest in signals featuring characteristics of the radar measurement model. In these, the physics of the problem guarantees minimal separation between ground truth distances of the reflectors in the scene. It hence matches the particular case for which our proposition is particularly efficient.

Entering Part III, we provided in **Chapter** 6 a transition introducing crucial concepts to extend from the framework of one measurement to the simultaneous observation of a scene by means of Q distinct sensors, each yielding one measurement. While the research on sensor network estimation algorithms is widely spread across multiple distinct communities, we aimed to unify their model into a single formalism that encompasses the properties shared by distinct applications. Next, extending a work performed during my Master thesis, we particularized our model to the factorized context appearing in FMCW MSR systems. Specifically, we deepened the study of the model mismatch that inevitably arises while applying this factorization. We thus defined a novel theoretical bound on the estimation error in a simple scenario. One may consider future work deriving a tighter formulation of the bounds, using, for instance, higher order bounding lemmas on the Parameter Projection Function (PPF). Still, this might not appear as a fertile direction since the current formulation of the bounds is already sufficient to relate the central tendency between the geometry of the system and the impact of the model mismatch.

We ended the technical part of this manuscript with **Chapter** 7 and the introduction of GH. Our technique is built upon the concepts and the models described in Chapters 5 and 6 and intends to accelerate single-step estimation methods for sensor networks. With extensive Monte-Carlo of distinct systems falling into the scope of Chapters 6, we validated the ability of GH to result in similar performance as standard single-step methods while exhibiting a computation time closer to the one from two-step methods. We also observed this effect while applying GH to actual measurement recorded by the system presented in Chapter 2. Finally, we studied its limitation by means of an additional upper bound on the estimation error it provides in a simple noiseless scenario. We confirmed our interpretation of the implication of the bound by observing its effect on the Monte-Carlo simulations. We leave the critical discussion on the content of this chapter to the next section, in which we suggest new directions for potential future research projects.

8.2 Perspectives of Future Research

Let us extract from our conclusions a few hints for future work targeting each of the three levels corresponding to the columns of Table 8.1.

Sparse Decomposition Level : Based on conclusions explained at the beginning of this chapter, we identify in that level two main research directions to follow up on our work: (i) the extension of peeling to more sophisticated scenarios and (ii) pursuing the work on improving Orthogonal Matching Pursuit (OMP) by injecting some BnB inspired tests within its iterations.

Sensor Network Level: Interestingly, the grid hopping estimation bound provided by Thm. 7.1 to evaluate the impact of the systems' properties on the precision loss caused by GH can be connected with the joint Cramèr-Rao Lower Bound (CRLB) computed in a similar context (*e.g.,* see [GHB10, IKY16]). That is, the CRLB is given either by the determinant or the eigenvalues of the inverse of the Fisher Information Matrix (FIM). In the single-step estimation framework, the evaluation of this latest in any $\boldsymbol{p} \in \mathcal{P}$ is proportional to $\boldsymbol{J}(\boldsymbol{p})^\top \boldsymbol{J}(\boldsymbol{p})$, *i.e.,* the same quantity we used to define the constant $\lambda_{\min}$ that drives the value of our GH bound. Still, it is worth noting that both the joint CRLB and our bound study are distinct but complementary aspects of the estimation. On the one hand, CRLB study the estimation error caused by the noise while considering an exact continuous single-step solver. On the other hand, our bound focuses on the impact of the model approximations of the discretization and the use of GH in a *noiseless* scenario. It is, therefore, of high interest to observe that optimizing the sensors' localizations of a given system is aligned with respect to these two distinct aspects. Our bound, however, enables novel considerations with respect to the CRLB: it conveys information about the impact of the choice of grid $\mathcal{P}_{\mathrm{G}}$ on the performance of GH. To sum up the considerations brought by this paragraph, we suggest that future work connecting our bound with the CRLB to optimize the design of sensor networks, as well as the grid $\mathcal{P}_{\mathrm{G}}$ discretizing $\mathcal{P}$, can be of high interest to increase the estimation performance in many localization scenarios.

Additionally, our GH bound suggests that working with signals of equivalent power yields more robust results. While this property can be ensured by first normalizing the power of each received signal, this may appear as

an imprudent decision when working in the real, noisy world. Indeed, this would amount to dramatically increase the noise power of signal observing the features of interest with low Signal-to-Noise Ratio (SNR), hence degrading the overall estimation performance of single-step methods. We propose instead to apply a first pre-processing step to separately denoise each signal prior to the above-mentioned power normalization at the input of a single-step algorithm. We emphasize that our suggestion is not equivalent to a two-step method since there is no objective of finding preliminary estimates. This denoising step may, for instance, be efficiently accomplished with a weakly regularized LASSO procedure at each receiver's level. On top of the desired increase in estimation performance, this could also lead to simplifications in the hardware requirements since one should only transmit —to the central computation unit— compressed versions of the received signals instead of their full-resolution versions.

Multistatic Radar Level : The radar signal model derived in this report and subsequently used to validate our methods possesses one main drawback: The assumption that a single-point scatterer well represents targets. In reality, the signal received by each antenna is a sum of the echos originating from the reflection of *multiple* salient scattering points of each target. Since the exact location of these salient points depends on the angle from which the target is observed, this has two main consequences on the application of single-step methods to MSR.

First, one needs to estimate an unknown number of estimates (of location and velocity) for each target instead of one. Studying the structure of the clusters of resulting estimates is of paramount interest for target recognition. While this aspect extends beyond the scope of our work, it was notably investigated by Dr. Zijian Wang in the framework of the LUMINET project I participated in. Combining the effort of this thesis on accelerating single-step procedures with the target recognition topic may lead to new research projects in the future.

The second consequence is that each sensor may not observe the same parameter of interest if one keeps the radar formulation we proposed in Chapter 6. This leads to two open questions: (i) Can we find a new definition of the PPF operator to take this mismatch into account, and (ii) Can we evaluate its impact using an extension of our estimation bound?

8.3 Final Words

We are now reaching the end of this dissertation. At the end of the day, seeking to enhance the estimation capabilities of multistatic radars has driven my research far beyond the specific scope of this application.

While some conclusions of this work apply to MSR alone, others quickly appeared to me as generalizable to other sensors. Investigating the acceleration of heavy sparse decomposition algorithms in that particular context eventually led me to briefly set aside the sensor applications to study, at a more abstract level, how seldom used decomposition problems can be brought back to the scene while maintaining tractable solving times. While these considerations might have diffused the focus of this thesis over many distinct topics, I aspire to have provided new connections between yet separated research fields.

In the end, bringing these algorithms forward to the scene as new credible approaches for real-time applications has led to more questions than directly applicable contributions in off-the-shelf sensor networks. It is thereby expected to have opened the way for a long list of future exciting research projects on a wide variety of related topics.

Bibliography

[AAM21] K. N. Aliyu, S. A. Alawsh, and A. H. Muqaibel. Doa-based localization using deep learning for wireless seismic acquisition. *PREPRINT (Version 1) available at Research Square [https://doi.org/10.21203/rs.3.rs-925575/v1*, 09 2021.

[ABM15] A. Alexandridis, G. Borboudakis, and A. Mouchtaris. Addressing the data-association problem for multiple sound source localization using doa estimates. In *2015 23rd European Signal Processing Conference (EUSIPCO)*, pages 1551–1555, 2015.

[ADGD17] J. Aleksandravicius, N. Decarli, A. Guerra, and D. Dardari. High-accuracy localization of backscattering uwb tags: Implementation and experimental results. In *2017 IEEE International Conference on RFID Technology and Application (RFID-TA)*, pages 34–39, 2017.

[AG20] A. Atamturk and A. Gomez. Safe screening rules for ℓ_0-regression, 04 2020.

[AM18] A. Alexandridis and A. Mouchtaris. Multiple sound source location estimation in wireless acoustic sensor networks using doa estimates: The data-association problem. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 26(2):342–356, 2018.

[BCWP21] D. Bertsimas, R. Cory-Wright, and J. Pauphilet. A unified approach to mixed-integer optimization problems with logical constraints. *SIAM Journal on Optimization*, 31(3):2340–2367, 2021.

[BD20a] R. Boora and S. Dhull. Iterative volumetric reduction (ivr) steered response power method for acoustic source localization. *International Journal of Sensors, Wireless Communications and Control*, 10, 06 2020.

[BD20b] R. Boora and S. Dhull. A tdoa-based multiple source localization using delay density maps. *Sādhanā*, 45, 12 2020.

[BD22] R. Boora and S. K. Dhull. Performance evaluation of iterative srp-phat techniques for acoustic source localization. In Sanyog Rawat, Arvind Kumar, Pramod Kumar, and Jaume Anguera, editors, *Proceedings of First International Conference on Computational Electronics for Wireless Communications*, pages 403–418, Singapore, 2022. Springer Nature Singapore.

[BdGD+23] M. Bodehou, G. Monnoyer de Galland, M. Drouguet, K. Al Khalifeh, L. Vandendorpe, and C. Craeye. Metasurface antennas for fmcw radar. *IEEE Antennas and Wireless Propagation Letters*, 22(5):1040–1044, 2023.

[Bec17] A. Beck. *First-Order Methods In Optimization*. SIAM, 2017.

[BERG15] A. Bonnefoy, V. Emiya, L. Ralaivola, and R. Gribonval. Dynamic screening: Accelerating first-order algorithms for the lasso and group-lasso. *IEEE Transactions on Signal Processing*, 63(19):5121–5132, 2015.

[BKM16] D. Bertsimas, A. King, and R. Mazumder. Best subset selection via a modern optimization lens. *The Annals of Statistics*, 44(2):813–852, 2016.

[BKY16] J. Bosse, O. Krasnov, and A. Yarovoy. Direct target localization with an active radar network. *Signal Processing*, 125:21–35, 2016.

[BM11] C.R. Berger and J.M.F. Moura. Noncoherent compressive sensing with application to distributed radar. *Annual Conference on Information Sciences and Systems*, 45(4):1–6, Mar. 2011.

[BMBN22] R. Ben Mhenni, S. Bourguignon, and J. Ninin. Global optimization for sparse solution of least squares problems. *Optimization Methods and Software*, 37(5):1740–1769, 2022.

[BNCM15] S. Bourguignon, J. Ninin, H. Carfantan, and M. Mongeau. Exact sparse approximation problems via mixed-integer programming: Formulations and computational performance. *IEEE Transactions on Signal Processing*, 64(6):1405–1419, 2015.

[BP07] A. N. Bishop and P. N. Pathirana. Localization of emitters via the intersection of bearing lines: A ghost elimination approach. *IEEE Transactions on Vehicular Technology*, 56(5):3106–3110, 2007.

[BV04] S. P. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[C.22] Gilles C. Theoretical analysis of beamforming steering vector formulations for acoustic source localization. *Journal of Sound and Vibration*, 517:116544, 2022.

[CAA+17] M. Cobos, F. Antonacci, A. Alexandridis, A. Mouchtaris, and B. Lee. A survey of sound source localization methods in wireless acoustic sensor networks. *Wireless Communications and Mobile Computing*, 2017:1–24, 08 2017.

[CBH06] J. Chen, J. Benesty, and Y. Huang. Time delay estimation in room acoustic environments: An overview. *EURASIP J. Appl. Signal Proces*, 26503, 12 2006.

[CFG14] E. J. Candès and C. Fernandez-Granda. Towards a mathematical theory of super-resolution. *Communications on pure and applied Mathematics*, 67(6):906–956, 2014.

[CGWY12] X. Chen, D. Ge, Z. Wang, and Y. Ye. Complexity of unconstrained l2-lp minimization. *Mathematical Programming*, 143(1-2):371–383, 2012. doi:10.1007/s10107-012-0613-0.

[CH19] F. Champagnat and C. Herzet. Atom selection in continuous dictionaries: Reconciling polar and svd approximations. In *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5516–5520, 2019.

[CH21] F. Champagnat and C. Herzet. Interpolating and translation-invariant approximations of parametric dictionaries. In *2020*

28th European Signal Processing Conference (EUSIPCO), pages 2011–2015, 2021.

[CKS89] J.A. Cadzow, Y.-S. Kim, and D.-C. Shiue. General direction-of-arrival estimation: a signal subspace approach. *IEEE Transactions on Aerospace and Electronic Systems*, 25(1):31–47, 1989.

[CML11] M. Cobos, A. Marti, and J. J. Lopez. A modified srp-phat functional for robust real-time sound source localization with scalable spatial sampling. *IEEE Signal Processing Letters*, 18(1):71–74, 2011.

[COH18] M. B. Coteli, O. Olgun, and H. Hacıhabiboglu. Multiple sound source localization with steered response power density and hierarchical grid refinement. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 26(11):2215–2229, 2018.

[Cpl09] IBM ILOG Cplex. V12. 1: User's manual for cplex. *International Business Machines Corporation*, 46(53):157, 2009.

[CRT90] J. D Camm, A. S Raturi, and S. Tsubakitani. Cutting big M down to size. *Interfaces*, 20(5):61–66, 1990.

[CT05] E. J. Candes and T. Tao. Decoding by linear programming. *IEEE transactions on information theory*, 51(12):4203–4215, 2005.

[CT18] R. Chellappa and S. Theodoridis. Chapter 5 - multistatic radar systems. In R. Chellappa and S. Theodoridis, editors, *Academic Press Library in Signal Processing, Volume 7*, pages 253–275. Academic Press, 2018.

[DBA07] J. Dmochowski, J. Benesty, and S. Affes. A generalized steered response power method for computationally viable source localization. *Audio, Speech, and Language Processing, IEEE Transactions on*, 15:2510 – 2526, 12 2007.

[DDG+20] D. Dardari, N. Decarli, A. Guerra, M. Fantuzzi, D. Masotti, A. Costanzo, D. Fabbri, A. Romani, M. Drouguet, T. Feuillen, C. Raucy, L. Vandendorpe, and C. Craeye. An ultra-low power ultra-wide bandwidth positioning system. *IEEE Journal of Radio Frequency Identification*, 4(4):353–364, 2020.

[DDP17] Q. Denoyelle, V. Duval, and G. Peyré. Support recovery for sparse super-resolution of positive measures. *Journal of Fourier Analysis and Applications*, 23:1153–1194, 2017.

[DDSvW21] T. Dietzen, E. De Sena, and T. van Waterschoot. Low-complexity steered response power mapping based on nyquist-shannon sampling. In *2021 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 206–210, 2021. doi:10.1109/WASPAA52581.2021.9632774.

[DDWB07] T. Derham, S. Doughty, K. Woodbridge, and C.J. Baker. Design and evaluation of a low-cost multistatic netted radar system. *Radar, Sonar and Navigation, IET*, 1:362 – 368, 11 2007.

[DE03] D. Donoho and M. Elad. Optimally sparse representation in general (nonorthogonal) dictionaries via l-1 minimization. *Proceedings of the National Academy of Sciences*, 100(5):2197–2202, 2003.

[Den18] Q. Denoyelle. *Theoretical and Numerical Analysis of Super-Resolution Without Grid*. PhD thesis, 07 2018.

[DET05] D. L. Donoho, M. Elad, and V. N. Temlyakov. Stable recovery of sparse overcomplete representations in the presence of noise. *IEEE Transactions on information theory*, 52(1):6–18, 2005.

[dG19] G. Monnoyer de Galland. *Sparsity-driven moving target detection in distributed multistatic FMCW radars*. PhD thesis, UCL - Ecole polytechnique de Louvain, 2019. URL: http://hdl.handle.net/2078.1/thesis:19469.

[dGFVJ21] G. Monnoyer de Galland, T. Feuillen, L. Vandendorpe, and L. Jacques. Sparse factorization-based detection of off-the-grid moving targets using fmcw radars. *in Proceedings of ICASSP 2021*, 06 2021.

[dGVJ20] G. Monnoyer de Galland, L. Vandendorpe, and L. Jacques. Factorization over interpolation: A fast continuous orthogonal matching pursuit. *in proceedings of iTwist'20*, 2020. arXiv:2007.01060.

[DHD19] C. Dorffer, C. Herzet, and A. Drémeau. Region-based relaxations to accelerate greedy approaches. In *2019 27th European*

Signal Processing Conference (EUSIPCO), pages 1–5. IEEE, 2019.

[DLJ+22] X. Ding, H. Liang, A. Jakobsson, X. Tu, and Y. Huang. High-resolution source localization exploiting the sparsity of the beamforming map. *Signal Processing*, 192:108377, 2022.

[DM22] D. Desai and N. Mehendale. A review on sound source localization systems. *Archives of Computational Methods in Engineering*, 29, 05 2022.

[Dog05] K. Dogan. Bearings-only target localization using total least squares. *Signal Processing*, 85(9):1695–1710, 2005.

[DOR08] B. Demissie, M. Oispuu, and E. Ruthotto. Localization of multiple sources with a moving array using subspace data fusion. In *2008 11th International Conference on Information Fusion*, pages 1–7, 2008.

[DSB01] J. H. DiBiase, H. F. Silverman, and M. S. Brandstein. *Robust Localization in Reverberant Rooms*, pages 157–180. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001.

[DSY07] H. Do, H. F. Silverman, and Y. Yu. A real-time srp-phat source location implementation using stochastic region contraction(src) on a large-aperture microphone array. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*, volume 1, pages I–121–I–124, 2007.

[EGSH21] C. Elvira, R. Gribonval, C. Soussen, and C. Herzet. When does OMP achieve exact recovery with continuous dictionaries? *Applied and Computational Harmonic Analysis*, 51:39, March 2021. URL: https://inria.hal.science/hal-02099464.

[EKB10] Y.C. Eldar, P. Kuppinger, and H. Bölcskei. Block-sparse signals: Uncertainty relations and efficient recovery. *IEEE Transaction on Signal Processing*, 58(6):3042 – 3054, Jul. 2010.

[ETS11] C. Ekanadham, D. Tranchina, and E. P. Simoncelli. Recovery of sparse translation-invariant signals with continuous basis pursuit. *IEEE Transactions on Signal Processing*, 59(10):4735–4744, Oct 2011. doi:10.1109/TSP.2011.2160058.

[FDD13] K. Fyhn, H. Dadkhahi, and M. F. Duarte. Spectral compressive sensing with polar interpolation. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 6225–6229, May 2013.

[FDJ15] K. Fyhn, M. F. Duarte, and S. H. Jensen. Compressive parameter estimation for sparse translation-invariant signals using polar interpolation. *IEEE Transactions on Signal Processing*, 63(4):870–881, Feb 2015.

[FHW11] Y. Fang, B. Huang, and J. Wu. 2d sparse signal recovery via 2d orthogonal matching pursuit. *Science China Information Sciences*, 55, 01 2011.

[FHW17] A. Frischen, J. Hasch, and C. Waldschmidt. A cooperative mimo radar network using highly integrated fmcw radar sensors. *IEEE Transactions on Microwave Theory and Techniques*, PP:1–12, 02 2017.

[FMV16] T. Feuillen, A. Mallat, and L. Vandendorpe. Stepped frequency radar for automotive application: Range-doppler coupling and distortions analysis. In *MILCOM 2016 - 2016 IEEE Military Communications Conference*, pages 894–899, Nov 2016.

[FR13] S. Foucart and H. Rauhut. *An Invitation To Compressive Sensing.* Springer, 2013.

[FRMTS17] J. Fresno, G. Robles, J. Martínez-Tarifa, and B. Stewart. Survey on the performance of source localization algorithms. *Sensors*, 17:2666, 11 2017.

[GHB10] H. Godrich, A. M. Haimovich, and R. S. Blum. Target localization accuracy gain in mimo radar-based systems. *IEEE Transactions on Information Theory*, 56(6):2783–2803, 2010.

[GHE21] T. Guyard, C. Herzet, and C. Elvira. Node-screening tests for l0-penalized least-squares problem with supplementary material, 10 2021.

[GHE22] T. Guyard, C. Herzet, and C. Elvira. Node-screening tests for the l0-penalized least-squares problem. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5448–5452. IEEE, 2022.

[GMEH23] T. Guyard, G. Monnoyer, C. Elvira, and C. Herzet. Safe peeling for l0-regularized least-squares with supplementary material, 2023. arXiv:2302.14471.

[GN11] S. Gogineni and A. Nehorai. Target estimation using sparse modeling for distributed mimo radar. *IEEE Transaction on Signal Processing*, 59:5315 – 5325, Dec. 2011.

[Gri01] R. Gribonval. Fast matching pursuit with a multiscale dictionary of gaussian chirps. *IEEE Transactions on Signal Processing*, 49(5):994–1001, 2001.

[GTK13] A. A. Gorji, R. Tharmarasa, , and T. Kirubarajan. Widely separated mimo versus multistatic radars for target localization and tracking. *IEEE Transactions on Aerospace and Electronics*, 49(4):2179–2194, 10 2013.

[Han87] E. Hanle. Survey of bistatic and multistatic radars. *IEE Proceedings*, 133(7), Jan. 1987.

[HBC08] A. M. Haimovich, R. S. Blum, and L. J. Cimini. Mimo radar with widely separated antennas. *IEEE Signal Processing Magazine*, 25(1):116–129, 2008.

[HM20] H. Hazimeh and R. Mazumder. Fast best subset selection: Coordinate descent and local combinatorial optimization algorithms. *Operations Research*, 68(5):1517–1537, 2020.

[HMS22] H. Hazimeh, R. Mazumder, and A. Saab. Sparse regression at scale: Branch-and-bound rooted in first-order optimization. *Mathematical Programming*, 196(1-2):347–388, 2022.

[Hof07] K. Hoffman. *Banach spaces of analytic functions*. Courier Corporation, 2007.

[HPKS12] D. E. Hack, L. K. Patton, A. D. Kerrick, and M. A. Saville. Direct cartesian detection, localization, and de-ghosting for passive multistatic radar. In *2012 IEEE 7th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pages 45–48, 2012.

[HTT17] T. Hastie, R. Tibshirani, and R. J Tibshirani. Extended comparisons of best subset selection, forward stepwise selection, and the lasso. *arXiv preprint arXiv:1707.08692*, 2017.

[IKY16] I.M. Ivashko, O.A. Krasnov, and A.G. Yarovoy. Sparsity-based optimization of the sensors positions in radar networks with separated transmit and receive nodes. *International Journal of Distributed Sensor Networks*, 2016, Jan. 2016.

[JFM16] L. Jing, L. Feng, and M. Mahendra. Distributed compressed sensing based joint detection and tracking for multistatic radar system. *Information Sciences*, 369:100–118, 2016.

[JV08] L. Jacques and C. De Vleeschouwer. A geometrical study of matching pursuit parametrization. *IEEE Transactions on Signal Processing*, 56(7):2835–2848, 2008.

[KC76] C. Knapp and G. Carter. The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 24(4):320–327, 1976.

[KGBR21] O. Kanhere, S. Goyal, M. Beluri, and T. S. Rappaport. Target localization using bistatic and multistatic radar with 5g nr waveform. In *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, pages 1–7, 2021.

[KGP20] S. Kitic, C. Gaultier, and G. Pallone. A comparative study of multilateration methods for single-source localization in distributed audio, 2020. arXiv:1910.10661.

[KSF14] Y. Khmou, S. Said, and M. Frikel. Comparative study between several direction of arrival estimation methods. *Journal of Telecommunications and Information Technology*, 2014:41–48, 03 2014.

[KTTG17] N. Keriven, N. Tremblay, Y. Traonmilin, and R. Gribonval. Compressive k-means. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6369–6373, 2017.

[KYHP15] K. Knudson, J. Yates, A. Huk, and J. Pillow. Inferring sparse representations of continuous signals with continuous orthogonal matching pursuit. *Advances in neural information processing systems*, 27, 04 2015.

[LJT+20] J. Liu, X. Jiang, X. Tian, M. Mallick, K. Huang, and C. Ma. Hybrid particle filter based dynamic compressed sensing for

signal-level multitarget tracking. *IEEE Access*, 8:17134–17148, 2020.

[LLC+16] J.J. Lin, Y.P. Li, W. C.Hsu, , and T.S. Lee. Design of an fmcw radar baseband signal processing system for automotive application. *Leuzzi F., Ferilli S. (eds) Traffic Mining Applied to Police Activities. TRAP 2017. Advances in Intelligent Systems and Computing*, 5(42), 01 2016.

[LMN+15] M. V. S. Lima, W. A. Martins, L. O. Nunes, L. W. Biscainho, T. N. Ferreira, M. V. M. Costa, and B. Lee. A volumetric srp with refinement step for sound source localization. *IEEE Signal Processing Letters*, 22(8):1098–1102, 2015.

[LP15] B. Li and A.P. Petropulu. Distributed mimo radar based on sparse sensing: Analysis and efficient implementation. *IEEE Transactions on Aerospace and Electronics Systems*, 51(4):3055–3070, Jan. 2015.

[LZL+18] Q. Liu, Q. Zhang, Y. Luo, K. Li, and L. Sun. Fast algorithm for sparse signal reconstruction based on off-grid model. *IET Radar, Sonar Navigation*, 12(4):390–397, 2018.

[MBM+20] R. B. Mhenni, S. Bourguignon, M. Mongeau, J. Ninin, and H. Carfantan. Sparse branch and bound for exact optimization of l0-norm penalized least squares. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5735–5739. IEEE, 2020.

[MCLE13] A. Marti, M. Cobos, J. Lopez, and J. Escolano. A steered response power iterative method for high-accuracy acoustic source localization. *The Journal of the Acoustical Society of America*, 134:2627–2630, 10 2013.

[MFD+23] G. Monnoyer, T. Feuillen, M. Drouguet, L. Jacques, and L. Vandendorpe. Grid hopping: Accelerating direct estimation algorithms for multistatic fmcw radar, 2023. arXiv:2307.16550.

[MFJV19] G. Monnoyer de Galland, T. Feuillen, L. Jacques, and L. Vandendorpe. Sparsity-driven moving target detection in distributed multistatic fmcw radars. In *2019 IEEE 8th Interna-*

tional Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), pages 151–155, 2019.

[Nat95] B. K. Natarajan. Sparse approximate solutions to linear systems. *SIAM journal on computing*, 24(2):227–234, 1995.

[NBVD12] R. Niu, R. S. Blum, P. K. Varshney, and A. L. Drozd. Target localization and tracking in noncoherent multiple-input multiple-output radar systems. *IEEE Transactions on Aerospace and Electronic Systems*, 48(2):1466–1489, 2012.

[NFGS16] E. Ndiaye, O. Fercoq, A. Gramfort, and J. Salmon. Gap safe screening rules for sparsity enforcing penalties. *NIPS*, 18, 11 2016.

[Nik16] M. Nikolova. Relationship between the optimal solutions of least squares regularized with l0-norm and constrained by k-sparsity. *Applied and Computational Harmonic Analysis*, 41(1):237–265, 2016.

[P.03] Aarabi P. The fusion of distributed microphone arrays for sound localization. *EURASIP Journal on Advances in Signal Processing*, 2003, 03 2003. doi:10.1155/S1110865703212014.

[PGPM13] D. Pavlidi, A. Griffin, M. Puigt, and A. Mouchtaris. Real-time multiple sound source localization and counting using a circular microphone array. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(10):2193–2206, 2013.

[PRK95] Y.C. Pati, R. Rezaiifar, and P.S. Krishnaprasad. Orthogonal matching pursuit: recursive function approximation with application to wavelet decomposition. *Asilomar Conf. On Signals, Systems and Comput*, page 40–44, Jun. 1995.

[SC19] M. Szirtes and R. Cselkó. Comparative examination of iterative and non- iterative algorithms for partial discharge location approximation in high voltage equipment. In *2019 16th Workshop on Positioning, Navigation and Communications (WPNC)*, pages 1–6, 2019.

[SCM18] F. Cuccoli S. Capobianco, L. Facheris and S. Marinai. Vehicle classification based on convolutional networks applied tofmcw

radar signals. *Leuzzi F., Ferilli S. (eds) Traffic Mining Applied to Police Activities. TRAP 2017. Advances in Intelligent Systems and Computing*, 728, 03 2018.

[SCWL15] B. Sun, H. Chen, X. Wei, and X. Li. Multitarget direct localization using block sparse bayesian learning in distributed mimo radar. *International Journal of Antennas and Propagation*, 2015:11–23, May 2015.

[SGGM13] P. Stinco, M. S. Greco, F. Gini, and M. La Manna. Noncooperative target recognition in multistatic radar systems. *Radar, Sonar and Navigation, IET*, 8(4):396–405, Jan. 2013.

[SH08] Sylvia Schulz and Thorsten Herfet. On the window-disjoint-orthogonality of speech source in reverberant humanoid scenarios. In ., 2008.

[SM03] E. Süli and D. F. Mayers. *An introduction to numerical analysis*. Cambridge university press, 2003.

[SN16] S. Saponara and B. Neri. Radar sensor signal acquisition and 3d fft processing for smart mobility surveillance systems. *2016 IEEE Sensors Applications Symposium (SAS)*, pages 1–6, 04 2016.

[SNH09] B. G. Stewart, A. Nesbitt, and L. Hall. Triangulation and 3d location estimation of rfi and partial discharge sources within a 400kv substation. In *2009 IEEE Electrical Insulation Conference*, pages 164–168, 2009.

[SRHG18] S. Semper, F. Roemer, T. Hotz, and G. Del Galdo. Grid-free direction-of-arrival estimation with compressed sensing and arbitrary antenna arrays. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3251–3255. IEEE, 2018.

[TA18] Y. Traonmilin and J-F. Aujol. The basins of attraction of the global minimizers of the non-convex sparse spikes estimation problem. *ArXiv*, abs/1811.12000, 2018.

[TAL19] Y. Traonmilin, J-F. Aujol, and A. Leclaire. Projected Gradient Descent for Non-convex Sparse Spike Estimation. work-

ing paper or preprint, October 2019. URL: https://hal.archives-ouvertes.fr/hal-02311624.

[TBSR13] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht. Compressed sensing off the grid. *IEEE Transactions on Information Theory*, 59(11):7465–7490, Nov 2013.

[Tib96] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996. doi:https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.

[Tro04] J. A. Tropp. Just relax: Convex programming methods for subset selection and sparse approximation. *ICES report*, 404, 2004.

[TW10] J. A Tropp and S. J Wright. Computational methods for sparse solution of linear inverse problems. *Proceedings of the IEEE*, 98(6):948–958, 2010.

[Wei04] A.J. Weiss. Direct position determination of narrowband radio frequency transmitters. *IEEE Signal Processing Letters*, 11(5):513–516, 2004.

[WL15] T. J. Wei and Y. H. Lun. An empirical exploration of a simple deghosting method for multistatic radars. In *2015 IEEE Radar Conference (RadarCon)*, pages 1440–1445, 2015.

[WLW22] Z. Wan, W. Liu, and P. Willett. Non-coherent source localization with distributed sensor array networks. In *2022 IEEE 12th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pages 86–90, 2022.

[Wol20] L. A Wolsey. *Integer Programming*. John Wiley & Sons, 2020.

[Wri15] S. J Wright. Coordinate descent algorithms. *Mathematical programming*, 151(1):3–34, 2015.

[WYL20] L. Wang, Y. Yang, and X. Liu. A direct position determination approach for underwater acoustic sensor networks. *IEEE Transactions on Vehicular Technology*, 69(11):13033–13044, 2020.

[XWR17] Z. J. Xiang, Y. Wang, and P. J. Ramadge. Screening tests for lasso problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(5):1008–1027, 2017. doi:10.1109/tpami.2016.2568185.

[YGD13] T. Yu, Z. Gong, and B. De. Joint sparse modeling for target parameter estimation in distributed mimo radar. *IET International Radar Conference*, 2013:719–719, Jan. 2013.

[YGL+19] L. Yang, H. Gao, B. Li, Y. Yang, and G. Ru. Joint position and velocity estimation of a moving target in multistatic radar by bistatic range, tdoa, and doppler shifts. *International Journal of Antennas and Propagation*, 2019:1–7, 2019.

[YLGD21] Z. Yu, J. Li, Q. Guo, and J. Ding. Efficient direct target localization for distributed mimo radar with expectation propagation and belief propagation. *IEEE Transactions on Signal Processing*, 69:4055–4068, 2021.

[YX14] Z. Yang and L. Xie. Continuous compressed sensing with a single or multiple measurement vectors. In *2014 IEEE Workshop on Statistical Signal Processing (SSP)*, pages 288–291, 2014.

[ZFZ08] C. Zhang, D. Florencio, and Z. Zhang. Why does phat work well in low noise, reverberative environments? In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pages 2565–2568, 03 2008.

[ZW13] S. Zubair and W. Wang. Tensor dictionary learning with sparse tucker decomposition. In *2013 18th international conference on Digital Signal Processing (DSP)*, pages 1–6. IEEE, 2013.