FEDERAL UNIVERSITY OF TECHNOLOGY - PARANÁ
GRADUATE PROGRAM IN ELECTRICAL AND COMPUTER ENGINEERING

THIAGO ALBERTO RIGO PASSARIN

SPARSE ULTRASOUND IMAGING VIA LINEAR APPROXIMATION OF THE ACQUISITION MANIFOLD AND NON-CONVEX GREEDY PURSUIT

DOCTORAL THESIS

CURITIBA

2019

THIAGO ALBERTO RIGO PASSARIN

SPARSE ULTRASOUND IMAGING VIA LINEAR APPROXIMATION OF THE ACQUISITION MANIFOLD AND NON-CONVEX GREEDY PURSUIT

Doctoral Thesis presented to the Graduate Program in Electrical and Computer Engineering of Federal University of Technology - Paraná as partial requirement for obtaining the degree of "Doctor in Sciences" – Research Area: Automation and Systems Engineering.

Supervisor: Prof. Dr. Daniel R. Pipa

Co-supervisor: Prof. Dr. Marcelo V. W. Zibetti

CURITIBA

2019

Dados Internacionais de Catalogação na Publicação

P286s Passarin, Thiago Alberto Rigo
Sparse ultrasound imaging via linear approximation of the acquisition manifold and non-convex greedy pursuit / [recurso eletrônico] / Thiago Alberto Rigo Passarin.-- 2019.
1 arquivo texto (67 f.) : PDF ; $c 1,17 MB.

Texto em inglês com resumo em português.
Tese (Doutorado) - Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial. Área de Concentração: Engenharia de Automação e Sistemas, Curitiba, 2019.
Bibliografia: f. 53-57.

1. Engenharia elétrica - Teses. 2. Ultrassonografia. 3. Algoritmos 4. Reconstrução de imagens. 5. Modelos Matemáticos. 6. Otimização matemática. 7. Teoria da aproximação. 8. Programação não-convexa. 9. Testes não-destrutivos. I. Pipa, Daniel Rodrigues, orient. II. Zibetti, Marcelo Victor Wüst, coorient. III. Universidade Tecnológica Federal do Paraná. Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial. IV. Título.

CDD: Ed. 23 -- 621.3

Biblioteca Central do Câmpus Curitiba – UTFPR
Bibliotecária: Luiza Aquemi Matsumoto CRB-9/794

Ministério da Educação
Universidade Tecnológica Federal do Paraná
Diretoria de Pesquisa e Pós-Graduação

TERMO DE APROVAÇÃO DE TESE Nº 187

A Tese de Doutorado intitulada **"*Sparse Ultrasound Imaging Via Linear Approximation of the Acquisition Manifold and Non-Convex Greedy Pursuit*",** defendida em sessão pública pelo(a) candidato(a) **Thiago Alberto Rigo Passarin**, no dia 15 de fevereiro de 2019, foi julgada para a obtenção do título de Doutor em Ciências, área de concentração Engenharia de Automação e Sistemas, e aprovada em sua forma final, pelo Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial.

BANCA EXAMINADORA:

Prof(a). Dr(a). Daniel Rodrigues Pipa - Presidente – (UTFPR)
Prof(a). Dr(a). Lucia Valeria Ramos de Arruda – (UTFPR)
Prof(a). Dr(a). Elizabeth Wegner Karas – (UFPR)
Prof(a). Dr(a). Solivan Arantes Valente – (UP)
Prof(a). Dr(a). Hugo Vieira Neto - (UTFPR)

A via original deste documento encontra-se arquivada na Secretaria do Programa, contendo a assinatura da Coordenação após a entrega da versão corrigida do trabalho.

Curitiba, 15 de fevereiro de 2019.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the supervisor of this work, Professor Daniel R. Pipa, for his guidance and support and for the insightful idea of approaching the "off-grid problem" in ultrasound imaging via manifold approximation.

I am also sincerely thankful to the co-supervisor of this work, Professor Marcelo V. W. Zibetti, for having introduced this problem to me during the course of my master's degree some years ago and for his support as a careful reviewer of the articles and the thesis developed in the course of this doctorate work.

I would like to thank the members of the evaluation board Professor Elizabeth W. Karas, Professor Solivan A. Valente, Professor Lucia V. R. de Arruda and Professor Hugo Vieira Neto for their valuable contribution.

I also would like to thank Petrobras for financial support through the project CENPES/PETROBRAS 5850.0106184.17.9.

All models are wrong, but some are useful.
George E. P. Box

ABSTRACT

PASSARIN, T. A. R.. SPARSE ULTRASOUND IMAGING VIA LINEAR APPROXIMATION OF THE ACQUISITION MANIFOLD AND NON-CONVEX GREEDY PURSUIT. 67 f. Doctoral Thesis – Graduate Program in Electrical and Computer Engineering, Federal University of Technology - Paraná. Curitiba, 2019.

Model-based image and signal reconstruction has brought important improvements in terms of contrast and spatial resolution to applications such as magnetic resonance imaging and emission computed tomography. However, their use for pulse-echo techniques like ultrasound imaging is limited by the fact that model-based algorithms assume a finite grid of possible locations of scatterers in a medium – an assumption that does not reflect the continuous nature of real world objects and creates a problem known as off-grid deviation. To cope with this problem, we present a method of dictionary expansion and constrained reconstruction that approximates the continuous manifold of all possible scatterer locations within a region of interest (ROI). The creation of the expanded dictionary is based on a highly coherent sampling of the ROI, followed by a rank reduction of the corresponding data that encompasses two possible approximation criteria: one based on singular-value decomposition (SVD) and one minimize-maximum (Minimax). Although we develop here a formulation for two-dimensional sparse imaging problems, it can be readily extended to any D dimensions. We develop a greedy algorithm, based on the Orthogonal Matching Pursuit (OMP), that uses a correlation-based non-convex constraint set that allows for the division of the ROI into cells of any size. To evaluate the performance of the proposed method, we present results of two-dimensional ultrasound image reconstructions with simulated data in a nondestructive testing application. The proposed method succeeds at reconstructing sparse images from noisy measurements and provides higher accuracy than previous approaches based on regular discrete models. Results also confirm a theoretical expectation that the Minimax dictionary outperforms the SVD dictionary on the estimation of the cardinality of the solution.

Keywords: ultrasonic imaging, image reconstruction, optimization methods, greedy algorithms

RESUMO

PASSARIN, T. A. R.. RECONSTRUÇÃO DE IMAGENS ESPARSAS DE ULTRASSOM ATRAVÉS DE APROXIMAÇÃO LINEAR DO *MANIFOLD* DE AQUISIÇÃO E BUSCA ITERATIVA NÃO CONVEXA. 67 f. Tese de Doutorado – Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná. Curitiba, 2019.

Nas últimas décadas, as técnicas de reconstrução de imagens e sinais baseadas em modelos possibilitaram importantes melhorias em termos de contraste e resolução espacial em aplicações como ressonância magnética e tomografia computadorizada. No entanto, o uso de técnicas desse tipo em aplicações de pulso-eco como ultrassom é limitado pelo fato delas pressuporem uma grade finita de possíveis localizações para os refletores existentes num meio – um pressuposto que vai contra a natureza contínua dos objetos do mundo real, o que cria um problema conhecido como desvio da grade. Com o objetivo de superar esse problema, este trabalho apresenta um método de expansão de dicionário e de reconstrução com restrições que aproxima a variedade (comumente referida como *manifold*) contínua dos dados de aquisição formada por todas as localizações possíveis de refletores ao longo de uma região de interesse (RDI). A criação do dicionário expandido baseia-se numa amostragem altamente coerente da RDI, seguida de uma redução de posto matricial nos dados correspondentes para a qual são propostos dois critérios: um baseado em decomposição em valores singulares (SVD) e um baseado em minimização do máximo (Minimax). Embora a formulação seja aqui desenvolvida para o caso de 2 dimensões, a mesma é extensível para quaisquer D dimensões. É proposto um algoritmo baseado no *Orthogonal Matching Pursuit* (OMP), que usa um conjunto de restrições não convexas baseadas em correlação e permite que a RDI seja dividida em células de qualquer tamanho. O método proposto é avaliado através da reconstrução de imagens de ultrassom em 2 dimensões a partir de dados simulados para uma aplicação de ensaios não destrutivos. O método proposto obteve êxito na reconstrução de imagens esparsas a partir de dados de aquisição ruidosos e possibilitou maior acurácia do que abordagens concorrentes baseadas em modelos discretos sem expansão do dicionário. Os resultados confirmaram também uma expectativa teórica de que o dicionário Minimax supera o dicionário SVD com relação à estimativa de cardinalidade da solução.

Palavras-chave: imagens de ultrassom, reconstrução de imagens, métodos de otimização, algoritmos gulosos

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AMRR	Alternate Minimax Rank Reduction
BP	Basis Pursuit
CBP	Continuous Basis Pursuit
COMP	Continuous Orthogonal Matching Pursuit
DAS	Delay-and-sum
ECT	Emission Computed Tomography
FE	Frequency Estimation
IBOMP	Interpolating Band-excluded Orthogonal Matching Pursuit
LS	Least Squares
MMLR	Multiple Most Likely Replacement
MP	Matching Pursuit
MRI	Magnetic Resonance Imaging
NDT	Nondestructive Testing
OMP	Orthogonal Matching Pursuit
OMPED	Orthogonal Matching Pursuit for Expanded Dictionaries
PSF	Point Spread Function
ROI	Region of Interest
SIR	Spatial Impulse Response
SVD	Singular Value Decomposition
TDE	Time Delay Estimation
TSVD	Truncated SVD

LIST OF SYMBOLS

$\cdot^+$	Moore-Penrose pseudoinverse
$\langle \cdot, \cdot \rangle$	inner product
$\| \cdot \|_p$	ℓ_p norm
$\| \cdot \|$	Euclidean (ℓ_2) norm
$\mathbf{0}$	all-zero column vector of proper size
$\mathbf{b}_k^{(n)}$	k-th column of matrix $\mathbf{B}^{(n)}$
$\mathbf{B}^{(n)}$	n-th basis matrix of the expanded model
$\mathbf{B}$	concise representation for $\mathbf{B}^{(n)}$
$\{\mathbf{B}^{(\mathbf{n})}\}$	concise representation for the set of N matrices $\{\mathbf{B}^{(n)}\}_{n=1}^{N}$
$\mathbf{c}$	any arbitrary acquired signal
$\mathcal{C}$	set of constraints imposed to an optimization problem
D	number of nonlinear parameters (usually spacial coordinates) of an acquisition model
e_{est}	estimated total residual norm
e_{noise}	expected ℓ_2 norm of the acquisition noise $\mathbf{w}$
$\mathbf{e}$	residual of a solution
$\mathbf{e}_{\text{rank}}$	estimate of the residual resulting from the low rank reconstruction
f_c	center frequency
f_s	sampling rate
$f_{k,i}^{(n)}$	k-th element of vector $\mathbf{f}_i^{(n)}$
$\mathbf{f}_i$	i-th column of matrix $\mathbf{F}$
$\mathbf{f}_i^{(n)}$	i-th column of matrix $\mathbf{F}^{(n)}$
$\mathbf{F}^{(n)}$	n-th modulating matrix for the expanded model
$\mathbf{F}$	concise representation for $\mathbf{F}^{(n)}$
$\{\mathbf{F}^{(\mathbf{n})}\}$	concise representation for the set of N matrices $\{\mathbf{F}^{(n)}\}_{n=1}^{N}$
$\mathbf{h}_n$	PSF of the n-th modelled position (n-th column of $\mathbf{H}$)
$\mathbf{H}$	regular discrete acquisition model
K	number of columns of each matrix $\mathbf{B}^{(n)}$
$\mathbf{m}_i$	i-th column of matrix $\mathbf{M}$
$\mathbf{M}^{(n)}$	matrix containing the R PSFs of the n-th local manifold
$\mathbf{M}$	concise representation for $\mathbf{M}^{(n)}$
$\tilde{\mathbf{M}}$	rank-K approximation of $\mathbf{M}$
M	number of time samples on an acquired signal
$\mathcal{M}$	manifold of all possible point spread functions within a ROI
$\mathcal{M}_n$	n-th local manifold (for the n-th local ROI)
$\{\mathbf{M}^{(\mathbf{n})}\}$	concise representation for the set of N matrices $\{\mathbf{M}^{(n)}\}_{n=1}^{N}$
$\mathbf{n}$	vector of residual norms
N	number of events events encoded in the acquired signal $\mathbf{c}$
N	number of pixels in an image
$\mathbf{r}_i^{(n)}$	i-th column from $\mathbf{R}^{(n)}$
$r_{\max}$	maximum residual norm

R	number of locations (total) sampled from the local ROI for the construction of the expanded dictionary
R_x	number of locations (lateral direction) sampled from the local ROI for the construction of the expanded dictionary
R_z	number of locations (axial direction) sampled from the local ROI for the construction of the expanded dictionary
$\mathbf{R}^{(n)}$	matrix of residuals from n-th rank reduction
$\mathbf{R}$	concise representation of $\mathbf{R}^{(n)}$
$\mathbb{R}^M$	M-dimensional real-valued vector space
$\mathbf{r}_i$	i-th column of matrix $\mathbf{R}$
s	k-th singular value from an SVD
S	support (set of indices whose elements have nonzero values) of the solution
S^c	complement of S
$\mathbf{S}$	diagonal matrix of singular values from an SVD decomposition
$\tilde{\mathbf{S}}$	truncated matrix of $\mathbf{\Sigma}$ (only the K largest singular values)
$.^T$	transpose of a matrix
u_i	lateral positions of ultrasound transducer during acquisition
$\mathbf{U}$	matrix of left singular values from an SVD decomposition
$\tilde{\mathbf{U}}$	truncated matrix of $\mathbf{U}$ (only the K left-most vectors)
$\mathbf{v}$	vector of scatterer amplitudes
$\hat{\mathbf{v}}$	vector $\mathbf{v}$ recovered by some reconstruction algorithm
v_n	the amplitude of the n-th scatterer or pixel
$\mathbf{V}$	matrix of right singular values from an SVD decomposition
$\tilde{\mathbf{V}}$	truncated matrix of $\mathbf{V}$ (only the K left-most vectors)
$\mathbf{w}$	vector containing additive white Gaussian noise
x	lateral direction
x_n	location of the n-th scatterer or pixel (lateral direction)
$x_k^{(n)}$	k-th element of vector $\mathbf{x}^{(n)}$
$\mathbf{x}^{(n)}$	vector of coefficients that modulates $\mathbf{B}^{(n)}$
$\mathbf{y}(\boldsymbol{\tau})$	PSF of a single unity amplitude event located at position $\boldsymbol{\tau}$
z	axial direction
z_n	location of the n-th scatterer or pixel (lateral direction)
δ_x	spacing between two consecutive positions in the fine sampling of a local ROI (lateral direction)
δ_z	spacing between two consecutive positions in the fine sampling of a local ROI (axial direction)
Δ_x	size of each local ROI (lateral direction)
Δ_z	size of each local ROI (axial direction)
Δ_μ	decreased performed in μ_c in case of infeasibility
ϵ	parameter that limits the energy of the residual of the solution for an optimization problem
ϵ_{AMRR}	parameter of stop criterion for the AMRR algorithm
λ	regularization parameter
μ_c	minimum correlation parameter
σ^2	variance of the noise
$\boldsymbol{\tau}$	vector of D spatial coordinates of a scatterer

SUMMARY

1 INTRODUCTION

1.1 MOTIVATION

Model-based image reconstruction methods provided important advances to imaging techniques such as magnetic resonance imaging (MRI) (FESSLER, 2010) and emission computed tomography (ECT) (OLLINGER; FESSLER, 1997) in the last decades. These methods rely on a known model which results in the captured signal being represented by a sum of N coefficient-weighted responses. These responses are usually Point Spread Functions (PSF), and coefficients are usually intensity of pixels at a modelled location. The discrete model is then fed to regression algorithms along with a vector of acquired data, and the intensity on each pixel is determined (CENSOR, 1983). The use of model-based techniques in pulse-echo imaging relies on a strong assumption: that all reflectors (or scatterers) are located on any of a finite grid of N modelled positions (LAVARELLO et al., 2006). Naturally, real-world inspected objects easily break this assumption and many scatterers may be located *off-grid*. Many previous studies with model-based algorithms for ultrasound imaging –including but not limited to the works of Lavarello et al. (2006), Zanin et al. (2011), Zanin et al. (2012), Desoky et al. (2003), Lingvall et al. (2003), Lingvall and Olofsson (2007), Olofsson and Wennerstrom (2007) and Viola et al. (2008)– have reported that resolution and contrast are substantially improved in comparison to delay-and-sum (DAS) algorithms when data comes from simulations with scatterers located strictly on a modelled grid. However, images are corrupted by artifacts when the grid is not respected, which is typical in data acquired from real measurements. Consequently, DAS beamforming algorithms remain as state-of-the-art for ultrasound imaging, despite having well understood physical limitations regarding spatial resolution (JR, 2014; SMITH; WEBB, 2010).

The present work proposes to address the problem of discretely representing the continuous object, with all possible scatterer locations, by a dictionary expansion. The framework of dictionary expansion was recently proposed for 1-dimensional, shift-invariant problems with focus on neural spike detection (EKANADHAM et al., 2011; KNUDSON

et al., 2014). We generalize the framework so that it can be used in D-dimensional, shift-variant problems. The expanded dictionaries then feed a sparsity-promoting greedy algorithm, which is a generalization of the Orthogonal Matching Pursuit (OMP). The sparsity assumption is in line with images expected in nondestructive testing (NDT) ultrasound, in which small dimension discontinuities from flaws like cracks and oxidation holes are searched.

The document is organized as follows: in Chapter 2, we review the foundations of model-based imaging for pulse-echo applications as well as techniques of reconstruction with regularization and sparsity promotion. Still in Chapter 2, we define off-grid scatterers and explain how dictionary expansion works as a means to cope with them. Two approaches of dictionary expansion (SVD and Minimax) are presented in Chapter 3, both based on a highly coherent sampling of the continuous manifold followed by a rank reduction procedure. We present them as two particular cases of a general formulation based on the optimization of a cost function over the residual created by the rank reduction. Also in Chapter 3, we present a generalization of the OMP (TROPP; GILBERT, 2007) algorithm that works with the proposed expanded dictionaries. The non-convex constraints imposed by the proposed algorithm work out a limitation of previous algorithms regarding the approximation of the region of interest (ROI) to be imaged. Chapter 4 presents results of our technique applied to the reconstruction of two-dimensional ultrasound images with simulated data from an NDT application. Both SVD and Minimax dictionaries are tested with several orders of approximation and under varying levels of noise. An example of resulting image is compared to competing model-based approaches. Finally, Chapter 5 contains a discussion of the results and of some aspects of the proposed technique, as well as directions for future works.

1.2 OBJECTIVES

1.2.1 GENERAL OBJECTIVE

To propose an expansion of the discrete linear acquisition model for ultrasound imaging as a means to minimize the degradation that is caused in images when there are scatterers located off the modelled grid.

1.2.2 SPECIFIC OBJECTIVES

1. To reproduce the results reported in (EKANADHAM et al., 2011) and (KNUDSON et al., 2014) with prototypical 1-dimensional data;
2. To extend the types of dictionary expansion proposed in (EKANADHAM et al., 2011) and (KNUDSON et al., 2014) to 2-dimensional problems or propose a type of dictionary expansion suitable for 2 dimensions;
3. To create representative computer simulations for the problem of detecting and locating point-like scattereres through 2-dimensional ultrasound NDT;
4. To choose an existing algorithm or develop an algorithm suitable for solving the optimization problem formulated using the expanded dictionaries;
5. To reconstruct 2-dimensional images from data simulating the ultrasound acquisition with point-like scatterers.

1.3 RELATED PUBLICATIONS

This work presents two dictionaries to be used with a proposed reconstruction algorithm. The first of them (the SVD expansion) and the reconstruction algorithm have been presented in (PASSARIN et al., 2018) along with some of the results, developments and discussions presented here. The article, available for open access at https://www.mdpi.com/1424-8220/18/12/4097, was published under the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/). The present work contains excerpts, adaptations and extensions of the text of the article.

A less generic method for the construction of Minimax dictionaries than the one presented in the present work was presented in (PASSARIN et al., 2017). The relation and equivalences between the two methods are formalized in Appendix A.

2 THEORETICAL FOUNDATIONS

2.1 MODEL-BASED IMAGING AND REGULARIZATION

Let $\mathbb{R}^M$ be the space of the data observed through an acquisition process. A single, unity amplitude event located at position $\boldsymbol{\tau} \in \mathbb{R}^D$ (in the D-dimensional continuous space) causes the discrete acquired signal $\mathbf{y}(\boldsymbol{\tau}) \in \mathbb{R}^M$, known as the PSF. The physical meaning of such event depends on the type of quantity being measured. In pulse-echo applications such as ultrasound and radar imaging, the event denotes a point-like reflexivity (also called a scatterer) (JENSEN, 1991, 2004), as represented in Fig. 1, and D typically equals 2 as the reflexivity is being mapped over a two-dimensional plane. The variation of the set of D parameters $\boldsymbol{\tau}$ within a region of interest describes a D-dimensional manifold

$$\mathcal{M} := \{\mathbf{y}(\boldsymbol{\tau}) : \boldsymbol{\tau} \in \text{ROI}\} \tag{1}$$

of all possible PSFs on $\mathbb{R}^M$. We will develop our notation for the two-dimensional case and consider the two parameters $\boldsymbol{\tau} = [x, z]^T$ (where $\cdot^T$ denotes the transpose) as the lateral and axial spatial dimensions respectively.

An acquired signal $\mathbf{c} \in \mathbb{R}^M$ is assumed to be composed by a sum of individual contributions from N events, or N *samples* from the continuous PSF manifold

$$\mathbf{c} = \sum_{n=1}^{N} v_n \mathbf{y}(x_n, z_n) + \mathbf{w}, \tag{2}$$

where v_n is the amplitude of the n-th event and the vector $\mathbf{w} \in \mathbb{R}^M$ accounts for acquisition noise, which we will assume to be Gaussian white noise with variance σ^2.

In a pulse-echo image with N pixels, v_n in (2) encodes the reflexivity of the n-th scatterer, located at position (x_n, z_n), and is represented as the brightness of the corresponding pixel. This naturally implies a sampling of the parameters (x, z) as a finite number N of possible scatterer locations (or pixels) is assumed.

Once we have defined the N coordinate pairs (x_n, z_n) to be considered by the

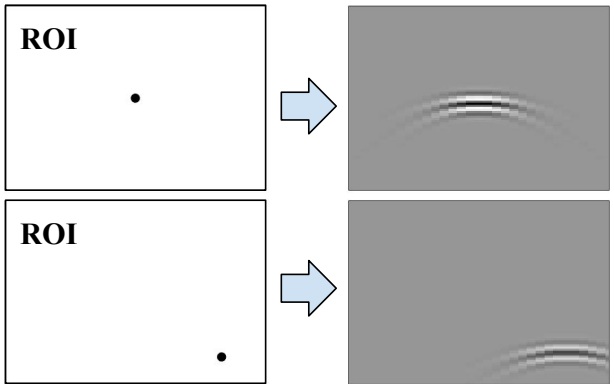


Figure 1: Acquisition of the point spread function (PSF). For each position (x, z) of the unity amplitude scatterer within the ROI (left side), an M-sample response $\mathbf{y}(x, z) \in \mathbb{R}^M$ (arranged as an M-pixel image, known as the B-scan, on the right side) is generated by the acquisition model. The set of all possible PSFs within the region of interest forms a manifold $\mathcal{M}$ onto the data space. This example is taken from the pulse-echo ultrasound model described in Section 3.3.
Source: (PASSARIN et al., 2018).

acquisition model, we make $\mathbf{h}_n = \mathbf{y}(x_n, z_n),\ n = 1, \ldots, N$, and define the model matrix $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_N] \in \mathbb{R}^{M \times N}$. Then (2) can be written in compact form as

$$\mathbf{c} = \mathbf{H}\mathbf{v} + \mathbf{w}, \tag{3}$$

where $\mathbf{v} = [v_1, \ldots, v_N]^T$ is the vector of scatterer amplitudes. This model has been used in B-mode (two-dimensional) (DESOKY et al., 2003; ZANIN et al., 2011, 2012; LINGVALL et al., 2003; LAVARELLO et al., 2006; LINGVALL; OLOFSSON, 2007), A-mode (one-dimensional) (MOR et al., 2010; CARCREFF et al., 2014), and three-dimensional (KRUIZINGA et al., 2017) ultrasound imaging.

The reconstruction of a vector of amplitudes $\hat{\mathbf{v}}$ from a given acquisition $\mathbf{c}$ in (3) is based on the minimization of a cost function, such as the Least Squares (LS) problem

$$\hat{\mathbf{v}} = \arg\min_{\mathbf{v}} \|\mathbf{c} - \mathbf{H}\mathbf{v}\|^2, \tag{4}$$

which has a closed solution with the Moore-Penrose pseudoinverse

$$\hat{\mathbf{v}} = \mathbf{H}^{+}\mathbf{c} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{c}. \tag{5}$$

This is a linear problem and, although the inversion on (5) may lead to instability when performed directly, there are well-known iterative methods to solve the problem (GOLUB; LOAN, 1996).

However, model matrices for real-world problems are often ill-conditioned, which

causes artifacts on the reconstructed signals in the presence of noise (HANSEN, 1998). This is an issue even in reconstructions with simulated data where all events are on grid, i.e., where the discrete acquisition model (3) is obeyed. The specific problem of poor conditioning of the ultrasound acquisition model has been addressed with linear regularization methods such as Truncated SVD (TSVD) (DESOKY et al., 2003) and Tikhonov regularizarion (ZANIN et al., 2011, 2012; LINGVALL et al., 2003), where the main goal is to stabilize the inverse operator.

Nonlinear, sparsity-promoting regularization penalties such as ℓ_p-(pseudo)norm minimization with $p \leq 1$ have shown successful results in ultrasound NDT, where the assumption of sparsity in the space domain reflects the nature of discontinuities in observed materials (LAVARELLO et al., 2006; LINGVALL; OLOFSSON, 2007; GUARNERI et al., 2015; CARCREFF et al., 2014).

Greedy algorithms effectively solve reconstruction problems in which the cost function involves the ℓ_0 pseudonorm. In (OLOFSSON; WENNERSTROM, 2007), sparsity is induced in the solution assuming that the presence of scatterers can be modelled by a Bernoulli process with a low value for the probability parameter. The problem is then solved with a greedy algorithm called Multiple Most Likely Replacement (MMLR) (CHI et al., 1985). In (MOR et al., 2010), a Gabor dictionary is used in the reconstruction of thickness with a Matching Pursuit (MP)-based algorithm that penalizes a relaxed support measure corresponding to the ℓ_p-pseudonorm with $0 < p < 1$.

2.2 OFF-GRID EVENTS AND DICTIONARY EXPANSION

Aside from poor matrix conditioning, another problem known as off-grid deviation (EKANADHAM et al., 2011) limits the applicability of model-based approaches on signal and image reconstruction. It derives from the fact that, in many applications, the existing scatterers may not be located strictly on the N positions modelled by (2) and (3), i.e., many scatterers may be off-grid. Fig. 2a illustrates a grid of $N = 9$ modelled positions, represented by gray dots. Since three scatterers (represented by black dots) are located on modelled positions, then the corresponding data vector $\mathbf{c}$ can be synthesized according to the acquisition models (2) and (3). The same does not hold when an off-grid scatterer (represented by a red dot) is added: attempts to reconstruct the locations and amplitudes for the corresponding scatterers may fail, causing artifacts and degradation on the reconstructed image.

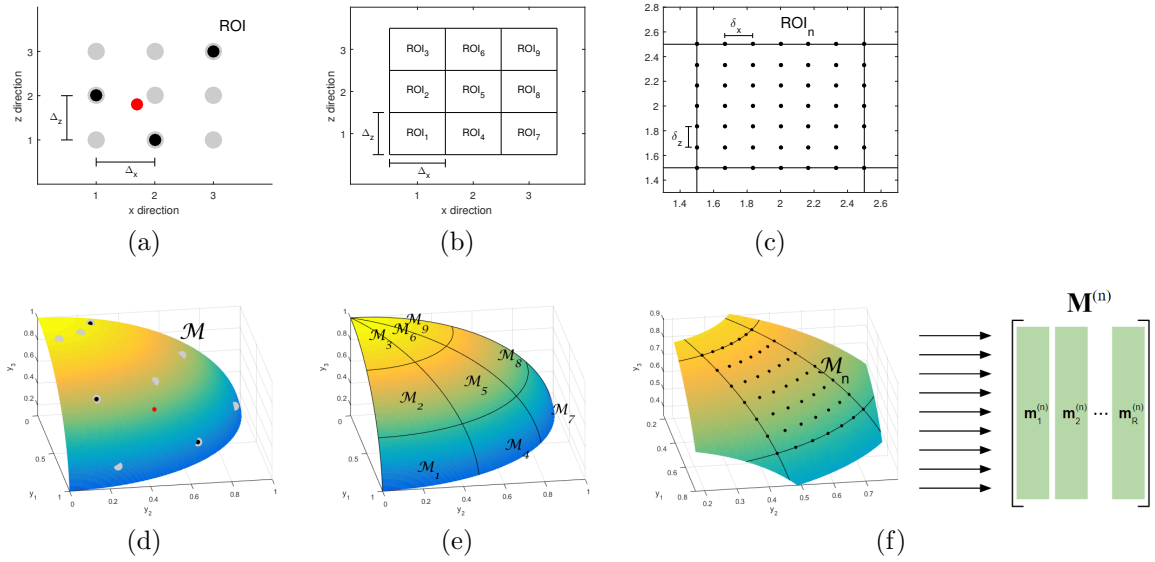


Figure 2: Local ROIs and sampling of local manifolds. (a) An illustrative discrete acquisition model with $N = 3 \times 3 = 9$ modelled positions, represented by the gray dots. The black dots represent 3 well located scatterers and the red dot represents an off-grid scatterer. Because of the latter, the corresponding acquisition data vector c cannot be synthesized as a linear combination of the columns of the discrete model matrix H. (b) The ROI is divided into N local ROIs with area $\Delta_x \times \Delta_z$. (c) Each local ROI is sampled with a fine grid with lateral and axial distances δ_x and δ_z. (d) On the space $\mathbb{R}^M$ of acquired data, the set of all possible PSFs within the ROI forms a D-manifold. The gray dots are the PSFs of the modelled positions of Fig. 2a. The black dots are on the grid, while the red dot is off-grid. (e) As the ROI is divided into N local ROIs (Fig. 2b), the manifold is divided into N corresponding local manifolds. (f) The acquisitions over the fine grid on each n-th local ROI create R samples from the corresponding local manifold. Those samples compose matrix $\mathbf{M}^{(n)} \in \mathbb{R}^{M \times R}$.
Source: (PASSARIN et al., 2018).

Some formulations have been proposed for off-grid signal reconstruction, mainly within the framework of Compressive Sensing. In (YANG et al., 2012), the acquisition model considers a perturbation matrix summed column-wise to the (here referred to as **H**) regular discrete model matrix. The formulation is applied to direction-of-arrival (DOA) estimation using the derivatives of the columns of **H** with respect to the sampled parameters as perturbation matrix. In (TEKE et al., 2014), an adaptation of the Orthogonal Matching Pursuit (OMP) algorithm is proposed, in which the columns of the model matrix are iteratively updated in order to accommodate variations in the parameters of the PSFs. The algorithm is applied to pulse-Doppler radar. In (TANG et al., 2013), the problem of continuous line spectral estimation is approached with an algorithm based on the atomic norm minimization, which is solved via semi-definite programming. Similarly to the ℓ_1 minimization, the atomic norm minimization promotes sparse solutions. In

(ZHU et al., 2011), the regression problem uses a Total Least Squares (TLS) penalization with sparsity constraints. The motivation is that the "errors-in-variables" assumption of the TLS regression might be able to capture the mismatch between the model matrix and the acquired data. The method is then applied to cognitive radio sensing and DOA estimation.

The framework of dictionary expansion was proposed in (EKANADHAM et al., 2011) as a means to overcome the problem of off-grid deviation in neuron spike detection. Each column $\mathbf{h}_n$ of the discrete model $\mathbf{H}$ of (3) is replaced by K columns $[\mathbf{b}_1^{(n)}, \ldots, \mathbf{b}_K^{(n)}] = \mathbf{B}^{(n)} \in \mathbb{R}^{(M \times K)}$ so that a data vector $\mathbf{c}$ resulting from the acquisition of an scatterer located in the neighborhood of an n-th modelled position can be approximated by some linear combination of $\mathbf{B}^{(n)}$, i.e., by $\mathbf{B}^{(n)}\mathbf{x}^{(n)}$, where $\mathbf{x}^{(n)} \in \mathbb{R}^K$. As a result, an arbitrarily acquired $\mathbf{c}$ might be approximated as

$$\mathbf{c} \approx \sum_{n=1}^{N} \mathbf{B}^{(n)}\mathbf{x}^{(n)} + \mathbf{w}. \tag{6}$$

In the 2-dimensional case, the neighborhood of the n-position is the region within $(x_n \pm 0.5\Delta_x, z_n \pm 0.5\Delta_z)$. This is represented in Fig. 2b, where the 9 modelled locations give place to 9 neighborhoods (local ROIs).

Two forms of approximation are proposed in (EKANADHAM et al., 2011) for 1-dimensional linear time-invariant (LTI) problems. The first one is the Taylor approximation, which relies on the fact that small shifts on a waveform can be well approximated by its Taylor expansion, i.e., by linearly combining the original waveform and its time derivatives. In this case, the column $\mathbf{b}_1^{(n)}$ is identical to the original atom $\mathbf{h}_n$ and the columns $\mathbf{b}_k^{(n)}$ for $k > 1$ correspond to its $(k-1)$-th time derivatives. The second is the Polar approximation, which is motivated by the fact that the continuous manifold $\mathcal{M}$ of an LTI system lies over a hypersphere on the M-dimensional data space (EKANADHAM et al., 2011). The PSFs of the neighborhood of each n-th modelled position are approximated by an arc of a circle and the the column $\mathbf{h}_n$ is replaced by three normal vectors with the directions of the center ($\mathbf{b}_1^{(n)}$) and the two trigonometric components ($\mathbf{b}_2^{(n)}$ and $\mathbf{b}_3^{(n)}$) of the circle. While the Taylor approximation can be done for any order K, in the Polar case K always equals 3.

An extension of the Basis Pursuit (BP) formulation (CHEN et al., 1998), referred to as Continuous Basis Pursuit (CBP), is proposed in (EKANADHAM et al., 2011) for the recovery of the expanded coefficients $\{\mathbf{x}^{(n)}\}_{1 \leq n \leq N}$. For the sake of conciseness, from this point on we will represent sets $\{\mathbf{x}^{(n)}\}_{1 \leq n \leq N}$ simply as $\{\mathbf{x}^{(n)}\}$. The formulation of

CBP is given by

$$\{\hat{\mathbf{x}}^{(n)}\} = \arg\min_{\{\mathbf{x}^{(n)}\}} \frac{1}{2\sigma^2}\|\mathbf{c} - \sum_{n=1}^{N} \mathbf{B}^{(n)}\mathbf{x}^{(n)}\|^2 + \lambda \sum_{n=1}^{N} |x_1^{(n)}| \text{ s.t. } \{\mathbf{x}^{(n)}\} \in \mathcal{C} \tag{7}$$

where the constraint set $\mathcal{C}$ prevents recovered expanded coefficients from having any arbitrary values that do not represent actual PSFs. The definition of the convex set $\mathcal{C}$ varies according to the type of approximation used. The sum of absolute values of the first element $x_1^{(n)}$ of each K-tuple $\mathbf{x}^{(n)}$ is added to the cost function to obtain sparse solutions. The parameter λ controls the weights between the fidelity of the solution to the sparsity prior and to the acquired data.

In (KNUDSON et al., 2014), a low-rank approximation of the PSFs within the neighborhood of each n-th modelled position is performed by means of a Singular Value Decomposition (SVD). The continuous manifold drawn by $\boldsymbol{\tau}$ in a local ROI is sampled with a very fine grid of R locations, generating R columns that form a matrix $\mathbf{M}^{(n)} \in \mathbb{R}^{M \times R}$, as represented in Fig. 2f. Each matrix $\mathbf{M}^{(n)}$ then undergoes an SVD decomposition and the K first left singular vectors compose the corresponding expanded coefficients $\mathbf{B}^{(n)}$ for the n-th local ROI.

An adaptation of the Orthogonal Matching Pursuit (OMP) (TROPP; GILBERT, 2007) algorithm, referred to as Continuous OMP (COMP), is also presented in (KNUDSON et al., 2014). It aims at solving the $\ell_2 - \ell_0$ problem

$$\{\hat{\mathbf{x}}^{(n)}\} = \arg\min_{\{\mathbf{x}^{(n)}\}} \|[x_1^{(1)}, \ldots, x_1^{(N)}]^T\|_0 \text{ s.t. } \begin{Bmatrix} \|\mathbf{c} - \sum_{n=1}^{N} \mathbf{B}^{(n)}\mathbf{x}^{(n)}\|^2 \leq \epsilon \\ \{\mathbf{x}^{(n)}\} \in \mathcal{C} \end{Bmatrix} \tag{8}$$

where the symbol $\|\cdot\|_0$ denotes the ℓ_0 pseudonorm, i.e., the cardinality (number of nonzero elements) of a vector and the parameter ϵ limits the energy of the residual of the solution.

In (PASSARIN et al., 2017), a minimize-maximum (Minimax) formulation is presented for the definition of the expanded set $\{\mathbf{B^{(n)}}\}$. The resulting approximation minimizes the maximum residual within the representation of each n-th local ROI. It is motivated by the assumption that the off-grid deviation from a discrete grid follows a uniform distribution, therefore the off-grid error should be as constant as possible, not privileging any distance from originally modelled positions.

3 DEVELOPMENTS

This chapter describes the formulations, methods and algorithms developed to cope with the problem presented as motivation for this work.

Section 3.1 describes how the framework of dictionary expansion was applied to the problem of 2-D pulse-echo ultrasound imaging and is divided in three subsections. Subsection 3.1.1 describes the division of the ROI into local ROIs and the fine sampling performed along the considered dimensions, which is the first step on the construction of both dictionary expansions considered. Subsection 3.1.2 presents the adaptation of the SVD expansion (KNUDSON et al., 2014) to our problem, while a new expansion method referred to as the Minimax expansion is presented in Subsection 3.1.3.

Section 3.2 presents the reconstruction algorithm. The motivation for the development of a new algorithm is presented in Subsection 3.2.1. In Subsection 3.2.2, the core idea of the proposed algorithm (the non-convex constraint set) is presented and formalized. The algorithm is then described in Subsections 3.2.3 and 3.2.4.

Section 3.3 presents the acquisition set that was simulated in order to generate the dataset used for validation and performance analysis of the proposed algorithm and dictionaries.

3.1 RANK-K APPROXIMATION OF LOCAL MANIFOLDS

The core idea of dictionary expansion is the substitution of each n-th column $\mathbf{h}_n$ from the discrete model $\mathbf{H}$ by K basis vectors $[\mathbf{b}_1^{(n)}, \ldots, \mathbf{b}_K^{(n)}] = \mathbf{B}^{(n)}$ whose column space approximates the n-th local PSF manifold $\mathcal{M}_n$. We develop two criteria to determine $\mathbf{B}^{(n)}$: SVD and Minimax. Both have been proposed for 1-dimension, shift-invariant problems (KNUDSON et al., 2014; PASSARIN et al., 2017). Their extension to D-dimension problems relies mainly on the first step, which is a fine sampling of each local manifold $\mathcal{M}_n$: here the regular, fine grid is defined for all D dimensions. As for the SVD expansion, this extension is possible due to the fact that its formulation is non-parametric, i.e., the

deviation from modelled positions is not mapped onto any independent variable and does not play any role on the definition on the bases. On the other hand, in the Taylor and Polar expansions (EKANADHAM et al., 2011), modelled positions are part of the dictionary and the off-grid deviation is a parameter from which the remaining elements are derived. Consequently, their extension to 2 or higher dimensions is not defined. While the original formulation of the Minimax expansion (PASSARIN et al., 2017) is parametric (the off-grid deviation is mapped onto a variable which is used to cast a polynomial regression), we generalize the formulation to eliminate the need of regression functions and make the corresponding weights also subject to optimization as a modulating matrix $\mathbf{F}^{(n)}$, as explained in Subsection 3.1.3.

3.1.1 HIGHLY COHERENT DISCRETE LOCAL MANIFOLDS

Fig. 2d shows an illustrative example of a D-manifold embedded in an M-dimension data space. In this case, $D = 2$ and $M = 3$. It can be, for instance, a PSF manifold resulting from the shifting of spatial coordinates x and z of a unitary amplitude event on the plane represented in Fig. 2a, followed by the acquisition of the data vector $\mathbf{c} \in \mathbb{R}^M$ through an arbitrary acquisition process. The nine D-dimension modelled positions shown in Fig. 2a correspond here to nine samples of the M-dimensional manifold, as well represented by gray dots in Fig. 2d. The red dot corresponds to the data caused by the off-grid reflector from Fig. 2a.

Fig. 2e shows the same manifold as Fig. 2d but, instead of having N modelled positions, it divides the manifold into N local manifolds. Given $\Delta_x, \Delta_z > 0$, for each (x_n, z_n) we define the n-th local manifold

$$\mathcal{M}_n := \{\mathbf{y}(x, z) : x \in [x_n - 0.5\Delta_x, x_n + 0.5\Delta_x], z \in [z_n - 0.5\Delta_z, z_n + 0.5\Delta_z]\}. \quad (9)$$

It shall be noted that no correspondence is required between the positions of a simple model and the local manifolds of an expanded model, but we keep it for unity of notation.

We start by performing a fine sampling on each local manifold $\mathcal{M}_n$, as represented in Fig. 2f. In practice, this means acquiring the PSF of a set of points from a fine grid of R points defined for each local ROI (Fig. 2c). The result is a matrix $\mathbf{M}^{(n)} \in \mathbb{R}^{M \times R}$ whose columns are local manifold samples. The finer this grid is, the better the approximation of the local manifold. For simplicity of notation, we keep regular spacing δ_x and δ_z for the lateral and axial directions respectively. The number of sampled points is $R = R_x \times R_z$,

where $R_x > 1$ and $R_z > 1$ are the number of locations defined on the lateral and axial directions respectively. In the example of Fig. 2c, $R_x = R_z = 7$, thus $R = 49$.

Our sampling includes the boundaries of the local ROIs. For this reason, the relation between the spacing and the number of locations on the lateral direction is given by

$$\delta_x = \frac{\Delta_x}{R_x - 1} \tag{10}$$

and the same holds for all other directions.

Once we have the local discrete manifolds $\{\mathbf{M}^{(n)}\}$, we create a rank-K approximation for each of them and define the sets of K basis vectors $\{\mathbf{B}^{(n)}\}$, which form orthonormal bases for such approximations, to be later used on inverse reconstruction problems such as (7) and (8).

For a clean notation throughout the remaining of Section 3.1, we suppress indices n and represent matrices related to the n-th local ROI such as $\mathbf{M}^{(n)}$ and $\mathbf{B}^{(n)}$ as simply $\mathbf{M}$, $\mathbf{B}$ and so forth. It shall be noted that the processes presented in those sections have to be independently performed for every n-th local ROI. Although the construction of expanded dictionaries is computationally demanding, it is an offline procedure that is carried only once for each given acquisition set.

For each local matrix $\mathbf{M}$, a rank-K approximation $\tilde{\mathbf{M}} \in \mathbb{R}^{M \times R}$ is to be defined and also factorized in the form of Eq. (11):

$$\tilde{\mathbf{M}} = \mathbf{BF}, \tag{11}$$

where $\mathbf{B}$ is an orthonormal basis matrix and $\mathbf{F} \in \mathbb{R}^{K \times R}$ modulates $\mathbf{B}$ to form $\tilde{\mathbf{M}}$. We assume that $\text{rank}(\mathbf{M}) > K$, tipically $\text{rank}(\mathbf{M}) = R$. Any approximation creates a residuals matrix $\mathbf{R} \in \mathbb{R}^{M \times R}$, as shown in Fig. 3. Our two approximation approaches are defined by the minimization of a cost function that operates on $\mathbf{R}$. Having $\mathbf{B}$ orthonormal is required by the algorithm presented in Section 3.2.

As can be seen in Fig. 3, each i-th column from $\mathbf{R}$ contains the residual for the approximation of the corresponding i-th local manifold sample from $\mathbf{M}$. To work each residual independently, we define in Eq. (12) a vector of residual norms $\mathbf{n} \in \mathbb{R}^R$ which will be used in the cost function

$$\mathbf{n} = [\|\mathbf{r}_1\|, \|\mathbf{r}_2\|, \ldots, \|\mathbf{r}_R\|]^T. \tag{12}$$

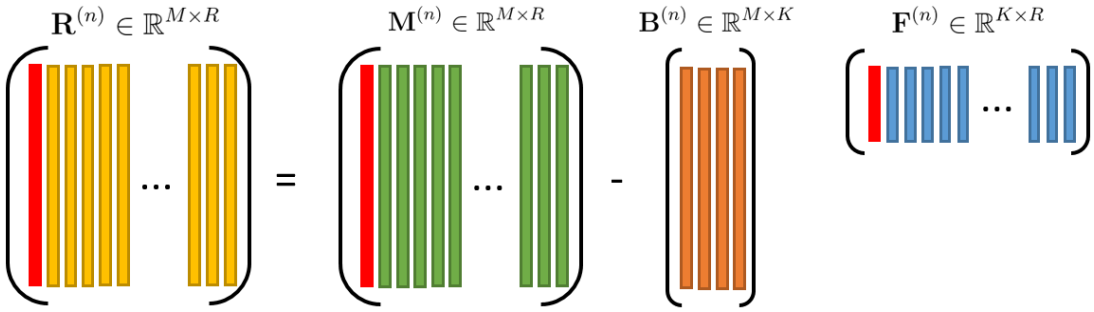


Figure 3: Structure of the matrix equations. M contains *R* local manifold samples (see Fig. 2f) and is approximated by the rank-*K* product BF, resulting in residuals R. The SVD expansion is defined by minimizing the Frobenius norm of R, while the Minimax expansion is defined by minimizing the maximum residual norm $\max_i \|\mathbf{r}_i\|$ among the columns of R. The first columns of R, M and F are highlighted to show that the minimization of the residual norm $\|\mathbf{r}_i\|$ with respect to F (for a given B) depends only on the corresponding column $\mathbf{f}_i$, making the problem separable into *R* linear problems.

Now we define the cost function as the ℓ_p norm of $\mathbf{n}$:

$$\hat{\mathbf{B}}, \hat{\mathbf{F}} = \arg\min_{\mathbf{B},\mathbf{F}} \|\mathbf{n}\|_p \text{ s.t. } \mathbf{R} = \mathbf{M} - \mathbf{BF} \tag{13}$$

Subsections 3.1.2 and 3.1.3 detail our two approximation criteria for the definition of the dictionary expansion, which follow from setting $p = 2$ and $p \to \infty$ on the problem of Eq. (13).

3.1.2 SVD EXPANSION

When $p = 2$ on Eq. (13), the problem boils down to finding the matrices $\mathbf{B}$ and $\mathbf{F}$ that minimize the Frobenius norm (GOLUB; LOAN, 1996) of the residuals matrix $\mathbf{R}$. According to the Eckart-–Young theorem, a solution for this case is achieved by a truncated SVD (ECKART; YOUNG, 1936). Consider the SVD of $\mathbf{M}$

$$\mathbf{M} = \mathbf{USV}^T, \tag{14}$$

where $\mathbf{U} \in \mathbb{R}^{M\times M}$ is the unitary matrix of left singular vectors, $\mathbf{S} \in \mathbb{R}^{M\times R}$ is the diagonal matrix of singular values and $\mathbf{V} \in \mathbb{R}^{R\times R}$ is the unitary matrix of right singular vectors (GOLUB; LOAN, 1996). The rank-K SVD truncation is obtained by using only the K largest singular values from $\mathbf{S}$ and the K corresponding vectors from $\mathbf{U}$ and $\mathbf{V}$. This low rank approximation is given by

$$\tilde{\mathbf{M}} = \tilde{\mathbf{U}}\tilde{\mathbf{S}}\tilde{\mathbf{V}}^T, \tag{15}$$

where $\tilde{\mathbf{M}} \in \mathbb{R}^{M\times R}$, $\tilde{\mathbf{U}} \in \mathbb{R}^{M\times K}$, $\tilde{\mathbf{S}} \in \mathbb{R}^{K\times K}$ and $\tilde{\mathbf{V}} \in \mathbb{R}^{R\times K}$.

The K columns of $\tilde{\mathbf{U}}$ form an orthonormal basis for $\tilde{\mathbf{M}}$ and compose the expanded set $\mathbf{B}$, while the product $\tilde{\mathbf{S}}\tilde{\mathbf{V}}^T$ composes the modulating matrix $\mathbf{F}$:

$$\mathbf{B} = \tilde{\mathbf{U}}, \tag{16a}$$

$$\mathbf{F} = \tilde{\mathbf{S}}\tilde{\mathbf{V}}^T. \tag{16b}$$

This criterion for the choice of the expanded set defines the order-K SVD Expansion of the discrete acquisition model. Naturally, large values for K mean more degrees of freedom in the approximation, which reduces the residuals. Fig. 4a shows how the value of K affects the average and maximum residual norms for the centermost local ROI of the acquisition set presented in Section 3.3.

3.1.3 MINIMAX EXPANSION

When $p \to \infty$, Eq. (13) becomes a minimize-maximum (Minimax) problem, as the ℓ_∞ norm extracts the maximum absolute value among the elements of a vector (BJÖRCK, 1996). In this case, the solution is the pair $\mathbf{B}, \mathbf{F}$ that minimizes the maximum residual norm among all $n_i = \|\mathbf{r}_i\|$. The values obtained for the other, nonmaximal elements of $\mathbf{n}$ are not taken into account in the problem.

As we will see in Section 3.2, the residuals from recovered events are used in each iteration of our greedy algorithm to decide whether the expected reconstruction residual has been achieved (causing the algorithm to stop) or not. Due to factors such as acquisition noise, modelling errors and the greedy nature of the algorithm, the recovery of event positions within their respective local ROIs is subject to inaccuracies. This becomes an important problem when the residual norm $\|\mathbf{r}_i\|$ has a high variance within the local ROI (as i varies from 1 to R) as in the example of Fig. 4b. That is the main motivation for the development of the Minimax expansion. The use of a cost function with $p \to \infty$ imposes a tendency of decreasing the variance among the residual norms $\|\mathbf{r}_i\|$. It has been used, for instance, in filter design, where regions of flat magnitude response are desired (LU, 2002; SHPAK, 2003).

To find a solution for the Minimax problem, we alternate between the optimization of $\mathbf{B}$ and $\mathbf{F}$ until a convergence criterion is met. As an initial guess for $\mathbf{F}$, we use the matrix which is optimal for the quadratic case, i.e. the product of Eq. (16b). We performed several tests where $\mathbf{F}$ was initialized with white Gaussian noise. In all cases

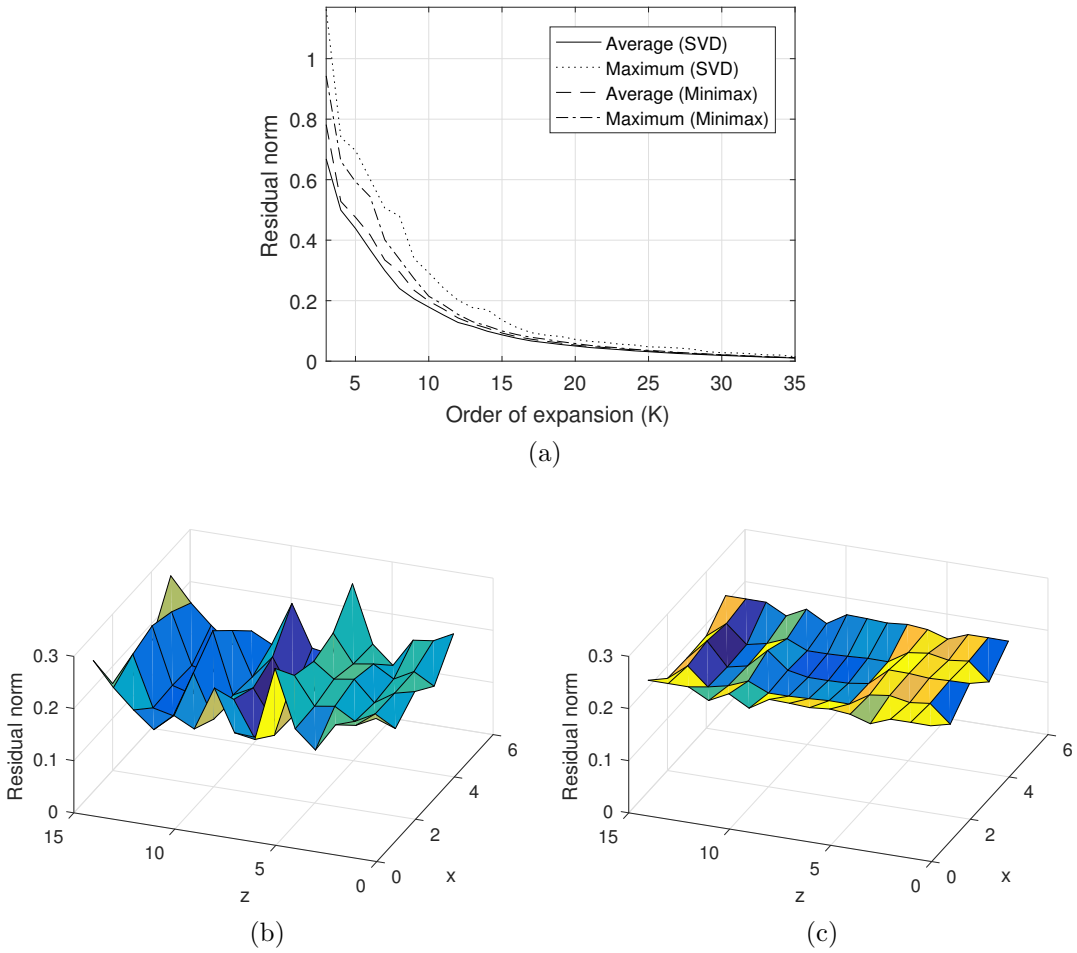


Figure 4: Residual norms $\|\mathbf{r}_i\|$ from the low rank approximation of the centermost local manifold of the ultrasound acquisition set presented in Section 3.3, with $R = 75$ ($R_x = 5$ and $R_z = 15$). (a) Average and maximum values from SVD and Minimax approximations with K ranging from 3 to 35. While the SVD expansion yields smaller average values, the maximum values are smaller with the Minimax expansion, as well as difference between the average and the maximum. (b) Individual residual norms from SVD approximation with $K = 10$, spatially arranged according the corresponding positions of the local ROI. (c) Individual residual norms from Minimax approximation with $K = 10$, spatially arranged according the corresponding positions of the local ROI. Notice that the surface formed by the residual norms from the Minimax expansion is significantly flatter than that from SVD expansion. A consequence is higher accuracy on the estimation of this residual by the reconstruction algorithm described in Subsection 3.2.3 when using the Minimax dictionaries.
Source: extended from (PASSARIN et al., 2018).

the resulting basis vectors $\mathbf{B}$ spanned the same column space as that obtained when $\mathbf{F}$ was initialized with (16b), but the algorithm took roughly twice the number of iterations to reach the convergence criterion.

For a given modulating matrix $\mathbf{F}$, we search for the basis matrix $\mathbf{B}$ that solves the problem of Eq. (17):

$$\hat{\mathbf{B}} = \arg\min_{\mathbf{B}} \|\mathbf{n}\|_\infty \text{ s.t. } \mathbf{R} = \mathbf{M} - \mathbf{BF}. \tag{17}$$

We recast the problem of Eq. (17) as a Second Order Cone Program (SOCP) (BOYD; VANDENBERGHE, 2004; ANTONIOU; LU, 2007) by introducing a scalar variable $r_{\max}$, which is then minimized while also constrained to be equal to or grater than each residual norm $\|\mathbf{m}_i - \mathbf{Bf}_i\|$:

$$\hat{\mathbf{B}} = \arg\min_{\mathbf{B}} r_{\max} \text{ s.t. } \|\mathbf{Bf}_i - \mathbf{m}_i\| \leq r_{\max},\ \forall i \in \{1, \ldots, R\}. \tag{18}$$

We solve the SOCP of Eq. (18) with SDPT3 package for Matlab (TOH et al., 1998). Appendix B details the equivalence between the elements of (18) and those of the standard form of SOCPs.

We now turn to the optimization of the modulating matrix $\mathbf{F}$ for a given basis $\mathbf{B}$, i.e.

$$\hat{\mathbf{F}} = \arg\min_{\mathbf{F}} \|\mathbf{n}\|_\infty \text{ s.t. } \mathbf{R} = \mathbf{M} - \mathbf{BF}. \tag{19}$$

Differently from the optimization of $\mathbf{B}$ for a given $\mathbf{F}$, this problem is linear. As can be observed in Fig. 3, for a given matrix $\mathbf{B}$, the problem of Eq. (19) can turn to the optimization of $\mathbf{F}$ according to Eq. (20):

$$\hat{\mathbf{F}} = \arg\min_{\mathbf{F}} \max_i \|\mathbf{r}_i\| \text{ s.t. } \mathbf{r}_i = \mathbf{m}_i - \mathbf{Bf}_i,\ \forall i \in \{1, 2, \ldots, R\}. \tag{20}$$

But the minimization of each residual norm $\|\mathbf{r}_i\|$ on the problem of Eq. (20) is a totally independent LS problem which can be solved with the pseudoinverse of $\mathbf{B}$:

$$\mathbf{f}_i = \mathbf{B}^+\mathbf{m}_i,\ \forall i \in \{1, 2, \ldots, R\}. \tag{21}$$

By simply grouping the R vector equations on Eq. (21) we obtain a compact form for the solution of the problem of Eq. (20):

$$\mathbf{F} = \mathbf{B}^+\mathbf{M}. \tag{22}$$

The inverted bar operator "\" in Matlab was used to solve Eq. (22). When the inversion of Eq. (22) is performed, every residual norm $\|\mathbf{r}_i\|$ achieves its lowest possible value (for a given $\mathbf{B}$). The resulting $\mathbf{F}$, when compared to the one obtained by Eq. (16b), either preserves or decreases the maximum residual norm.

We perform an alternate optimization of $\mathbf{B}$ and $\mathbf{F}$ solving the problems of Eq. (17) and (20) respectively. The alternate method, named Alternate Minimax Rank Reduction (AMRR), is summarized in Algorithm 1.

Algorithm 1 Alternate Minimax Rank Reduction (AMRR)

Input: $\mathbf{M}$, K, ϵ_{AMRR}

1: $[\tilde{\mathbf{U}}, \tilde{\mathbf{S}}, \tilde{\mathbf{V}}] \leftarrow \text{TSVD}(\mathbf{M}, K)$
2: $\mathbf{F} \leftarrow \tilde{\mathbf{S}}\tilde{\mathbf{V}}^T$
3: $\mathbf{R} \leftarrow \mathbf{M} - \tilde{\mathbf{U}}\mathbf{F}$
4: $r_{\text{new}} \leftarrow \max_i \|\mathbf{r}_i\|$
5: **repeat**
6: $\quad r_{\text{old}} \leftarrow r_{\text{new}}$
7: $\quad \mathbf{B} \leftarrow \arg\min_{\mathbf{B}} r_{\max}$ s.t. $\|\mathbf{B}\mathbf{f}_i - \mathbf{m}_i\| \leq r_{\max}\ \forall i \in \{1, \ldots, R\}$
8: $\quad \mathbf{F} \leftarrow \mathbf{B}^{+}\mathbf{M}$
9: $\quad \mathbf{R} \leftarrow \mathbf{M} - \mathbf{B}\mathbf{F}$
10: $\quad r_{\text{new}} \leftarrow \max_i \|\mathbf{r}_i\|$
11: **until** $(r_{\text{old}} - r_{\text{new}})/r_{\text{old}} < \epsilon_{\text{AMRR}}$
12: $\mathbf{B} \leftarrow \text{GramSchmidt}(\mathbf{B})$
13: $\mathbf{F} \leftarrow \mathbf{B}^T\mathbf{M}$

Output: $\mathbf{B}$, $\mathbf{F}$

The function TSVD on line 1 performs the SVD of matrix $\mathbf{M}$ and returns the K principal singular vectors and singular values. Matrix $\mathbf{F}$ is then initialized with the product $\tilde{\mathbf{S}}\tilde{\mathbf{V}}^T$ following Eq. (16b). The alternated optimizations of $\mathbf{B}$ (line 7) and $\mathbf{F}$ (line 8) are performed until the variation of the maximum residual norm meets the stop criterion of line 11. For the creation of the Minimax dictionaries used on the reconstructions of Chapter 4, we used $\epsilon_{\text{AMRR}} = 10^{-4}$.

Both operations of line 7 and line 8 either decrease or maintain the value of the cost function of Eq. (17) with $p \to \infty$, which guarantees the convergence of AMRR. Yet, there is no guarantee that the minimum found is the global minimum.

The reconstruction algorithm described in Section 3.2 requires that all bases $\mathbf{B}$ are orthonormal, a condition that is not guaranteed after the alternating steps of AMRR. This condition is met by submitting $\mathbf{B}$ to a Gram-Schmidt process on line 12. Matrix $\mathbf{F}$ is then updated on line 13 to match the new orthonormal $\mathbf{B}$.

Figs. 4a to 4c show an example of how the residual norms $\|\mathbf{r}_i\|$ behave for the SVD and Minimax expansions. The residuals are taken from the approximations for the centermost local ROI on the ultrasound acquisition model described in Section 3.3. In Fig. 4a, the average and maximal values of $\|\mathbf{r}_i\|$ are taken from the collection of $R = 75$ residual norms, for K varying from 3 to 35. Notice that the Minimax expansion yields smaller values for the maximum residual norm for all cases, while also narrowing the difference between the average and the maximal values, at the cost of larger average values.

In Figs. 4b and 4c, all the $R = 75$ residual norms from the expansions with $K = 10$ are plotted with a spatial disposition corresponding the respective points in the centermost local ROI. Fig. 4b shows the result for the SVD and Fig. 4c for the Minimax expansion. Notice that the surface formed by the residual norms from the Minimax expansion is significantly flatter than its SVD counterpart.

As an intermediate step towards the formulation of the Minimax expansion presented in this section, a polynomial Minimax expansion was formulated and evaluated for a prototypical 1-dimensional problem and reported in (PASSARIN et al., 2017). Appendix A describes the formulation of the polynomial Minimax expansion. Section A.1 details how the Minimax expansion presented here generalizes the polynomial Minimax.

3.2 RECONSTRUCTION ALGORITHM

3.2.1 LIMITATIONS OF CONIC CONSTRAINTS

Two main algorithms were proposed to work with expanded dictionaries: the convex CBP (EKANADHAM et al., 2011) and the greedy COMP (KNUDSON et al., 2014). The first one aims at solving problem (7) while the second attempts to solve problem (8). A hybrid approach called Interpolating Band-excluded Orthogonal Matching Pursuit (IBOMP) was also proposed and applied to frequency estimation (FE) and time delay estimation (TDE) (FYHN et al., 2015). Basically, it performs a rough greedy estimation of the support of the solution, followed by a refining convex optimization.

In order to implement a constraint set $\mathcal{C}$, all the aforementioned algorithms have at least one step involving a constrained convex optimization where the constraints define either first-order (SVD, Minimax and Taylor) or second-order (Polar) cones. Fig. 5a illustrates an example of a first-order cone for $K = 2$. The black curved line represents the projection onto the basis $\mathbf{B}^{(n)}$ of a continuous 1-dimensional PSF manifold. The R vectors

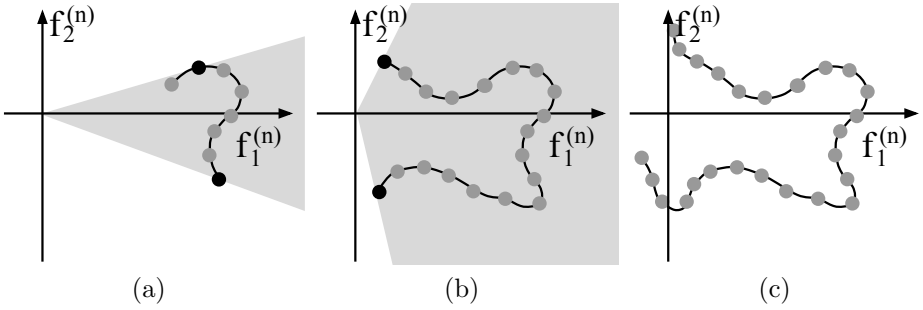


Figure 5: (a) Illustrative case of projection of local manifold samples $\mathbf{M}^{(n)}$ on a basis $\mathbf{B}^{(n)}$, for $K = 2$. The curved line represents the projection of a continuous 1-dimensional manifold, while the dots represent the projection of the samples (columns of $\mathbf{M}^{(n)}$) on $\mathbf{B}^{(n)}$. When $\boldsymbol{\Delta}$ is sufficiently small, the projections have single-signed, relatively large values on the first component $\boldsymbol{f}_1^{(n)}$ and smaller values on the remaining components. In this case, the definition of a first-order cone (represented by the shadowed region) is possible and can be used in the reconstruction algorithm combined with a non-negativity constraint for the first component, ensuring that the recovered coefficients represent weighted copies of the local manifold, rather than other arbitrary combinations. The upper and lower angles of the cone depend on $\max_i(\boldsymbol{f}_{2,i}^{(n)}/\boldsymbol{f}_{1,i}^{(n)})$ and $\min_i(\boldsymbol{f}_{2,i}^{(n)}/\boldsymbol{f}_{1,i}^{(n)})$ respectively. (b) As $\boldsymbol{\Delta}$ increases, the angle of the cone may as well increase, making the constraint less effective, as a broader area is allowed for the recovered coefficients $\mathbf{f}^{(n)}$. (c) An example where the definition of a convex cone is no longer possible. This imposes a limit on the definition of $\boldsymbol{\Delta}$. Source: (PASSARIN et al., 2018).

that compose a local manifold matrix $\mathbf{M}^{(n)}$, when projected onto $\mathbf{B}^{(n)}$, result in vectors $\mathbf{f}^{(n)}$, represented by the dots, which compose the columns of $\mathbf{F}^{(n)}$. When a reconstruction is performed, the recovered coefficients set $\mathbf{x}^{(n)} \in \mathbb{R}^2$ for this n-th local ROI is constrained to lie within a first-order cone, represented by the shadowed area (which extends indefinitely to the right). This cone is defined by two linear constraints that impose upper and lower bounds for the relation $x_2^{(n)}/x_1^{(n)}$, combined with a non-negativity constraint for the first component $x_1^{(n)}$. This constraint set aims to avoid arbitrary combinations for $\mathbf{x}^{(n)}$ that do not represent positively-weighted copies of actual manifold samples. The upper black dot dictates the upper angle of the cone, and is defined from the modulating matrix $\mathbf{F}^{(n)}$ as

$$(f_{1,\hat{i}}^{(n)}, f_{2,\hat{i}}^{(n)}), \text{ where } \hat{i} = \max_i(f_{2,i}^{(n)}/f_{1,i}^{(n)}),$$

i.e., the maximum relation between the first and second components found among the projections of $\mathbf{M}^{(n)}$. Similarly, the lower black dot is defined by

$$(f_{1,\hat{i}}^{(n)}, f_{2,\hat{i}}^{(n)}), \text{ where } \hat{i} = \min_i(f_{2,i}^{(n)}/f_{1,i}^{(n)}),$$

and defines the lower angle of the cone. For higher orders of K, such a cone is defined for all $K-1$ relations between each k-th ($k \geq 2$) component and the first one. The resulting

linear constraint set is defined as (KNUDSON et al., 2014; PASSARIN et al., 2017)

$$\min_{1\leq i\leq R}\left(\frac{f_{k,i}^{(n)}}{f_{1,i}^{(n)}}\right)\leq\frac{x_k^{(n)}}{x_1^{(n)}}\leq\max_{1\leq i\leq R}\left(\frac{f_{k,i}^{(n)}}{f_{1,i}^{(n)}}\right),f_{1,i}^{(n)}\geq 0,\ \forall k\in\{2,\ldots,K\},\ n\in\{1,2,\ldots,N\}, \tag{23}$$

where $f_{k,i}^{(n)}$ denotes the element on the k-th line and i-th column on $\mathbf{F}^{(n)}$. The principle is similar for the Polar expansion, though in that case the cones are of second order (EKANADHAM et al., 2011).

Notice that the cone-based convex constraints assume that the projection of $\mathbf{M}^{(n)}$ on the K components of $\mathbf{B}^{(n)}$ yields relatively large, positive, small-variance values for the first component and small values for the remaining, yielding relatively small values for minimum and maximum relations of (23). If this assumption is broken, the cone will span too large an area of the right half-plane, i.e., it will constrain less, being less effective, as represented in Fig. 5b. This means that a large set of arbitrary combinations of coefficients can be recovered, many of them not representing actual weighted PSFs. In some cases, defining the cone is not even possible, like in the example depicted in Fig. 5c.

Assuring a well behaved relation between the first and the remaining components, as shown in Fig. 5a, implies choosing considerably small values for Δ_x and Δ_z, what limits the applicability of recovery algorithms based on conic constraints. For instance, on the simulated acquisition set of Section 3.3, choosing $\Delta_x = \Delta_z = 0.2$mm still causes the first component to have both positive and negative values on certain local manifolds.

3.2.2 NON-CONVEX CONSTRAINTS

The problem described in Subsection 3.2.1 is the main reason why our algorithm does not rely on conic constraints. Instead, it attempts to constrain each K-tuple of recovered coefficients $\mathbf{x}^{(n)}$ to be similar to any column of the modulating matrix $\mathbf{F}^{(n)}$. The current section explains the motivation for this constraint and how it is formalized.

As expressed in Eq. (2), an acquired signal $\mathbf{c}$ should ideally be composed by a linear combination of samples from the PSF manifold $\mathcal{M}$. This means, recalling Fig. 2f, that $\mathbf{c}$ should be composed by a linear combination of vectors that are very similar to the columns $\mathbf{m}_i^{(n)}$ of the matrices $\mathbf{M}^{(n)}$, i.e.,

$$\mathbf{c}\approx\sum_{n=1}^{N}\mathbf{M}^{(n)}\mathbf{v}^{(n)}+\mathbf{w}, \tag{24}$$

where vectors $\{\mathbf{v}^{(n)}\}$ are the weights of the linear combination. But our low-rank dictio-

nary $\{\mathbf{B}^{(n)}\}$ is only capable of reproducing the low-rank approximation $\tilde{\mathbf{M}}$ instead of $\mathbf{M}$, so we go one step further on the approximation of $\mathbf{c}$

$$\mathbf{c} \approx \sum_{n=1}^{N} \tilde{\mathbf{M}}^{(n)} \mathbf{v}^{(n)} + \mathbf{w}. \tag{25}$$

The factorization of Eq. (11) allows us to rewrite Eq. (25) as

$$\mathbf{c} \approx \sum_{n=1}^{N} \mathbf{B}^{(n)} \mathbf{F}^{(n)} \mathbf{v}^{(n)} + \mathbf{w}. \tag{26}$$

If we assume that each n-th local ROI contains at most one scatterer, then the number non-zero elements of each $\mathbf{v}^{(n)}$ is at most one, and the product $\mathbf{F}^{(n)}\mathbf{v}^{(n)}$ in Eq. (26) results in a scaled copy of one column of $\mathbf{F}^{(n)}$. This means that, when $\mathbf{c}$ is written as approximately a combination of the bases $\{\mathbf{B}^{(n)}\}$

$$\mathbf{c} \approx \sum_{n=1}^{N} \mathbf{B}^{(n)} \mathbf{x}^{(n)} + \mathbf{w}, \tag{27}$$

each K-tuple of coefficients $\{\mathbf{x}^{(n)}\}$ should ideally be a scaled copy of any of the columns of the corresponding matrix $\mathbf{F}^{(n)}$. Since our fine sampling of the manifold (Fig. 2f) is not continuous, we relax this ideal restriction and require the K-tuples $\{\mathbf{x}^{(n)}\}$ to be *similar* to a scaled copy of any column of $\mathbf{F}^{(n)}$. We translate "similarity" as high correlation, as formalized in the non-convex constraint set

$$\mathcal{C} := \left\{ \{\mathbf{x}^{(n)}\} : \left(\max_{1 \le i \le R} \frac{\langle \mathbf{x}^{(n)}, \mathbf{f}_i^{(n)} \rangle}{\|\mathbf{x}^{(n)}\| \|\mathbf{f}_i^{(n)}\|} \right) \ge \mu_{\mathrm{c}}, \ \forall n \in \{1, 2, \ldots, N\} \right\}, \tag{28}$$

where $\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}^T \mathbf{b}$ denotes the inner product of two vectors.

The minimum correlation parameter μ_{c} controls how similar to any of the manifold samples on $\mathbf{M}^{(n)}$ a recovered event must be. If a given $\mathbf{x}^{(n)}$ passes the test (28), proving to be sufficiently similar to some i-th modulating vector $\mathbf{f}_i^{(n)}$, then the approximation

$$\tilde{\mathbf{m}}_i^{(n)} \frac{\|\mathbf{x}^{(n)}\|}{\|\mathbf{f}_i^{(n)}\|} = \mathbf{B}^{(n)} \mathbf{f}_i^{(n)} \frac{\|\mathbf{x}^{(n)}\|}{\|\mathbf{f}_i^{(n)}\|} \approx \mathbf{B}^{(n)} \mathbf{x}^{(n)} \tag{29}$$

is assumed and the product $\mathbf{B}^{(n)}\mathbf{x}^{(n)}$ is considered as a valid weighted copy of a PSF within the n-th local ROI, rather than an arbitrary combination of the n-th basis vectors. This constraint is imposed by our greedy algorithm on the decision of which expanded set $\mathbf{B}^{(n)}$ will be added to the reconstruction problem at each iteration.

3.2.3 OMP FOR EXPANDED DICTIONARIES

The proposed algorithm, summarized in Algorithm 2 (page 37), is an extension of the OMP algorithm, referred to as OMP for Expanded Dictionaries (OMPED).

It attempts to solve a problem similar to (8) (i.e., to explain an acquired data vector $\mathbf{c}$ with the expanded dictionary $\{\mathbf{B}^{(n)}\}$) with the non-convex constraint set $\mathcal{C}$ defined in (28). The stop criterion is based on the residual yielded by the LS solution with a given cardinality, yet instead of comparing the residual to a fixed parameter ϵ, we compare it to an estimate of the current residual that takes into account the expected acquisition noise and the estimated residuals resulting from the reduced-rank approximation.

The input parameter e_{noise} contains the expected ℓ_2 norm of the acquisition noise. In practice, this value can be obtained from acquisitions with samples of the inspected material known to have neither discontinuities nor other sort of scatterers. For our simulations, we use the relation

$$e_{\text{noise}}^2 = \|\mathbf{w}\|^2 \approx M\sigma^2, \tag{30}$$

which holds if the noise vector $\mathbf{w}$ contains white Gaussian noise with variance σ^2. The approximation of (30) becomes an equality as $M \to \infty$. We assume the equality and use $e_{\text{noise}} = \sqrt{M\sigma^2}$.

We define the support S of the solution, which is initialized as the empty set, and its complement $S^c = \{1, \ldots, N\} \setminus S$. The solution residual $\mathbf{e} \in \mathbb{R}^M$ is initialized with the vector of acquired data $\mathbf{c}$ on line 2.

At each iteration, an index $j \in S^c$ is added to S as we choose the expanded set $\mathbf{B}^{(j)}$ which is capable of causing the maximal decrease on the energy of the residual, as represented on the left side of (31). Since the columns of each $\mathbf{B}^{(n)}$ are orthonormal, the identity

$$\hat{j} = \arg\min_j \|\mathbf{e} - \mathbf{B}^{(j)}{\mathbf{B}^{(j)}}^T\mathbf{e}\| = \arg\max_j \|{\mathbf{B}^{(j)}}^T\mathbf{e}\| \tag{31}$$

holds as a consequence of Parseval's relation (BARRETT; MYERS, 2004), which allows us to perform the simpler operation of taking the norm of each product ${\mathbf{B}^{(j)}}^T\mathbf{e}$.

This operation is a generalization of the measurement of maximum correlation on the original OMP (TROPP; GILBERT, 2007). A constraint based on (28) is imposed to prune candidates that do not accomplish the minimum correlation required. The resulting

criterion is formalized as

$$\hat{j} = \arg\max_{j \in S^c} \left\| \mathbf{B}^{(j)^T} \mathbf{e} \right\| \text{ s.t. } \max_{1 \le i \le R} \frac{\langle \mathbf{B}^{(j)^T} \mathbf{e}, \mathbf{f}_i^{(j)} \rangle}{\|\mathbf{B}^{(j)^T} \mathbf{e}\| \|\mathbf{f}_i^{(j)}\|} \geq \mu_c. \tag{32}$$

To clarify the equivalence between the constraint sets of (32) and (28), we recall that, since every $\mathbf{B}^{(n)}$ is orthonormal, the product $\mathbf{B}^{(j)^T}\mathbf{e}$ yields the coefficients that reconstruct the portion of the residual $\mathbf{e}$ that lies on the column space of $\mathbf{B}^{(j)}$. Were all the bases orthogonal to each other, formally

$$\mathbf{B}^{(i)^T}\mathbf{B}^{(j)} = \mathbf{0}, \forall i, j \in \{1, \dots N\}, i \neq j, \tag{33}$$

then the equality $\mathbf{B}^{(j)^T}\mathbf{e} = \mathbf{x}^{(j)}$ would apply exactly.

The constraint in (32) allows to recover only positive-amplitude events. It can be adapted to consider both positive and negative amplitudes by simply replacing the inner product by its absolute value $|\langle \mathbf{B}^{(j)^T}\mathbf{e}, \mathbf{f}_i^{(j)} \rangle|$.

The algorithm must consider the case in which no index meets the correlation criterion of (32). This case is treated from line 5 to line 8: while problem (11) remains infeasible, a decrease of Δ_μ is made on the parameter μ_c and a new attempt to compute the index j is performed.

The support S is then updated to include the new index j (line 9) and used to compute the coefficients

$$\{\hat{\mathbf{x}}^{(n)}\} = \arg\min_{\{\mathbf{x}^{(n)}\}} \|\mathbf{c} - \sum_{n=1}^{N} \mathbf{B}^{(n)}\mathbf{x}^{(n)}\|^2 \text{ s.t. } \mathbf{x}^{(n)} = \mathbf{0}, \ \forall n \in S^c \tag{34}$$

(where $\mathbf{0} \in \mathbb{R}^K$ is the zero vector), which then yield a residual

$$\mathbf{e} = \mathbf{c} - \sum_{n \in S} \mathbf{B}^{(n)}\mathbf{x}^{(n)}. \tag{35}$$

An estimate of the residual resulting from the low rank approximation is computed on vector $\mathbf{e}_{\text{rank}} \in \mathbb{R}^M$ as

$$\mathbf{e}_{\text{rank}} = \sum_{n \in S} \mathbf{r}_{\hat{i}}^{(n)} \frac{\|\mathbf{x}^{(n)}\|}{\|\mathbf{f}_{\hat{i}}^{(n)}\|}, \tag{36a}$$

$$\text{where } \hat{i} = \arg\max_{1 \le i \le R} \frac{\langle \mathbf{x}^{(n)}, \mathbf{f}_i^{(n)} \rangle}{\|\mathbf{x}^{(n)}\| \|\mathbf{f}_i^{(n)}\|}, \tag{36b}$$

and $\mathbf{r}_i^{(n)}$ denotes the i-th column from $\mathbf{R}^{(n)}$. Based on (28), the index i in (36b) is a function of n: for every index n in the current support S, the correlations performed in (36b) estimate which i-th PSF within the n-th local manifold best explains the recovered coefficients $\mathbf{x}^{(n)}$ (see Figs. 2c and 2f).

The residual $\mathbf{r}_i^{(n)}$, from the dictionary low-rank approximation, is then used as template for the estimation of the current approximation residual. The amplitude estimate is taken from the ratio between the norms of the recovered coefficients $\mathbf{x}^{(n)}$ and of the similar modulating vector $\mathbf{f}_i^{(n)}$.

The current total residual norm is estimated as

$$e_{\text{est}} = (\|\mathbf{e}_{\text{rank}}\|^2 + e_{\text{noise}}^2)^{\frac{1}{2}}, \tag{37}$$

where the addition is performed under the assumption that the acquisition noise and the vector $\mathbf{e}_{\text{rank}}$ have negligible correlation.

Algorithm 2 OMP for Expanded Dictionaries (OMPED)

Input: $\{\mathbf{B}^{(n)}\}$, $\{\mathbf{F}^{(n)}\}$, $\{\mathbf{R}^{(n)}\}$, $\mathbf{c}$, e_{noise}, μ_{c}, Δ_μ

1: $S \leftarrow \emptyset$
2: $\mathbf{e} \leftarrow \mathbf{c}$
3: **repeat**
4: $\quad j \leftarrow$ Compute from (32)
5: $\quad$ **while** $j = \emptyset$ **do**
6: $\quad\quad \mu_{\text{c}} \leftarrow \mu_{\text{c}} - \Delta_\mu$
7: $\quad\quad j \leftarrow$ Compute from (32)
8: $\quad$ **end while**
9: $\quad S \leftarrow S \bigcup \{j\}$
10: $\quad \{\mathbf{x}^{(n)}\} \leftarrow$ Compute from (34)
11: $\quad \mathbf{e} \leftarrow$ Compute from (35)
12: $\quad \mathbf{e}_{\text{rank}} \leftarrow$ Compute from (36)
13: $\quad e_{\text{est}} \leftarrow$ Compute from (37)
14: **until** $e_{\text{est}} \geq \|\mathbf{e}\|$ or $S^c = \emptyset$

Output: S, $\{\mathbf{x}^{(n)}\}_{n \in S}$

The algorithm greedily increases the support until the estimated residual norm e_{est} reaches the norm $\|\mathbf{e}\|$ of the actual residual yielded by the LS or all indices $n = 1, \ldots, N$ have been added to the support S. Notice that the expected noise energy e_{noise}^2 (known a priori) is taken into account in both sides of the condition $e_{\text{est}} \geq \|\mathbf{e}\|$ on line 14: while its contribution to e_{est} is explicit in Eq. (37), the contribution to $\mathbf{e}$ is implicit and

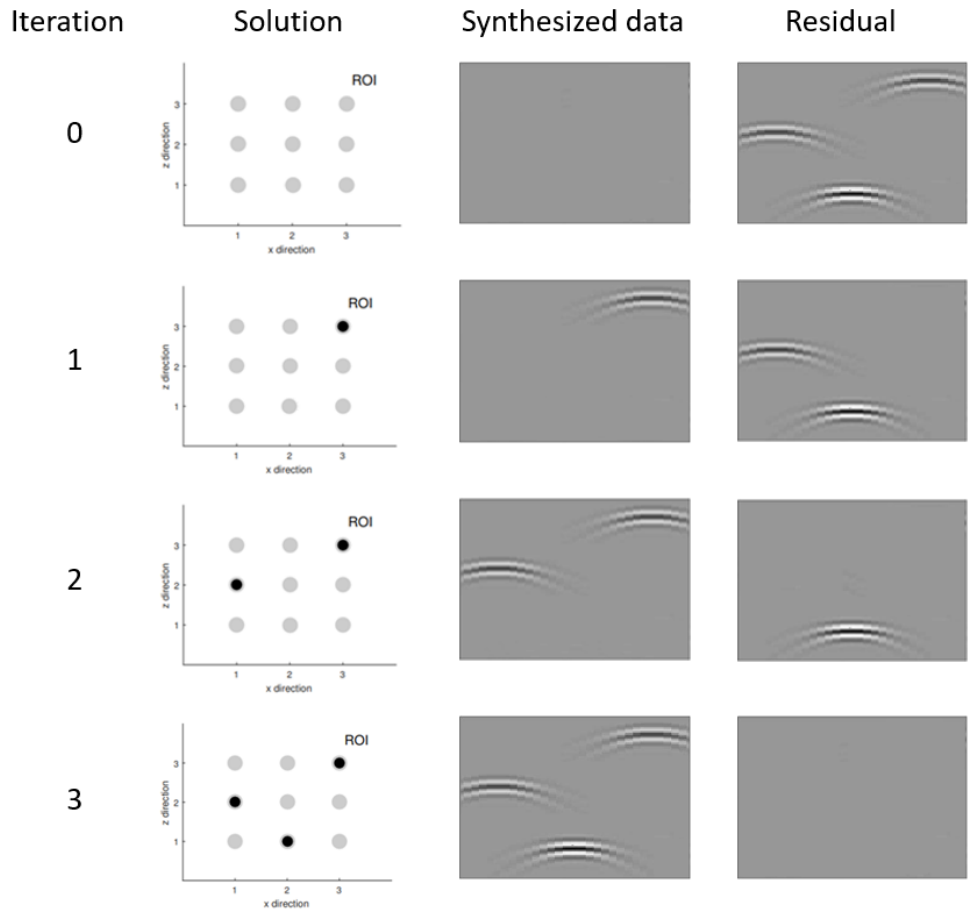


Figure 6: Illustrative example of OMP and OMPED algortihms.
Source: own authorship.

is based on the assumption that $\mathbf{c}$ in Eq. (35) contains additive noise that cannot be represented by the bases $\{\mathbf{B}^{(n)}\}$.

Figure 6 illustrates the iterative nature of OMPED, which follows the same logic of OMP. The initial value of the residual is the acquired vector $\mathbf{c}$ itself. Then, the index $\mathbf{j}$ that satisfies Eq. (32) is determined and included to the support. A partial solution with cardinality 1 is obtained from Eq. (34) and its respective residual is determined through Eq. 35. On the second iteration, the residual is used to define the new index $\hat{j}$ to be added to the support. A partial solution with cardinality 2 and its respective residual are determined. The process repeats until the stop criterion is met and the last partial solution is delivered as final solution.

3.2.4 RECOVERY OF LOCATIONS AND AMPLITUDES

OMPED yields a support S as well as the sets of expanded coefficients $\{\mathbf{x}^{(n)}\}_{n\in S}$. The computation of the locations and amplitudes follows the same principle used on (36a)

and (36b): each event is located inside an n-th local ROI; its high resolution location is assigned the same as that of the i-th response $\mathbf{m}_i^{(n)}$ within the R responses of the fine grid (Fig. 2c) which most correlates to $\mathbf{x}^{(n)}$. Recalling the approximation $\mathbf{m}_i^{(n)} \approx \mathbf{B}^{(n)}\mathbf{f}_i^{(n)}$, we determine i by finding out which $\mathbf{f}_i^{(n)}$ most correlates to $\mathbf{x}^{(n)}$:

$$\hat{i}(n) = \arg\max_{1\leq i\leq R} \frac{\langle \mathbf{x}^{(n)}, \mathbf{f}_i^{(n)} \rangle}{\|\mathbf{x}^{(n)}\| \|\mathbf{f}_i^{(n)}\|}, \ \forall n \in S. \tag{38}$$

The amplitude estimations v_n result from the ratios between the norms of $\mathbf{x}^{(n)}$ and of the chosen template $\mathbf{f}_i^{(n)}$:

$$v_n = \frac{\|\mathbf{x}^{(n)}\|}{\|\mathbf{f}_{\hat{i}}^{(n)}\|}, \ \forall n \in S, \ \hat{i} \text{ as in (38)}. \tag{39}$$

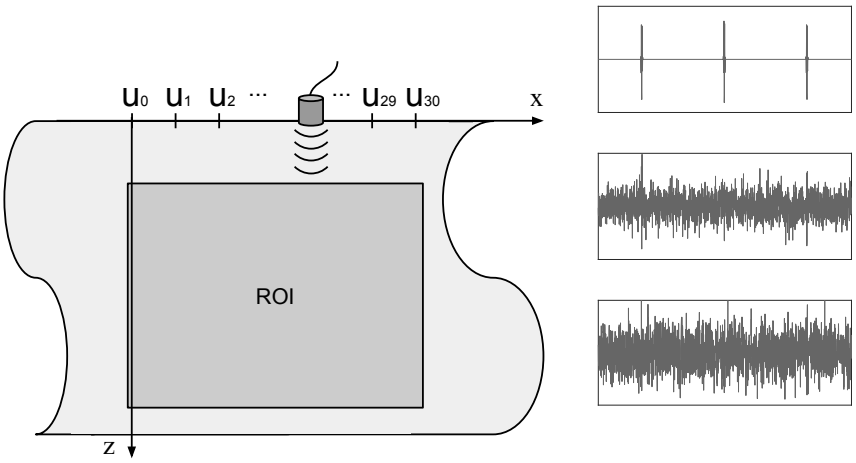
As consequence, the spatial resolution of the reconstructed events is the same as that of the fine sampling represented in Fig.2c, i.e., δ_x and δ_z for the lateral and axial axes respectively.

3.3 SIMULATED ACQUISITION SET

To simulate the ultrasound NDT acquisition set from (GUARNERI et al., 2015), represented in Fig. 7a, we used the Field II package for Matlab (JENSEN, 2004). A piston transducer with 3mm radius (125μm mathematical element size) interrogates a steel sample object (sound speed 5680m/s). The excitation pulse has center frequency f_c = 5MHz and 6dB fractional bandwidth of 100%. The simulated transducer slides horizontally along the surface of the object, acquiring scanlines from 31 lateral positions u_i, from u_0 = 0mm to u_{30} = 30mm (center of transducer), with a distance of 1mm between consecutive lateral positions. The 31 scanlines are sampled with sampling rate f_s = 25MHz and concatenated to form the acquisition vector $\mathbf{c}$.

Following (GUARNERI et al., 2015), the model grid has $31 \times 41 = 1271$ modelled locations distributed with regular spacing of 1mm on both x and z directions. On x direction, the locations are the same as the transducer positions, i.e. $x = 0\text{mm}, 1\text{mm}, \ldots, 30\text{mm}$. On z direction, 41 locations are modelled regularly between 18mm and 58mm, i.e., $z = 18\text{mm}, 19\text{mm}, \ldots, 58\text{mm}$.

As explained in Subsection 3.1.1, in the expanded acquisition model, the grid locations give place to local ROIs. Our expanded model has 1271 local ROIs with $\Delta_x = \Delta_z$ = 1mm, with centers corresponding to the modelled locations of the regular model.



(a) (b)

Figure 7: (a) Simulated set. The transducer, fixed vertically at $z = 0$, slides horizontally over the surface of the interrogated object, acquiring scanlines at 31 positions $x = \{u_0, \ldots, u_{30}\}$, corresponding to 0mm up to 31mm with 1mm step. The scanlines are concatenated to form the acquired vector c. A PSF y(x, z) is determined by placing a unity amplitude scatterer on position (x, z) and acquiring the corresponding c. (b) Extracts from the acquired data for the three center-most transducer positions, with a unity amplitude scatterer located at the center of the ROI. Witouht noise (top) and with additive White Gaussian noise with $\sigma = 0.08$ (middle) and $\sigma = 0.12$ (bottom).
Source: (PASSARIN et al., 2018) adapted from (GUARNERI et al., 2015).

Consequently, our ROI extends from $x = -0.5$mm to $x = 30.5$mm and from $z = 17.5$mm to $z = 58.5$mm. The highly coherent local manifolds were created with $R_x = 5$ and $R_z = 15$, thus $R = 75$. Therefore, $\delta_x = 250\mu$m and $\delta_z = 71.4\mu$m.

We simulated the acquisition for 200 cases of 5 unity amplitude scatterers randomly distributed over the ROI. The scatterers positions were not forced over any kind of grid. White Gaussian noise with three different levels (no noise, $\sigma = 0.08$ and $\sigma = 0.12$) was added to each simulated acquisition. Since the energy of the acquired signal (without noise) varies according to factors such as distance to transducer and constructive/destructive interference, we consider that the parametrization of noise in terms of its standard deviation σ is more appropriate than signal-to-noise ratio (SNR). To provide a visual notion of the noise levels, Fig. 7b shows an extract of acquired data for the three noise levels from an acquisition where a single scatterer was placed on the center of the ROI. Scanlines from the three center-most positions of the transducer are concatenated.

4 RESULTS

The current chapter presents the results obtained from image reconstructions with both OMP and OMPED algorithms from data acquired from simulations with the acquisition set presented in Section 3.3.

The recovery accuracy of OMPED with respect to count of misses and spatial accuracy, with both SVD and Minimax dictionaries, is evaluated in Section 4.1. In Section 4.2, the performance of both dictionaries with respect to cardinality estimation is compared and briefly discussed. A comparison between OMP using finer model grids and OMPED is performed in Section 4.3. For a visual notion of the advantage brought by the proposed methods, Section 4.4 displays the results of the reconstruction of an image with OMPED as well as other model-based algorithms.

4.1 RECOVERY ACCURACY

To compute the accuracy on the recovery of scatterers, OMPED was ran with a fixed number of 5 iterations, with $\mu_c = 0.8$ and $\Delta_\mu = 0.1$, using SVD and Minimax dictionaries with K varying from 2 to 10 for the 200 simulated acquisitions with the three levels of noise. Each recovered scatterer whose distance to the closest original simulated scatterer was equal to or smaller than 0.5mm in both axial and lateral directions was computed as a hit – otherwise it was computed as a miss. Fig. 8 shows the percentage of misses from 1000 recovered scatterers for all 9 values of K and 3 noise levels. Even for the highest level of noise, misses are kept below 10% for $6 \leq K \leq 10$.

A small increase in the count of misses is observed for values of $K \geq 8$. This is possibly explained by the fact that, for $K \geq 8$, increasing K adds few useful information to the dictionary at the cost of increasing coherence. For the SVD basis, the decay of the singular values s_k can be used as a measure of useful information. Fig. 9 shows how s_k behave for the centermost local manifold $\mathbf{M}^{(636)}$. Notice that values of s_k for $k \geq 8$ are significantly smaller than the previous ones. In general, in terms of misses count, the best

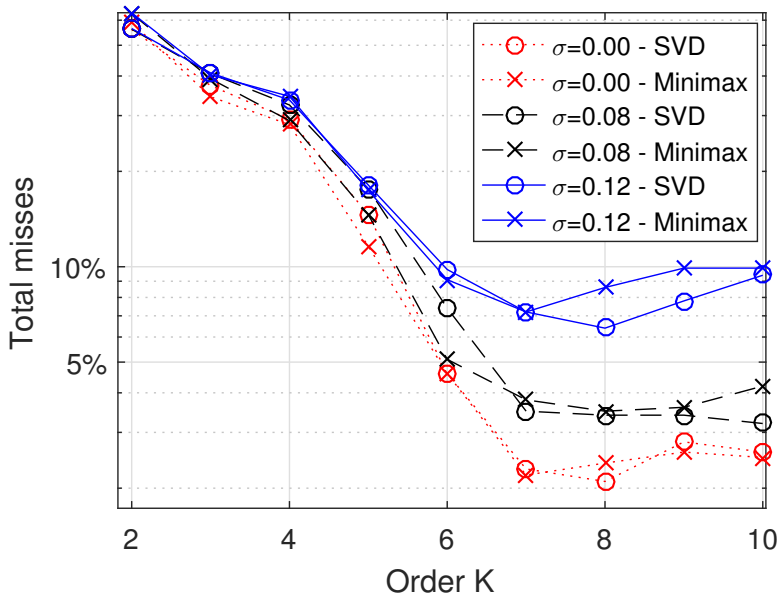


Figure 8: Percentage of misses (from 1000 simulated events) as a function of K, for three levels of noise, with OMPED running with a fixed number of 5 iterations (each of the 200 simulated acquisition had 5 scatterers). Each recovered scatterer whose distance to the closest original simulated scatterer was greater than 0.5mm in any direction (axial or lateral) was computed as a miss. Regarding this criterion, the performance of the expanded dictionaries peaks near $K = 8$. For $K > 8$, few useful information is added to the dictionary (see Fig. 9) at the expense of increased coherence.
Source: extended from (PASSARIN et al., 2018)

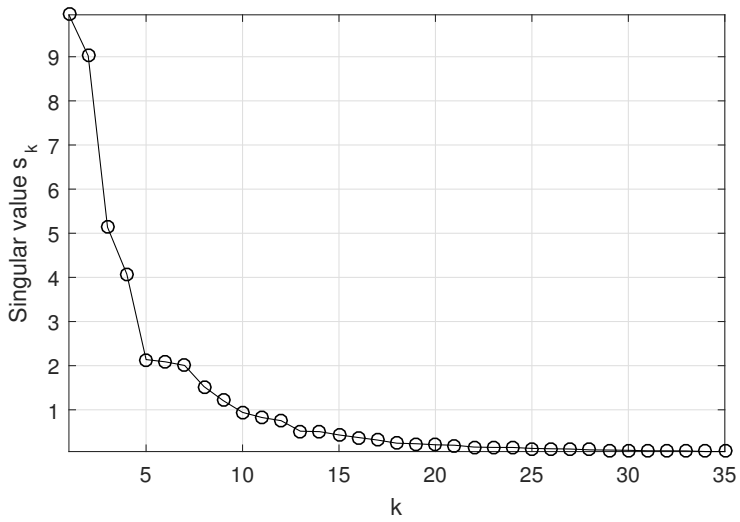


Figure 9: 35 first singular values s_k from the SVD of $\mathbf{M}^{(n)}$ for the center-most local ROI of the ultrasound acquisition set escribed in Section 3.3. In this case, increasing the order of expansion to $K \geq 8$ adds relatively few useful information to the dictionary, at the cost of increased coherence.
Source: (PASSARIN et al., 2018).

performing dictionary was Minimax for $K \leq 6$ and SVD for $K > 6$.

For every hit, the distance between the original and the recovered scatterers was computed. The average distances are shown in Fig. 10a. Also for this criterion the Minimax dictionaries showed better performance for low values of K while the SVD dictionaries performed better for larger values of K.

The computation of hits and misses does not take into account the amplitude of recovered scatterers, i.e., recovered scatterers are implicitly considered as having unity amplitude. To endorse this assumption, the average amplitudes of recovered events are shown in fig. 10b. Notice that, for all cases, the average amplitudes are between 0.98 and 1.01, i.e., the average amplitude error is less than 2%. The average absolute amplitude resulting from the reconstructions with OMP using the regular dictionary $\mathbf{H}$ was 0.70, 0.70 and 0.71 for noise levels $\sigma = 0$ (no noise), $\sigma = 0.08$ and $\sigma = 0.12$ respectively.

4.2 ESTIMATION OF RESIDUAL AND STOP CRITERION

The residual norm $\|\mathbf{e}_{\text{rank}}\|$, caused by the dictionary rank reduction, is estimated at each iteration of OMPED and then combined with e_{noise} to form the estimated regression residual e_{est} used on the stop decision. That computation involves the estimation of the high resolution locations of each recovered scatterer and their corresponding ap-

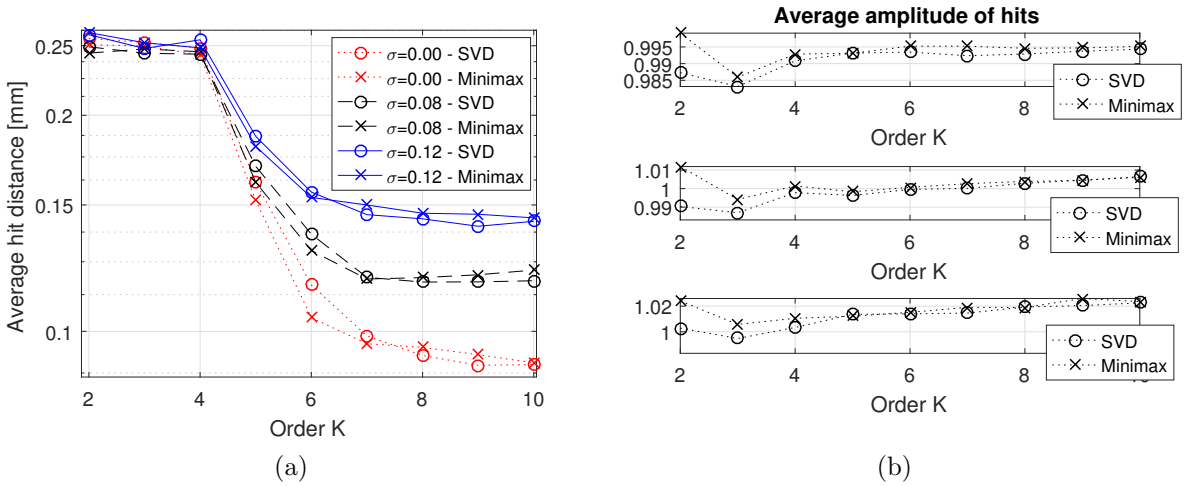


Figure 10: Accuracy of SVD and Minimax dictionaries. (a) Distance between recovered events (hits) and their corresponding simulated true event. (b) Average amplitude of the events computed as hits, without noise (top) and with noise levels $\sigma = 0.08$ (middle) and $\sigma = 0.12$ (bottom). All simulated events have unity amplitude.
Source: extended from (PASSARIN et al., 2018)

proximation residuals $\mathbf{r}_i^{(n)}$ (Eq. (36b)). Inaccuracies on these estimations may degrade the computation of $\|\mathbf{e}_{\text{rank}}\|$ and consequently e_{est}. As can be seen in Figs. 4b and 4c, this degradation is less significant with the Minimax than with the SVD dictionaries. A practical result expected is an increased accuracy on the computation of e_{est} with Minimax dictionaries compared to the SVD, what provides an increased accuracy on the guess of the cardinality of the final solution. To verify this result, we simulated 200 noiseless acquisitions, each with a single scatterer randomly located over the ROI, and ran a single iteration of OMPED for each acquisition. For this controlled case, e_{est} should ideally equal the reconstruction residual norm $\|\mathbf{e}\|$ returned by OMPED. Fig. 11 shows the average of the absolute difference $|e_{\text{est}} - \|\mathbf{e}\||$. For all values of K from 2 to 10, the Minimax bases yielded a more accurate residual prediction, as expected.

OMPED was executed one more time on the 5-scatterer dataset of Section 3.3, this time with the residual-based stop criterion defined on line 14 of Algorithm 2, with a maximum of 10 iterations. Because all images contained 5 scatterers, the algorithm was expected to stop on iteration 5, i.e, yield a cardinality of 5. Since the approximation residual estimated with the Minimax bases is more accurate than with the SVD bases, reconstructions with Minimax dictionaries should more frequently stop at the expected (5th) iteration. The histograms of Fig. 12a show this outcome: the peak of stops at iteration 5 is slightly greater with Minimax. In the neighboring final iterations 4 and 6, Minimax is also more frequent than SVD. The maximum iteration allowed was 10, at

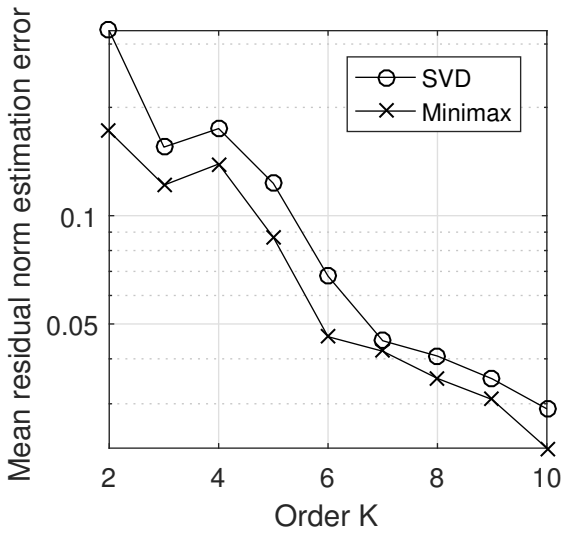


Figure 11: Average of residual norm prediction errors $|e_{\mathbf{est}} - \|\mathbf{e}\||$ from reconstructions of 200 simulated noiseless acquisitions with one single scatterer each. The Minimax dictionaries yield a more accurate prediction of the residual.
Source: own authorship.

which the algorithm stopped when e_{est} failed to reach $\|\mathbf{e}\|$, a phenomenon that was less frequent with the Minimax bases. The results for values of K from 2 to 10 and for the 3 noise levels are summed on the histograms of Fig. 12a. A total of 5400 reconstruction (3 noise levels $\times$ 200 images $\times$ 9 orders K) per type of dictionary are computed for each type of expansion. The individual histograms for the reconstructions without noise and with $\sigma = 0.08$ and $\sigma = 0.12$ are shown in Figs. 13a, 13b and 13c respectively.

Fig. 12b shows an example of the evolution of the regression residual norm $\|\mathbf{e}\|$ and the estimated residual norm e_{est}. As new events are iteratively added to the solution, the latter decreases while the former increases. On iteration 5, $\|\mathbf{e}\|$ drops below e_{est} and OMPED correctly meets the stop criterion, yielding a final solution with cardinality 5. White Gaussian noise with $\sigma = 0.08$ was added to the data. OMPED ran with Minimax ($K = 8$) dictionary.

4.3 COMPARISON TO OMP WITH FINER GRID

The reconstruction problem solved by OMPED with an expanded dictionary of order K is NK-large, while the problem posed with a corresponding regular model is N-large. A question that naturally arises is: would the use of a K-finer regular grid (creating an NK-large reconstruction problem) with the original OMP be as effective as using an order-K expanded dictionary with OMPED?

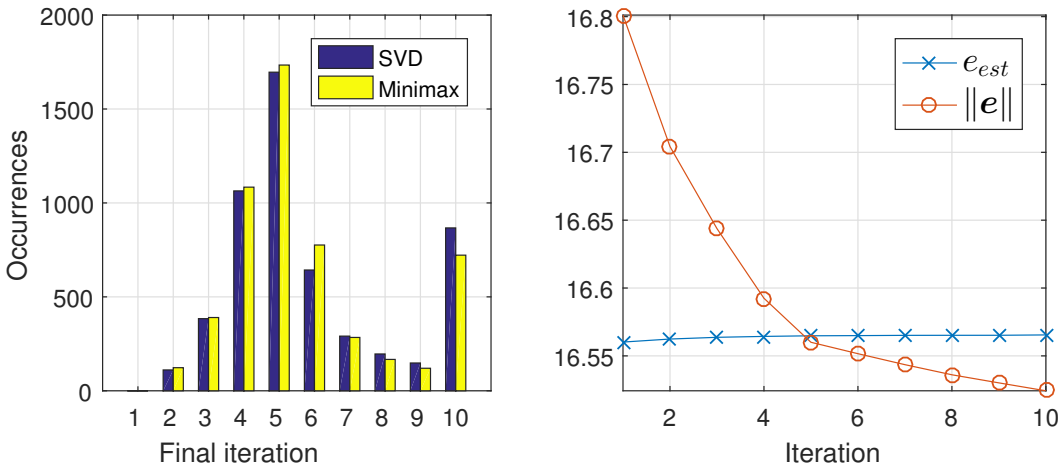
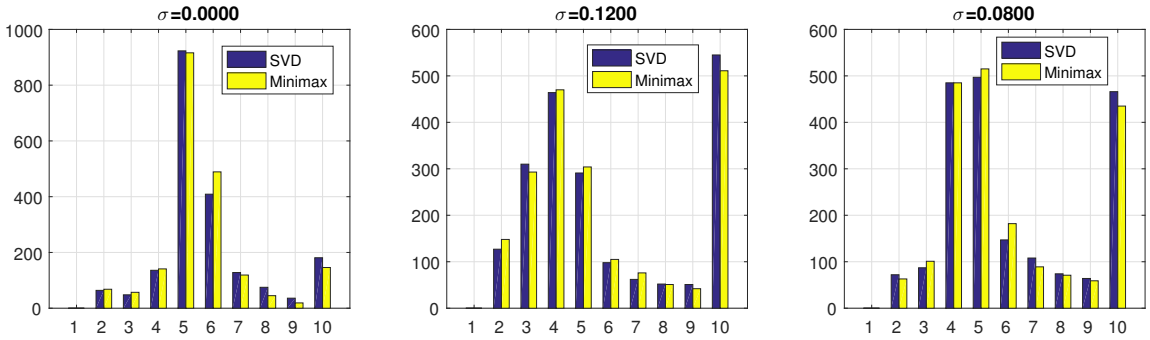


Figure 12: (a) Histograms of final iteration (when e_{est} reaches $\|e\|$) for OMPED running with SVD and Minimax dictionaries, for K varying from 2 to 10. The Minimax dictionaries allow OMPED to stop more frequently at the correct iteration. Results from all values of K are summed. The total number of reconstructions is 5400 for each type of expanded dictionary. (b) Example of evolution of e_{est} and $\|e\|$ along the iterations of OMPED. In this case, e_{est} dropped below $\|e\|$ at the 5th iteration, which was correctly identified as the final iteration. The simulated object contained 5 scatterers. White Gaussian noise with $\sigma = 0.08$ was added to the acquired data. OMPED ran with Minimax dictionary with $K = 8$.
Source: extended from (PASSARIN et al., 2018).

To answer that question, we ran OMP with three different models. The first one follows the original grid described in Section 3.3 ($\Delta_x = \Delta_z = 1$mm and $N = 31 \times 41 = 1271$ modelled locations). The second covers the same ROI but makes the grid twice as dense in each dimension, i.e., $\Delta_x = \Delta_z = 0.5$mm. For the third model, the ROI is also the same and the grid is thrice as dense in each dimension, that is, $\Delta_x = \Delta_z = 1/3$mm. These three models are referred to as $K = 1$, $K = 4$ and $K = 9$ respectively. The same simulated dataset of 200 images $\times$ 5 scatterers described in Section 4.1 was used.

For comparison to the regular model ($K = 1$), the Minimax and SVD dictionaries with $K = 1$ were created and fed to OMPED to run on the same dataset. Note that, although these dictionaries have the same dimensions as a regular model, they are not composed of PSFs, but of single-vector approximations of local manifolds according to the two criteria. The results for $K = 4$ and $K = 9$ are taken from Section 4.1 (Fig. 8).

The performance results in terms of percentage of misses are shown in Fig. 14a (without noise), Fig. 14b ($\sigma = 0.08$ noise) and Fig. 14c ($\sigma = 0.12$ noise). Notice that the increment of the order K decreases the number of misses on both algorithms, but the decrease is more significant on OMPED. In all cases (even $K = 1$), OMPED performed



(a) (b) (c)

Figure 13: Histograms of final iteration (when e_{est} reaches $\|\mathbf{e}\|$) for OMPED running with SVD and Minimax dictionaries, for K varying from 2 to 10, separated by noise level. (a) No noise. (b) $\sigma = 0.08$. (c) $\sigma = 0.12$.

better than OMP.

4.4 RECONSTRUCTED IMAGES: EXAMPLES

Fig. 15a shows the ground truth for a simulation from the dataset of Section 4.1. The scatterers are labelled 1 to 5 and their locations are displayed in Table 1. The deviations from the nearest modelled locations in both axes are also displayed in Table 1. Gaussian noise was added to the acquired data with $\sigma = 0.08$. The reconstructed image using OMPED with a Minimax dictionary ($K = 5$) is shown in Fig. 15b. No limit was imposed on the number of iterations, i.e., the algorithm correctly stopped at the 5^{th} iteration based on the values of the estimated and actual residuals. The activated pixels are the same in the ground truth of Fig. 15a and the OMPED result of Fig. 15b. While all simulated scatterers had unity-amplitude, the recovered amplitudes ranged from 0.9398 to 1.0387. Both Figs. 15a and 15b have 41×31 pixels corresponding to the local ROIs of the expanded model.

The result of the reconstruction using OMP with the regular dictionary model $\mathbf{H}$ is shown in Fig. 15c. We ran 10 iterations of the algorithm in order to show that some scatterers are not found even when the number of iterations doubles the previously known number of scatterers.

Fig. 15d shows the image yielded by the LS (unregularized) solution of Eq. (4). As is common in unregularized model-based solutions, the image is dominated by noise (BOVIK, 2000). We also applied ℓ_1 regularization to the LS problem, which corresponds

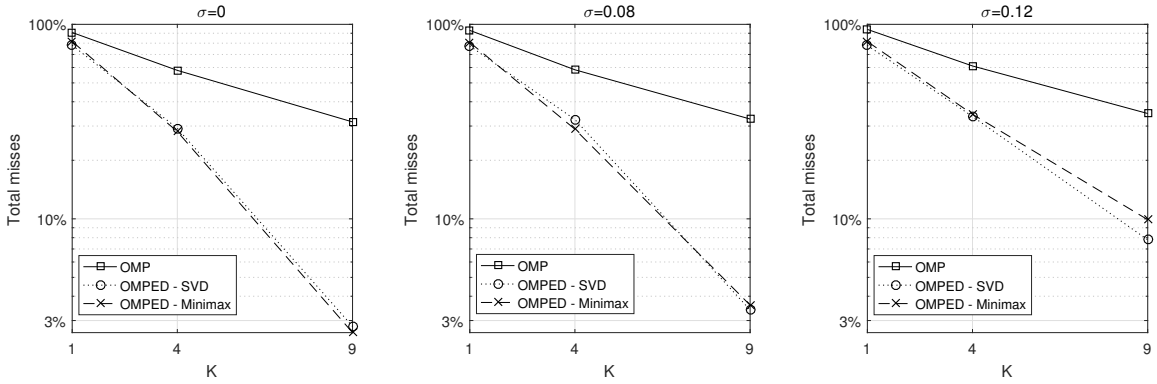


Figure 14: Comparison between OMP with finer grid OMPED: percentage of misses obtained with for $K = 1$ (a), $K = 4$ (b) and $K = 9$ (c) with algorithm OMP and OMPED, the latter with SVD and Minimax dictionaries. In the case of OMP, the order K implies a finer discretization of the model grid. In all cases, OMPED performed better than OMP. The decrease in the number of misses as K increases is greater in OMPED than in OMP.
Source: own authorship.

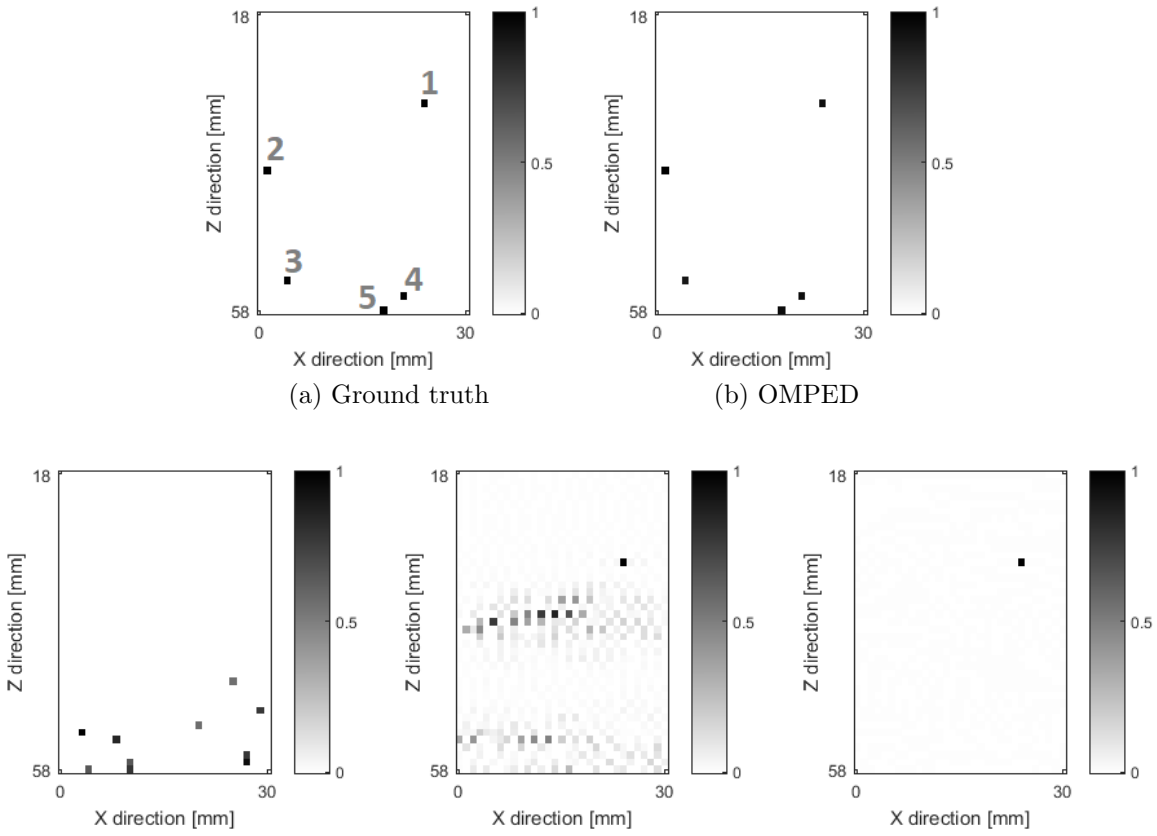
to the Basis Pursuit (BP) formulation:

$$\hat{\mathbf{v}} = \arg\min_{\mathbf{v}} \|\mathbf{c} - \mathbf{H}\mathbf{v}\|^2 + \lambda\|\mathbf{v}\|_1. \tag{40}$$

Scatterer	x (m)	x deviation (mm)	z (mm)	z deviation (mm)
1	25.0020	0.0020	31.0036	0.0036
2	2.4417	0.4417	39.8512	-0.1488
3	4.5961	-0.4039	54.6730	-0.3270
4	22.1781	0.1781	56.8436	-0.1564
5	19.1593	0.1593	59.3316	0.3316

Table 1: Locations and off-grid deviations of the points in the example image

The ℓ_1-regularized formulation was solved with the L1_LS package for Matlab (KIM et al., 2007). The resulting image is shown in Fig. 15e. While a small value for λ yields an image dominated by noise, such as that of Fig. 15d, larger values cause the image to be too sparse, suppressing some features. This is a consequence of the penalization of recovered amplitudes on (40). The chosen regularization parameter $\lambda = 2.0691$ minimizes the norm $\|\mathbf{v} - \hat{\mathbf{v}}\|$, where $\mathbf{v}$ is the ground truth and $\hat{\mathbf{v}}$ is the BP result. Notice that the only recovered scatterer corresponds to scatterer 1, which has the least deviation from the grid, as displayed in Table 1.



(a) Ground truth (b) OMPED

(c) OMP (d) Least Squares (e) ℓ_1-regularized

Figure 15: Example of image simulated and reconstructed, from the dataset described in Section 3.3. All images are normalized by the maximum absolute pixel value. (a) Ground truth, with 5 unity-amplitude scatterers randomly distributed over the ROI. (b) Result from OMPED with Minimax ($K = 5$) dictionary. The algorithm correctly identified the 5^{th} iteration as the final one. (c) Result from OMP with regular model H. 10 iterations were run to show that some scatterers are not found even when the number of iterations doubles the number of scatterers present on the ground true image. (d) Solution of the unregularized LS problem of Eq. (4). The image is dominated by noise. (e) Solution of the ℓ_1-regularized problem (40). The penalization of the recovered amplitudes causes the suppression of most points on the resulting image. The chosen regularization parameter $\lambda = 2.0691$ minimizes the norm $\|\mathbf{v} - \hat{\mathbf{v}}\|$, where $\mathbf{v}$ is the ground truth.
Source: extended from (PASSARIN et al., 2018).

5 CONCLUSION

To cope with the problem of off-grid deviation in image reconstruction from pulse-echo data, we developed a technique of dictionary expansion based on a highly coherent sampling of the PSF manifold followed by a rank reduction procedure. The resulting dictionaries feed a greedy algorithm developed to work with non-convex constraints that avoid the recovery of arbitrary coefficients combination that do not represent actual PSFs. For the rank reduction step, two criteria were presented and developed. The first one is a purely LS-based criterion (SVD), based on (KNUDSON et al., 2014), whose implementation is relatively simple and which yielded the best global results on our experiments with simulated ultrasound NDT data both in terms of hits/misses and average hit distance. The second one is a nonlinear (Minimax) cirterion, based on (PASSARIN et al., 2017), with a more complex implementation, which provided better results in terms or accuracy of residual estimation and, consequently, stop decision and cardinality estimation on the OMPED algorithm. While in (PASSARIN et al., 2017) the Minimax formulation is based on a polynomial regression, we generalized the optimization procedure so that the regression functions are implicitly optimized as a modulating matrix $\mathbf{F}$. For not constraining the matrix $\mathbf{F}$ to any family of functions, our formulation of Minimax rank reduction yields a solution at least as optimal as that of the Polynomial Minimax expansion formulation, as explained in Appendix A.

Since no assumption is made regarding the geometry of the continuous PSF manifold, our expansion formulations are applicable to both shift-invariant and shift-variant problems. On the other hand, for instance, the Polar expansion (EKANADHAM et al., 2011) is conceived based on the fact that the PSF manifold of any shift-invariant system lies over a hypersphere. In two-dimensional ultrasound (our main motivating application), the fact that the Spatial Impulse Response (SIR) is spatially variant (TUPHOLME, 1969; JENSEN, 2004) puts the direct acquisition model in the class of shift-variant systems.

The criterion for definition of the order K of expansion may vary according to each application. In cases where it is possible to carry out simulations (as presented here) or

a relevant amount of data with accessible ground truth is available, K can be determined empirically. Moreover, in our case, a minimum in the number of misses is identifiable and lies near to a transition on the baseline of singular values shown in Fig. 9. A suggestion for future studies is the development of a generalized criterion for the definition of K. The behavior of the singular values yielded by an SVD decomposition of matrices $\mathbf{M}^{(n)}$ is potentially a starting point for such investigation.

The original OMP algorithm (TROPP; GILBERT, 2007) is a particular case of OMPED where $K = 1$ and the parameter μ_c (Eqs. (28) and (32)) is set to zero. Running OMP with a K times finer (denser) grid treats a problem with the same size as running OMPED with an order-K dictionary, but OMPED consistently outperformed OMP on the cases simulated in this work.

In both OMP and OMPED, the residual vector $\mathbf{e}$ on each iteration is orthogonal to all active elements of the dictionary, what places OMPED in the family of *Orthogonal* Matching Pursuit algorithms. The same does not hold for the COMP algorithm presented in (KNUDSON et al., 2014): the fact that the LS regression performed at each iteration contains linear constraints may result in eventual coherence between the residual and the active elements of the dictionary.

Another particularity of OMPED in regard to previously proposed algorithms for expanded dictionaries (EKANADHAM et al., 2011; KNUDSON et al., 2014; FYHN et al., 2015) is that it is not based on conic constraints, which removes any restrictions on the choice of the sizes Δ_x and Δ_z (and further dimensions if that is the case) for the division of the ROI into local ROIs.

5.1 CONTRIBUTIONS

The original contributions of the present work are listed below.

1. Application of the framework of dictionary expansion (EKANADHAM et al., 2011; KNUDSON et al., 2014) to the problem of two-dimensional ultrasound imaging;
2. Extension of the SVD expansion (KNUDSON et al., 2014) to a two-dimensional problem;
3. Development of the Minimax expansion, which creates a dictionary where the maximum residual norm is minimal;

4. Development of the OMPED algorithm that, contrary to the algorithms presented in the works of Ekanadham et al. (2011), Knudson et al. (2014) and (FYHN et al., 2015) and thanks to a non-convex constraint set, allows for the use of local ROIs of any size.

5.2 ACCOMPLISHMENT OF OBJECTIVES

We conclude that the general and specific objectives declared in Section 1.2 were accomplished, as detailed below.

As an intermediate step towards the formulation of the generalized Minimax expansion described in Subection 3.1.3, a polynomial Minimax expansion was formulated and evaluated for a prototypical one-dimensional problem. The results of a comparison to the types of expansion proposed in (EKANADHAM et al., 2011) and (KNUDSON et al., 2014) were reported in (PASSARIN et al., 2017), accomplishing the specific objective 1. The formulation of the polynomial Minimax expansion and its relation to the generalized Minimax expansion presented in Subsection 3.1.3 are described in Appendix A.

The specific objective 2 was accomplished in Section 3.1 both by the extension of the SVD expansion to two dimensions (Section 3.1.2) and by the proposition of of the generalized Minimax expansion for two dimensions (Subsection 3.1.3).

To accomplish specific objective 3, simulations were carried using Field II package (JENSEN, 2004), as described in Section 3.3. The simulations reproduced the acquisition environment of (GUARNERI et al., 2015) for ultrasound NDE.

The convex constraints used in previously proposed algorithms impose restrictions that limit the creation of acquisition models, as explained in Subsection 3.2.2. To work out that limitation, an extension of the OMP algorithm which uses non-convex constraints was formulated in Subsection 3.2.3. The results reported in Chapter 4 proved the algorithm effective, accomplishing specific objective 4.

Also on Chapter 4 results were presented from the image reconstruction from 200 simulated cases representing ultrasound acquisition with point-like scatterers randomly distributed over a ROI. Performance of the proposed algorithm was assessed in terms of cardinality of the solution (Fig. 12a), rate of event recovery (Fig. 8), recovered location accuracy (Fig. 10a) and recovered amplitude accuracy (Fig. 10b). A reconstrcuted image was also displayed as example and compared to the results obtained with other model-based reconstruction algorithms which do not account for off-grid deviations (Fig. 15).

Therefore, specific objective 5 was achieved, as well the general objective.

5.3 FUTURE WORK

The adaptation of OMP into OMPED, with a constraint imposed on the selection of the index added to the support at each iteration, might be replicable to other greedy search algorithms. The class of forward-backward algorithms is of special interest in signal and image recovery because of its capacity of later "correction" of "wrong" choices made on the selection of indices to add to the support (MILLER, 2002; SOUSSEN et al., 2011), which constitutes a motivation for future investigation.

One limitation of our technique is that one single point-like event is identifiable inside each local ROI. The search for a means to overcome this limitation, allowing for the recovery of several scatterers inside the same local ROI is a relevant topic for further investigation and may broaden the applicability of the proposed technique.

Our simulated data considered point-like reflectors, with spatial coordinates (x, z) as the only nonlinear parameters. The ultrasound NDT literature contains parametric reflection models for more complex discontinuity structures, such as spherical voids and circular cracks, where the distortion of ultrasound waves is modelled as a nonlinear function of parameters like diameter and angle to the surface (SCHMERR; SONG, 2007; VELICHKO et al., 2017). The proposed method is applicable to those cases as long as those parameters are comprised in the parameter set $\boldsymbol{\tau}$ in (1) and sampled like the parameters of spatial location. In this case, characterization of discontinuities could be performed along with location. Classification of discontinuities could also be jointly performed if dictionaries for several types of discontinuities are combined. An equivalent principle has been used in the joint detection and identification of neuron activity using using SVD (KNUDSON et al., 2014) and Taylor (EKANADHAM et al., 2014) expanded dictionaries.

Finally, the application of the proposed method to real acquired data and to simulated data representing more complex structures (e.g. using the CIVA environment (CALMON et al., 2006)) is an important next step.

REFERENCES

ANTONIOU, A.; LU, W.-S. **Practical Optimization: Algorithms and Engineering Applications**. 1st. ed. [S.l.]: Springer Publishing Company, Incorporated, 2007. ISBN 0387711066, 9780387711065.

BARRETT, H.; MYERS, K. **Foundations of image science**. [S.l.]: Wiley-Interscience, 2004. (Wiley series in pure and applied optics). ISBN 9780471153009.

BJÖRCK, A. **Numerical methods for least squares problems**. [S.l.]: SIAM, 1996. ISBN 0898713609.

BOVIK, A. **Handbook of Image & Video Processing**. [S.l.]: Academic Press, 2000. (Academic Press Series in Communications, Networking and Multimedia). ISBN 9780121197902.

BOYD, S.; VANDENBERGHE, L. **Convex Optimization**. New York, NY, USA: Cambridge University Press, 2004. ISBN 0521833787.

CALMON, P.; MAHAUT, S.; CHATILLON, S.; RAILLON, R. Civa: An expertise platform for simulation and processing ndt data. **Ultrasonics**, Elsevier, v. 44, p. e975–e979, 2006.

CARCREFF, E.; BOURGUIGNON, S.; IDIER, J.; SIMON, L. A linear model approach for ultrasonic inverse problems with attenuation and dispersion. **IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control**, v. 61, n. 7, p. 1191–1203, July 2014. ISSN 0885-3010.

CENSOR, Y. Finite series-expansion reconstruction methods. **Proceedings of the IEEE**, IEEE, v. 71, n. 3, p. 409–419, 1983.

CHEN, S. S.; DONOHO, D. L.; SAUNDERS, M. A. Atomic decomposition by basis pursuit. **SIAM Journal on Scientific Computing**, Society for Industrial and Applied Mathematics, v. 20, n. 1, p. 33–61, 1998.

CHI, C.-Y.; GOUTSIAS, J.; MENDEL, J. A fast maximum-likelihood estimation and detection algorithm for Bernoulli-Gaussian processes. In: **ICASSP '85. IEEE International Conference on Acoustics, Speech, and Signal Processing**. [S.l.: s.n.], 1985. v. 10, p. 1297–1300.

DESOKY, H.; YOUSSEF, A.-B. M.; KADAH, Y. M. Reconstruction using optimal spatially variant kernel for b-mode ultrasound imaging. In: **SPIE Medical Imaging 2003: Ultrasonic Imaging and Signal Processing**. [S.l.: s.n.], 2003. p. 147–153.

ECKART, C.; YOUNG, G. The approximation of one matrix by another of lower rank. **Psychometrika**, v. 1, n. 3, p. 211–218, Sep 1936. ISSN 1860-0980.

EKANADHAM, C.; TRANCHINA, D.; SIMONCELLI, E. P. Recovery of sparse translation-invariant signals with continuous basis pursuit. **IEEE Transactions on Signal Processing**, v. 59, n. 10, p. 4735–4744, Oct 2011. ISSN 1053-587X.

EKANADHAM, C.; TRANCHINA, D.; SIMONCELLI, E. P. A unified framework and method for automatic neural spike identification. **J. Neuroscience Methods**, v. 222, p. 47–55, Jan 2014.

FESSLER, J. A. Model-based image reconstruction for MRI. **IEEE Signal Processing Magazine**, IEEE, v. 27, n. 4, p. 81–89, 2010.

FYHN, K.; MARCO, F.; HOLDT, S. et al. Compressive parameter estimation for sparse translation-invariant signals using polar interpolation. **IEEE Transactions on Signal Processing**, IEEE, v. 63, n. 4, p. 870–881, 2015.

GOLUB, G. H.; LOAN, C. F. V. **Matrix Computations (Johns Hopkins Studies in Mathematical Sciences)(3rd Edition)**. 3rd. ed. [S.l.]: The Johns Hopkins University Press, 1996. Paperback. ISBN 0801854148.

GRANT, M.; BOYD, S. Graph implementations for nonsmooth convex programs. In: BLONDEL, V.; BOYD, S.; KIMURA, H. (Ed.). **Recent Advances in Learning and Control**. [S.l.]: Springer-Verlag Limited, 2008, (Lecture Notes in Control and Information Sciences). p. 95–110.

GRANT, M.; BOYD, S. **CVX: Matlab Software for Disciplined Convex Programming, version 2.1**. mar. 2014.

GUARNERI, G. A.; PIPA, D. R.; JUNIOR, F. N.; ARRUDA, L. V. R. de; ZIBETTI, M. V. W. A sparse reconstruction algorithm for ultrasonic images in nondestructive testing. **Sensors**, v. 15, n. 4, p. 9324–9343, 2015. ISSN 1424-8220.

HANSEN, C. **Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion**. Philadelphia: SIAM, 1998.

JENSEN, J. Simulation of advanced ultrasound systems using field ii. In: **Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on**. [S.l.: s.n.], 2004. p. 636–639 Vol. 1.

JENSEN, J. A. A model for the propagation and scattering of ultrasound in tissue. **Acoustical Society of America**, v. 89, p. 182–190, 1991.

JR, L. W. S. **Fundamentals of ultrasonic phased arrays**. [S.l.]: Springer, 2014.

KIM, S.-J.; KOH, K.; LUSTIG, M.; BOYD, S.; GORINEVSKY, D. An interior-point method for large-scale l1-regularized least squares. **IEEE journal of selected topics in signal processing**, IEEE, v. 1, n. 4, p. 606–617, 2007.

KNUDSON, K. C.; YATES, J.; HUK, A.; PILLOW, J. W. Inferring sparse representations of continuous signals with continuous orthogonal matching pursuit. In: **Advances in Neural Information Processing Systems 27**. [S.l.: s.n.], 2014. p. 1215–1223.

KRUIZINGA, P.; MEULEN, P. van der; FEDJAJEVS, A.; MASTIK, F.; SPRINGELING, G.; JONG, N. de; BOSCH, J. G.; LEUS, G. Compressive 3d ultrasound imaging using a single sensor. **Science Advances**, American Association for the Advancement of Science, v. 3, n. 12, 2017.

LAVARELLO, R.; KAMALABADI, F.; O'BRIEN, W. D. A regularized inverse approach to ultrasonic pulse-echo imaging. **Medical Imaging, IEEE Transactions on**, IEEE, v. 25, n. 6, p. 712–722, 2006.

LINGVALL, F.; OLOFSSON, T. On time-domain model-based ultrasonic array imaging. **Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on**, IEEE, v. 54, n. 8, p. 1623–1633, 2007.

LINGVALL, F.; OLOFSSON, T.; STEPINSKI, T. Synthetic aperture imaging using sources with finite aperture: Deconvolution of the spatial impulse response. **The Journal of the Acoustical Society of America**, v. 114, p. 225–234, 07 2003.

LU, W. S. Minimax design of nonlinear-phase FIR filters: a least-pth approach. In: **2002 IEEE International Symposium on Circuits and Systems.** [S.l.: s.n.], 2002. v. 1, p. I–409–I–412 vol.1.

MILLER, A. **Subset selection in regression**. [S.l.]: CRC Press, 2002.

MOR, E.; AZOULAY, A.; ALADJEM, M. A matching pursuit method for approximating overlapping ultrasonic echoes. **IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control**, v. 57, n. 9, p. 1996–2004, September 2010. ISSN 0885-3010.

OLLINGER, J. M.; FESSLER, J. A. Positron-emission tomography. **IEEE Signal Processing Magazine**, IEEE, v. 14, n. 1, p. 43–55, 1997.

OLOFSSON, T.; WENNERSTROM, E. Sparse deconvolution of b-scan images. **IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control**, v. 54, n. 8, p. 1634–1641, August 2007. ISSN 0885-3010.

PASSARIN, T. A. R.; PIPA, D. R.; ZIBETTI, M. V. W. A minimax dictionary expansion for sparse continuous reconstruction. In: **2017 25th European Signal Processing Conference (EUSIPCO)**. [S.l.: s.n.], 2017. p. 2136–2140.

PASSARIN, T. R.; ZIBETTI, M. W.; PIPA, D. R. Sparse ultrasound imaging via manifold low-rank approximation and non-convex greedy pursuit. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 18, n. 12, p. 4097, 2018.

SCHMERR, L. W.; SONG, S.-J. **Ultrasonic nondestructive evaluation systems: models and measurements.** [S.l.]: Springer, 2007.

SHPAK, D. Design of mixed-norm fir filters using an unconstrained least-pth algorithm. In: **2003 IEEE Pacific Rim Conference on Communications Computers and Signal Processing (PACRIM 2003)**. [S.l.: s.n.], 2003. v. 1, p. 253–255 vol.1.

SMITH, N.; WEBB, A. **Introduction to Medical Imaging: Physics, Engineering and Clinical Applications**. [S.l.]: Cambridge University Press, 2010. (Cambridge Texts in Biomedical Engineering). ISBN 9780521190657.

SOUSSEN, C.; IDIER, J.; BRIE, D.; DUAN, J. From Bernoulli Gaussian deconvolution to sparse signal restoration. **IEEE Transactions on Signal Processing**, v. 59, n. 10, p. 4572–4584, Oct 2011. ISSN 1053-587X.

STURM, J. F. Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones. **Optimization methods and software**, Taylor & Francis, v. 11, n. 1-4, p. 625–653, 1999.

TANG, G.; BHASKAR, B. N.; SHAH, P.; RECHT, B. Compressed sensing off the grid. **IEEE Transactions on Information Theory**, v. 59, n. 11, p. 7465–7490, Nov 2013. ISSN 0018-9448.

TEKE, O.; GURBUZ, A. C.; ARIKAN, O. A robust compressive sensing based technique for reconstruction of sparse radar scenes. **Digital Signal Processing**, Elsevier, v. 27, p. 23–32, 2014.

TOH, K. C.; TODD, M.; TÜTÜNCÜ, R.; TUTUNCU, R. H. SDPT3 - a Matlab software package for semidefinite programming. **Optimization Methods and Software**, v. 11, p. 545–581, 1998.

TROPP, J. A.; GILBERT, A. C. Signal recovery from random measurements via orthogonal matching pursuit. **IEEE Transactions on Information Theory**, v. 53, n. 12, p. 4655–4666, Dec 2007. ISSN 0018-9448.

TUPHOLME, G. E. Generation of acoustic pulses by baffled plane pistons. **Mathematika**, v. 16, p. 209–224, 12 1969. ISSN 2041-7942.

VELICHKO, A.; BAI, L.; DRINKWATER, B. Ultrasonic defect characterization using parametric-manifold mapping. **Proc. R. Soc. A**, The Royal Society, v. 473, n. 2202, p. 20170056, 2017.

VIOLA, F.; ELLIS, M. A.; WALKER, W. F. Time-domain optimized near-field estimator for ultrasound imaging: initial development and results. **Medical Imaging, IEEE Transactions on**, IEEE, v. 27, n. 1, p. 99–110, 2008.

WILLIAMS, H. P. **Model building in mathematical programming**. [S.l.]: John Wiley & Sons, 2013.

YANG, Z.; ZHANG, C.; XIE, L. Robustly stable signal recovery in compressed sensing with structured matrix perturbation. **IEEE Transactions on Signal Processing**, IEEE, v. 60, n. 9, p. 4658–4671, 2012.

ZANIN, L.; ZIBETTI, M.; SCHNEIDER, F. Conjugate gradient and regularized inverse problem-based solutions applied to ultrasound image reconstruction. In: **2011 IEEE International Ultrasonics Symposium (IUS)**. [S.l.: s.n.], 2011. p. 377–380. ISSN 1948-5719.

ZANIN, L. G. S.; K., S. F.; ZIBETTI, M. V. W. Regularized reconstruction of ultrasonic imaging and the regularization parameter choice. In: **International Conference on Bio-inspired Systems and Signal Processing**. [S.l.: s.n.], 2012. p. 438–442.

ZHU, H.; LEUS, G.; GIANNAKIS, G. B. Sparsity-cognizant total least-squares for perturbed compressive sampling. **IEEE Transactions on Signal Processing**, v. 59, n. 5, p. 2002–2016, May 2011. ISSN 1053-587X.

APPENDIX A – POLYNOMIAL MINIMAX EXPANSION

As an initial step on the development of the Minimax expansion described in Subsection 3.1.3, a Minimax expansion approach based on a polynomial regression was derived for a prototypical 1-dimensional linear time-invariant (LTI) case. We refer to this approach as Polynomial Minimax. The derivations are presented in this appendix. Comparisons of performance between the Polynomial Minimax and other expansion approaches were reported in (PASSARIN et al., 2017). Here we refer to the Minimax expansion presented in Subsection 3.1.3 as Generalized Minimax

Consider a one-dimensional linear time-invariant (LTI) system whose impulse response $h(t)$ so that the response to a delayed impulse $\delta(t-\tau)$ is $h(t-\tau)$. The delayed response $h(t-\tau)$ is time sampled, generating the vector $\mathbf{y}(\tau) \in \mathbb{R}^M$ Recalling the manifold definition of (1), in this case the vector of nonlinear parameters $\boldsymbol{\tau}$ contains only the time delay τ

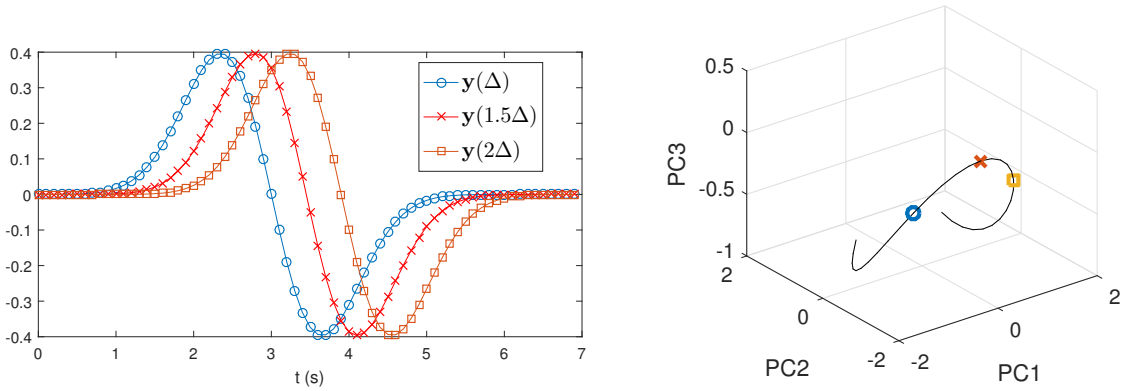
$$\boldsymbol{\tau} = [\tau]. \tag{41}$$

The blue circle curve in Fig. 16a shows $\mathbf{y}(\Delta)$, i.e., the response delayed by an arbitrary real-valued Δ for τ.

The problem of deconvolving an arbitrary signal that has been output by the system consists in finding a set of N impulses at the input, each having amplitude v_n and delay τ_n, such that the time-sampled output $\mathbf{c}$ can be represented as a finite combination of impulse responses:

$$\mathbf{c} = \sum_{n=1}^{N} v_n \mathbf{y}(\tau_n) + \mathbf{e} \tag{42}$$

where $\mathbf{e} \in \mathbb{R}^M$ represents additive Gaussian measurement noise with zero mean and variance σ^2. This formulation is found in many applications of inverse problems such as neuron spike detection (EKANADHAM et al., 2014) and 1-dimensional ultrasound nondestructive testing (NDT) (CARCREFF et al., 2014; MOR et al., 2010), where the signal to be reconstructed is modelled as a limited sum of shifted impulses with arbitrary amplitudes.



(a) (b)

Figure 16: Sampling of 1-dimensional LTI manifold. (a) Responses $\mathbf{y}(\boldsymbol{\Delta})$, $\mathbf{y}(\mathbf{2\Delta})$ and off-grid response $\mathbf{y}(\mathbf{1.5\Delta})$ (b) Black line: 3-component PCA view of the manifold drawn by the variation of $\boldsymbol{\tau}$ in the LTI acquisition model $\mathbf{y}(t-\boldsymbol{\tau})$. The circle, X and square marks correspond to $\boldsymbol{\tau} = \boldsymbol{\Delta}$, $\boldsymbol{\tau} = \mathbf{1.5\Delta}$ (off-grid) and $\boldsymbol{\tau} = \mathbf{2\Delta}$ respectively. Source: adapted from (PASSARIN et al., 2017).

In order to build the acquisition model, a discretization of the delay τ is required. Without loss of generality, we define a regular sampling of τ at intervals of size Δ so that each n-th modelled delay τ_n is defined as $n\Delta$. The resulting discrete model $\mathbf{H}$ is defined as

$$\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_N,] = [\mathbf{y}(\Delta), \mathbf{y}(2\Delta), \ldots, \mathbf{y}(N\Delta)], \tag{43}$$

which allows for writing (42) in compact the form of (3).

The problem of off-grid deviation arises when some component of the data vector $\mathbf{c}$ corresponds to a response $\mathbf{y}(\tau)$ with a value for τ not contemplated by the model $\mathbf{H}$, i.e., not a multiple of Δ. To cope with that phenomenon, the acquisition model (42) can be adapted to

$$\mathbf{c} = \sum_{n=1}^{N} v_n \mathbf{y}([n + \tau_n]\Delta) + \mathbf{e}, \tag{44}$$

where τ_n can vary continuously within the interval $[-0.5, 0.5]$, defining a time bin with length Δ. Each column $\mathbf{h}_n$ in $\mathbf{H}$ is replaced by a matrix $\mathbf{B}^{(n)} \in \mathbb{R}^{M \times K}$ whose columns combine linearly to approximate the n-th corresponding local manifold (time bin) $\mathbf{y}([n + \tau_n]\Delta)$, $\tau_n = [-0.5, 0.5]$. The resulting approximated acquisition model is expressed in (6).

The polynomial Minimax criterion for the definition of $\mathbf{B}^{(n)}$ is based on a polynomial regression, with τ_n as the independent variable. We want to approximate deviations within the n-th time bin as an order-k polynomial composition of the columns $\mathbf{b}_k^{(n)}$ of

$\mathbf{B}^{(n)}$:

$$\mathbf{y}([n+\tau_n]\Delta) \approx \sum_{k=1}^{K} \tau_n^{k-1}\mathbf{b}_k^{(n)} \tag{45a}$$

$$\mathbf{y}([n+\tau_n]\Delta) \approx \mathbf{b}_1^{(n)} + \tau_n\mathbf{b}_2^{(n)} + \tau_n^2\mathbf{b}_3^{(n)} + \cdots + \tau_n^{K-1}\mathbf{b}_K^{(n)} \tag{45b}$$

$$\mathbf{y}([n+\tau_n]\Delta) \approx \mathbf{B}^{(n)}[1, \tau_n, \tau_n^2, \ldots, \tau_n^{K-1}]^T. \tag{45c}$$

In order to cast the regression problem, a fine sampling of the off-grid deviation τ_n is performed uniformly within the interval $[-0.5, 0.5]$, yielding a set of R different deviations $\tau_{n,i},\ i = 1, \ldots, R$. Let

$$\mathbf{m}_i^{(n)} = \mathbf{y}([n+\tau_{n,i}]\Delta)\ \forall i \in \{1, \ldots, R\} \tag{46}$$

be the i-th (out of R) manifold sample within the n-th time bin, and let

$$\mathbf{f}_i^{(n)} = [1, \tau_{n,i}, \tau_{n,i}^2, \ldots, \tau_{n,i}^{K-1}]^T\ \forall i \in \{1, \ldots, R\} \tag{47}$$

be the vector formed by the values yielded from taking all the $K-1$ powers of $\tau_{n,i}$. The R manifold samples yield R instances of approximation (45c):

$$\mathbf{m}_i^{(n)} \approx \mathbf{B}^{(n)}\mathbf{f}_i^{(n)}\ \forall i \in \{1, \ldots, R\}, \tag{48}$$

each of which yield a vector of residuals $\mathbf{r}_i^{(n)} \in \mathbb{R}^M$:

$$\mathbf{r}_i^{(n)} = \mathbf{m}_i^{(n)} - \mathbf{B}^{(n)}\mathbf{f}_i^{(n)}\ \forall i \in \{1, \ldots, R\}. \tag{49}$$

Our goal now is the definition of a set $\{\mathbf{d}_n^{(k)}\}$ that minimizes the residual norms $\|\mathbf{r}_{n,i}\|$ for every τ_i drawn from our oversampling of the time bin. One simple approach would be solving a simple Least Squares (LS) problem:

$$\{\hat{\mathbf{B}}^{(n)}\} = \arg\min_{\{\mathbf{B}^{(n)}\}} \sqrt{\sum_{i=1}^{R} \|\mathbf{r}_i^{(n)}\|^2}\ \text{ s.t. } \{\mathbf{r}_i^{(n)} = \mathbf{m}^{(n)} - \mathbf{B}^{(n)}\mathbf{f}_i^{(n)}\}\ \forall i \in \{1, \ldots, T\}. \tag{50}$$

In order to create an egalitarian approximation, i.e. an approximation that distributes the residuals along the delays $\tau_n \in [-0.5, 0.5]$, the problem is formulated as a minimize-maximum (Minimax) residual:

$$\hat{\mathbf{B}}^{(n)} = \arg\min_{\mathbf{B}^{(n)}} r_{max}\ \text{ s.t. } \left\{ \begin{matrix} \|\mathbf{r}_i^{(n)}\| \geq r_{max} \\ \mathbf{r}_i^{(n)} = \mathbf{m}^{(n)} - \mathbf{B}^{(n)}\mathbf{f}_i^{(n)} \end{matrix} \right\}\ \forall i \in \{1, \ldots, R\}. \tag{51}$$

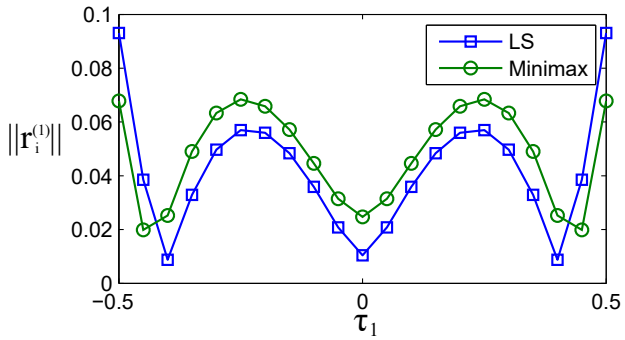


Figure 17: Residual norms $\|\mathbf{r}_i^{(1)}\|$ for the waveform of Fig. 16a. Least Squares solution is yielded by (50) and Minimax solution is yielded by (51).
Source: (PASSARIN et al., 2017).

The real-valued scalar r_{max} in (51) is a slack variable introduced in the problem. It is minimized in the cost function and, through the set of inequalities $\|\mathbf{r}_i^{(n)}\| \geq r_{max}$, this minimization is propagated to the maximum residual norm $\max_i \|\mathbf{r}_i^{(n)}\|$. As consequence, problem (51) is equivalent to:

$$\hat{\mathbf{B}}^{(n)} = \arg\min_{\mathbf{B}^{(n)}} \max_i \ \|\mathbf{r}_i^{(n)}\| \ \text{ s.t. } \{\mathbf{m}^{(n)} - \mathbf{B}^{(n)}\mathbf{f}_i^{(n)} = \mathbf{r}_i^{(n)}\} \ \forall i \in 1, \ldots, R. \tag{52}$$

Problem (51) a is Second-Order Cone Program (SOCP) which can be solved by convex optimization packages such as SeDuMi (STURM, 1999), SDPT3 (TOH et al., 1998) and CVX (GRANT; BOYD, 2014, 2008).

Fig. 17 shows the residual norms $\|\mathbf{r}_i^{(1)}\|$ ($n = 1$ denotes expansion of the first time bin) for the waveform of Fig. 16a as yielded by the LS (50) and Minimax (51) expansions.

A.1 RELATION BETWEEN POLYNOMIAL MINIMAX AND GENERALIZED MINIMAX EXPANSIONS

We refer to the Minimax expansion presented in Subsection 3.1.3 as Generalized Minimax, in opposition to the Polynomial Minimax expansion presented in Appendix A. The term Generalized derives from the fact that the modulating matrices $\{\mathbf{F}^{(n)}\}$ are also subject to optimization along with the basis matrices $\{\mathbf{B}^{(n)}\}$, whereas in the Polynomial Minimax the matrices $\{\mathbf{F}^{(n)}\}$ contain samples of the polynomial terms τ^{k-1}. The correspondence between the two approaches is presented in the current appendix. As in Subsection 3.1.3, we suppress indices n and represent matrices related to the n-th local ROI such as $\mathbf{M}^{(n)}$ and $\mathbf{B}^{(n)}$ as simply $\mathbf{M}$, $\mathbf{B}$ and so forth.

The search for the Polynomial Minimax basis is performed via a vector regression

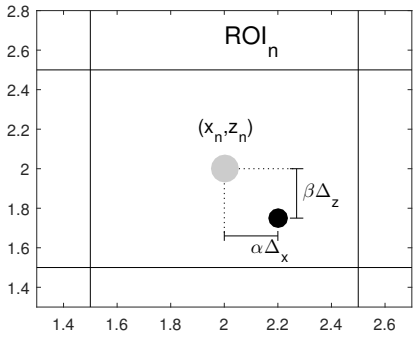


Figure 18: To cast a regression problem, the off-grid deviation of a given point within the the n-th ROI is mapped onto D independent variables (two in this example): α and β denote the deviation of the point from the corresponding modelled position in the lateral (x) and axial (z) axis respectively. Their values vary within the range $[-0.5, 0.5]$, normalized over Δ_x and Δ_z, so that the coordinates of a point are $(x, z) = (x_n + \alpha\Delta_x, z_n + \beta\Delta_z)$.

problem. To cast the regression problem, it is necessary to map the off-grid deviation of each of the R points sampled within the local ROI onto D independent variables, one for each dimension. In our case $D = 2$, then two variables are required to locate any given point relative to a modelled position: α for the x direction and β for the z direction, as represented in Fig. 18. The values of α and β are constrained within the normalized range $[-0.5, 0.5]$ so that any point witihn a local ROI has coordinates

$$[x, z]^T = [x_n + \alpha\Delta_x, z_n + \beta\Delta_z]^T.$$

Now each i-th point from the fine sampling of the n-th local ROI (see Fig. 2c) can have its position expressed in terms of its off-grid deviation α_i and β_i:

$$[x, z]^T = [x_n + \alpha_i\Delta_x, z_n + \beta_i\Delta_z]^T.$$

The basis of the regression problem is expressed in Eq. (53). Each column $\mathbf{m}_i$ from $\mathbf{M}$ will have a rank-K approximation $\tilde{\mathbf{m}}_i$, composed by a linear combination of the basis vectors $\mathbf{b}_k$. Such linear combination is determined by K regression functions $f_k(\alpha, \beta)$.

$$\mathbf{m}_i \approx \tilde{\mathbf{m}}_i = \sum_{k=1}^{K} \mathbf{b}_k f_k(\alpha_i, \beta_i) \ \forall i \in \{1, 2, \ldots, R\} \tag{53}$$

If the K functions $f_k(\alpha, \beta)$ are determined, we can define a minimax optimization

problem for the basis vectors $\mathbf{b}_k$:

$$\hat{\mathbf{B}} = \arg\min_{\mathbf{B}} \max_{i} \|\mathbf{r}_i\| \text{ s.t. } \left\{ \mathbf{r}_i = \mathbf{m}_i - \sum_{k=1}^{K} \mathbf{b}_k f_k(\alpha_i, \beta_i) \right\} \forall\ i \in \{1, 2, \ldots, R\} \tag{54}$$

Now let vectors $\mathbf{f}_i \in \mathbb{R}^K$ be defined in Eq. (55):

$$\mathbf{f}_i = [f_1(\alpha_i, \beta_i), \ldots, f_K(\alpha_i, \beta_i)]^H \ \forall i \in \{1, \ldots, R\} \tag{55}$$

We group all the R vectors $\mathbf{f}_i$ to form the matrix $\mathbf{F}$, as shown in Eq. 56.

$$\mathbf{F} = \begin{bmatrix} f_1(\alpha_1, \beta_1) & f_1(\alpha_2, \beta_2) & \ldots & f_1(\alpha_R, \beta_R) \\ f_2(\alpha_1, \beta_1) & f_2(\alpha_2, \beta_2) & \ldots & f_2(\alpha_R, \beta_R) \\ \vdots & \vdots & \ddots & \vdots \\ f_K(\alpha_1, \beta_1) & f_K(\alpha_2, \beta_2) & \ldots & f_K(\alpha_R, \beta_R) \end{bmatrix} \tag{56a}$$

$$\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_K] \tag{56b}$$

The definition of $\mathbf{F}$ allows for writing the system of R equations of problem (54) in compact form:

$$\hat{\mathbf{B}} = \arg\min_{\mathbf{B}} \max_{i} \|\mathbf{r}_i\| \text{ s.t. } \mathbf{R} = \mathbf{M} - \mathbf{B}\mathbf{F}. \tag{57}$$

The resulting Eq. (57) presents the same optimization problem as Eq. (17).

On (PASSARIN et al., 2017), a prototypical one-dimension problem is approached, thus only one independent variable τ is defined for the off-grid deviation. A simple choice is then made as the family of polynomials $f_k(\tau) = \tau^{k-1}$ is used. This set of functions may be sufficiently versatile for one-dimension problems, but defining its best extension for two or higher dimensions is not straightforward. Rather than choosing a set of analytically determined functions $f_k(\alpha, \beta)$ and sampling them over the set $(\alpha_i, \beta_i) : i \in \{1, 2, \ldots, R\}$ to obtain $\mathbf{F}$, we make $\mathbf{F}$ also subject to optimization.

APPENDIX B – CASTING THE MINIMAX PROBLEM AS SOCP

In (18), we introduce a variable $r_{\max}$, which is the only term on the cost function, and constrain it to be equal to or grater than every residual norm $\|\mathbf{r}_i\| = \|\mathbf{B}\mathbf{f}_i - \mathbf{m}_i\|$. For any candidate solution, the minimization of $r_{\max}$ forces the minimization of solely the largest residual norm among all $\|\mathbf{r}_i\|$. This is a common choice in the formulation of minimax problems in linear programming (WILLIAMS, 2013). The resulting problem of Eq. (18) can be translated into a Second Order Cone Program (SOCP) (BOYD; VANDENBERGHE, 2004)(ANTONIOU; LU, 2007).

The standard form of SOCPs (BOYD; VANDENBERGHE, 2004)(ANTONIOU; LU, 2007) is presented in Eq. (58).

$$\min_{\mathbf{y}} \mathbf{w}^T\mathbf{y} \text{ s.t. } \|\mathbf{A}_i\mathbf{y} + \mathbf{g}_i\| \leq \mathbf{c}_i^T\mathbf{y} + s_i \ \forall i = \{1, \ldots, R\} \tag{58}$$

To translate the problem of Eq. (18) into the SOCP (58), we start by defining the variables vector $\mathbf{y} \in \mathbb{R}^{KM+1}$:

$$\mathbf{y} = [\mathbf{b}_1^T, \ldots, \mathbf{b}_K^T, r_{\max}]^T. \tag{59}$$

We then define the weights vector $\mathbf{w} \in \mathbb{R}^{KM+1}$:

$$\mathbf{w} = [\mathbf{0}^T, \ldots, \mathbf{0}^T, 1]^T, \tag{60}$$

where $\mathbf{0} \in \mathbb{R}^M$ is an all-zero vector, so that

$$\mathbf{w}^T\mathbf{y} = r_{\max}. \tag{61}$$

The product $\mathbf{B}\mathbf{f}_i$ can be replaced with the following identity:

$$(\mathbf{f}_i^T \otimes \mathbf{I})[\mathbf{b}_1^T, \ldots, \mathbf{b}_K^T]^T = \mathbf{B}\mathbf{f}_i, \tag{62}$$

where $\otimes$ denotes the Kronecker product and $\mathbf{I} \in \mathbb{R}^{M \times M}$ is the identity matrix. This allows

for the definition of $\mathbf{A}_i \in \mathbb{R}^{M\times(KM+1)}$:

$$\mathbf{A}_i = [\mathbf{f}_i^T \otimes \mathbf{I}, \mathbf{0}] \ \forall i \in \{1, \ldots, R\}, \tag{63}$$

Finally we define vectors $\mathbf{g}_i \in \mathbb{R}^M$:

$$\mathbf{g}_i = -\mathbf{b}_i \ \forall i \in \{1, \ldots, R\}. \tag{64}$$

Since the affine terms s_i are not used, they are set to zero:

$$\mathbf{s}_i = 0 \ \forall i \in \{1, \ldots, R\}. \tag{65}$$

When we replace the members of the SOCP (58) with the equivalences (59)-(65), we obtain the problem of (18).