

Ultrasound Imaging: From Physical Modeling to Deep Learning

Présentée le 8 juillet 2021

Faculté des sciences et techniques de l'ingénieur
Laboratoire de traitement des signaux 5
Programme doctoral en génie électrique

pour l'obtention du grade de Docteur ès Sciences

par

Dimitris PERDIOS

Acceptée sur proposition du jury

Prof. D. N. A. Van De Ville, président du jury
Prof. J.-Ph. Thiran, directeur de thèse
Prof. M. Tanter, rapporteur
Prof. J. A. Jensen, rapporteur
Prof. M. Unser, rapporteur

*To my parents, Thomas and Gabrielle,
for their unconditional love and support*

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my academic advisor, Prof. Jean-Philippe Thiran, for giving me the opportunity to undertake and for having supported me throughout this fascinating journey that was my thesis at the Signal Processing Laboratory 5 (LTS5), École polytechnique fédérale de Lausanne (EPFL). I am particularly grateful for the trust and freedom you gave me during all these years and through the multiple research directions explored, as well as for your numerous scientific, professional, and personal advices which enabled me to unlock many situations. Thank you also for giving me the opportunity to work with a wonderful team that contributed in an extraordinary way to the accomplishment of this thesis.

I am grateful to my thesis committee, presided by Prof. Dimitri Van De Ville, and composed of my thesis advisor and three internationally renowned experts: Prof. Mickaël Tanter, Prof. Jørgen A. Jensen, and Prof. Michaël Unser. I am deeply thankful for their interest in my research work and for the memorable hours-long discussion during my private defense. It was truly an honor to defend my thesis in front of the three experts who inspired this thesis the most. I would like to express my special thanks to each of them. Dear Prof. Mickaël Tanter, thank you for your early interest and enthusiasm in my work, and for the discussions we had at the Institute of Physics for Medicine Paris about the concrete needs and challenges in ultrafast ultrasound (US) imaging. These exciting discussions were extremely helpful and oriented my research considerably. Dear Prof. Jørgen A. Jensen, your pioneer work on US modeling inspired me greatly and was foundational to my thesis. I am extremely grateful that you were able to serve on my thesis committee. Dear Prof. Michaël Unser, it is an understatement to say that your work on biomedical imaging and on splines was central to this thesis. Thank you for the exciting discussions at the Imaging@EPFL seminars, and for transmitting your true love of splines to me; it is definitely also a perfect fit for US modeling and imaging.

During my thesis, I had the chance to meet, at various conferences and events, many brilliant minds who inspired and supported me, and whom I would like to thank warmly. On two occasions, I visited the Institute of Physics for Medicine Paris, during which I had the great opportunity to discuss many applications of ultrafast US imaging with Prof. Mickaël Tanter, Prof. Olivier Couture, Prof. Mathieu Pernot, Dr. Béatrice Walker, Dr. Thomas Defieux, and Dr. Jonathan Porée, to whom I am very grateful for the precious insights received. Many thanks to Béatrice for her invitation to present my work there and for creating this opportunity. I am also deeply grateful to Prof. Muyinatu A. Lediju Bell for her continuous interest in my research, and for inviting me to present at the Johns

Hopkins University, for the visit of her laboratory, and for the inspiring discussions. Many thanks to Arun A. Nair for welcoming me there and for the fruitful discussions. I would also like to express my sincere gratitude to all the members and people connected to the CREATIS laboratory I had the chance to meet, but space prevents me from mentioning them all. Thank you Prof. Hervé Liebgott, Prof. Olivier Bernard, Prof. Adrian Basarab, Prof. Damien Garcia, Prof. François Varray, Prof. Barbara Nicolas, Prof. Thomas Grenier, Dr. Sébastien Salles, and many others, for the memorable moments, your help, and your support. Special thanks to Hervé and Olivier for having followed and encouraged me from the very beginning, and from whom I learned a lot. Many thanks to Thomas and Olivier for giving me the opportunity to host a workshop at the first edition of the Deep Learning for Medical Imaging School, that fostered many interactions and ideas. I am grateful to Lauren Pflugrath, Dr. Ron Daigle, Brian Pflugrath, and Bryan Cunitz from Verasonics® for the many technical discussions to specify, customize, and optimize the use of our acquisition system, and to Mike Vega for inviting us to present demos in their exhibit booth at the 2019 IEEE International Ultrasonic Symposium (IUS). Also, many thanks to Dominique Braconnier and Dr. Ewen Carcreff from The Phased Array Company for their interest and exciting discussions. Finally, I would like to thank the Imaging@EPFL initiative for the opportunity to enjoy extended discussions with Prof. Mathias Fink, Prof. Yonina C. Eldar, and Prof. Florian Jug, whose outstanding research work significantly inspired me.

This thesis would never have been possible without my precious colleagues, close collaborators, and former students. I am forever indebted to Dr. Marcel Arditi, with whom I started to collaborate at the LTS5 before I even thought of undertaking doctoral studies. Marcel, you not only convinced me to start this thesis, you were an integral part of it, even if your modesty never allowed you to admit it. You have been by my side at every moment and every step of this thesis, shared with me your immense knowledge, and taught me a scientific rigor I am extremely proud of. Thank you for everything you have done for me during these extraordinary years, and for making me a better scientist. I am also deeply grateful to Dr. Adrien Besson, with whom the “US group” adventure started when he joined the LTS5 as a Ph.D. student. Adrien, working with you for all these years was an incredible opportunity in so many ways: the working pair we formed so quickly was pivotal in starting my doctoral studies; you introduced me to the wonders of inverse problems, convex optimization, and signal processing; and our early work as well as the countless passionate discussions greatly influenced the direction of this thesis. Thank you so much for having been such an amazing colleague and a dear friend, and for all the memorable moments we spent together at multiple conferences around the globe. I want to express my sincere gratitude to two of my former M.Sc. students who became pillars of our group, but also major contributors to this thesis, and most importantly dear friends, namely, Florian Martinez and Manuel Vonlanthen. Florian, I cannot thank you enough for your unwavering efforts and dedication to improving and accelerating crucial algorithms. Your incredible work simply made this thesis possible. Manuel, I am immensely grateful for your countless scientific contributions to this thesis, for your tremendous help in writing many important papers related it, and for your unrelenting support. I truly enjoyed every moment of our work together. To the both of you, I hope

you know that part of this thesis is also yours. Many thanks also to Dr. Rafael E. Carrillo for his initial guidance and help on sparse regularization. I would also like to thank all other former students I had the opportunity to supervise or co-supervise, namely, Antoine, Arthur, Benjamin, Eric, Julie, Jérémie, Louis, Malo, Philippe, Quentin, Saleh, Silas, Yeray, and Yuliang. Special thanks to Arthur, Eric, Jérémie, Malo, Philippe, and Quentin whose help contributed particularly at different stages of my research work. I am also thankful to the newest members of our group, Sandra, Samuel, and Simon, with whom I had a lot of pleasure to collaborate, and to whom I wish the best for the future. I also had the chance to collaborate closely with passionate researchers and wonderful persons from other institutions whom I would like to thank: Dr. Matthieu Simeoni and Lucien Roquette from IBM / LCAV, and Dr. Pascal Hager from ETHZ.

It goes without saying that accomplishing this thesis required much more than academic support, and I am sincerely grateful to every person who helped me through this journey, in one way or another. I am thankful to all the remarkable colleagues I met during my thesis at the EPFL: Anil, Anne, Carlos, Christina, Christophe, Damien, David, Didrik, Elda, Gabriel, Jonathan, Marina, Mina, Muhamed, Murat, Rosie, Saeed, Saleh, Sasan, Tom, Vijay, and so many more from the LTS5; Eric, Lucien, Matthieu, and Sepand from the LCAV; and Apostolos, Evangelos, Kirell, and Michaël and from the SPLabs. I have been extremely lucky to have many close friends on whom I could count at every moment. Thank you Abid, Adrien, Audrey, Alex, Baptiste, Grégoire, Manile, Marc, Patrick, Philippe, Rroni, Sean, and Vanessa for your support and your tremendous efforts in getting me out of the lab. I am especially grateful to Louis, Esther, and their exceptional daughter Saül, for their infinite kindness and for welcoming me several times in their own family cocoon; you have brought so much happiness in my life. To Pauline, writing these lines are particularly emotional. Thank you for the few but extraordinary years spent together, for all the love, support, and encouragement you gave me, and for having made me a better person. You did great sacrifices to support me through difficult and exhausting times, and for that I am forever indebted to you. I also want to thank you for your precious help in typesetting this thesis.

To conclude, I would like to thank my entire family and close relatives from Switzerland, France, and Greece. Despite the distance, I have had the chance to feel and receive tremendous support throughout these years. To my brother, Théodore, having you in my life is a true blessing. Thank you deeply for having always been there for me during these intense years, and for having been so supportive when I needed it the most. To my parents, Thomas and Gabrielle, I am eternally grateful to you for having nurtured my intellectual curiosity ever since my early childhood, and for your unconditional love and support. You have been behind me at every moment and every crucial step, always providing me with the extra push I needed. I owe you a lot for accomplishing this thesis, and I dedicate it to both of you.

Dimitris Perdios
January 2021

Foreword

*“For every complex problem, there is an answer that is
clear, simple, and wrong.”*

H. L. Mencken

When Prof. Jean-Philippe Thiran offered in 2015 that I join his research group at the Signal Processing Laboratory 5 (LTS5), École polytechnique fédérale de Lausanne (EPFL), as a scientific advisor in medical ultrasound, I thought I would share my (dated) understanding of the “state of the art.” Little did I know that, under his guidance, a very small team of young researchers and students would so quickly make contributions to ultrasound image reconstruction capable of revolutionizing medical ultrasound. In less than two years, their initial work attracted the attention of prestigious research groups, and was followed by multiple seminal publications.

Dimitris and his colleagues succeeded in demonstrating that, far from being a mature and sclerotic field of investigation, ultrasound imaging had, on the contrary, enormous possibilities for overcoming its limitations in different directions. Through his inventiveness, curiosity and perseverance, Dimitris revisited the reconstruction of ultrasound images from the ground up, in the framework of inverse problems; on his way, he proposed a highly efficient and novel approach to the calculation of ultrasound fields, he applied in an impressive and impeccable way the principles of deep learning and perfected ultrafast ultrasound by greatly reducing artifacts previously considered as inevitable. He thus also opened up whole swathes of new applications in 2-D and 3-D imaging. Examples could be in 3-D shear-wave elastography, 3-D vector flow mapping, low-cost and high-quality portable systems, etc.

His doctoral thesis is rich in lessons for those who will follow up on his research, and I have no doubt that Dimitris will continue to pollinate other fields of research with his ideas, whether in medical imaging or elsewhere. In any case, it was a real pleasure for me to be a part of this adventure.

Marcel Arditi, PhD
January 2021

Abstract

Among the medical imaging modalities, ultrasound (US) imaging is one of the safest, most widespread, and least expensive method used in medical diagnosis. In the past decades, several technological advances enabled the advent of ultrafast US imaging, an acquisition technique capable of imaging large tissue zones at very high frame rates of multiple kilohertz. Achieving such frame rates on large tissue zones enables the analysis of complex physical phenomena occurring in the human body. For instance, such capability enables estimating both very fast and very slow flows occurring in the cardiovascular system, with high sensitivity. High frame rates also enable estimating micrometer tissue displacements induced by naturally occurring or externally induced shear waves that propagate through tissue at a few meters per second. Ultrafast US imaging is already at the origin of several breakthrough imaging modes such as shear-wave elastography and functional neuroimaging.

One of the main advantages of pulse-echo US imaging is that it is a dynamic imaging modality. Conventional US images reconstructed using the well-known delay-and-sum algorithm are characterized by speckle patterns. Despite being an “illusion” of the imaging system, these patterns react coherently to underlying physical phenomena, thus containing positional information of the tissue being imaged that can be exploited by displacement estimation techniques. Because ultrafast acquisitions are performed using unfocused wavefronts, resulting images are of low quality, characterized by broad main lobes (low resolution) and high diffraction artifacts (low contrast) caused by grating lobes, side lobes, and edge waves. Such artifacts can be detrimental to both lesion detection and displacement estimation techniques, the latter being the core objectives of most ultrafast US imaging modes. A natural way of increasing the image quality of consecutive frames consists of averaging coherently multiple low-quality images obtained from differently steered unfocused wavefronts, at the expense of a reduced frame rate and possible motion artifacts.

This thesis aims at answering the increasing need for US image reconstruction methods capable of producing high-quality images from single ultrafast acquisitions, may it be to improve the accuracy and robustness of ultrafast imaging modes such as shear-wave elastography, to reduce the cost and complexity of 3-D ultrasound scanners, or to mitigate the power and data transfer rate requirements of portable systems. This work builds in the context of inverse problems, with an efficient modeling of the physical measurement process (forward model) involved in ultrasound acquisitions. It leverages recent deep-learning-based projection methods to overcome a crucial limitation of regularized

ultrasound imaging: conventional image-processing regularizers are not well suited to the high dynamic range and statistical properties of radio frequency ultrasound images, especially in the presence of speckle patterns. Physical modeling is fundamental to this work. It was crucial to derive a computationally tractable forward model for image reconstruction, but also to develop a highly efficient, spline-based, spatial impulse response ultrasound simulator that allowed generating sufficiently large datasets to train convolutional neural networks. Applications were carried out in single-plane-wave imaging, ultrafast displacement estimation, and sparse-array imaging.

Keywords Biomedical imaging, B-splines, convolutional neural networks, deep learning, diffraction artifacts, displacement estimation, image reconstruction, inverse problems, physical modeling, speckle tracking, ultrafast ultrasound imaging.

Résumé

De toutes les modalités d'imagerie médicale, l'échographie est parmi les plus sûres, les plus répandues et les moins onéreuses utilisées pour le diagnostic médical. Dans les dernières décennies, plusieurs avancées technologiques ont permis l'avènement de l'échographie ultrarapide, une technique d'acquisition permettant d'imager de grandes régions tissulaires à des cadences d'imagerie de plusieurs kilohertz. En atteignant de telles cadences sur de larges zones, cela a ouvert la voie à l'analyse de phénomènes physiques complexes dans le corps humain. Par exemple, cette capacité permet l'estimation de flux très rapides ou très lents présents dans le système cardiovasculaire, avec une haute sensibilité. Ces cadences élevées permettent également d'estimer des déplacements tissulaires micrométriques induits par des ondes de cisaillement survenant naturellement ou induites extérieurement, et se propageant dans les tissus à quelques mètres par seconde. L'échographie ultrarapide est déjà à l'origine de plusieurs percées en imagerie, telles que l'élastographie par ondes de cisaillement ou la neuroimagerie fonctionnelle.

L'un des principaux avantages de l'échographie est qu'il s'agit d'une modalité d'imagerie dynamique, dite « en temps réel ». Les images d'échographie reconstruites par la technique conventionnelle du « delay-and-sum » se caractérisent par la présence de « speckle » (motifs tachetés). Bien que ces motifs ne soient qu'apparents, ils réagissent de façon cohérente aux phénomènes physiques sous-jacents, portant ainsi une information utile sur la position relative du tissu visualisé, qui peut être exploitée par des techniques d'estimation de déplacements. Du fait que les acquisitions ultrarapides sont réalisées à l'aide de fronts d'ondes non-focalisés, elles produisent en général des images de mauvaise qualité, en raison de lobes principaux larges (faible résolution latérale), et entachées d'artéfacts de diffraction (faible contraste), causés par des franges d'interférences, les lobes latéraux et les ondes de bord. Ces artéfacts portent préjudice aussi bien à la détection de lésions qu'aux techniques de mesure de déplacement, ces dernières étant au cœur des objectifs de la plupart des méthodes d'échographie ultrarapide. Une approche naturelle pour améliorer la qualité d'images successives consiste à moyenner de façon cohérente une multitude d'images de faible qualité, obtenues à partir de fronts d'ondes d'angles différents, au détriment d'une cadence d'imagerie réduite et de possibles artéfacts de mouvement.

L'objectif de cette thèse est d'apporter une réponse au besoin grandissant de méthodes de reconstruction d'images, capable de produire des images de haute qualité à partir d'acquisitions ultrarapides individuelles, que ce soit pour améliorer la précision et la robustesse de modes d'imagerie ultrarapide comme l'élastographie, pour réduire le coût

et la complexité de systèmes 3D, ou pour réduire les exigences en puissance et taux de transferts de systèmes portables. Ce travail s'est construit dans le contexte des problèmes inverses, avec une modélisation efficace du processus de mesure physique (modèle « forward ») impliqué dans l'acquisition par ultrasons. Il exploite des méthodes récentes de projection basées sur l'apprentissage profond afin de surmonter une limitation cruciale des techniques de régularisation en échographie : les régularisateurs conventionnels de traitement d'images ne se prêtent pas bien à la très grande dynamique et aux propriétés statistiques des images échographiques radiofréquences, particulièrement en présence de « speckle ». La modélisation physique a ainsi été fondamentale dans ce travail. Il a été crucial afin de dériver un modèle « forward » pour la reconstruction, se prêtant à une implémentation numérique pour la reconstruction d'images, mais aussi de développer un simulateur hautement efficace de réponse impulsionnelle spatiale, basé sur des B-splines, ayant ainsi permis de générer des jeux de données d'apprentissage de très grande ampleur.

Mots clés Imagerie biomédicale, B-splines, réseaux de neurones convolutifs, apprentissage profond, artéfacts de diffraction, estimation de déplacements, reconstruction d'images, problèmes inverses, modélisation physique, traçage de speckle, échographie ultrarapide.

Contents

Acknowledgments	v
Foreword	ix
Abstract	xi
Résumé	xiii
Contents	xvii
List of Figures	xix
List of Tables	xxix
List of Abbreviations	xxxix
1. Introduction	1
1.1. Context and Motivation	1
1.2. Contributions and Organization of this Thesis	4
2. A Linear Pulse-Echo Model for Weak Scattering	7
2.1. Introduction	7
2.2. Scalar Wave Equation for Weak Scattering	8
2.3. Generic Solutions of the Scalar Wave Equation	10
2.3.1. Green's Functions for the Scalar Wave Equation	10
2.3.2. Typical Boundary Conditions	11
2.3.3. Homogeneous Wave Equation with Inhomogeneous Boundary Conditions	12
2.3.4. Inhomogeneous Wave Equation with Homogeneous Boundary Conditions	13
2.4. Specifying Green's Function	14
2.4.1. Single Volume Source in an Unbounded Domain	14
2.4.2. Infinite Planar Boundaries	16
2.4.3. Time-Dependent Rayleigh-Sommerfeld Equations	17
2.4.4. Corresponding Spatial Impulse Response Formulations	20
2.5. The Pulse-Echo Model for Weak Scattering	22
2.5.1. Born Approximation	23
2.5.2. Specifying Green's Function	24
2.5.3. Transmit Field	24
2.5.4. Receive Field	25
2.5.5. Generic Pulse-Echo Model	30
2.6. Conclusion	31

CONTENTS

3. Efficient Modeling for Ultrafast Ultrasound Imaging	33
3.1. Introduction	33
3.2. Proposed Methods	34
3.2.1. Generic Approximated Pulse-Echo Model for Imaging	34
3.2.2. Inverse Problem and Imaging Methods	36
3.2.3. Implementation and Interpretation Notes on the Approximated Model	38
3.3. Experiments	40
3.4. Results	41
3.5. Discussion	42
3.5.1. Generic Approximated Pulse-Echo Model	42
3.5.2. Sparse Regularization for Enhanced Ultrafast Imaging	43
3.6. Conclusion	43
4. A Spline-Based Spatial Impulse Response Simulator	45
4.1. Introduction	45
4.2. Proposed Approach	48
4.2.1. Numerical Quadrature of the Spatial Impulse Response	49
4.2.2. Non-Uniform Rational B-Spline Surface Representations	51
4.2.3. Spatial Impulse Response in B-Spline Bases	53
4.2.4. Implementation Details	58
4.2.5. Extension to Arrays	59
4.3. Experiments and Results	61
4.3.1. Convergence Order of Various Basis Functions	62
4.3.2. Validation Against Analytic Solutions	63
4.4. Discussion	66
4.4.1. Benefit of High-Order Basis Functions	68
4.4.2. Non-Uniform Rational B-Spline Representation of Surfaces	69
4.4.3. Gaussian Quadrature Rules	69
4.4.4. Comparison to Other Strategies	69
4.4.5. Hypothetical Perspective for Analytic Expressions	70
4.5. Conclusion	71
5. CNN-Based Image Reconstruction Method for Ultrafast Ultrasound Imaging	73
5.1. Introduction	73
5.2. Background	74
5.2.1. Pulse-Echo Modeling and Imaging	74
5.2.2. Statistical Considerations of Fully Developed Speckle	76
5.3. Methods	79
5.3.1. Proposed Image Reconstruction Method	79
5.3.2. Convolutional Neural Network Architecture	80
5.3.3. Training Loss for High-Dynamic-Range Radio-Frequency Data	82
5.4. Experimental Setup	84
5.4.1. Imaging Configurations	84
5.4.2. Element Raw-Data Generation	87
5.4.3. Simulation Phantoms	88

5.4.4.	Simulated Dataset	89
5.5.	Training and Hyperparameter Search	90
5.5.1.	Training and Validation Setup	90
5.5.2.	Hyperparameter Search	91
5.5.3.	Summary	98
5.6.	Experiments and Results	99
5.6.1.	Numerical Test Phantom	99
5.6.2.	Experimental Evaluations	104
5.7.	Discussion	107
5.7.1.	Performance in Ideal Conditions	107
5.7.2.	Performance in Experimental Conditions	108
5.7.3.	Potential Improvements	109
5.7.4.	Application Perspectives	111
5.8.	Conclusion	111
6.	CNN-Based Ultrasound Image Reconstruction for Ultrafast Displacement Tracking	113
6.1.	Introduction	113
6.2.	Materials and Methods	114
6.2.1.	Imaging Configurations	114
6.2.2.	CNN-Based Image Reconstruction Method	115
6.2.3.	Comparative Image Reconstruction Methods	116
6.2.4.	Speckle Tracking Algorithm	118
6.2.5.	Metrics	119
6.3.	Experiments and Results	120
6.3.1.	Numerical Experiment	120
6.3.2.	<i>In Vivo</i> Experiment	124
6.4.	Discussion	126
6.4.1.	Performance in Numerical Conditions	127
6.4.2.	Performance in Physical Conditions	128
6.4.3.	Potential, Perspectives, and Limitations	129
6.5.	Conclusion	130
7.	CNN-Based Ultrafast Ultrasound Imaging with Sparse Linear Arrays	131
7.1.	Introduction	131
7.2.	Materials and Methods	132
7.2.1.	Imaging Configurations	132
7.2.2.	CNN-Based Image Reconstruction Method	133
7.2.3.	Numerical Experiment	134
7.3.	Results and Discussion	135
7.4.	Conclusion	137
8.	Concluding Remarks	139
A.	Supplementary Material to Chapter 4	143
	References	153

List of Figures

- 2.1. Generic representation of the Green's function method to determine the velocity potential at an observer point due to the presence of a source distribution embedded in a homogeneous volume enclosed by a boundary surface onto which boundary conditions are specified. 10
- 2.2. Method of images to specify suitable Green's functions when considering an infinite planar boundary: (a) both the observer and the source are within the volume; (b) the observer is on the boundary plane and the source within the volume; (c) the observer is within the volume and the observer on the boundary plane; (d) both the observer and the source are on the boundary plane. 16
- 2.3. Schematic representation of a pulse-echo experiment in the presence of a weakly scattering zone for the derivation of the linear pulse-echo model. The scattering zone is considered to be embedded in a homogeneous medium and is represented by a source distribution. The pulse-echo experiment involves a transmitter that insonifies the volume including the weakly scattering zone and a receiver that senses the echo signals backscattered by the weakly scattering zone. Both the transmitter and the receiver are assumed to be embedded in an infinite planar boundary onto which boundary conditions are specified. 23
- 3.1. Illustration of a 2-D plane-wave (PW) imaging configuration. The PW wavefront is generated by applying suitable delays (and potential apodization) to the transducer elements composing the array. Backscattered echoes (here from a single scatterer) are received on all transducer elements (i.e., full aperture). (Adapted from [25], [30].) 36
- 3.2. Two-dimensional interpretation of the approximated measurement model as the curvilinear integral over a 1-D conic for a specific time instant of the signal acquired on the j -th receiver from the i -th transmit wavefront (i.e., transmitter). The time instant is represented by a filled circle in the corresponding signal. The corresponding 1-D conic is a parabola in the case of 2-D plane wave (PW) imaging. (Adapted from [25], [30].) 39
- 3.3. B-mode image representations (60-dB range) of the numerical plane-wave imaging challenge in medical ultrasound (PICMUS) phantom reconstructed using (from left to right) the proposed backprojection-based delay-and-sum (DAS) operator from a single plane wave (PW) acquisition, the proposed sparse regularization (SR) approach from a single PW acquisition, and the proposed backprojection-based DAS operator from 75 PW acquisitions (considered as reference). (Modified from [30].) 41

LIST OF FIGURES

- 3.4. B-mode image representations (60-dB range) of an *in vivo* carotid (longitudinal view) reconstructed using (from left to right) the proposed backprojection-based delay-and-sum (DAS) operator from a single plane wave (PW) acquisition, the proposed sparse regularization (SR) approach from a single PW acquisition, and the proposed backprojection-based DAS operator from 75 PW acquisitions (considered as reference). (Modified from [30].) 42
- 4.1. Example of (a) a spatial impulse response (SIR) and (b) corresponding frequency spectrum for a rectangular transducer aperture. The rectangle has a width of one wavelength at 5 MHz and a height-to-width ratio of 1.6. The field point at which the SIR was evaluated is positioned relatively to the element center at $(\lambda/5, \lambda/5, 2.5\lambda)$ in the width, height, and depth directions. Characteristic abrupt slope changes are circled in (a). The first one represents the first time of arrival of a spherical shell centered at the position of the evaluation field point. The eight following ones represent time instants at which the spherical shell intersects the edges and vertices of the rectangular surface. 48
- 4.2. Typical examples of B-spline basis functions of different degrees defined on uniform knot vectors with no interior knots: (a) linear basis functions defined on the knot vector $(0, 0, 1, 1)$; (b) quadratic basis functions defined on the knot vector $(0, 0, 0, 1, 1, 1)$; (c) cubic basis functions defined on the knot vector $(0, 0, 0, 0, 1, 1, 1, 1)$. 51
- 4.3. Example of a non-uniform rational B-spline (NURBS) surface representing a spherical cap, decomposed into four rational biquadratic Bézier smooth patches. The NURBS mapping from the parametric space to the physical space is illustrated with a parent element comprising (5×3) Gauss-Legendre quadrature points. The two (unit) tangent vectors and the (unit) normal vector at each quadrature point in the physical space are depicted by green, orange and red arrows, respectively. 53
- 4.4. Example of a non-uniform rational B-spline (NURBS) surface representing a cylindrical shell that is also a smooth rational Bézier of degree $(1, 2)$, that is, linear in u and quadratic in v directions. The NURBS mapping from the parametric space to the physical space is illustrated with a parent element comprising (3×5) Gauss-Legendre quadrature points. The two (unit) tangent vectors and the (unit) normal vector at each quadrature point in the physical space are depicted by green, orange and red arrows, respectively. 54
- 4.5. Illustration of the generalized interpolation strategy deployed for approximating the signal radiated by a transducer element. This strategy involves the convolution of coefficients, evaluated from the excitation waveform, with shifted basis functions. Three interpolating basis functions are depicted, namely nearest-neighbor, linear, and (quadratic) Keys. One non-interpolating basis function is also shown, namely cubic B-spline (B-spline3). 58
- 4.6. Illustration of all steps composing the proposed approach for the approximation of field signals using the spatial impulse response (SIR) of radiating surfaces. (Please refer to the associated text for a summary of each step involved.) 59
- 4.7. Normalized amplitude of (a) the excitation waveform and (b) corresponding frequency spectrum of the pulse model considered for the numerical experiments. 61

- 4.8. Results of the numerical experiment performed to validate the theoretical orders of convergence for different basis functions. Different types of basis functions were considered: three interpolating basis functions, namely nearest-neighbor (degree zero), linear (degree one), and quadratic Keys (degree two); four non-interpolating B-spline basis functions of degree 2, 3, 4, and 5; and the non-interpolating optimal-maximal-order-minimal-support (O-MOMS) basis function of degree three. The theoretical convergence order of a basis function is equal to its degree plus one. 63
- 4.9. Comparison of the spatial impulse responses (SIRs) and field signals radiated at different field points by a spherically focused transducer element with a rigid baffle condition, excited by a windowed-sinusoidal waveform. The excitation waveform is a differentiated log-normal-windowed sine wave, with a characteristic (center) wavelength λ . The geometry of the spherical cap is defined by an active diameter of 20λ and a spherical radius of 240λ . The three field points (A, B, C) lie in the same revolution plane at a depth of 10λ and a lateral coordinate of 0, 8.1λ (i.e., projection on the edge of the surface), and 16.2λ , respectively. The proposed approach was implemented with a B-spline basis function of degree five, and was evaluated at two sampling rates of (first row) 30 MHz and (second row) 80 MHz. The reference SIRs and field signals were evaluated at a sampling rate of 20 THz. They are depicted with the same colors as the approximated counterparts, with a lower opacity. 65
- 4.10. Comparison of the spatial impulse responses (SIRs) and field signals radiated at different field points by a rectangular transducer element with a rigid baffle condition, excited by a windowed-sinusoidal waveform. The excitation waveform is a differentiated log-normal-windowed sine wave, with a characteristic (center) wavelength λ . The geometry of the rectangular element is defined by a width of λ and a height of 10λ . The three field points (A, B, C) lie in a plane parallel to the element. They were positioned at a depth of $\lambda/2$ and an elevation of $\lambda/2$, with lateral coordinates of 0, $\lambda/2$, and λ , respectively. The proposed approach was implemented with a B-spline basis function of degree five, and was evaluated at two sampling rates of (first row) 30 MHz and (second row) 80 MHz. The reference SIRs and field signals were evaluated at a sampling rate of 20 THz. They are depicted with the same colors as the approximated counterparts, with a lower opacity. 67
- 4.11. Comparison of the spatial impulse responses (SIRs) and field signals radiated at different field points by a rectangular transducer element with a soft baffle condition, excited by a windowed-sinusoidal waveform. The remaining settings relevant to the results depicted are identical to those detailed in Figure 4.10. 68
- 5.1. Lower and upper confidence bounds (in decibels) with respect to the confidence level (in percent) of a Rayleigh random variable normalized by its expected value (i.e., mean). 79

LIST OF FIGURES

- 5.2. Proposed residual convolutional neural network (CNN) architecture, adapted from U-Net [145], and from [31], [137]: (a) generic overall CNN architecture; convolutional blocks considered, namely (b) conventional fully convolutional block (FCB) and (c) proposed residual convolutional block (RCB). Connections and tensor operations (i.e., layers) are represented as straight, colored arrows (legend in the upper right corner). The 3-D tensors are represented as colorized rectangles. (Their width and height are depicted in proportion to the number of channels and image dimension, respectively. One image dimension is not represented for readability reasons.) 81
- 5.3. B-mode image representations (98-dB range) of simulated point spread function (PSF) examples: (a) point reflector positions in which the PSFs were evaluated; images reconstructed using each imaging configuration considered (Table 5.1), namely (b) low-quality (LQ) configuration, (c) high-quality (HQ) configuration (i.e., gold-standard image for the physical transducer array), and (d) ultra-high-quality (UQ) configuration (i.e., gold-standard image for the spatially oversampled virtual version of the transducer array, considered as ground-truth). Some zones dominated by grating lobe (GL), side lobe (SL), and edge wave (EW) artifacts are highlighted by colorized arrows and associated annotations. 87
- 5.4. Representation (in the transducer plane) of the spatial domains used for generating the simulation phantoms for the imaging configurations considered (Table 5.1). 88
- 5.5. B-mode image representations (98-dB range) of a simulated dataset sample: (a) the phantom mask composed of elliptical inclusions; images reconstructed using each imaging configuration considered (Table 5.1), namely (b) low-quality (LQ) configuration, (c) high-quality (HQ) configuration (i.e., gold-standard image for the physical transducer array), and (d) ultra-high-quality (UQ) configuration (i.e., gold-standard image for the spatially oversampled virtual version of the transducer array, considered as ground-truth). Some zones dominated by grating lobe (GL), side lobe (SL), and edge wave (EW) artifacts are highlighted by colorized arrows and associated annotations. 90
- 5.6. Validation metric curves (SSIM evaluated on B-mode representations) of training experiments performed using different image representations (i.e., RF, IQ, envelope, and B-mode) and different reference images (i.e., HQ and UQ). All training experiments were performed on identical instances of the proposed residual CNN (Figure 5.2) with 16 initial expansion channels, RCBs, and additive intrinsic skip connections. 92
- 5.7. B-mode image representations (98-dB range) of a numerical test phantom sample (extracted from the simulated dataset): (a) the phantom mask; images reconstructed using each imaging configuration considered (Table 5.1), namely (b) low-quality (LQ), (c) high-quality (HQ) (i.e., gold-standard image for the physical transducer array), and (d) ultra-high-quality (UQ) (i.e., reference image); images recovered from the LQ input image using the proposed approach with different convolutional neural networks (CNNs), deployed with 16 initial expansion channels, residual convolutional blocks (RCBs), and additive intrinsic skip connections, trained on different image representations and image references, namely (e) UQ + B-mode + mean absolute error (MAE), (f) UQ + envelope + mean signed logarithmic absolute error (MSLAE), (g) HQ + in-phase quadrature (IQ) + MSLAE, and (h) UQ + IQ + MSLAE. 93

- 5.8. Validation metric curves (SSIM evaluated on B-mode representations) of training experiments performed using different training losses, namely mean squared error (MSE), mean absolute error (MAE), and mean signed logarithmic absolute error (MSLAE). All training experiments were performed on identical instances of the proposed residual CNN (Figure 5.2) with 16 initial expansion channels, RCBs, and additive intrinsic skip connections. 94
- 5.9. Validation metric curves (SSIM evaluated on B-mode representations) of training experiments performed using different combinations of initial channel expansion numbers (i.e., 8, 16, and 32), convolutional blocks (i.e., RCBs and FCBs), and intrinsic skip connections (i.e., additive and concatenated). Each training experiment was performed using the MSLAE as loss function. 95
- 5.10. Training and validation loss curves of training experiments performed using different training set sizes. All training experiments were performed on identical instances of the proposed residual CNN (Figure 5.2) with 16 initial expansion channels, RCBs, and additive intrinsic skip connections, using the MSLAE as loss function. 97
- 5.11. Validation metric curves (SSIM evaluated on B-mode representations) of training experiments performed using different kernel initializers (i.e., Glorot uniform, Glorot normal, He uniform, and He normal). Each training experiment was performed on the proposed residual CNN (Figure 5.2) with 16 initial expansion channels, RCBs, and additive intrinsic skip connections, using the MSLAE as loss function. 98
- 5.12. Validation metric curves (SSIM evaluated on B-mode representations) of training experiments performed using different learning rates (i.e., 1×10^{-5} , 5×10^{-5} , 1×10^{-4} , 5×10^{-4} , and 1×10^{-3}). Each training experiment was performed on the proposed residual CNN (Figure 5.2) with 16 initial expansion channels, RCBs, and additive intrinsic skip connections, using the MSLAE as loss function. 99
- 5.13. B-mode image representations (98-dB range) of a numerical test phantom sample: (a) the phantom mask and annotated zones in which the metrics were evaluated; images reconstructed using each imaging configuration considered (Table 5.1), namely (b) low-quality (LQ) configuration, (c) high-quality (HQ) configuration (i.e., gold-standard image for the physical transducer array), and (d) ultra-high-quality (UQ) configuration (i.e., reference image); images recovered from the low-quality (LQ) input image using the proposed approach with each of the trained convolutional neural networks (CNNs) considered (Section 5.6.1.1), namely (e) MSE-16, (f) MAE-16, (g) MSLAE-16, and (h) MSLAE-32. 102
- 5.14. Mean amplitude responses (averaged along the axial dimension) of the horizontal gradient zone in the numerical phantom [Ω_{LG} in Figure 5.13(a)], averaged over 300 independent realizations. 104

LIST OF FIGURES

- 5.15. B-mode image representations (98-dB range) of the incoherent average (performed after envelope detection) of all images reconstructed from the 300 independent realizations (random scatterers) of the numerical test phantom: (a) the phantom mask; images reconstructed using each imaging configuration considered (Table 5.1), namely (b) low-quality (LQ) configuration, (c) high-quality (HQ) configuration (i.e., gold-standard image for the physical transducer array), and (d) ultra-high-quality (UQ) configuration (i.e., reference image); images recovered from the low-quality (LQ) input image using the proposed approach with each of the trained convolutional neural networks (CNNs) considered (Section 5.6.1.1), namely (e) MSE-16, (f) MAE-16, (g) MSLAE-16, and (h) MSLAE-32. Some remaining side lobe (SL) and edge wave (EW) artifacts are highlighted by colored arrows and associated annotations. 105
- 5.16. B-mode image representations of an *in vitro* test phantom (top row, 78-dB range) and an *in vivo* carotid sample (bottom row, 50-dB range): (a) and (d) single plane wave (PW) low-quality (LQ) images; (b) and (e) images recovered from LQ using the proposed convolutional neural network (CNN)-based image reconstruction method with the selected trained CNN (i.e., MSLAE-16); (c) and (f) reference high-quality (HQ) images reconstructed from the complete set of synthetic aperture (SA) acquisitions. 106
- 6.1. Sketch of the acquisition schemes deployed for the different plane wave (PW) imaging configurations considered. To form a single frame, a sequence of echo-signals from differently steered PWs is acquired at a pulse repetition frequency (PRF) of 9 kHz. The number of PWs composing each acquisition sequence depends on the imaging configuration (Table 6.2). The resulting frame rate is determined by the time interval between consecutive acquisition sequences, and is limited by the duration of a single acquisition sequence. 118
- 6.2. B-mode image representations (80-dB range) of a numerical test phantom sample: (a) the 2-D geometry of the deployed numerical phantoms, composed of four cylinders (A, B, C, and D) filled with dense point-scatterers rotating at constant angular velocity around their respective cylinder center; (b) image reconstructed by delay-and-sum (DAS) beamforming a single plane-wave (PW) acquisition (CPWC-1), simultaneously representing the convolutional neural network (CNN) input image for the proposed method; (c) image reconstructed using CNN-based reconstruction; images reconstructed by coherent plane wave compounding (CPWC) using nine steered PW acquisitions (CPWC-9): (d) small displacement range and (e) large displacement range. The frame rate and displacement range for each image reconstruction method considered are given in Table 6.3. 121

- 6.3. Local relative endpoint error (REPE), averaged over 50 independent realizations, of the 2-D displacement estimates inside each of the numerical phantom zones [A, B, C, and D in Figure 6.2(a)], obtained by applying the deployed 2-D speckle tracking algorithm (Section 6.2.4) on two consecutive frames for the two inter-frame displacement ranges considered: (a) large displacement range (from 33 μm to 600 μm); (b) small displacement range (from 3.3 μm to 60 μm). Consecutive frames were reconstructed either by coherent plane wave compounding (CPWC) from 1, 3, 9, and 15 differently steered PWs, or using the proposed convolutional neural network (CNN)-based image reconstruction method from single PWs. The frame rate and displacement range for each image reconstruction method considered are given in Table 6.3. The displayed REPE range is limited to 100 %. Local REPE values were interpolated onto a fine grid for display purposes. 123
- 6.4. Examples of displacement estimates, mean reference displacement magnitude, and mean relative endpoint error (MREPE), obtained using the displacement estimation methods considered, in a fully developed speckle zone above the carotid artery: images of a longitudinal view of the carotid artery, are shown for (a) CPWC-1 (also CNN input), (b) CPWC-15, (c) CNN, and (d) CPWC-87 (reference); the bottom row shows (e) the mean reference displacement magnitude and (f) the MREPE along the entire *in vivo* sequence for each method considered. In each B-mode image of the top row, the square region of interest is highlighted and the corresponding magnified inset displays the 2-D displacement estimates. B-mode images are displayed using a dynamic range of 50 dB. The mean value (through time) of each quantity represented by a colorized solid line in (e) and (f) is represented by a horizontal dashed line of the same color. These mean values were computed ignoring samples at 2.2 s due to the resulting extreme MREPE values. An animation of the figure and the corresponding slideshow are provided as supplementary material to [34]. 126
- 7.1. B-mode image representations (98-dB range) of a numerical test phantom sample, reconstructed using each imaging configuration considered (Table 7.2): (a) low-quality (LQ) configuration (i.e., full array on receive) and (f) ultra-high-quality (UQ) configuration (i.e., reference image), providing image quality comparisons; (b) to (e) U2, U3, U4, and U6 uniformly undersampled configurations (by a factor of 2, 3, 4, and 6 on receive only); (g) to (j) images recovered from U2 U3, U4, and U6 using the proposed approach with the corresponding trained convolutional neural network (CNN). 135
- 7.2. Mean amplitude responses (averaged along the axial dimension) of the lateral gradient zone in the numerical phantom [Figure 7.1(f)], averaged over 300 independent realizations. 136

LIST OF FIGURES

- A.1. Comparison of the spatial impulse responses (SIRs) and field signals radiated at different field points by a spherically focused transducer element with a rigid baffle condition, excited by a windowed-sinusoidal waveform. The excitation waveform is a differentiated log-normal-windowed sine wave, with a characteristic (center) wavelength λ . The geometry of the spherical cap is defined by an active diameter of 20λ and a spherical radius of 240λ . The three field points (A, B, C) lie in the same revolution plane at a depth of 10λ and a lateral coordinate of 0 , 8.1λ (i.e., projection on the edge of the surface), and 16.2λ , respectively. The proposed approach was implemented with a nearest-neighbor basis function (degree zero), and was evaluated at two sampling rates of (first row) 30 MHz and (second row) 80 MHz. The reference SIRs and field signals were evaluated at a sampling rate of 20 THz. They are depicted with the same colors as the approximated counterparts, with a lower opacity. 143
- A.2. *Ibid.* for a linear basis function (degree one). 144
- A.3. *Ibid.* for a quadratic Keys basis function. 144
- A.4. *Ibid.* for a cubic B-spline basis function. 145
- A.5. *Ibid.* for a cubic O-MOMS basis function. 145
- A.6. Comparison of the spatial impulse responses (SIRs) and field signals radiated at different field points by a rectangular transducer element with a rigid baffle condition, excited by a windowed-sinusoidal waveform. The excitation waveform is a differentiated log-normal-windowed sine wave, with a characteristic (center) wavelength λ . The geometry of the rectangular element is defined by a width of λ and a height of 10λ . The three field points (A, B, C) lie in a plane parallel to the element. They were positioned at a depth of $\lambda/2$ and an elevation of $\lambda/2$, with lateral coordinates of 0 , $\lambda/2$, and λ , respectively. The proposed approach was implemented with a nearest-neighbor basis function (degree zero), and was evaluated at two sampling rates of (first row) 30 MHz and (second row) 80 MHz. The reference SIRs and field signals were evaluated at a sampling rate of 20 THz. They are depicted with the same colors as the approximated counterparts, with a lower opacity. 146
- A.7. *Ibid.* for a linear basis function (degree one). 147
- A.8. *Ibid.* for a quadratic Keys basis function. 147
- A.9. *Ibid.* for a cubic B-spline basis function. 148
- A.10. *Ibid.* for a cubic O-MOMS basis function. 148
- A.11. Comparison of the spatial impulse responses (SIRs) and field signals radiated at different field points by a rectangular transducer element with a soft baffle condition, excited by a windowed-sinusoidal waveform. The remaining settings relevant to the results depicted are identical to those detailed in Figure A.6. The proposed approach was implemented with a nearest-neighbor basis function (degree zero), and was evaluated at two sampling rates of (first row) 30 MHz and (second row) 80 MHz. The reference SIRs and field signals were evaluated at a sampling rate of 20 THz. They are depicted with the same colors as the approximated counterparts, with a lower opacity. 149
- A.12. *Ibid.* for a linear basis function (degree one). 150
- A.13. *Ibid.* for a quadratic Keys basis function. 150
- A.14. *Ibid.* for a cubic B-spline basis function. 151

A.15. <i>Ibid.</i> for a cubic O-MOMS basis function.	151
---	-----

List of Tables

3.1.	Image Quality Metrics Computed on the PICMUS Numerical Phantom	42
4.1.	Relative Two-Norm Errors of Field Signals Radiated by a Spherically Focused Transducer Element with a Rigid Baffle Condition	65
4.2.	Relative Two-Norm Errors of Field Signals Radiated by a Rectangular Transducer Element with a Rigid Baffle Condition	66
4.3.	Relative Two-Norm Errors of Field Signals Radiated by a Rectangular Transducer Element with a Soft Baffle Condition	67
5.1.	Specifications of the Imaging Configurations Considered	85
5.2.	Average Inference Time for Different Image Representations and Initial Channel Expansion Numbers	96
5.3.	Numerical Test Phantom Metrics	103
5.4.	Experimental Test Phantom Metrics	107
6.1.	Specifications of the Imaging Configurations Considered	115
6.2.	Plane Wave Imaging Acquisition Sequences Considered	117
6.3.	Displacement and Velocity Ranges Considered for the Numerical Experiment	122
6.4.	Global Evaluation Metrics of the Numerical Experiment	123
7.1.	Acquisition System Specifications	133
7.2.	Transmit and Receive Array Configurations	133
7.3.	Numerical Test Phantom Metrics	136

List of Abbreviations

ACF	autocovariance function
CAD	computer-aided design
CDF	cumulative distribution function
CL	convolutional layer
CNN	convolutional neural network
CNR	contrast-to-noise ratio
CPU	central processing unit
CPWC	coherent plane wave compounding
CT	computed tomography
DAS	delay-and-sum
DC	direct current
DRA	dynamic range alteration
DW	diverging wave
EPE	endpoint error
EW	edge wave
FCB	fully convolutional block
FFT	fast Fourier transform
FISTA	fast iterative shrinkage-thresholding algorithm
FWHM	full width at half maximum
GL	grating lobe
GPU	graphics processing unit
HDR	high dynamic range
HQ	high-quality
IEEE	Institute of Electrical and Electronics Engineers
IGA	isogeometric analysis
IIR	infinite impulse response
IQ	in-phase quadrature

LIST OF ABBREVIATIONS

IUS	International Ultrasonic Symposium
LQ	low-quality
MAE	mean absolute error
MMUAE	mean μ -law absolute error
MOMS	maximal-order-minimal-support
MREPE	mean relative endpoint error
MRI	magnetic resonance imaging
MSE	mean squared error
MSLAE	mean signed logarithmic absolute error
NURBS	non-uniform rational B-spline
O-MOMS	optimal-maximal-order-minimal-support
PCC	Pearson correlation coefficient
PDF	probability density function
PICMUS	plane-wave imaging challenge in medical ultrasound
PIV	particle image velocimetry
PRF	pulse repetition frequency
PSF	point spread function
PSNR	peak signal-to-noise ratio
PW	plane wave
RCB	residual convolutional block
ReLU	rectified linear unit
REPE	relative endpoint error
RF	radio frequency
RVE	ratio of valid estimates
SA	synthetic aperture
SA-VFI	synthetic aperture vector flow imaging
SIR	spatial impulse response
SL	side lobe
SNR	signal-to-noise ratio
SR	sparse regularization
SSIM	structural similarity
TGC	time gain compensation
UQ	ultra-high-quality
US	ultrasound

LIST OF ABBREVIATIONS

USSR	ultrasound sparse regularization framework
WSS	wide-sense stationary
ZNCC	zero-normalized cross-correlation

1. Introduction

1.1. Context and Motivation

Ultrasound (US) imaging is one of the most widely used medical imaging modalities, thanks to being non-ionizing and having a greater cost-effectiveness and portability compared with X-ray computed tomography (CT) or magnetic resonance imaging (MRI). While US imaging might be limited in its ability to image deep tissue compared with X-ray CT or MRI, it can produce high-resolution images of soft tissue in real time. Recently, US imaging has been revolutionized by the development of ultrafast US imaging, an acquisition technique enabling frame rates of multiple kilohertz, thereby unlocking the analysis of the most complex physical phenomena occurring in the human body.

Pulse-echo US imaging is typically performed by transmitting short acoustic pulses through a medium of interest, namely soft biological tissue in the context of medical US imaging. These short acoustic pulses are generated using an US transducer (also referred to as an US probe), composed of multiple transducer elements arranged as an array. Such (bidirectional) transducer elements can convert electric signals into acoustic pressure and *vice versa*. While propagating through (inhomogeneous) tissue, acoustic pulses generated by the US transducer are being scattered in all directions by local variations in acoustic impedance. Some of these echoes propagate back to the face of the transducer (backscattered echoes) and are sensed (received) by each transducer element as radio frequency (RF) electric signals. From these signals (measurements), corresponding RF images are conventionally reconstructed using the well-known delay-and-sum (DAS) algorithm. Because amplitude backscattering coefficients are sensitive to slight variations in acoustic impedance, backscattered echoes, and consequently RF images, span a high dynamic range (HDR) of several decibels that allow distinguishing between different types of tissue. As both RF and HDR properties of US images complicate their interpretation, gray-scale B-mode images are usually displayed. These B-mode images are obtained from RF images by envelope detection and compression. Thanks to an acoustic wave velocity (speed of sound) of about 1500 m/s in soft tissue, as well as high-quality acquisition components and efficient implementations of DAS algorithms, interpretable B-mode images can be displayed in real time.

To date, most commercial US scanners rely on an acquisition process called line-by-line scanning to reconstruct full-view images. This process, as indicated by its name, consists of performing sequential pulse-echo acquisitions using focused transmit beams, suitably steered for the different image scan lines. To obtain high-quality images, several tens or

1. INTRODUCTION

few hundreds of scan lines are conventionally used in 2-D imaging, enabling frame rates of about 10 Hz to 30 Hz that are compatible for real-time visualization. In the last two decades, ultrafast US imaging has drawn a lot of interest in the research community for its ability to increase the achievable frame rate by orders of magnitude [1]. Ultrafast US imaging relies on the insonification of the entire field of view at once by transmitting an unfocused wavefront, such as a plane wave (PW) or a diverging wave (DW), thus breaking with the trade-off between field of view and frame rate inherent to conventional transmit-focused line-by-line scanning. This strategy, introduced in the late 1970s [2]–[4], enables imaging large tissue regions at very high frame rates of multiple kilohertz [1], limited only by the round-trip propagation time of single acoustic waves.

Imaging large tissue regions at such high frame rates is necessary for studying the most rapidly changing physical phenomena in the human body, such as tracking the propagation of naturally occurring or externally induced shear waves [5]–[9]. In the cardiovascular system, where a frame rate of several hundred hertz is needed for resolving tissue motion and flow patterns accurately [10]–[13], ultrafast imaging enables increased ensemble lengths, improving the robustness and sensitivity of displacement estimates significantly [13]. Coupled with advances in electronics and software-based DAS beam-forming, ultrafast US imaging unlocked several breakthrough US imaging modes such as shear-wave elastography [5], high-frame-rate vector flow imaging [14], ultrasensitive flow imaging [13], and functional US neuroimaging [15]. All of these imaging modes rely on motion estimation within a large field of view, performed on frames acquired at ultrafast rates.

The main disadvantage of ultrafast US imaging compared with conventional line-by-line scanning is a serious decrease in image quality. Indeed, compared with a focused transmit beam which concentrates most of its energy in a limited region of interest, the energy of an unfocused wavefront is spread over the entire field of view, resulting in backscattered echoes of lower amplitude and measurements with lower signal-to-noise ratio (SNR). The absence of transmit focusing also results in a broader main lobe of the point spread function (PSF), consequently degrading the lateral image resolution. Furthermore, diffraction artifacts, such as the ones caused by grating lobes (GLs), side lobes (SLs), and edge waves (EWs), are more pronounced in ultrafast US imaging, consequently degrading the image contrast. As a result, images obtained from ultrafast acquisitions are of low quality, suffering heavily from poor lateral resolution and low contrast [6]–[8], [11], [16], [17], that can hamper lesion detectability. Naturally, low-quality images also limit the accuracy of subsequent displacement estimation methods involved in ultrafast US imaging modes [6], [8], [12].

The state-of-the-art strategy to increase the image quality in ultrafast US imaging consists of coherently compounding low-quality images obtained from multiple, differently steered, unfocused transmit wavefronts [1], [6], [8], [16], [18]. This method successfully improves the image quality by increasing the number of steered acquisitions, potentially even surpassing the quality of conventional focused imaging. In [6], an image quality surpassing state-of-the-art multi-focus imaging was obtained by compounding 71 PW acquisitions, while increasing the frame rate by a factor of approximately seven. While in-

creasing the number of compounded acquisition improves the resulting imaging quality, it inevitably comes at the cost of lower frame rates, larger data transfers, and increased computational requirements, as multiple transmit-receive events and image reconstruction processes are required. Synthetic aperture (SA) imaging is another acquisition strategy based on the coherent compounding of multiple low-quality images. Instead of using the entire transducer aperture to transmit unfocused wavefronts, each transducer element is used to transmit a widely diffracting DW sequentially [19]. Thanks to its ability to focus synthetically both on transmit and on receive in each pixel of the image, SA is often considered as the gold-standard in US imaging. Yet, due to a lower transmitted energy than when using the entire aperture, it can suffer from increased SNR issues and remains difficult to deploy for imaging deep tissue without using multi-element sub-apertures [19].

However, for analyzing motion at very high frame rates, coherent compounding suffers from two considerable disadvantages. Firstly, the increase in image quality is directly linked to the number of compounded acquisitions, which in turn is limited by the minimum frame rate necessary to analyze the underlying physical phenomenon of interest. Secondly, coherent compounding assumes, similarly to line-by-line scanning, that the region of interest is stationary for the duration of an acquisition sequence used to reconstruct a single frame. This assumption does not hold when imaging fast-moving tissue regions or complex flows, for which coherent compounding suffers from strong motion artifacts [17], [20].

The first issue is well exemplified in [6], in which Montaldo *et al.* demonstrated, in the context of shear-wave elastography, that the quality of estimated elasticity maps is directly linked to the number of compounded acquisitions, which in turn was limited to a maximum of twelve acquisitions to ensure a minimum frame rate of 1 kHz. In particular, displacement estimation in highly heterogeneous tissue regions, where the aforementioned diffraction artifacts were dominant, was a major obstacle. Issues due to diffraction artifacts hindering accurate displacement estimates have been reported for several methods, all of them suffering from the trade-off between image quality (i.e., number of compounded acquisitions) and frame rate [6], [8], [21].

The occurrence of severe motion artifacts when compounding multiple acquisitions of rapidly evolving physical phenomena (inter-frame displacement close to the effective wavelength) was discussed in [17], [20], [22], and motion compensation techniques were proposed to tackle this problem. They consist of estimating inter-acquisition displacement, using either conventional Doppler [20], [22] or 1-D correlation methods [17], and compensate for it before compounding all acquisitions to produce a motion-compensated high-quality image. However, these motion compensation techniques can also suffer from strong diffraction artifacts [17], as they are themselves based on displacement estimation from low-quality images, obtained from unfocused wavefronts. It thus remains unclear if such methods could help improve motion estimation in regions plagued by such artifacts.

Consequently, there exists a great need for image reconstruction methods capable of extracting more information from ultrafast acquisitions to provide high-quality images from a minimum number of transmit-receive acquisitions; a need that gave rise to the

1. INTRODUCTION

plane-wave imaging challenge in medical ultrasound (PICMUS) [23]. Promising results were obtained using regularized inverse problems, where the image reconstruction problem is expressed by means of a measurement model, and solved using regularized iterative algorithms. Many regularizers were proposed, such as elastic net [24], sparsity in wavelet bases [25], or a weighted sum of multiple regularization terms [26]. However, as opposed to other imaging techniques which can rely on robust regularizers (e.g., CT), common image-processing regularizers are not well suited to the HDR and statistical properties of RF US images. This imposes an image-dependent fine-tuning of hyperparameters which does not generalize well, diminishing the appeal of these approaches except in specific cases. Also, due to the necessity of iterative solvers, real-time deployment is difficult to achieve. Most importantly, the difficulties encountered by regularization techniques in US imaging are particularly pronounced in the presence of speckle patterns. Even if these patterns are an “illusion” (or artifact) resulting from deterministic interferences of multiple sub-resolved diffusive scatterers, they carry positional information of the underlying physical phenomena. In particular, the coherence of their motion through time is fundamental to ultrafast US imaging modes, and more precisely to the corresponding motion estimation techniques. As such, these patterns not only need to be preserved on a (first-order) statistical point of view, but with a high positional accuracy, typically of few micro meters.

In short, any image reconstruction method intending to improve ultrafast US imaging modes should not only be capable of producing high-quality images from a minimum number of acquisitions, but also to preserve the information of underlying physical phenomena that can be further exploited by suitable motion estimation techniques. One means to achieve this goal is by being able to remove diffraction artifacts inherent to ultrafast acquisitions while restoring accurate speckle patterns initially shadowed by such artifacts. This is the approach proposed in this thesis.

1.2. Contributions and Organization of this Thesis

This thesis aims at answering the increasing need for US image reconstruction methods capable of producing high-quality images from single ultrafast acquisitions, primarily to improve the accuracy and robustness of ultrafast imaging modes such as shear-wave elastography, but also to reduce the cost and complexity of 3-D US scanners, or to mitigate the power and data transfer rate requirements of portable systems. To achieve this goal, this work builds in the context of inverse problems. It leverages recent deep-learning-based projection methods to overcome a crucial limitation of regularized US imaging: conventional image-processing regularizers are not well suited to the HDR and statistical properties of RF US images, especially in the presence of speckle patterns. Physical modeling is fundamental to this work. It was crucial to derive a computationally tractable forward model for image reconstruction, but also to develop a highly efficient, spline-based, spatial impulse response (SIR) US simulator that allowed generating sufficiently large datasets to train convolutional neural networks (CNNs). The different developments

and works involved led to several original contributions that are presented in the following chapters, namely from Chapters 2 to 7. These chapters are arranged in a natural order, namely, from physical modeling to deep learning. This order does not necessarily reflect the chronological order in which these contributions were realized.

- Because physical modeling played a central role both for developing suitable image reconstruction methods and for simulating realistic US signals, we start with a complete derivation from the scalar wave equation to a linear pulse-echo model for weak scattering (Chapter 2). The main contribution of this chapter comes from the fact that derivations were carried out on the velocity potential to obtain a generic model accounting for both Neumann (i.e., hard) and Dirichlet (i.e., soft) boundary conditions. Most results presented in this chapter were already known [27] (Neumann case) or proven experimentally to be adequate (Dirichlet case). Note that these derivations were carried out last, mainly because I have never been able to find them in the literature.
- Chapter 3 presents our attempt at reconstructing high-quality US images by means of sparse regularization (SR). In particular, it presents a fully generic derivation of a computationally tractable forward model developed throughout the years in our laboratory. Special cases of this model were successfully used to address some inverse problems in the context of US imaging that led to several publications such as [25], [28]–[30]. This chapter also highlights a crucial limitation we faced in regularized US imaging: conventional image-processing regularizers are not well suited to the HDR and the statistical properties of RF US images containing speckle patterns.
- Before being able to dive into CNN-based image reconstruction methods, we needed a way to generate sufficiently large datasets in a realistic time frame. Because no other option was available to us, we developed a spline-based SIR US simulator. It ended being a wonderful adventure into B-spline approximation, both for representing bandpass field signals efficiently and for parametrizing radiating surfaces as non-uniform rational B-spline (NURBS). This crucial contribution, which is currently in preparation for submission, is presented in Chapter 4.
- Chapter 5 presents all methods that enabled us to develop our CNN-based US image reconstruction methods, heavily inspired by our previous attempts using SR techniques and by [31], [32]. In particular, we introduce a re-weighted backprojection-based DAS operator to obtain low-quality estimates further processed by a CNN trained for the purpose of reducing diffraction artifacts and restoring accurate speckle patterns. Great efforts were invested into the design of simulated-image datasets composed of low-quality and high-quality image pairs, as well as numerical test phantoms for evaluation purposes. *In vitro* and *in vivo* experiments confirmed that trainings performed on simulated images translate well to physical conditions. This chapter is the core image reconstruction method proposed in this thesis and is also presented in [33], currently under review for potential publication.
- Recall that our goal is to develop an US image reconstruction method that could be used for improving ultrafast US imaging modes, and that these imaging modes rely

1. INTRODUCTION

on accurate displacement estimation at ultrafast rates. In Chapter 6, we present a displacement estimation method that relies on single ultrafast acquisitions to reconstruct consecutive frames by means of our CNN-based image reconstruction method, and on only two consecutive frames to obtain 2-D displacement estimates using a speckle tracking algorithm. Promising results were obtained on estimating accurate displacements in zones initially hindered by GL and SL artifacts, in both numerical and *in vivo* experiments. The content of this chapter is also presented in [34], which has recently been accepted for publication in the IEEE Transactions on Medical Imaging.

- Chapter 7 is a direct application of our CNN-based image reconstruction method to sparse-array imaging. We investigated the case of uniformly undersampled linear arrays, which suffer from even more severe diffraction artifacts, and in particular from GL ones. The content of this chapter is also presented in [35].

As most chapters contain in-depth discussions on the results obtained, the limitations observed, and the perspectives envisaged, concluding remarks (Chapter 8) are restricted to prospective notes on some aspects and potential applications related to the different contributions that are deemed deserving further investigations.

2. A Linear Pulse-Echo Model for Weak Scattering

2.1. Introduction

As for most (medical) imaging modalities, US scanners do not acquire images of a tissue of interest directly, but rather acquire signals on sensors that needs further processing to produce an image that can be interpreted by its user. Such additional processing is called an *image reconstruction method*. This rather general fact implies that, to develop suitable image reconstruction methods, one should at the very least understand how the signals received on the sensors composing the acquisition system were acquired. And since all signals acquired on real sensors are the result of (multiple) physical phenomena, it is no wonder that a good knowledge of the physical principles involved in the acquisition process can only help to develop more efficient image reconstruction methods.¹ Hence, the physical modeling of acquisition processes is at the heart of image reconstruction; may it be for deriving efficient model-based image reconstruction methods, or to develop filters for the removal (or reduction) of image artifacts. Such artifacts can typically be related to the acquisition process, but also to approximations made by the subsequent image reconstruction method. Obviously, being able to model a physical phenomenon also implies that it can be approximated by means of numerical simulations, for instance to optimize some components of an acquisition system.

The goal of this chapter is to provide a fairly detailed derivation to model the acquisition of US signals in the context of pulse-echo imaging. As most biological tissues are of extraordinary complexity, substantial approximations must be made to obtain tractable models. Throughout the following derivation, assumptions typically made in the context of pulse-echo imaging should be made clear. An important one being that linear modeling will be considered. Most of these derivations are available in the literature (sometimes quite scattered), and are as such inspired by it. The main source of inspiration of this chapter is the work of Jensen [27], in which a complete linear pulse-echo model for weak scattering was derived for transducer elements assumed to be embedded in a rigid baffle (Neumann boundary condition); a work partially based on the SIR model introduced by Tupholme [36] and Stepanishen [37]. Another major source of inspiration was found in the reference books by Morse and Feshbach [38] and by Morse and Ingard [38]. This chapter not only provides a complete (and augmented) derivation of the linear pulse-

¹In the sense that it is most probably never a waste to spend the necessary time on studying the underlying physics.

2. A LINEAR PULSE-ECHO MODEL FOR WEAK SCATTERING

echo model for weak scattering introduced in [27], it also includes all derivations and specific approximations to derive such scattering model for soft baffles (Dirichlet boundary condition), which, to the best of my knowledge, has never been reported. In particular, the derivation is carried out in the time domain using the velocity potential rather than the pressure. It starts from the linearized scalar wave equation as derived by Chernov [40, Part II] all the way down to a generic pulse-echo model for the weak scattering under different boundary conditions. Therefore, it includes a heavy use of Green's functions that are suitable for solving the scalar wave equation in specific conditions, leading to well-known intermediary results such as the Rayleigh-Sommerfeld equations.

2.2. Scalar Wave Equation for Weak Scattering

Let us consider a scattering volume V embedded in a nondispersive homogeneous medium (i.e., dispersive attenuation neglected) with constant mean density ρ_0 and mean acoustic wave velocity c_0 (mean sound speed). Assuming that both the density and the acoustic wave velocity deviate only slightly from their mean values, namely

$$\rho(\mathbf{r}) = \rho_0 + \Delta\rho(\mathbf{r}), \quad (2.1)$$

$$c(\mathbf{r}) = c_0 + \Delta c(\mathbf{r}), \quad (2.2)$$

where $\Delta\rho(\mathbf{r}) \ll \rho_0$ and $\Delta c(\mathbf{r}) \ll c_0$, and restricting to first order terms with respect to $\Delta\rho$ and Δc , Chernov [40, Part II] derived a (linearized) scalar wave equation to describe the propagation of (longitudinal) acoustic waves in a weakly inhomogeneous medium. This scalar wave equation can be expressed as

$$\nabla_{\mathbf{r}}^2 p(\mathbf{r}, t) - \frac{\partial^2 p(\mathbf{r}, t)}{\partial t^2} = -\frac{2\Delta c(\mathbf{r})}{c_0^3} \frac{\partial^2 p(\mathbf{r}, t)}{\partial t^2} + \frac{1}{\rho_0} \nabla_{\mathbf{r}}[\Delta\rho(\mathbf{r})] \cdot \nabla_{\mathbf{r}} p(\mathbf{r}, t), \quad (2.3)$$

where $p(\mathbf{r}, t)$ is the acoustic pressure at position \mathbf{r} (field point) and at time t , $\mathbf{a} \cdot \mathbf{b}$ represents the scalar product (or dot product) between the vectors \mathbf{a} and \mathbf{b} , $\nabla_{\mathbf{r}}$ is the gradient operator with respect to \mathbf{r} ,² and $\nabla_{\mathbf{r}}^2 := \nabla_{\mathbf{r}} \cdot \nabla_{\mathbf{r}}$ is the Laplacian operator with respect to \mathbf{r} . The right-hand side of (2.3) represents the effect of fluctuations in density and acoustic velocity, namely, the scattering of acoustic waves.

Equation (2.3) was also derived by Jensen [27] and formed the basis of a linear model for the propagation and scattering of US in weakly inhomogeneous tissue. It can be noted that for a homogeneous medium in which $\Delta\rho(\mathbf{r}) = \Delta c(\mathbf{r}) = 0$, $\forall \mathbf{r} \in V$, the scattering term becomes null, resulting in the homogeneous scalar wave equation for the pressure. An equivalent equation, for which the scattering term was expressed with respect to the compressibility and density, was derived in [39, Sec. 8.1] and used in [41].

Let us introduce the velocity potential φ , a scalar field from which the particle velocity

²Not to be confused with the directional derivative which is sometimes defined using the same symbols.

\mathbf{v} derives³ as

$$\mathbf{v}(\mathbf{r}, t) = -\nabla_{\mathbf{r}}\varphi(\mathbf{r}, t). \quad (2.4)$$

Under the assumption of an inviscid and irrotational flow, the velocity potential can be related to the pressure as

$$p(\mathbf{r}, t) = \rho_0 \frac{\partial \varphi(\mathbf{r}, t)}{\partial t}. \quad (2.5)$$

It is of particular interest to solve the scalar wave equation for the velocity potential instead of the pressure or particle velocity separately, as once a solution is found for the velocity potential, both the particle velocity and the pressure can be found using either (2.4) or (2.5), respectively. From (2.5), it can be noted that (2.3) is also satisfied by the velocity potential,⁴ namely

$$\nabla_{\mathbf{r}}^2 \varphi(\mathbf{r}, t) - \frac{\partial^2 \varphi(\mathbf{r}, t)}{\partial t^2} = -\frac{2\Delta c(\mathbf{r})}{c_0^3} \frac{\partial^2 \varphi(\mathbf{r}, t)}{\partial t^2} + \frac{1}{\rho_0} \nabla_{\mathbf{r}}[\Delta \rho(\mathbf{r})] \cdot \nabla_{\mathbf{r}} \varphi(\mathbf{r}, t). \quad (2.6)$$

Following the notations introduced in [27], which were retained in subsequent reference works [42], let us define the (linear) scattering operator⁵ as

$$f_{\text{op}}\{\varphi\}(\mathbf{r}, t) = \left[\frac{2\Delta c(\mathbf{r})}{c_0^3} \frac{\partial^2}{\partial t^2} - \frac{1}{\rho_0} \nabla_{\mathbf{r}}[\Delta \rho(\mathbf{r})] \cdot \nabla_{\mathbf{r}} \right] \varphi(\mathbf{r}, t). \quad (2.7)$$

Using (2.7) and defining a generic source function $f(\mathbf{r}, t) = f_{\text{op}}\{\varphi\}(\mathbf{r}, t)$, we can rewrite (2.6) in the generic form for the (inhomogeneous) scalar wave equation as [38, Sec. 7.3]

$$\nabla_{\mathbf{r}}^2 \varphi(\mathbf{r}, t) - \frac{\partial^2 \varphi(\mathbf{r}, t)}{\partial t^2} = -f(\mathbf{r}, t). \quad (2.8)$$

In the absence of volume sources, namely if $f(\mathbf{r}, t) = 0, \forall \mathbf{r} \in V, \forall t \geq 0$, (2.8) reduces to the homogeneous scalar wave equation, expressed as

$$\nabla_{\mathbf{r}}^2 \varphi(\mathbf{r}, t) - \frac{\partial^2 \varphi(\mathbf{r}, t)}{\partial t^2} = 0. \quad (2.9)$$

³The negative sign is conventionally used in electrical engineering.

⁴It can be noted that the scalar wave equation (2.3) is also satisfied by the particle velocity.

⁵The scattering operator is defined here with an opposite sign with respect to the original work [27] to better match the generic form of the scalar wave equation defined next in (2.8).

2.3. Generic Solutions of the Scalar Wave Equation

2.3.1. Green's Functions for the Scalar Wave Equation

The Green's function technique is a well-known and convenient approach to solve partial differential equations, such as the inhomogeneous scalar wave equation (2.8), for a particular set of boundary conditions, volume source distribution, and initial conditions. Let us consider that a volume V containing a distribution of sources with some initial conditions is enclosed by a surface S onto which boundary conditions are specified (Figure 2.1). Note that the (unit) normal vector to the surface is defined to point in the direction *outward* to the enclosed volume, to obtain the conventional *outflow* integrals. Such a Green's function $g(\mathbf{r}, t; \mathbf{r}', t')$ must satisfy (2.6) for a single source located at $\mathbf{r} = \mathbf{r}'$ with an impulse starting at $t = t'$, namely

$$\nabla_{\mathbf{r}}^2 g(\mathbf{r}, t; \mathbf{r}', t') - \frac{\partial^2 g(\mathbf{r}, t; \mathbf{r}', t')}{\partial t^2} = -\delta(\mathbf{r} - \mathbf{r}')\delta(t - t'), \quad (2.10)$$

where δ is the Dirac delta function. In other words, a Green's function must be a solution to a problem that is homogeneous everywhere except at $\mathbf{r} = \mathbf{r}'$ and $t = t'$.

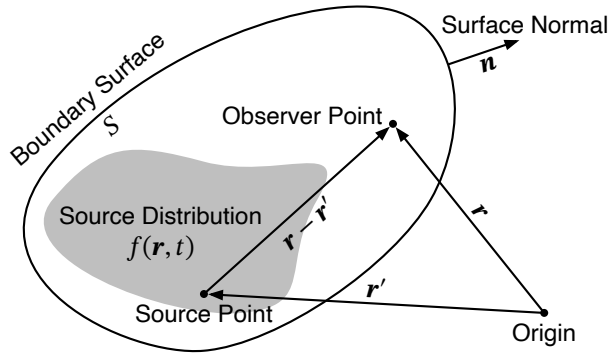


Figure 2.1. Generic representation of the Green's function method to determine the velocity potential at an observer point due to the presence of a source distribution embedded in a homogeneous volume enclosed by a boundary surface onto which boundary conditions are specified.

With the help of Green's Theorem and subtle manipulations of (2.6) and (2.10), including multiplication and rearrangements, it can be shown [38, Sec. 7.3] that the velocity potential can be evaluated $\forall \mathbf{r} \in S$, $\forall \mathbf{r} \in V$, and $\forall t \geq 0$ as

$$\begin{aligned} \varphi(\mathbf{r}, t) = & \int_T \int_S \left[g(\mathbf{r}, t; \mathbf{r}', t') \frac{\partial \varphi(\mathbf{r}', t')}{\partial \mathbf{n}'} - \varphi(\mathbf{r}', t') \frac{\partial g(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}'} \right] d\sigma(\mathbf{r}') dt' \\ & + \int_T \int_V g(\mathbf{r}, t; \mathbf{r}', t') f(\mathbf{r}', t') dv(\mathbf{r}') dt' \\ & - \frac{1}{c_0^2} \int_V \left[\frac{\partial g(\mathbf{r}, t; \mathbf{r}', t')}{\partial t'} \varphi(\mathbf{r}', t') - g(\mathbf{r}, t; \mathbf{r}', t') \frac{\partial \varphi(\mathbf{r}', t')}{\partial t'} \right]_{t'=0} dv(\mathbf{r}'). \end{aligned} \quad (2.11)$$

The first term of (2.11) accounts for the boundary conditions imposed over the enclosing surface S , the second term accounts for the distribution of sources present in the volume V (enclosed by S), and the last term accounts for the initial conditions of the problem. Boundary conditions typically enforce values taken by the velocity potential φ (Dirichlet boundary conditions), the normal derivative $\partial\varphi/\partial\mathbf{n}$ (Neumann boundary conditions), or a linear composition of the two (mixed boundary conditions). Provided that there exists a suitable Green's function that satisfies the boundary conditions and initial conditions of the original problem under consideration, the velocity potential can be evaluated at any point within the volume considered *and* on the enclosing surface.⁶

As we will restrict ourselves to problems involving media initially at rest, namely for which $\varphi(\mathbf{r}, t) = (\partial/\partial t)\varphi(\mathbf{r}, t) = 0$, $\forall \mathbf{r} \in V$ at $t = 0$, the last term of (2.11) becomes null and the solution to the scalar wave equation simplifies to

$$\begin{aligned} \varphi(\mathbf{r}, t) = & \iint_{T S} \left[g(\mathbf{r}, t; \mathbf{r}', t') \frac{\partial \varphi(\mathbf{r}', t')}{\partial \mathbf{n}'} - \varphi(\mathbf{r}', t') \frac{\partial g(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}'} \right] d\sigma(\mathbf{r}') dt' \\ & + \iiint_{T V} g(\mathbf{r}, t; \mathbf{r}', t') f(\mathbf{r}', t') dv(\mathbf{r}') dt'. \end{aligned} \quad (2.12)$$

2.3.2. Typical Boundary Conditions

Boundary conditions imposed on the enclosing surface S are of primary importance in the search for an appropriate Green's function, as the latter must satisfy them. The most common ones are Dirichlet and Neumann boundary conditions, which will result in the so-called “soft” and “hard” baffle conditions, respectively.

The Dirichlet boundary conditions prescribe the values taken by the velocity potential on the surface, imposing the corresponding Green's function to be zero on the surface. Such boundary conditions can be expressed for both φ_D and g_D , $\forall t \geq 0$, as

$$\begin{cases} \varphi_D(\mathbf{r}, t) = f_D(\mathbf{r}, t), & \forall \mathbf{r} \in S, \\ g_D(\mathbf{r}, t; \mathbf{r}', t') = 0, & \forall \mathbf{r} \in S, \\ g_D(\mathbf{r}, t; \mathbf{r}', t') = 0, & \forall \mathbf{r}' \in S, \end{cases} \quad (2.13)$$

where f_D is a scalar function. In cases where f_D is not zero everywhere, it is often referred to as *inhomogeneous* Dirichlet boundary conditions. In cases where $f_D = 0$ everywhere, it is often referred to (by opposition) as *homogeneous* Dirichlet boundary conditions.

The Neumann boundary conditions prescribe the values taken by the normal derivative of the velocity potential on the surface, imposing the normal derivative of the corresponding Green's function to be zero on the surface. Such boundary conditions can be expressed

⁶The latter being of particular interest in the case of pulse-echo imaging systems such as US transducers.

2. A LINEAR PULSE-ECHO MODEL FOR WEAK SCATTERING

for both φ_N and g_N , $\forall t \geq 0$, as

$$\begin{cases} \frac{\partial \varphi_N(\mathbf{r}, t)}{\partial \mathbf{n}} = f_N(\mathbf{r}, t), & \forall \mathbf{r} \in S, \\ \frac{\partial g_N(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}} = 0, & \forall \mathbf{r} \in S, \\ \frac{\partial g_N(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}'} = 0, & \forall \mathbf{r}' \in S, \end{cases} \quad (2.14)$$

where f_N is a scalar function. Similarly, the terminology *inhomogeneous* Neumann boundary conditions is used when f_N is not everywhere zero, and *homogeneous* Neumann boundary conditions when $f_N = 0$ everywhere. In the case of inhomogeneous Neumann boundary conditions (i.e., f_N not everywhere zero), it is interesting to note that the normal derivative of φ_N can be expressed as

$$\frac{\partial \varphi_N(\mathbf{r}, t)}{\partial \mathbf{n}} = \nabla_{\mathbf{r}} \varphi_N(\mathbf{r}, t) \cdot \mathbf{n}, \quad (2.15)$$

which is the directional derivative with respect to \mathbf{r} taken along \mathbf{n} . Hence, using the definition of the velocity potential in (2.4), imposing inhomogeneous Neumann boundary conditions corresponds to prescribing the *inward* normal component of the particle velocity on the surface.

Equipped with these definitions for the boundary conditions, two special cases are of particular interest, namely the homogeneous (scalar) wave equation (2.9) with inhomogeneous boundary conditions and the inhomogeneous (scalar) wave equation (2.8) with homogeneous boundary conditions.

2.3.3. Homogeneous Wave Equation with Inhomogeneous Boundary Conditions

In the absence of volume sources (i.e., $f(\mathbf{r}, t) = 0$), the second term of (2.12) becomes null and the solution for the homogeneous scalar wave equations with inhomogeneous boundary conditions can be expressed $\forall \mathbf{r} \in S$, $\forall \mathbf{r} \in V$, and $\forall t \geq 0$, as

$$\varphi(\mathbf{r}, t) = \iint_{TS} \left[g(\mathbf{r}, t; \mathbf{r}', t') \frac{\partial \varphi(\mathbf{r}', t')}{\partial \mathbf{n}'} - \varphi(\mathbf{r}', t') \frac{\partial g(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}'} \right] d\sigma(\mathbf{r}') dt'. \quad (2.16)$$

This means that, provided a suitable Green's function exist, the velocity potential can be evaluated anywhere within the volume or on the enclosing surface by an integration over the surface only.

In the case of inhomogeneous Dirichlet boundary conditions defined in (2.13), (2.16) reduces to

$$\varphi_D(\mathbf{r}, t) = - \iint_{TS} \varphi_D(\mathbf{r}', t') \frac{\partial g_D(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}'} d\sigma(\mathbf{r}') dt', \quad (2.17)$$

which can be interpreted as having a doublet layer (i.e., normal derivative of Green's function) to satisfy the boundary conditions. In the case of inhomogeneous Neumann

boundary conditions defined in (2.14), (2.16) reduces to

$$\varphi_N(\mathbf{r}, t) = \iint_{T S} g_N(\mathbf{r}, t; \mathbf{r}', t') \frac{\partial \varphi_N(\mathbf{r}', t')}{\partial \mathbf{n}'} d\sigma(\mathbf{r}') dt', \quad (2.18)$$

corresponding to having a single (charge) layer to satisfy the boundary conditions.

2.3.4. Inhomogeneous Wave Equation with Homogeneous Boundary Conditions

Imposing homogeneous boundary conditions on the velocity potential implies that the first term of (2.16) becomes null, thus the solution for the inhomogeneous scalar wave equations (i.e., $f(\mathbf{r}, t)$ not everywhere zero in V) with inhomogeneous boundary conditions can be expressed $\forall \mathbf{r} \in S, \forall \mathbf{r} \in V$, and $\forall t \geq 0$, as

$$\varphi(\mathbf{r}, t) = \iint_{T V} g(\mathbf{r}, t; \mathbf{r}', t') f(\mathbf{r}', t') dv(\mathbf{r}') dt'. \quad (2.19)$$

Here again a distinction between inhomogeneous Dirichlet and Neumann boundary conditions is of interest.

For homogeneous Dirichlet boundary conditions defined in (2.13) with $f_D = 0$ everywhere, (2.19) is simply satisfied using (2.13) as

$$\varphi_D(\mathbf{r}, t) = \iint_{T V} g_D(\mathbf{r}, t; \mathbf{r}', t') f(\mathbf{r}', t') dv(\mathbf{r}') dt', \quad (2.20)$$

where both sides of the equation are zero. Obviously, (2.20) has no practical use except to confirm that the boundary conditions are satisfied. Yet, by taking the normal derivative at point $\mathbf{r} \in S$ we obtain

$$\frac{\partial \varphi_D(\mathbf{r}, t)}{\partial \mathbf{n}} = \iint_{T V} \frac{\partial g_D(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}} f(\mathbf{r}', t') dv(\mathbf{r}') dt', \quad (2.21)$$

which, from the definition of the velocity potential given in (2.4), corresponds to the *inward* normal component of the particle velocity on the surface.

For homogeneous Neumann boundary conditions defined in (2.14) with $f_N = 0$ everywhere, (2.19) is expressed as

$$\varphi_N(\mathbf{r}, t) = \iint_{T V} g_N(\mathbf{r}, t; \mathbf{r}', t') f(\mathbf{r}', t') dv(\mathbf{r}') dt'. \quad (2.22)$$

By taking the normal derivative at point $\mathbf{r} \in S$, we can confirm that the boundary conditions are satisfied, namely

$$\frac{\partial \varphi_N(\mathbf{r}, t)}{\partial \mathbf{n}} = \iint_{T V} \frac{\partial g_N(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}} f(\mathbf{r}', t') dv(\mathbf{r}') dt', \quad (2.23)$$

is indeed zero on both sides.

2.4. Specifying Green's Function

In Section 2.3.1, we showed that solutions to the inhomogeneous scalar wave equation (2.8) can be found using the Green's function technique. In particular, provided that there exists a suitable Green's function satisfying the boundary conditions (may them be of Neumann or Dirichlet type) and initial conditions of the original problem under consideration, the velocity potential can be evaluated anywhere within the volume as well as on the enclosing surface using (2.11), or using (2.12) in the absence of initial conditions (special case of interest). Therefore, specifying such Green's functions is crucial to finding solutions of the wave equation for weak scattering defined in (2.6).

2.4.1. Single Volume Source in an Unbounded Domain

One of the simplest inhomogeneous problem to consider consists of a single volume source embedded in an unbounded (infinite) domain, namely no reflections of the acoustic wave will occur on the surface boundaries. The Green's function for this unbounded problem is well-known and expressed as

$$g_U(\mathbf{r}, t; \mathbf{r}', t') = \frac{\delta\left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right)}{4\pi\|\mathbf{r} - \mathbf{r}'\|_2}, \quad \|\mathbf{r} - \mathbf{r}'\|_2, t - t' > 0. \quad (2.24)$$

It represents a point source located at \mathbf{r}' with an impulse at time t' resulting in a spherical *outgoing* "wave,"⁷ which expands in the radial direction with an amplitude decaying proportionally to $1/\|\mathbf{r} - \mathbf{r}'\|_2$. It should be noted that g_U is singular at the source point (i.e., $\mathbf{r} = \mathbf{r}'$) when the impulse is initiated (i.e., $t = t'$), hence the condition $\|\mathbf{r} - \mathbf{r}'\|_2, t - t' > 0$ in (2.24).

As the normal derivative of the Green's function for the unbounded domain with respect to both \mathbf{r} and \mathbf{r}' will be required for further derivations, they are already provided hereafter for convenience. Let us first consider the case of a source point \mathbf{r}' on S . Recall that the normal derivative can be expressed as the directional derivative in the direction of the normal, namely

$$\frac{\partial g_U(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}'} = \nabla_{\mathbf{r}'} g_U(\mathbf{r}, t; \mathbf{r}', t') \cdot \mathbf{n}' \quad (2.25)$$

where the gradient of g_U with respect to \mathbf{r}'

$$\nabla_{\mathbf{r}'} g_U(\mathbf{r}, t; \mathbf{r}', t') = \frac{1}{4\pi} \left[\frac{1}{\|\mathbf{r} - \mathbf{r}'\|_2} \nabla_{\mathbf{r}'} + \nabla_{\mathbf{r}'} \left(\frac{1}{\|\mathbf{r} - \mathbf{r}'\|_2} \right) \right] \delta\left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right). \quad (2.26)$$

⁷Strictly speaking, it is a spherical shell.

Using the fact that

$$\nabla_{\mathbf{r}'} \left(\frac{1}{\|\mathbf{r} - \mathbf{r}'\|_2} \right) = \frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|_2^3}, \quad (2.27)$$

$$\nabla_{\mathbf{r}'} \delta \left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right) = \frac{\mathbf{r} - \mathbf{r}'}{c_0 \|\mathbf{r} - \mathbf{r}'\|_2} \delta' \left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right), \quad (2.28)$$

where δ' is the derivative of the Dirac delta function,

$$\nabla_{\mathbf{r}'} g_U(\mathbf{r}, t; \mathbf{r}', t') = \frac{\mathbf{r} - \mathbf{r}'}{4\pi \|\mathbf{r} - \mathbf{r}'\|_2} \left[\frac{\delta' \left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right)}{c_0 \|\mathbf{r} - \mathbf{r}'\|_2} + \frac{\delta \left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right)}{\|\mathbf{r} - \mathbf{r}'\|_2^2} \right]. \quad (2.29)$$

Since \mathbf{n}' is a unit vector,

$$\frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|_2} \cdot \mathbf{n}' = \frac{(\mathbf{r} - \mathbf{r}') \cdot \mathbf{n}'}{\|\mathbf{r} - \mathbf{r}'\|_2 \|\mathbf{n}'\|_2} = \cos(\mathbf{n}', \mathbf{r} - \mathbf{r}'), \quad (2.30)$$

we can finally express the normal derivative of g_U with respect to \mathbf{r}' as

$$\frac{\partial g_U(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}'} = \left[\frac{\delta' \left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right)}{4\pi c_0 \|\mathbf{r} - \mathbf{r}'\|_2} + \frac{\delta \left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right)}{4\pi \|\mathbf{r} - \mathbf{r}'\|_2^2} \right] \cos(\mathbf{n}', \mathbf{r} - \mathbf{r}'). \quad (2.31)$$

Following the exact same derivation for the normal derivative of g_U with respect to \mathbf{r} and noting that

$$\nabla_{\mathbf{r}} \left(\frac{1}{\|\mathbf{r} - \mathbf{r}'\|_2} \right) = -\frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|_2^3}, \quad (2.32)$$

$$\nabla_{\mathbf{r}} \delta \left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right) = -\frac{\mathbf{r} - \mathbf{r}'}{c_0 \|\mathbf{r} - \mathbf{r}'\|_2} \delta' \left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right), \quad (2.33)$$

$$\frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|_2} \cdot \mathbf{n} = \cos(\mathbf{n}, \mathbf{r} - \mathbf{r}'), \quad (2.34)$$

we obtain

$$\frac{\partial g_U(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}} = - \left[\frac{\delta' \left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right)}{4\pi c_0 \|\mathbf{r} - \mathbf{r}'\|_2} + \frac{\delta \left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right)}{4\pi \|\mathbf{r} - \mathbf{r}'\|_2^2} \right] \cos(\mathbf{n}, \mathbf{r} - \mathbf{r}'). \quad (2.35)$$

Both expressions (2.31) and (2.35) for the two distinct normal derivatives are thus very similar except for the inverted sign, which results from the direction of propagation $\mathbf{r} - \mathbf{r}'$ with respect to the corresponding normal. The amplitude is proportional to the cosine of the angle between the direction of propagation and the (outward) surface normal.

2.4.2. Infinite Planar Boundaries

As already mentioned, the presence of boundaries in the domain (e.g., an US transducer) has a direct impact on the Green's function as such a function must satisfy the boundary conditions imposed by the problem considered. Approximating US transducers as infinite planar baffles has a long history in the field of transient radiation modeling of US transducers [36], [37], [43]–[45]. Under such an approximation, the so-called “method of images” is a convenient technique for finding suitable Green's functions. This method consists of mirroring the impulsing source at point \mathbf{r}' with respect to the infinite boundary plane in an image source at point $\tilde{\mathbf{r}}'$ (Figure 2.2) while keeping the same impulsing time $\tilde{t}' = t'$ for the image and physical sources.

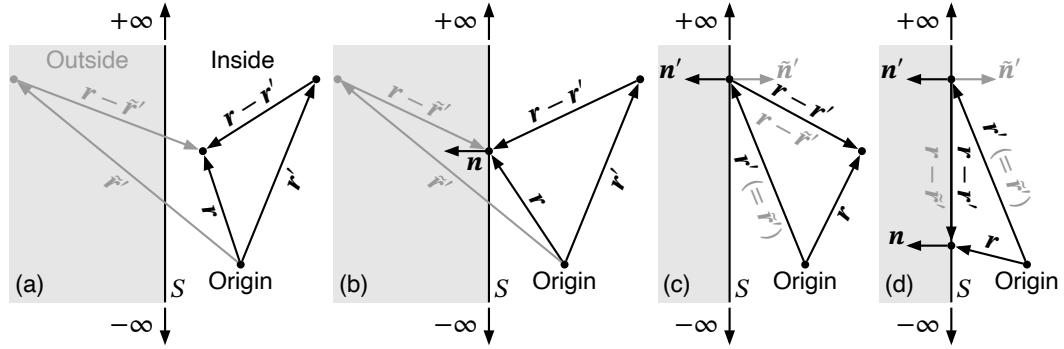


Figure 2.2. Method of images to specify suitable Green's functions when considering an infinite planar boundary: (a) both the observer and the source are within the volume; (b) the observer is on the boundary plane and the source within the volume; (c) the observer is within the volume and the observer on the boundary plane; (d) both the observer and the source are on the boundary plane.

In the case of Dirichlet boundary conditions, a suitable Green's function can be found by subtracting two Green's functions for the unbounded domain, namely

$$g_D(\mathbf{r}, t; \mathbf{r}', t') = g_U(\mathbf{r}, t; \mathbf{r}', t') - g_U(\mathbf{r}, t; \tilde{\mathbf{r}}', \tilde{t}'). \quad (2.36)$$

Using the definition (2.24), we can see that the boundary conditions for the Green's function (2.13) are satisfied for $\mathbf{r} \in S$ and $\mathbf{r}' \in V$ [Figure 2.2(b)], or for $\mathbf{r} \in V$ and $\mathbf{r}' \in S$ [Figure 2.2(c)], or for $\mathbf{r} \in V$ and $\mathbf{r}' \in S$ [Figure 2.2(d)]. Indeed, $t' = \tilde{t}'$, $\|\mathbf{r} - \mathbf{r}'\|_2 = \|\mathbf{r} - \tilde{\mathbf{r}}'\|_2$, hence we have

$$g_D(\mathbf{r}, t; \mathbf{r}', t') = \frac{\delta\left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right)}{4\pi\|\mathbf{r} - \mathbf{r}'\|_2} - \frac{\delta\left(t - \tilde{t}' - \frac{\|\mathbf{r} - \tilde{\mathbf{r}}'\|_2}{c_0}\right)}{4\pi\|\mathbf{r} - \tilde{\mathbf{r}}'\|_2} = 0. \quad (2.37)$$

For the normal derivative with respect to \mathbf{r}' , we have $\cos(\mathbf{n}', \mathbf{r} - \mathbf{r}') = -\cos(\tilde{\mathbf{n}}', \mathbf{r} - \tilde{\mathbf{r}}')$

[Figures 2.2(c) and 2.2(d)]. Hence, using (2.31), we obtain

$$\begin{aligned} \frac{\partial g_D(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}'} &= \frac{\partial g_U(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}'} - \frac{\partial g_U(\mathbf{r}, t; \tilde{\mathbf{r}}', \tilde{t}')}{\partial \tilde{\mathbf{n}}'} = 2 \frac{\partial g_U(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}'} \\ &= \left[\frac{\delta' \left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right)}{2\pi c_0 \|\mathbf{r} - \mathbf{r}'\|_2} + \frac{\delta' \left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right)}{2\pi \|\mathbf{r} - \mathbf{r}'\|_2^2} \right] \cos(\mathbf{n}', \mathbf{r} - \mathbf{r}'). \end{aligned} \quad (2.38)$$

For the normal derivative with respect to \mathbf{r} , since $\cos(\mathbf{n}, \mathbf{r} - \mathbf{r}') = -\cos(\mathbf{n}, \mathbf{r} - \tilde{\mathbf{r}}')$ [Figures 2.2(b) and 2.2(d)], we obtain using (2.35)

$$\begin{aligned} \frac{\partial g_D(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}} &= \frac{\partial g_U(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}} - \frac{\partial g_U(\mathbf{r}, t; \tilde{\mathbf{r}}', \tilde{t}')}{\partial \mathbf{n}} = 2 \frac{\partial g_U(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}} \\ &= - \left[\frac{\delta' \left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right)}{2\pi c_0 \|\mathbf{r} - \mathbf{r}'\|_2} + \frac{\delta' \left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right)}{2\pi \|\mathbf{r} - \mathbf{r}'\|_2^2} \right] \cos(\mathbf{n}, \mathbf{r} - \mathbf{r}'). \end{aligned} \quad (2.39)$$

In the case of Neumann boundary conditions, a suitable Green's function can be found by summing two Green's functions for the unbounded domain, namely

$$g_N(\mathbf{r}, t; \mathbf{r}', t') = g_U(\mathbf{r}, t; \mathbf{r}', t') + g_U(\mathbf{r}, t; \tilde{\mathbf{r}}', \tilde{t}'). \quad (2.40)$$

We can see that the boundary conditions (2.14), which are specified on the normal derivatives, are also satisfied in all cases. Indeed, using the same properties as in the case of Dirichlet boundary conditions, we have for the two normal derivatives

$$\frac{\partial g_N(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}'} = \frac{\partial g_U(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}'} + \frac{\partial g_U(\mathbf{r}, t; \tilde{\mathbf{r}}', \tilde{t}')}{\partial \tilde{\mathbf{n}}'} = 0, \quad (2.41)$$

$$\frac{\partial g_N(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}} = \frac{\partial g_U(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}} + \frac{\partial g_U(\mathbf{r}, t; \tilde{\mathbf{r}}', \tilde{t}')}{\partial \mathbf{n}} = 0. \quad (2.42)$$

If $\mathbf{r} \in S$ or $\mathbf{r}' \in S$, $\|\mathbf{r} - \mathbf{r}'\|_2 = \|\mathbf{r} - \tilde{\mathbf{r}}'\|_2$, and (2.40) can be expressed as

$$g_N(\mathbf{r}, t; \mathbf{r}', t') = 2g_U(\mathbf{r}, t; \mathbf{r}', t') = \frac{\delta \left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right)}{2\pi \|\mathbf{r} - \mathbf{r}'\|_2}. \quad (2.43)$$

2.4.3. Time-Dependent Rayleigh-Sommerfeld Equations

Using the Green's function derived in Section 2.4.2 for Dirichlet and Neumann boundary conditions, it is possible to derive the well-known Rayleigh-Sommerfeld diffraction equations for transient problems (i.e., time-dependent). These equations were originally derived in the frequency domain (i.e., continuous waves) by Rayleigh [46] and Sommerfeld [47, Sec. 34]. They are extensively used for estimating pressure fields generated by US transducers with “hard” and “soft” boundary conditions, respectively, assuming that

2. A LINEAR PULSE-ECHO MODEL FOR WEAK SCATTERING

the transducers are embedded in an infinite planar baffle.

To derive the Rayleigh-Sommerfeld equations, we will therefore consider a radiating surface embedded in an infinite planar boundary and a homogeneous medium. The velocity potential can thus be evaluated using (2.16). Inhomogeneous boundary conditions are specified on the radiating surface and homogeneous boundary conditions (of the same kind) are specified on the remaining infinite boundary surface, a similar situation that is depicted in Figure 2.2(c). We will also rely on the convolution property of the Dirac delta function, also known as the “sifting” property, which can be expressed as

$$\int_{-\infty}^{+\infty} f(\tau) \delta(t - \tau) d\tau = f(t), \quad (2.44)$$

as well as on the convolution property of its derivative δ' ,

$$\int_{-\infty}^{+\infty} f(\tau) \delta'(t - \tau) d\tau = f'(t), \quad (2.45)$$

for which it is assumed that f is a compactly supported smooth function and where f' is the derivative of f .

Let us first consider the case of Dirichlet boundary conditions (i.e., inhomogeneous on the radiating surface and homogeneous on the remaining of the infinite planar boundary). Using the normal derivative with respect to \mathbf{r}' of the Green's function for inhomogeneous Dirichlet boundary conditions from (2.38), and injecting it in the solution for the velocity potential under such conditions given in (2.17), we obtain

$$\begin{aligned} \varphi_D(\mathbf{r}, t) = & - \iint_{TS} \varphi_D(\mathbf{r}', t') \frac{\delta' \left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right)}{2\pi c_0 \|\mathbf{r} - \mathbf{r}'\|_2} \cos(\mathbf{n}', \mathbf{r} - \mathbf{r}') d\sigma(\mathbf{r}') dt' \\ & - \iint_{TS} \varphi_D(\mathbf{r}', t') \frac{\delta \left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right)}{2\pi \|\mathbf{r} - \mathbf{r}'\|_2^2} \cos(\mathbf{n}', \mathbf{r} - \mathbf{r}') d\sigma(\mathbf{r}') dt'. \end{aligned} \quad (2.46)$$

Using the convolution properties (2.44) and (2.45) of the Dirac delta function and its derivative,

$$\begin{aligned} \varphi_D(\mathbf{r}, t) = & - \frac{1}{2\pi} \int_S \frac{\partial}{\partial t} \left[\varphi_D \left(\mathbf{r}', t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right) \right] \frac{\cos(\mathbf{n}', \mathbf{r} - \mathbf{r}')}{c_0 \|\mathbf{r} - \mathbf{r}'\|_2} d\sigma(\mathbf{r}') \\ & - \frac{1}{2\pi} \int_S \varphi_D \left(\mathbf{r}', t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right) \frac{\cos(\mathbf{n}', \mathbf{r} - \mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|_2^2} d\sigma(\mathbf{r}'). \end{aligned} \quad (2.47)$$

Assuming that all sources (on the radiating surface) are sufficiently far from the observer point such that $\|\mathbf{r} - \mathbf{r}'\|_2 \gg \lambda/(2\pi)$, the second term of (2.47) can be neglected,⁸ reducing

⁸It should be noted that the wavelength involved in the $\|\mathbf{r} - \mathbf{r}'\|_2 \gg \lambda/(2\pi)$ condition depends on the

to

$$\varphi_D(\mathbf{r}, t) \approx -\frac{1}{2\pi} \int_S \frac{1}{c_0} \frac{\partial}{\partial t} \left[\varphi_D\left(\mathbf{r}', t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right) \right] \frac{\cos(\mathbf{n}', \mathbf{r} - \mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|_2} d\sigma(\mathbf{r}'), \quad (2.48)$$

which is commonly referred to as the time-dependent Sommerfeld (diffraction) integral [47, Sec. 34]. It can be seen that the velocity potential decays proportionally to $1/\|\mathbf{r} - \mathbf{r}'\|_2$ and to the cosine between the propagation direction and the (outward) normal to the surface (i.e., soft baffle). Recalling from (2.5) that $p(\mathbf{r}, t) = \rho_0(\partial/\partial t)\varphi(\mathbf{r}, t)$, and using a more “standard” *inward* normal $\tilde{\mathbf{n}}' = -\mathbf{n}'$ [Figure 2.2(c)] such that $\cos(\tilde{\mathbf{n}}', \mathbf{r} - \mathbf{r}') = -\cos(\mathbf{n}', \mathbf{r} - \mathbf{r}')$, the Sommerfeld integral (2.48) can be expressed in terms of physical quantities as

$$\varphi_D(\mathbf{r}, t) \approx \frac{1}{2\pi} \int_S \frac{1}{\rho_0 c_0} p_S\left(\mathbf{r}', t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right) \frac{\cos(\tilde{\mathbf{n}}', \mathbf{r} - \mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|_2} d\sigma(\mathbf{r}'), \quad (2.49)$$

where $p_S(\mathbf{r}, t)$, $\mathbf{r} \in S$, is the pressure distribution on the radiating surface. Equation (2.49) is also commonly referred to as the solution to the “pressure release surface problem.”

In the case of Neumann boundary conditions (i.e., inhomogeneous on the radiating surface and homogeneous on the remaining of the infinite planar boundary), the solution for the velocity potential is given by (2.18). Using the Green’s function corresponding to this type of boundary conditions as expressed in (2.43), we can rewrite (2.18) as

$$\varphi_N(\mathbf{r}, t) = \iint_{TS} \frac{\partial \varphi_N(\mathbf{r}', t')}{\partial \mathbf{n}'} \frac{\delta\left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right)}{2\pi\|\mathbf{r} - \mathbf{r}'\|_2} d\sigma(\mathbf{r}') dt'. \quad (2.50)$$

Using the convolution property (2.44) of the Dirac delta function, we obtain the time-dependent Rayleigh (diffraction) integral [48, Vol 2., p. 107]

$$\varphi_N(\mathbf{r}, t) = \frac{1}{2\pi} \int_S \frac{\partial}{\partial \mathbf{n}'} \left[\varphi_N\left(\mathbf{r}', t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right) \right] \frac{1}{\|\mathbf{r} - \mathbf{r}'\|_2} d\sigma(\mathbf{r}'), \quad (2.51)$$

which, similarly to the Sommerfeld integral, decays proportionally to $1/\|\mathbf{r} - \mathbf{r}'\|_2$, but independently to the cosine between the (outward) surface normal and the propagation direction (i.e., hard baffle). Note that contrary to the Sommerfeld integral, the Rayleigh integral is exact (i.e., no additional assumption was made on the distance $\|\mathbf{r} - \mathbf{r}'\|_2$). Recalling from the definition of the velocity potential (2.4) that $\mathbf{v}(\mathbf{r}, t) = -\nabla_{\mathbf{r}}\varphi(\mathbf{r}, t)$, let us define the *inward* normal velocity distribution as $v_n(\mathbf{r}, t) = -\mathbf{v}(\mathbf{r}, t) \cdot \mathbf{n}$, $\mathbf{r} \in S$, such that $v_n(\mathbf{r}, t) = (\partial/\partial \mathbf{n})\varphi_N(\mathbf{r}, t)$. Using such a definition, the Rayleigh integral (2.51) can be expressed in terms of physical quantities as

$$\varphi_N(\mathbf{r}, t) = \frac{1}{2\pi} \int_S v_n\left(\mathbf{r}', t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right) \frac{1}{\|\mathbf{r} - \mathbf{r}'\|_2} d\sigma(\mathbf{r}'). \quad (2.52)$$

bandwidth property of $(\partial/\partial t)\varphi_D$, a quantity proportional to the surface pressure distribution, and should therefore be analyzed with care.

2.4.4. Corresponding Spatial Impulse Response Formulations

In Section 2.4.3 we saw that both Rayleigh (2.52) and Sommerfeld (2.49) equations are expressed as surface integrals involving a surface waveform distribution (i.e., scaled pressure and normal velocity, respectively). A common approximation [27] consists of assuming separability of time and spatial variables of these surface waveform distributions. By doing so, it is possible to express the velocity potential with respect to the SIR, which relates the surface geometry of the radiator (i.e., transducer) to the acoustic fields [27], [37].

Let us start with the Sommerfeld equation (i.e., Dirichlet case) and assume separation of spatial and time variables for the surface pressure distribution such that $p_S(\mathbf{r}, t) = \xi_D(\mathbf{r})\bar{p}_S(t)$, where ξ_D is the surface amplitude distribution of the surface pressure, sometimes referred to as (surface) apodization function. By introducing a retarded Dirac delta function such that

$$p_S\left(\mathbf{r}, t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right) = \xi_D(\mathbf{r}) \int_{-\infty}^{+\infty} \bar{p}_S(\tau) \delta\left(t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} - \tau\right) d\tau \quad (2.53)$$

$$= \xi_D(\mathbf{r}) \bar{p}_S(t) *_t \delta\left(t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right), \quad (2.54)$$

we can express (2.49) as

$$\varphi_D(\mathbf{r}, t) \approx \frac{\bar{p}_S(t)}{\rho_0 c_0} *_t \int_S \frac{\xi_D(\mathbf{r}') \delta\left(t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right)}{2\pi \|\mathbf{r} - \mathbf{r}'\|_2} \cos(\tilde{\mathbf{n}}', \mathbf{r} - \mathbf{r}') d\sigma(\mathbf{r}'). \quad (2.55)$$

The surface integral term of (2.55) is the SIR of a (3-D) planar surface S with inhomogeneous Dirichlet boundary conditions. This term is expressed as

$$h_D^S(\mathbf{r}, t) = \int_S \frac{\xi_D(\mathbf{r}') \delta\left(t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right)}{2\pi \|\mathbf{r} - \mathbf{r}'\|_2} \cos(\tilde{\mathbf{n}}', \mathbf{r} - \mathbf{r}') d\sigma(\mathbf{r}'), \quad (2.56)$$

such that (2.55) can be rewritten in a compact form as

$$\varphi_D(\mathbf{r}, t) \approx \frac{\bar{p}_S(t)}{\rho_0 c_0} *_t h_D^S(\mathbf{r}, t). \quad (2.57)$$

By also assuming separation of spatial and time variables for the normal velocity distribution in the Neumann case such that $v_n(\mathbf{r}, t) = \xi_N(\mathbf{r})\bar{v}_n(t)$, and introducing the same retarded Dirac as before, we can express the Rayleigh integral (2.52) as

$$\varphi_N(\mathbf{r}, t) = \bar{v}_n(t) *_t \int_S \frac{\xi_N(\mathbf{r}') \delta\left(t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right)}{2\pi \|\mathbf{r} - \mathbf{r}'\|_2} d\sigma(\mathbf{r}'). \quad (2.58)$$

The surface integral term of (2.58) is the SIR of a (3-D) planar surface S with inhomoge-

neous Neumann boundary conditions. This term is expressed as

$$h_N^S(\mathbf{r}, t) = \int_S \frac{\xi_N(\mathbf{r}') \delta\left(t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right)}{2\pi\|\mathbf{r} - \mathbf{r}'\|_2} d\sigma(\mathbf{r}'), \quad (2.59)$$

such that (2.58) can be rewritten compactly as

$$\varphi_N(\mathbf{r}, t) = \bar{v}_n(t) *_{\mathbf{t}} h_N^S(\mathbf{r}, t). \quad (2.60)$$

In both Dirichlet (soft) and Neumann (hard) boundary conditions, one can note that the resulting SIRs defined in (2.56) and (2.59), respectively, are extremely similar, with an additional cosine term in the case of Dirichlet boundary conditions. Therefore, it is sometimes convenient to express them in a generic form as

$$h_S(\mathbf{r}, t) = \int_S \frac{\xi(\mathbf{r}') \delta\left(t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right)}{2\pi\|\mathbf{r} - \mathbf{r}'\|_2} \beta(\tilde{\mathbf{n}}', \mathbf{r} - \mathbf{r}') d\sigma(\mathbf{r}'), \quad (2.61)$$

where $\tilde{\mathbf{n}}'$ is the *inward* surface normal, ξ represents the spatial distribution of surface velocity or pressure amplitudes over the radiating surface, and β is a term depending on the boundary conditions, expressed as

$$\beta(\tilde{\mathbf{n}}', \mathbf{r} - \mathbf{r}') = \begin{cases} 1, & \text{for a hard baffle (Neumann type),} \\ \cos(\tilde{\mathbf{n}}', \mathbf{r} - \mathbf{r}'), & \text{for a soft baffle (Dirichlet type).} \end{cases} \quad (2.62)$$

Moreover, both resulting expressions for the velocity potential (2.57) and (2.60) consist of the time convolution of a waveform and a SIR, each of which depend on the type of boundary conditions considered. Again, we can introduce a generic definition for the waveform imposed as boundary condition on the surface S as

$$v_S(t) = \begin{cases} \bar{v}_n(t), & \text{for a hard baffle (Neumann type),} \\ \bar{p}_S(t)/(\rho_0 c_0), & \text{for a soft baffle (Dirichlet type),} \end{cases} \quad (2.63)$$

such that the velocity potential can be expressed in a generic form as⁹

$$\varphi(\mathbf{r}, t) = v_S(t) *_{\mathbf{t}} h_S(\mathbf{r}, t), \quad (2.64)$$

that is the time convolution of a time-dependent waveform v_S and a spatially varying SIR h_S .

It can be noted that if the radiating surface is assumed to launch a “quasi”-planar wavefront, the generic waveform v_S defined in (2.63) is equal in both Neumann and Dirichlet cases [49]. Under such plane-wave assumption, other generalizations may

⁹The approximation sign from the Dirichlet case was removed for convenience.

2. A LINEAR PULSE-ECHO MODEL FOR WEAK SCATTERING

be derived as shown in [49], [50]. Also, since the pressure can be evaluated from the velocity potential using (2.5) as $p(\mathbf{r}, t) = \rho_0(\partial/\partial t)\varphi(\mathbf{r}, t)$, it can be seen from (2.64) that the acoustic propagation from a radiating surface will amount to a time convolution of the time derivative of a waveform prescribed on the surface and the SIR corresponding to the boundary conditions specified.

It is important to keep in mind that these expressions are valid for 3-D *planar* radiating surfaces, assumed to be embedded in an *infinite* planar boundary surface. Obviously, US transducers are not embedded in an infinite planar baffle. Yet, this approximation has proved to be realistic considering the amount of work deriving from it, for instance [27], [37], [51]–[53]. Also, transducers are not necessarily planar, such as spherically focused “mono-elements” or elements of 1-D arrays that are typically slightly focused in elevation. It has also been proved that the Rayleigh-Sommerfeld equations remained accurate in the case of slightly focused transducers [43], [52], [54], namely with a curvature of multiple times the characteristic wavelength.

Under these approximations, the Rayleigh integral is exact (hard baffle), whereas the Sommerfeld integral (soft baffle) has an additional assumption on the distance between the radiating surface and the observer point at which the velocity potential is to be evaluated. The latter assumption, namely $\|\mathbf{r} - \mathbf{r}'\|_2 \gg \lambda/(2\pi)$, depends on the properties of the pressure waveform specified on the radiating surface, where the wavelength λ considered should be the “maximum” wavelength of the pressure waveform (corresponding to the lower frequency of its spectrum).

2.5. The Pulse-Echo Model for Weak Scattering

Let us now consider the generic pulse-echo set-up depicted in Figure 2.3. We assume a (3-D) volume V bounded by an infinite planar surface S onto which boundary conditions are specified. The volume is considered to be weakly scattering such that the (inhomogeneous) scalar wave equation for the velocity potential defined in (2.8) is satisfied. The source distribution $f(\mathbf{r}, t)$ accounts for the (weak) scattering. The enclosing surface consists of a transmitting part S_{tx} (i.e., transmitter), onto which inhomogeneous boundary conditions are specified. Homogeneous boundary conditions are specified on the remaining boundary, including the receiving part S_{rx} (i.e., receiver) onto which the backscattered field is to be measured. We also assume that there are no initial conditions, namely that the scattering volume is initially at rest.

Under these considerations, solutions to this problem may be found using the general expression of the solution for the inhomogeneous scalar wave equation with homogeneous initial conditions, defined in (2.12). By directly injecting the definition of the source term $f(\mathbf{r}, t) = f_{\text{op}}\{\varphi\}(\mathbf{r}, t)$ that involves the scattering operator f_{op} defined in (2.7), we can

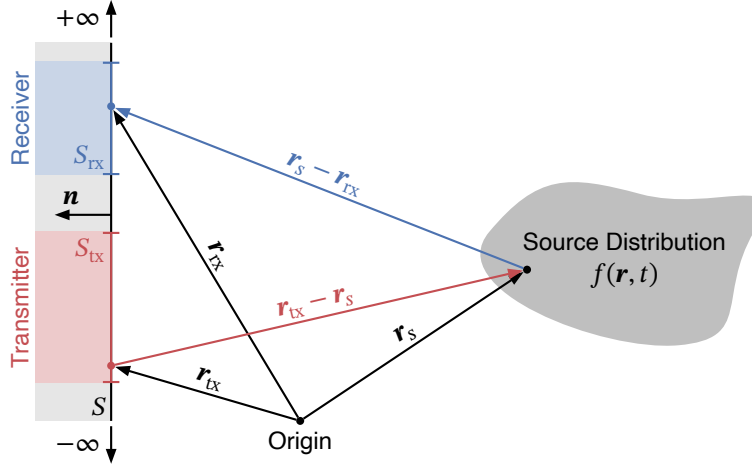


Figure 2.3. Schematic representation of a pulse-echo experiment in the presence of a weakly scattering zone for the derivation of the linear pulse-echo model. The scattering zone is considered to be embedded in a homogeneous medium and is represented by a source distribution. The pulse-echo experiment involves a transmitter that insonifies the volume including the weakly scattering zone and a receiver that senses the echo signals backscattered by the weakly scattering zone. Both the transmitter and the receiver are assumed to be embedded in an infinite planar boundary onto which boundary conditions are specified.

rewrite (2.12), $\forall \mathbf{r} \in V$ and $\forall \mathbf{r} \in S$, as

$$\begin{aligned} \varphi(\mathbf{r}, t) = & \int_T \int_S \left[g(\mathbf{r}, t; \mathbf{r}', t') \frac{\partial \varphi(\mathbf{r}', t')}{\partial \mathbf{n}'} - \varphi(\mathbf{r}', t') \frac{\partial g(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}'} \right] d\sigma(\mathbf{r}') dt' \\ & + \int_T \int_V g(\mathbf{r}, t; \mathbf{r}', t') f_{\text{op}}\{\varphi\}(\mathbf{r}', t') d\nu(\mathbf{r}') dt'. \end{aligned} \quad (2.65)$$

2.5.1. Born Approximation

Since the velocity potential is present on both sides of (2.65) because of the volume integral that applies to the scattering operator, it is not possible to find a solution for the velocity potential directly. In other words, this means that the velocity potential at some point in the volume depends on the integral over the entire volume of a term that itself depends on the velocity potential. This is a typical instance of a rather complex scattering process called *multiple scattering*. One way to solve such a problem consists of relying on the Born-Neumann expansion technique [27], [55]. As we are limiting ourselves to *weak* scattering, we can apply the well-known Born approximation, which is the first order of the Born-Neumann expansion. The Born approximation basically neglects multiple

2. A LINEAR PULSE-ECHO MODEL FOR WEAK SCATTERING

scattering such that (2.65) reduces to

$$\begin{aligned} \varphi(\mathbf{r}, t) = & \int_T \int_S \left[g(\mathbf{r}, t; \mathbf{r}', t') \frac{\partial \varphi(\mathbf{r}', t')}{\partial \mathbf{n}'} - \varphi(\mathbf{r}', t') \frac{\partial g(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}'} \right] d\sigma(\mathbf{r}') dt' \\ & + \int_T \int_V g(\mathbf{r}, t; \mathbf{r}', t') f_{\text{op}\{\bar{\varphi}\}}(\mathbf{r}', t') dv(\mathbf{r}') dt', \end{aligned} \quad (2.66)$$

where the velocity potential φ has been replaced in the volume integral by the *incoming* velocity potential $\bar{\varphi}$. This greatly simplifies the solution as $\bar{\varphi}$ is therefore solution of the homogeneous wave equation with inhomogeneous boundary conditions, namely as if the weakly scattering zone was absent (Section 2.5.3).

2.5.2. Specifying Green's Function

To specify the Green's function with respect to the boundary conditions specified on the boundary surface, we rely on the method of images described in Section 2.4.2. This means that we can use all derivations carried out for Dirichlet and Neumann boundary conditions, in particular the Rayleigh-Sommerfeld equations (Section 2.4.3), as well as the expressions for the corresponding SIRs.

It is however important to note that, by using the method of images described in Section 2.4.2, we are actually neglecting any “secondary” scattering arising from reflections on the surface boundary (i.e., transducer face) that propagate back into the medium and are scattered again by the tissue. This is probably a weaker assumption than neglecting multiple scattering (i.e., Born approximation, Section 2.5.1), first because we assume weak scattering, which means that the scattered field is much smaller than the incoming one, and second because the amplitude of such field decreases proportional to inverse of the propagation distance. Though, the method of images could also be deployed to account for multiple reflections as described in [38, Sec. 7.2], but this is beyond the scope of this work.

2.5.3. Transmit Field

To solve (2.66), we need to evaluate the incoming (or transmit) velocity potential $\bar{\varphi}$, which is, as already mentioned, solution of the homogeneous wave equation with inhomogeneous boundary conditions. It can therefore be evaluated using (2.16), as a surface integral over the transmitter only (since the remaining boundary surface is specified with homogeneous boundary conditions), namely

$$\bar{\varphi}(\mathbf{r}, t) = \int_T \int_{S_{\text{tx}}} \left[g(\mathbf{r}, t; \mathbf{r}', t') \frac{\partial \varphi(\mathbf{r}', t')}{\partial \mathbf{n}'} - \varphi(\mathbf{r}', t') \frac{\partial g(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}'} \right] d\sigma(\mathbf{r}') dt'. \quad (2.67)$$

Using the derivations carried out for the Rayleigh-Sommerfeld equations in Section 2.4.3, it is straightforward to obtain expressions for the transmit field for both soft and hard

baffles.

In the case of Dirichlet boundary conditions (i.e., soft baffle), we can therefore use the (compact) Sommerfeld integral (2.57), namely

$$\bar{\varphi}_D(\mathbf{r}, t) = \frac{\bar{p}_S(t)}{\rho_0 c_0} *_t h_D^{\text{tx}}(\mathbf{r}, t), \quad (2.68)$$

where h_D^{tx} is the SIR for a planar soft baffle as defined in (2.56) and $\bar{p}_S(t)$ is the pressure waveform¹⁰ applied to the transmitter. In the case of Neumann boundary conditions (i.e., hard baffle), we can therefore use the (compact) Rayleigh integral (2.60), namely

$$\bar{\varphi}_N(\mathbf{r}, t) = \bar{v}_n(t) *_t h_N^{\text{tx}}(\mathbf{r}, t), \quad (2.69)$$

where h_N^{tx} is the SIR for a planar hard baffle as defined in (2.59) and $\bar{v}_n(t)$ is the (inward) normal velocity waveform¹⁰ applied to the transmitter.

2.5.4. Receive Field

Equipped with expressions (2.68) and (2.69) to evaluate the transmit velocity potential for both Dirichlet and Neumann boundary conditions, respectively, we are now interested in the evaluation of the scattered field onto the receiver. To do so, we rely on (2.66) for an observer point on the receiver surface, namely $\mathbf{r} \in S_{\text{rx}}$. One can note that, for both Dirichlet and Neumann boundary conditions, the surface integral of (2.66) becomes null, reducing to

$$\varphi(\mathbf{r}, t) = \iint_{TV} g(\mathbf{r}, t; \mathbf{r}', t') f_{\text{op}}\{\bar{\varphi}\}(\mathbf{r}', t') dv(\mathbf{r}') dt', \quad \forall \mathbf{r} \in S_{\text{rx}}. \quad (2.70)$$

This means that the transmit (or incoming) field $\bar{\varphi}$ arising from the inhomogeneous boundary conditions specified on S_{tx} has no influence on S_{rx} .¹¹

2.5.4.1. Dirichlet Boundary Conditions

In the case of Dirichlet boundary conditions (i.e., soft baffle), and as already mentioned in Section 2.3.4, (2.70) is not “usable” as it is zero on both sides (due to the boundary conditions). Therefore, we rely on the normal derivative with respect to \mathbf{r} instead, as in (2.21), namely

$$\frac{\partial \varphi_D(\mathbf{r}, t)}{\partial \mathbf{n}} = \iint_{TV} \frac{\partial g_D(\mathbf{r}, t; \mathbf{r}', t')}{\partial \mathbf{n}} f_{\text{op}}\{\bar{\varphi}_D\}(\mathbf{r}', t') dv(\mathbf{r}') dt'. \quad (2.71)$$

From the expression of $\bar{\varphi}_D$ (2.68) and recalling that f_{op} is a linear operator, as defined in

¹⁰Recall that we assumed separability of time and space variables on the boundary surface.

¹¹This is obviously a crucial aspect for pulse-echo imaging as the scattered field is much weaker than the incoming one.

2. A LINEAR PULSE-ECHO MODEL FOR WEAK SCATTERING

(2.7), one can note that

$$f_{\text{op}}\{\bar{\varphi}_D\}(\mathbf{r}, t) = \frac{1}{\rho_0 c_0} f_{\text{op}}\{\bar{p}_S *_t h_D^{\text{tx}}\}(\mathbf{r}, t) = \frac{\bar{p}_S(t)}{\rho_0 c_0} *_t f_{\text{op}}\{h_D^{\text{tx}}\}(\mathbf{r}, t), \quad (2.72)$$

where the approximation sign from the derivation of the Sommerfeld integral (2.48) has been dropped to lighten notations as further approximations will be introduced hereafter. Using the normal derivative of the Green's function g_D given in (2.39), immediately neglecting the term proportional to $1/\|\mathbf{r} - \mathbf{r}'\|_2^2$ similarly to the Sommerfeld integral derivation, (2.71) becomes

$$\frac{\partial \varphi_D(\mathbf{r}, t)}{\partial \mathbf{n}} = - \int_T \int_V \frac{\delta' \left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right)}{2\pi c_0 \|\mathbf{r} - \mathbf{r}'\|_2} f_{\text{op}}\{\bar{\varphi}_D\}(\mathbf{r}', t') \cos(\mathbf{n}, \mathbf{r} - \mathbf{r}') dv(\mathbf{r}') dt' \quad (2.73)$$

$$= - \int_V \frac{\cos(\mathbf{n}, \mathbf{r} - \mathbf{r}')}{2\pi c_0 \|\mathbf{r} - \mathbf{r}'\|_2} \frac{\partial}{\partial t} \left[f_{\text{op}}\{\bar{\varphi}_D\} \left(\mathbf{r}', t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right) \right] dv(\mathbf{r}'), \quad (2.74)$$

where we used the convolution property of the Dirac delta function derivative (2.45). Introducing a time convolution with a retarded Dirac delta function, we obtain

$$\frac{\partial \varphi_D(\mathbf{r}, t)}{\partial \mathbf{n}} = - \int_V \frac{\delta \left(t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right)}{2\pi c_0 \|\mathbf{r} - \mathbf{r}'\|_2} *_t \frac{\partial}{\partial t} [f_{\text{op}}\{\bar{\varphi}_D\}(\mathbf{r}', t)] \cos(\mathbf{n}, \mathbf{r} - \mathbf{r}') dv(\mathbf{r}'). \quad (2.75)$$

Using (2.72), the normal derivative of the velocity potential can be expressed as

$$\frac{\partial \varphi_D(\mathbf{r}, t)}{\partial \mathbf{n}} = - \frac{\partial}{\partial t} \frac{\bar{p}_S(t)}{\rho_0 c_0^2} *_t \int_V \frac{\delta \left(t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0} \right)}{2\pi \|\mathbf{r} - \mathbf{r}'\|_2} \cos(\mathbf{n}, \mathbf{r} - \mathbf{r}') *_t f_{\text{op}}\{h_D^{\text{tx}}\}(\mathbf{r}', t) dv(\mathbf{r}'). \quad (2.76)$$

The received pulse-echo voltage e_D^{rx} is the surface integral over the receive surface S_{rx} of the time convolution of the *inward* normal velocity and the acousto-electric impulse response of the transducer h_D^{ae} , namely

$$e_D^{\text{rx}}(t) = - \int_{S_{\text{rx}}} h_D^{\text{ae}}(\mathbf{r}, t) *_t \frac{\partial \varphi_D(\mathbf{r}, t)}{\partial \mathbf{n}} d\sigma(\mathbf{r}). \quad (2.77)$$

Assuming separability of variables for the acousto-electric impulse response such that $h_D^{\text{ae}}(\mathbf{r}, t) = \xi_D(\mathbf{r}) \bar{h}_D^{\text{ae}}(t)$, we obtain

$$e_D^{\text{rx}}(t) = - \bar{h}_D^{\text{ae}}(t) *_t \int_{S_{\text{rx}}} \xi_D(\mathbf{r}) \frac{\partial \varphi_D(\mathbf{r}, t)}{\partial \mathbf{n}} d\sigma(\mathbf{r}). \quad (2.78)$$

Using (2.76), the pulse-echo voltage becomes

$$e_D^{\text{rx}}(t) = \bar{h}_D^{\text{ae}}(t) *_{\frac{\partial}{\partial t}} \frac{\bar{p}_S(t)}{\rho_0 c_0^2} *_{\frac{\partial}{\partial t}} \left[\int_V \int_{S_{\text{rx}}} \frac{\xi_D(\mathbf{r}) \delta\left(t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right)}{2\pi \|\mathbf{r} - \mathbf{r}'\|_2} \cos(\mathbf{n}, \mathbf{r} - \mathbf{r}') d\sigma(\mathbf{r}) \right] *_{\frac{\partial}{\partial t}} f_{\text{op}}\{h_D^{\text{tx}}\}(\mathbf{r}', t) dv(\mathbf{r}'). \quad (2.79)$$

Noting that $\cos(\mathbf{n}, \mathbf{r} - \mathbf{r}') = \cos(\tilde{\mathbf{n}}, \mathbf{r}' - \mathbf{r})$, where $\tilde{\mathbf{n}}$ the *inward* normal (i.e., $\tilde{\mathbf{n}} = -\mathbf{n}$), and that $\|\mathbf{r} - \mathbf{r}'\|_2 = \|\mathbf{r}' - \mathbf{r}\|_2$, it is apparent that the surface integral (in brackets) is equal to the SIR in the case of Dirichlet boundary conditions, as defined in (2.56), if an “observer” point were located at \mathbf{r}' and “sources” at $\mathbf{r} \in S_{\text{rx}}$.¹² Therefore,

$$e_D^{\text{rx}}(t) = \bar{h}_D^{\text{ae}}(t) *_{\frac{\partial}{\partial t}} \frac{\bar{p}_S(t)}{\rho_0 c_0^2} *_{\frac{\partial}{\partial t}} \int_V h_D^{\text{rx}}(\mathbf{r}', t) *_{\frac{\partial}{\partial t}} f_{\text{op}}\{h_D^{\text{tx}}\}(\mathbf{r}', t) dv(\mathbf{r}') \quad (2.80)$$

$$= \bar{h}_D^{\text{ae}}(t) *_{\frac{\partial}{\partial t}} \frac{\bar{p}_S(t)}{\rho_0 c_0^2} *_{\frac{\partial}{\partial t}} \int_V f_{\text{op}}\{h_D^{\text{pe}}\}(\mathbf{r}', t) dv(\mathbf{r}'), \quad (2.81)$$

where $h_D^{\text{pe}}(\mathbf{r}, t) = h_D^{\text{tx}}(\mathbf{r}, t) *_{\frac{\partial}{\partial t}} h_D^{\text{rx}}(\mathbf{r}, t)$ is the pulse-echo SIR corresponding to Dirichlet boundary conditions. To further simplify the volume integral following the approximations introduced in [27], it is assumed that the pulse-echo SIR is slowly varying over a finite volume such that

$$\begin{aligned} & \int_V f_{\text{op}}\{h_D^{\text{pe}}\}(\mathbf{r}', t) dv(\mathbf{r}') \\ &= \int_V \left[\frac{2\Delta c(\mathbf{r}')}{c_0^3} \frac{\partial^2 h_D^{\text{pe}}(\mathbf{r}', t)}{\partial t^2} - \frac{1}{\rho_0} \nabla_{\mathbf{r}'}[\Delta \rho(\mathbf{r}')] \cdot \nabla_{\mathbf{r}'} h_D^{\text{pe}}(\mathbf{r}', t) \right] dv(\mathbf{r}') \end{aligned} \quad (2.82)$$

$$= \int_V \left[\frac{2\Delta c(\mathbf{r}')}{c_0^3} \frac{\partial^2 h_D^{\text{pe}}(\mathbf{r}', t)}{\partial t^2} - \frac{\Delta \rho(\mathbf{r}')}{\rho_0} \nabla_{\mathbf{r}'}^2 h_D^{\text{pe}}(\mathbf{r}', t) \right] dv(\mathbf{r}'). \quad (2.83)$$

Assuming only small variations in propagation velocity, the Laplacian operator of the pulse-echo SIR can be related to the second order time derivative as [27]

$$\nabla_{\mathbf{r}'}^2 h_D^{\text{pe}}(\mathbf{r}, t) = \frac{1}{c_0^2} \frac{\partial^2 h_D^{\text{pe}}(\mathbf{r}, t)}{\partial t^2}, \quad (2.84)$$

further simplifying the volume integral as

$$\int_V f_{\text{op}}\{h_D^{\text{pe}}\}(\mathbf{r}', t) dv(\mathbf{r}') = \int_V \left[\frac{2\Delta c(\mathbf{r}')}{c_0} - \frac{\Delta \rho(\mathbf{r}')}{\rho_0} \right] \frac{1}{c_0^2} \frac{\partial^2 h_D^{\text{pe}}(\mathbf{r}', t)}{\partial t^2} dv(\mathbf{r}'). \quad (2.85)$$

¹²This is a typical instance of the acoustic reciprocity theorem stated by Lord Rayleigh [48, Vol. 2, p. 145] as: “If in a space filled with air which is partly bounded by finitely extended fixed bodies and is partly unbounded, sound waves be excited at any point A, the resulting velocity-potential at a second point B is the same both in magnitude and phase, as it would have been at A, had B been the source of the sound.”

2. A LINEAR PULSE-ECHO MODEL FOR WEAK SCATTERING

Injecting (2.85) in (2.81), and grouping all time derivative operations, the pulse-echo voltage can be expressed as

$$e_D^{\text{rx}}(t) = \bar{h}_D^{\text{ae}}(t) *_{\frac{1}{\rho_0 c_0^4}} \frac{\partial^3 \bar{p}_S(t)}{\partial t^3} *_{\frac{1}{\rho_0 c_0^4}} \int_V f_m(\mathbf{r}') h_D^{\text{pe}}(\mathbf{r}', t) d\mathbf{v}(\mathbf{r}'), \quad (2.86)$$

where $f_m(\mathbf{r}) = 2\Delta c(\mathbf{r})/c_0 - \Delta\rho(\mathbf{r})/\rho_0$ is a term accounting for the (weak) fluctuations in density and propagation velocity, giving rise to the scattered fields.

As the pressure waveform specified on the transmitter is the result of the time convolution of the excitation voltage waveform $e_{\text{tx}}(t)$ and the electro-acoustic impulse response of the transducer \bar{h}_D^{ea} , (2.86) can also be expressed as

$$e_D^{\text{rx}}(t) = \frac{1}{\rho_0 c_0^4} \frac{\partial^3 e_{\text{tx}}(t)}{\partial t^3} *_{\frac{1}{\rho_0 c_0^4}} \bar{h}_D^{\text{ea}}(t) *_{\frac{1}{\rho_0 c_0^4}} \bar{h}_D^{\text{ae}}(t) *_{\frac{1}{\rho_0 c_0^4}} \int_V f_m(\mathbf{r}') h_D^{\text{pe}}(\mathbf{r}', t) d\mathbf{v}(\mathbf{r}'). \quad (2.87)$$

2.5.4.2. Neumann Boundary Conditions

The derivation in the case of Neumann boundary conditions (i.e., hard baffle) is performed similarly to the case of Dirichlet boundary conditions, while being sensibly lighter. Contrary to the Dirichlet case, (2.70) can be exploited directly, and can be expressed as

$$\varphi_N(\mathbf{r}, t) = \int_T \int_V g_N(\mathbf{r}, t; \mathbf{r}', t') f_{\text{op}}\{\bar{\varphi}_N\}(\mathbf{r}', t') d\mathbf{v}(\mathbf{r}') dt'. \quad (2.88)$$

Using the Green's function for the Neumann case (2.43), exploiting the convolution property of the Dirac delta function (2.44), and introducing time convolution with a retarded Dirac, we successively obtain

$$\varphi_N(\mathbf{r}, t) = \int_T \int_V \frac{\delta\left(t - t' - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right)}{2\pi\|\mathbf{r} - \mathbf{r}'\|_2} f_{\text{op}}\{\bar{\varphi}_N\}(\mathbf{r}', t') d\mathbf{v}(\mathbf{r}') dt' \quad (2.89)$$

$$= \int_V \frac{1}{2\pi\|\mathbf{r} - \mathbf{r}'\|_2} f_{\text{op}}\{\bar{\varphi}_N\}\left(\mathbf{r}', t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right) d\mathbf{v}(\mathbf{r}') \quad (2.90)$$

$$= \int_V \frac{\delta\left(t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right)}{2\pi\|\mathbf{r} - \mathbf{r}'\|_2} *_{\frac{1}{\rho_0 c_0^4}} f_{\text{op}}\{\bar{\varphi}_N\}(\mathbf{r}', t) d\mathbf{v}(\mathbf{r}'). \quad (2.91)$$

From the expression of $\bar{\varphi}_N$ (2.69), we can expand the scattering operator as

$$f_{\text{op}}\{\bar{\varphi}_N\}(\mathbf{r}, t) = f_{\text{op}}\{\bar{v}_n *_{\frac{1}{\rho_0 c_0^4}} h_N^{\text{tx}}\}(\mathbf{r}, t) = \bar{v}_n(t) *_{\frac{1}{\rho_0 c_0^4}} f_{\text{op}}\{h_N^{\text{tx}}\}(\mathbf{r}, t), \quad (2.92)$$

such that (2.91) becomes

$$\varphi_N(\mathbf{r}, t) = \bar{v}_n(t) *_t \int_V \frac{\delta\left(t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right)}{2\pi\|\mathbf{r} - \mathbf{r}'\|_2} *_t f_{\text{op}}\{h_N^{\text{tx}}\}(\mathbf{r}', t) d\mathbf{v}(\mathbf{r}'). \quad (2.93)$$

The received pulse-echo voltage e_N^{rx} is the surface integral over the receive surface S_{rx} of the time convolution of the pressure distribution and the acousto-electric impulse response of the transducer h_N^{ae} , namely

$$e_N^{\text{rx}}(t) = \int_{S_{\text{rx}}} h_N^{\text{ae}}(\mathbf{r}, t) *_t \rho_0 \frac{\partial \varphi_N(\mathbf{r}, t)}{\partial t} d\sigma(\mathbf{r}), \quad (2.94)$$

where we used the relation (2.5) between the velocity potential and the pressure. Assuming separability of variables for the acousto-electric impulse response such that $h_N^{\text{ae}}(\mathbf{r}, t) = \xi_N(\mathbf{r})\bar{h}_N^{\text{ae}}(t)$, and using (2.94), we obtain

$$e_N^{\text{rx}}(t) = \bar{h}_N^{\text{ae}}(t) *_t \rho_0 \frac{\partial \bar{v}_n(t)}{\partial t} *_t \int_V \left[\int_{S_{\text{rx}}} \frac{\xi_N(\mathbf{r}) \delta\left(t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right)}{2\pi\|\mathbf{r} - \mathbf{r}'\|_2} d\sigma(\mathbf{r}) \right] *_t f_{\text{op}}\{h_N^{\text{tx}}\}(\mathbf{r}', t) d\mathbf{v}(\mathbf{r}'), \quad (2.95)$$

where the surface integral (in brackets) can be identified as the SIR in the case of Neumann boundary conditions, as defined in (2.59), if an “observer” point were located at \mathbf{r}' and “sources” at $\mathbf{r} \in S_{\text{rx}}$.¹³ Therefore,

$$e_N^{\text{rx}}(t) = \bar{h}_N^{\text{ae}}(t) *_t \rho_0 \frac{\partial \bar{v}_n(t)}{\partial t} *_t \int_V h_N^{\text{rx}}(\mathbf{r}', t) *_t f_{\text{op}}\{h_N^{\text{tx}}\}(\mathbf{r}', t) d\mathbf{v}(\mathbf{r}') \quad (2.96)$$

$$= \bar{h}_N^{\text{ae}}(t) *_t \rho_0 \frac{\partial \bar{v}_n(t)}{\partial t} *_t \int_V f_{\text{op}}\{h_N^{\text{pe}}\}(\mathbf{r}', t) d\mathbf{v}(\mathbf{r}'), \quad (2.97)$$

where $h_N^{\text{pe}}(\mathbf{r}, t) = h_N^{\text{tx}}(\mathbf{r}, t) *_t h_N^{\text{rx}}(\mathbf{r}, t)$ is the pulse-echo SIR corresponding to Neumann boundary conditions. Relying on the same approximations on the pulse-echo SIR as in the Dirichlet case to simplify the volume integral using (2.85), and grouping all time derivative operations in the same manner, one can obtain a similar expression for the pulse-echo voltage in the Neumann case, namely

$$e_N^{\text{rx}}(t) = \bar{h}_N^{\text{ae}}(t) *_t \frac{\rho_0}{c_0^2} \frac{\partial^3 \bar{v}_n(t)}{\partial t^3} *_t \int_V f_m(\mathbf{r}') h_N^{\text{pe}}(\mathbf{r}', t) d\mathbf{v}(\mathbf{r}'), \quad (2.98)$$

where $f_m(\mathbf{r}) = 2\Delta c(\mathbf{r})/c_0 - \Delta\rho(\mathbf{r})/\rho_0$ is the same term accounting for the (weak) fluctuations in density and propagation velocity as in (2.86).

As the Neumann boundary conditions case was derived by Jensen in his foundation

¹³Another typical instance of the acoustic reciprocity theorem stated by Lord Rayleigh [48, Vol. 2, p. 145].

2. A LINEAR PULSE-ECHO MODEL FOR WEAK SCATTERING

work [27], two slight differences can be noted. First, there is an additional 1/2 factor in [27], which comes from the selection of the unbounded Green's function to derive the receive field. This choice implies that the boundary condition due to the presence of the transducer is neglected on receive but not on transmit. Second, there is a sign difference in the resulting inhomogeneity function f_m , which most probably comes from different sign conventions for the source function of the inhomogeneous scalar wave equation (2.8).

As the normal velocity waveform specified on the transmitter is the result of the time convolution of the excitation voltage waveform e_{tx} and the electro-acoustic impulse response of the transducer \bar{h}_N^{ea} , (2.98) can also be expressed as

$$e_N^{rx}(t) = \frac{\rho_0}{c_0^2} \frac{\partial^3 e_{tx}(t)}{\partial t^3} *_t \bar{h}_N^{ea}(t) *_t \bar{h}_N^{ae}(t) *_t \int_V f_m(\mathbf{r}') h_N^{pe}(\mathbf{r}', t) d\mathbf{v}(\mathbf{r}'). \quad (2.99)$$

2.5.5. Generic Pulse-Echo Model

Inspection of the relations (2.87) and (2.99) derived in the case of Dirichlet (soft) and Neumann (hard) boundary conditions clearly reveals their overall similarity. Let us abstract the specifics of the electro-acoustic and acousto-electric impulse responses of the transmit and receive transducers involved, and assume that the scaling constants as well as the third-order time derivative are accounted for by generic impulse responses defined as h_{ae} and h_{ea} . Using the generic expression for the SIR defined in (2.61), one can rewrite both (2.87) and (2.99) as

$$e_{rx}(t) = e_{tx}(t) *_t h_{ea}(t) *_t h_{ae}(t) *_t \int_V \left[h_{tx}(\mathbf{r}, t) *_t h_{rx}(\mathbf{r}, t) \right] f_m(\mathbf{r}) d\mathbf{v}(\mathbf{r}), \quad (2.100)$$

for which the pulse-echo SIR has been split back to a transmit-receive convolutional form, and where we replaced \mathbf{r}' by \mathbf{r} in the volume integral to lighten notations as there is no more ambiguity at this stage between (volume) sources and (evaluation) field points (i.e., observers).

As highlighted in [27], the signal of interest that should be displayed by an US scanner is

$$f_m(\mathbf{r}) = \frac{2\Delta c(\mathbf{r})}{c_0} - \frac{\Delta \rho(\mathbf{r})}{\rho_0}, \quad (2.101)$$

as it represents the local fluctuations in density and propagation velocity, and thus gives rise to echo signals. However, the measurements obtained of f_m are blurred and deformed by the spatial and time convolutions involved in (2.100). In particular, due to the time convolution with the pulse-echo waveform¹⁴

$$v_{pe}(t) = e_{tx}(t) *_t h_{ea}(t) *_t h_{ae}(t), \quad (2.102)$$

¹⁴Originally called pulse-echo wavelet in [27].

and the spatial convolution with the spatially varying pulse-echo SIR

$$h_{\text{pe}}(\mathbf{r}, t) = h_{\text{tx}}(\mathbf{r}, t) *_{\mathbf{r}} h_{\text{rx}}(\mathbf{r}, t). \quad (2.103)$$

It is important to keep in mind that these generalizations require careful usage of suitable impulse responses, and especially the SIRs, depending on the boundary conditions assumed for the transducers. Hard baffles (i.e., Neumann boundary conditions) were extensively used in deriving analytic expressions for the corresponding SIRs of transducer (radiators) with specific geometries [43], [51]–[54], [56], [57]. Good agreements with respect to measurements were obtained when considering hard baffle conditions for transducers with characteristic lengths of many wavelengths [27], [58]. Fewer analytic expressions were derived in the case of soft baffles (i.e., Dirichlet boundary conditions) for the corresponding SIRs [59], maybe due to the added complexity of the cosine term in (2.56). Yet, thanks to that additional cosine term, soft baffle conditions were proved to better fit the radiation pattern of small elements with respect to the characteristic wavelength [49], [60]–[63]. Transducer arrays deployed in current US scanners are typically composed many “small” elements, for instance tens to hundreds of them compose 1-D phased or linear arrays for 2-D imaging, and thousands of them compose 2-D matrix arrays for 3-D imaging. Therefore, the soft baffle condition is of particular interest to this work.

2.6. Conclusion

In this chapter, we have derived a linear pulse-echo model accounting for the diffraction effect of infinite planar baffles (i.e., SIR model) and for the scattering of acoustic waves propagating through weakly inhomogeneous tissue. The derivation is based on the work of Jensen [27] and was generalized to account for both Dirichlet (soft) and Neumann (hard) baffle boundary conditions. For this purpose, the derivation was carried out in the transient case (i.e., time domain) on the velocity potential rather than the pressure. The resulting expression for the Neumann boundary condition match the one derived in [27].

Most importantly, this model is at the core of most (if not all) approaches presented in this thesis. As such, a deep understanding of the physical assumptions made to arrive to such a model is of primary importance. For instance, this model is the source of a further-approximated, computationally tractable, pulse-echo model presented in Chapter 3 for the purpose of US imaging. Also, and as it will be discussed in Chapter 3, efficient and accurate computer simulations were needed for the generation of large datasets to train CNN-based image reconstruction methods (Chapters 5 to 7). The SIR US simulator developed for this purpose is presented in Chapter 4 and directly derives from the pulse-echo model for weak scattering presented in the present chapter.

3. Efficient Modeling for Ultrafast Ultrasound Imaging

3.1. Introduction

As introduced in Section 1.1, images reconstructed from ultrafast US acquisitions are of low quality compared with conventional line-by-line scanning. This is obviously expected because ultrafast US images are reconstructed from single unfocused transmit wavefronts as opposed to many focused beams in line-by-line scanning. As a result, images reconstructed from ultrafast US acquisitions using a conventional DAS beamforming algorithm are characterized by PSFs with broad main lobes (poor lateral resolution) and high secondary lobes (low contrast) [6]–[8], [11], [16], [17]. These secondary lobes are typically the result of GL, SL, and EW artifacts, and will be referred to as *diffraction artifacts*. Because these artifacts may shadow tissue regions of interest, they can be detrimental to lesion detection. Moreover, they can reduce the accuracy of displacement estimation techniques that are fundamental to ultrafast US imaging modes such as shear-wave elastography [6]. Coherent compounding has emerged as the state-of-the-art technique to improve the quality of ultrafast US images [1], [6], [8], [16], [18]. This process consists of reconstructing a sequence of low-quality images from differently steered unfocused wavefronts that are coherently compounded to form an image of higher quality, at the expense of a lower achievable frame rate.

This chapter presents our initial approach, based on SR, to reconstructing high-quality US images from unfocused wavefronts. In particular, we present a generalization of the forward model developed throughout the years for the purpose of ultrafast US imaging. The proposed model derives from the generic pulse-echo model for weak scattering presented in Chapter 2. Because the latter involves complex surface integrals for the evaluation of the SIR of each transducer element composing an US transducer array, further approximations were required to obtain a computationally tractable forward model suitable for image reconstruction methods. The approximated model presented in this chapter relies on far-field assumptions for transducer elements and on ideal wavefront assumption for the transmitted unfocused wavefronts used in ultrafast US imaging, allowing to reduce the complexity drastically.

This model has been successfully applied to different US imaging problems such as sparsely regularized beamforming for enhanced ultrafast US imaging [25], [64, Sec. 4.2], compressed US beamforming [25], [64, Sec. 4.3], and deconvolution of US images [29], [64, Sec. 4.4]. Of particular interest to this thesis is the SR approach for enhanced ultrafast US imaging, for which a sparsity-promoting term in a concatenation of wavelet bases was

used as regularizer. This regularizer coupled to the proposed forward model enabled us to obtain promising images with sensibly reduced diffraction artifacts. The experiments presented in this chapter also highlight the main issues observed with this approach. In particular, the fact that we did not find a conventional image-processing regularizer capable of reducing diffraction artifacts while restoring and preserving accurate speckle patterns.

3.2. Proposed Methods

3.2.1. Generic Approximated Pulse-Echo Model for Imaging

Let us recall from Chapter 2 that under the first-order Born approximation, assuming (longitudinal) linear acoustics, and neglecting dispersive attenuation, the signal received (e.g., by a transducer element) from a weakly scattering medium Ω embedded in a homogeneous medium and insonified by a transmitter (e.g., wavefront) can be (compactly) expressed from (2.100) as

$$y_{i,j}(t) = v_{pe}(t) *_t \int_{\Omega} \left[h_i^{tx}(\mathbf{r}, t) *_t h_j^{rx}(\mathbf{r}, t) \right] x(\mathbf{r}) d\mathbf{v}(\mathbf{r}), \quad (3.1)$$

where $*_t$ denotes time convolution. The terms h_i^{tx} and h_j^{rx} represent the SIRs of the transmitter and receiver, respectively. The pulse-echo waveform,¹ as defined in (2.102), v_{pe} includes both electro-acoustic (transmit) and acousto-electric (receive) impulse responses as well as the electric excitation waveform (assumed identical for all transducer elements). The signal of interest (i.e., image) is defined from (2.101) as

$$x(\mathbf{r}) = \frac{2\Delta c(\mathbf{r})}{c_0} - \frac{\Delta \rho(\mathbf{r})}{\rho_0}, \quad (3.2)$$

and accounts for local fluctuations in density and propagation velocity, which induce the scattered echo signal (i.e., scattering term). Note that we used the symbol x instead of f_m for the scattering term, and y instead of e_{rx} for the electric signal received. One can also note from (3.2) that because the scattering term x is a composition of physical quantities (i.e., local fluctuations in density and propagation velocity), an image reconstruction method based on (3.1) will not be able to provide physical quantities. This is why conventional (B-mode) US imaging is not a quantitative method but rather a contrast imaging modality.

The main (and only) issue of (3.1) in the context of imaging, in particular when considering real-time imaging, is the high computational complexity of the SIRs involved.² Therefore, to obtain a computationally tractable (forward) model for use in image reconstruction methods, further assumptions must be made. We limit ourselves to array imaging and in particular to ultrafast imaging, namely performed by insonifying the re-

¹Originally called pulse-echo wavelet in [27].

²This will be studied in detail in Chapter 4.

gion of interest using unfocused wavefronts such as PWs or DWs. Ultrasound transducer arrays are composed of many “small” transducer elements, typically close to the excitation wavelength. By small, it is meant that their characteristic length in the imaging dimension is small. For instance, transducer elements composing 1-D arrays are characterized by a width close to the wavelength (i.e., characteristic length) and a height than can attain few tens of wavelengths. Since imaging is usually performed at few tens of wavelengths, we will assume that a receiving element (i.e., receiver) is sufficiently far away from all field points composing the image of interest such that far-field approximation holds. For the transmitter, we will assume that the transmit wavefront generated by the transducer array by applying suitable delays (and potential apodization) to its elements is an ideal wavefront.

With these assumptions in mind, we propose to (drastically) simplify the complexity of the transmit and receive SIRs of (3.1), such that they can be approximated generically as

$$h_S(\mathbf{r}, t) \approx \bar{h}_S(\mathbf{r})\delta(t - \tau_S(\mathbf{r})), \quad (3.3)$$

where δ is the Dirac delta function. The term \bar{h}_S is a scalar function representing the (far-field) acoustic diffraction effects of the aperture S (i.e., transmitter or receiver) to a field point \mathbf{r} . Similarly, the term τ_S is a scalar function representing the acoustic wave propagation time from the aperture to a field point \mathbf{r} . Equation (3.3) is obviously a drastic approximation of the complex surface integral over a radiating surface involved in the evaluation of the SIR, as derived in Chapter 2 and expressed in (2.61). As such, it requires careful usage, and it will be made clear to why this approximation was considered. Assuming that such an approximation is acceptable, (3.3) reduces the evaluation of the (approximated) SIR to some weighting function \bar{h}_S and a time delay τ_S . Injecting these definitions in (3.1), we obtain

$$y_{i,j}(t) \approx v_{pe}(t) * \int_{\Omega} \bar{h}_i^{tx}(\mathbf{r})\bar{h}_j^{rx}(\mathbf{r})\delta(t - \tau_i^{tx}(\mathbf{r}) - \tau_j^{rx}(\mathbf{r}))x(\mathbf{r}) dv(\mathbf{r}), \quad (3.4)$$

where \bar{h}_i^{tx} and \bar{h}_j^{rx} are scalar functions representing the (far-field) acoustic diffraction effects of the transmitter and the receiver to and from a field point \mathbf{r} , respectively. The terms τ_i^{tx} and τ_j^{rx} represent the acoustic wave propagation times from the transmitter and the receiver to a field point \mathbf{r} , respectively. Their sum represents what is often referred to as the round-trip time-of-flight. An illustration of the proposed (forward) pulse-echo model in the case of 2-D PW imaging is depicted in Figure 3.1.

Equation (3.4) is a generalization of the forward model developed throughout the years in our laboratory for the purpose of US imaging. It has been successfully applied to different US imaging problems such as sparsely regularized beamforming for enhanced ultrafast US imaging [25], [64, Sec. 4.2], compressed US beamforming [25], [64, Sec. 4.3], and deconvolution of US images [29], [64, Sec. 4.4]. Thanks to a versatility that goes beyond the scope of the present thesis, we termed the reconstruction framework deriving from it the ultrasound sparse regularization framework (USSR) [28]. From these different works conducted, one is of particular interest to the present thesis, namely the sparsely

3. EFFICIENT MODELING FOR ULTRAFAST ULTRASOUND IMAGING

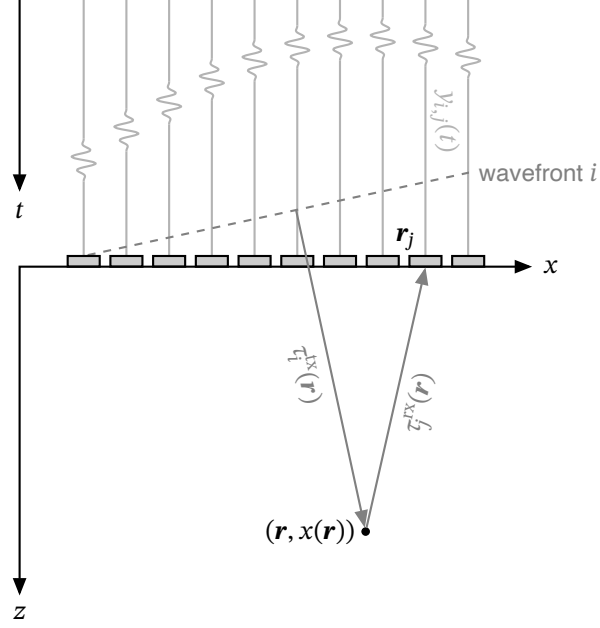


Figure 3.1. Illustration of a 2-D plane-wave (PW) imaging configuration. The PW wavefront is generated by applying suitable delays (and potential apodization) to the transducer elements composing the array. Backscattered echoes (here from a single scatterer) are received on all transducer elements (i.e., full aperture). (Adapted from [25], [30].)

regularized beamforming with its promise to enhance US images from few insonifications.

3.2.2. Inverse Problem and Imaging Methods

For the purpose of conventional US imaging, and by that we mean to reconstruct images with proper speckle patterns suitable for displacement estimation, we will further assume the pulse-echo waveform v_{pe} to be a Dirac delta function such that (3.4) can be rewritten as

$$y_{i,j}(t) \approx \int_{\Omega} \bar{h}_i^{tx}(\mathbf{r}) \bar{h}_j^{rx}(\mathbf{r}) \delta(t - \tau_i^{tx}(\mathbf{r}) - \tau_j^{rx}(\mathbf{r})) x(\mathbf{r}) d\mathbf{v}(\mathbf{r}). \quad (3.5)$$

Ultrasound transducers typically comprise a set of n_r receivers (i.e., transducer elements) arranged in a regular array (e.g., Figure 3.1), and echo signals (i.e., measurements) are sampled at discrete time intervals. As images are composed of discrete (pixel) values, the (approximated) pulse-echo model defined in (3.5) can be conveniently expressed as a discretized operation for all receivers as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (3.6)$$

where $\mathbf{H}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is the measurement (matrix) operator, $\mathbf{x} \in \mathbb{R}^n$ is the (vectorized) image we seek to recover, $\mathbf{y} \in \mathbb{R}^m$ are the (vectorized) transducer elements measurements, and $\mathbf{n} \in \mathbb{R}^m$ is the measurement noise. Note that the transmitter index i has been omitted

in the matrix expressions to lighten notations. It is thus assumed that \mathbf{H} accounts for all transmit-receive events that led to the acquisition of the corresponding (vectorized) signals $\mathbf{y} \in \mathbb{R}^m$. Recovering \mathbf{x} from \mathbf{y} is a well-known inverse problem in medical imaging and may be addressed using various strategies (see, e.g., [32] for CT imaging).

Classical backprojection algorithms, which exploit the adjoint operator $\mathbf{H}^* : \mathbb{R}^m \rightarrow \mathbb{R}^n$, may be used to address such an inverse problem. Using (3.5), one can express the adjoint operation in the continuous domain for all receivers and for the i -th transmitter as [29]

$$\tilde{x}_i(\mathbf{r}) = \bar{h}_i^{\text{tx}}(\mathbf{r}) \sum_{j=1}^{n_r} \bar{h}_j^{\text{rx}}(\mathbf{r}) y_{i,j} (\tau_i^{\text{tx}}(\mathbf{r}) + \tau_j^{\text{rx}}(\mathbf{r})), \quad (3.7)$$

which is an instance of the well-known DAS algorithm, where \bar{h}_i^{tx} and \bar{h}_j^{rx} can be interpreted as weighting functions, often referred to as apodization functions of conventional DAS algorithms. This observation is not “magical,” as all the approximations made from (3.1) to (3.5) were precisely intended to obtain an explicit backprojection-based formulation for the well-known DAS operator. As the resulting RF image \tilde{x}_i is obtained from a single insonification (i.e., i -th transmitter), it usually suffers from strong diffraction artifacts. To improve the resulting image quality, it is common to average coherently multiple (low-quality) RF images reconstructed using DAS-based algorithms from measurements acquired by insonifying the medium of interest differently, for instance using steered PWs or DWs in a process called coherent compounding [65], or using SA imaging [19]. In the frame of the proposed formulation, this process amounts to obtaining an image estimate from n_t transmit wavefronts as

$$\tilde{\mathbf{x}}(\mathbf{r}) = \frac{1}{n_t} \sum_{i=1}^{n_t} \bar{h}_i^{\text{tx}}(\mathbf{r}) \sum_{j=1}^{n_r} \bar{h}_j^{\text{rx}}(\mathbf{r}) y_{i,j} (\tau_i^{\text{tx}}(\mathbf{r}) + \tau_j^{\text{rx}}(\mathbf{r})). \quad (3.8)$$

As an alternative to using multiple insonifications to improve the image quality, regularization techniques may be used by solving

$$\hat{\mathbf{x}} = \underset{\mathbf{x} \in \mathbb{R}^n}{\operatorname{argmin}} \{ \mathcal{D}(\mathbf{H}\mathbf{x}, \mathbf{y}) + \kappa \mathcal{R}(\mathbf{x}) \}, \quad (3.9)$$

where $\mathcal{D} : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}_+$ is a data fidelity term (e.g., the ℓ_2 -norm), and $\mathcal{R} : \mathbb{R}^n \rightarrow \mathbb{R}_+$ is a regularizer used to infer prior knowledge on the expected image. The regularization parameter $\kappa \in \mathbb{R}_+$ controls the weighting of the regularization and is typically adjusted manually. One solution to (3.9), when using the ℓ_2 -norm as data fidelity, may be found using the well-known proximal gradient descent iteration [66]

$$\mathbf{x}^{(k+1)} = \operatorname{prox}_{\gamma \kappa \mathcal{R}}(\mathbf{x}^{(k)} - \gamma \mathbf{H}^* \mathbf{H} \mathbf{x}^{(k)} + \mathbf{H}^* \mathbf{y}), \quad (3.10)$$

where \mathbf{H}^* is the adjoint of \mathbf{H} , $\gamma \in \mathbb{R}_+$ is the gradient step size (that may also be iteration-dependent), and the proximity operator $\operatorname{prox}_{\mu \mathcal{R}} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is defined as

$$\operatorname{prox}_{\mu \mathcal{R}}(\mathbf{z}) = \underset{\mathbf{z} \in \mathbb{R}^n}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{x} - \mathbf{z}\|_2^2 + \mu \mathcal{R}(\mathbf{x}), \quad (3.11)$$

3. EFFICIENT MODELING FOR ULTRAFAST ULTRASOUND IMAGING

for some parameter $\mu \in \mathbb{R}_+$. The proximity operator in (3.10) acts as a projection to “denoise” each estimate based on some prior knowledge of \mathbf{x} . It is thus of primary importance and should be selected with care depending on known properties of \mathbf{x} and on the expected “noise.”

Among the many regularizers benchmarked in the context of US imaging, we observed that a promising regularizer for the purpose of reducing diffraction artifacts (i.e., noise) while preserving speckle patterns (i.e., signal of interest) is to use a sparsity-promoting term in a transformed domain

$$\mathcal{R}(\mathbf{x}) = \|\Psi^* \mathbf{x}\|_1. \quad (3.12)$$

For the sparsity-promoting transform $\Psi: \mathbb{R}^d \rightarrow \mathbb{R}^n$, $d \geq n$, we relied on the sparsity averaging model [67]. It is expressed as a concatenation of p Parseval frames

$$\Psi = \frac{1}{\sqrt{p}}[\Psi_1, \dots, \Psi_p]. \quad (3.13)$$

From the extensive benchmarks conducted, it was observed that setting p equal to eight and using for Ψ_k the k -th Daubechies multi-level inverse wavelet transform with two decomposition levels, provided good results in the context of US imaging. Note that because the sparsity-promoting transform (3.13) is a concatenation of normalized tight frames (Daubechies wavelet transforms), the proximity operator corresponding to the resulting sparsity-promoting term (3.12) amounts to soft-thresholding the transformed image samples (i.e., wavelet coefficients).

3.2.3. Implementation and Interpretation Notes on the Approximated Model

Inspection of (3.5) reveals that at each time t (i.e., at each time sample), a volume integral over Ω needs to be evaluated. This would obviously be prohibitive on a computational perspective. Provided that the implicit function $g_{i,j}(\mathbf{r}, t) = t - \tau_i^{\text{tx}}(\mathbf{r}) - \tau_j^{\text{rx}}(\mathbf{r})$ is continuously differentiable such that $|\nabla_{\mathbf{r}} g_{i,j}(\mathbf{r}, t)| \neq 0$, $\forall \mathbf{r} \in \Omega$, $\forall t \geq 0$, (3.5) may be rewritten as

$$y_{i,j}(t) \approx \int_{\Gamma_{i,j}(t)} \frac{\bar{h}_i^{\text{tx}}(\mathbf{r}) \bar{h}_j^{\text{rx}}(\mathbf{r})}{|\nabla_{\mathbf{r}} g_{i,j}(\mathbf{r}, t)|} x(\mathbf{r}) d\sigma(\mathbf{r}). \quad (3.14)$$

This step, crucial on a computational perspective, reduces the volume integral to a surface integral over the zero-level set of $g_{i,j}(\mathbf{r}, t)$ defined as $\Gamma_{i,j}(t) = \{\mathbf{r} \in \Omega : g_{i,j}(\mathbf{r}, t) = 0\}$. It is interesting to note that these surfaces are quadric surfaces, and that they are 1-D conics when considering 2-D imaging. Thus, the set of these surfaces (or curves) $\{\Gamma_{i,j}\}$ onto which integrals are to be performed can be efficiently parametrized. An illustration in 2-D is depicted in Figure 3.2 and many derivations can be found in [64, Sec. 3.4 to 3.6]. For instance, in the case of 2-D PW imaging, the set of 1-D conics is a set of parabolas.

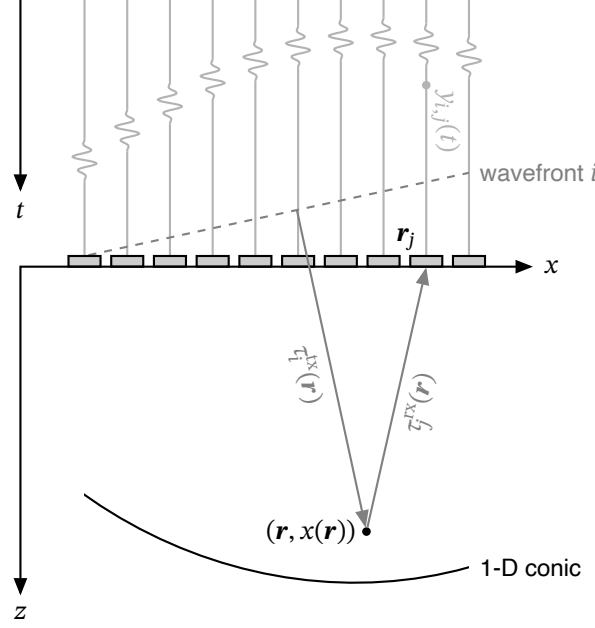


Figure 3.2. Two-dimensional interpretation of the approximated measurement model as the curvilinear integral over a 1-D conic for a specific time instant of the signal acquired on the j -th receiver from the i -th transmit wavefront (i.e., transmitter). The time instant is represented by a filled circle in the corresponding signal. The corresponding 1-D conic is a parabola in the case of 2-D plane wave (PW) imaging. (Adapted from [25], [30].)

Furthermore, applying numerical integration (quadrature) to (3.14), we obtain

$$y_{i,j}(t) \approx \sum_{q=1}^{n_q} \frac{\bar{h}_i^{\text{tx}}(\mathbf{r}_q) \bar{h}_j^{\text{rx}}(\mathbf{r}_q)}{|\nabla_{\mathbf{r}} g_{i,j}(\mathbf{r}, t)|} w_q |\det \mathbf{J}_{\Gamma}(\mathbf{r}_q, t)| x(\mathbf{r}_q), \quad (3.15)$$

where $\{w_q\}$, $\{\det \mathbf{J}_{\Gamma}(\mathbf{r}_q, t)\}$, and $\{\mathbf{r}_q\}$, are the corresponding quadrature weights, Jacobian determinants, and quadrature points, respectively. The numerical evaluation of (3.14) by means of (3.15) amounts to a weighted sum of values over some surface (or curve in 1-D). It is thus akin to a DAS operator in some sense, but applied on the image rather than on the signals. With these derivations, one can note that we end up with a forward model \mathbf{H} and its adjoint \mathbf{H}^* that are similar operators. As such, they have a similar computational complexity, and both involve operations that can be performed in parallel such that efficient graphics processing unit (GPU)-based implementations are possible. An important consideration when implementing (3.7) and (3.15) is the interpolation that needs to be performed between time samples for the adjoint operator \mathbf{H}^* and between spatial samples for the forward model \mathbf{H} (i.e., at quadrature point positions). Here again, they can be performed similarly, and efficiently, using generalized interpolation [68], such that B-spline interpolation can be leveraged with a minimal overhead in computational cost (i.e., by means of pre-filtering). Generalized interpolation will be addressed in detail in Chapter 4.

3.3. Experiments

We are interested in assessing the image quality of the SR approach described in Section 3.2 from single-PW insonifications with normal incidence. For this purpose, we considered the imaging configuration and corresponding acquisition data provided by the PICMUS³ [23]. Two different sets of data were considered, namely data acquired (i.e., simulated) on the numerical PICMUS phantom, and data acquired on an *in vivo* carotid (longitudinal view). We compared images reconstructed using the proposed SR approach from single-PW insonifications to images reconstructed by the corresponding backprojection-based DAS operator from single-PW and 75-PW insonifications (considered as reference).

Since the proposed approximated pulse-echo model assumes ideal wavefronts, in case of PW insonifications the corresponding weighting function $\bar{h}_i^{\text{tx}}(\mathbf{r}) = 1$. The diffraction effect of a transducer element (receiver) evaluated at a field point \mathbf{r} in the transducer plane can be derived from the far-field approximation for continuous waves of a plane rectangle. Considering a soft baffle boundary condition and proceeding similarly to [49], [61], it can be expressed, using [69], as

$$\bar{h}_j^{\text{rx}}(\mathbf{r}) = \frac{A \operatorname{sinc}(d/\lambda \sin(\theta))}{2\pi \|\mathbf{r} - \mathbf{r}_j\|_2} \cos(\theta), \quad (3.16)$$

where \mathbf{r}_j is the position (center) of the transducer element, θ is the angle between normal vector at the element center and the vector $\mathbf{r} - \mathbf{r}_j$, A and d are the surface area and width of the transducer element, respectively, and $\operatorname{sinc}(x) := \sin(\pi x)/(\pi x)$. The delay function representing the acoustic wave propagation time of the PW transmit wavefront τ_i^{tx} is straightforward to evaluate (Figure 3.2). It is equal to the distance between the field point \mathbf{r} and its projection onto the initial planar wavefront (initial reference time), divided by the mean sound speed of the medium c_0 . The receive delay function is also straightforward and can be expressed as $\tau_j^{\text{rx}}(\mathbf{r}) = (1/c_0) \|\mathbf{r} - \mathbf{r}_j\|_2$.

For the SR approach, we relied on the well-known fast iterative shrinkage-thresholding algorithm (FISTA) [70], an accelerated version of the proximal gradient descent iteration (3.10) proposed by Beck and Teboulle. The algorithm was stopped when the mean squared error (MSE) of two consecutive solution candidates was smaller than 10^{-3} . The regularization parameter χ was optimized by grid search for each case considered.

Image reconstruction quality was assessed quantitatively on the numerical PICMUS phantom (Figure 3.3). We evaluated a subset of the image quality metrics provided on the numerical PICMUS phantom. In particular, we considered the contrast-to-noise ratio (CNR), evaluated on the circular anechoic inclusion. We also computed both the average lateral and average axial resolution, namely full width at half maximum (FWHM), in the near field (i.e., at a depth of 14 mm) and in the far field (i.e., at a depth of 45 mm). Both US-specific quality tests were considered, namely statistical speckle tests performed in

³All specifications are detailed on the PICMUS evaluation platform <https://www.creatis.insa-lyon.fr/EvaluationPlatform/picmus/index.html>.

different fully developed zones of the phantom, and the log-linear-intensity test performed on the rectangular inclusion spanning the width of the image. For each RF US image reconstructed, the corresponding B-mode image was obtained by applying conventional envelope detection, normalization, and log-compression.

3.4. Results

B-mode image representations of images reconstructed using the three image reconstruction approach considered, the proposed backprojection-based DAS operator from a single-PW acquisition (1 PW + DAS), the proposed SR approach from a single-PW acquisition (1 PW + SR), and the proposed backprojection-based DAS operator from 75 PW acquisitions (75 PWs + DAS, considered as reference) are shown in Figure 3.3. The corresponding image quality metrics are reported in Table 3.1. Visual comparison of 1 PW + DAS and 1 PW + SR shows that the proposed SR approach is capable of reducing diffraction artifacts, especially in the circular anechoic zone. This results in a CNR improved by 3.4 dB (Table 3.1). Yet, it cannot properly deal with the large GL artifact, resulting from the high-echogenic (left) part of the rectangular gradient. Resolution metrics demonstrate that SR is capable of improving both the lateral and the axial resolutions, and even outperform the reference method (75 PWs + DAS). On the other hand, SR results in the failing of three speckle tests (i.e., half of them). The linearity also test fails with SR. This can be observed on the corresponding B-mode image on which an important portion of the rectangular gradient starts to be completely thresholded (at approximately 5 mm in the lateral dimension).

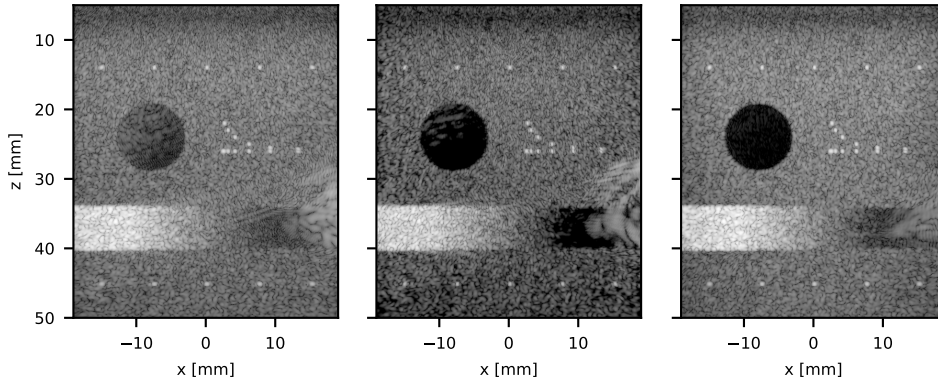


Figure 3.3. B-mode image representations (60-dB range) of the numerical plane-wave imaging challenge in medical ultrasound (PICMUS) phantom reconstructed using (from left to right) the proposed backprojection-based delay-and-sum (DAS) operator from a single plane wave (PW) acquisition, the proposed sparse regularization (SR) approach from a single PW acquisition, and the proposed backprojection-based DAS operator from 75 PW acquisitions (considered as reference). (Modified from [30].)

Figure 3.4 shows B-mode image representations for the three different methods on the longitudinal view of an *in vivo* carotid. We can observe that SR is capable of reducing diffraction artifacts compared with DAS beamforming from a single-PW acquisition. This

3. EFFICIENT MODELING FOR ULTRAFAST ULTRASOUND IMAGING

TABLE 3.1
Image Quality Metrics Computed on the PICMUS Numerical Phantom

Method	CNR (dB)	Lat. Res. (mm)		Ax. Res. (mm)		Speckle Tests	Linearity Test
		14 mm	45 mm	14 mm	45 mm	Passed	Passed
1 PW + DAS	7.2	0.36	0.53	0.38	0.41	6/6	Yes
1 PW + SR	10.6	0.23	0.31	0.33	0.37	3/6	No
75 PWs + DAS	16.4	0.32	0.46	0.39	0.40	6/6	Yes

effect is particularly visible on the many SL artifacts originating from the high-echogenic carotid walls. It seems that overall, it performs well provided that these diffraction artifacts are not too extreme.

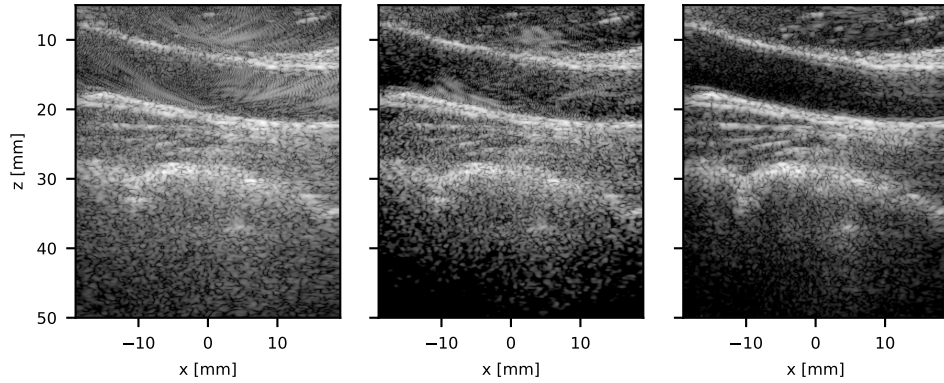


Figure 3.4. B-mode image representations (60-dB range) of an *in vivo* carotid (longitudinal view) reconstructed using (from left to right) the proposed backprojection-based delay-and-sum (DAS) operator from a single plane wave (PW) acquisition, the proposed sparse regularization (SR) approach from a single PW acquisition, and the proposed backprojection-based DAS operator from 75 PW acquisitions (considered as reference). (Modified from [30].)

3.5. Discussion

3.5.1. Generic Approximated Pulse-Echo Model

In Section 3.2.1, we derived a generic approximation of the linear pulse-echo model for weak scattering (Chapter 2) by assuming a specific set of (sometimes drastic) approximations. These approximations, namely splitting the SIR as the product of a spatially dependent weighting function and a shifted Dirac delta function (3.3), and assuming the pulse-echo waveform to be a Dirac delta function, enabled us to derive a forward model whose adjoint (backprojection) is similar to a conventional DAS operator. When interpreted as a DAS operator, the weighting functions \bar{h}_i^{tx} and \bar{h}_j^{rx} of the proposed backprojection-based DAS operator are apodization functions that are generally used to shape the resulting PSF of the imaging system. In particular, they are generally optimized to attain a satisfactory trade-off between the lateral resolution of the PSF (main lobe) and secondary lobes

such as those resulting from GL, SL, and EW artifacts. These artifacts are significant in ultrafast US imaging, in particular when using linear transducer arrays. With the proposed backprojection-based DAS operator, these weighting functions are not optimized “manually”, but are rather derived from far-field approximations. In general, this results in corresponding apodization functions that are less restrictive than conventional ones, with corresponding system PSFs that have tighter main lobes (better lateral resolution) but greater diffraction artifacts.

We also showed that the proposed forward model (3.5) can be reduced to surface integrals (curves in 2-D) that are quadric surfaces (conics in 2-D), as derived in (3.14). Moreover, after numerical integration (3.15), it results in operations that are akin to a DAS operator. Thus, both forward and adjoint operators are very similar and can benefit from similar implementations.

3.5.2. Sparse Regularization for Enhanced Ultrafast Imaging

The images reconstructed using the proposed SR approach from single-PW acquisitions are promising in the sense that diffraction artifacts are indeed reduced, resulting for instance in an improved CNR (Table 3.1). However, these improvements were obtained at the cost of a major issue, that is an altered speckle pattern. For instance, half of the speckle tests failed. A close visual inspection of the speckle patterns obtained using SR (Figure 3.3) reveals that these patterns are sensibly different from those obtained with a DAS-based image reconstruction both from single-PW and 75-PW insonifications (the latter two having very similar patterns). This can obviously be detrimental to displacement estimation techniques that rely on proper speckle patterns.

Another issue that initially appeared as an advantage of SR is the improvement of both axial and lateral resolutions (Table 3.1), even outperforming 75 PWs + DAS. While the improvement in lateral resolution is a great outcome, the improvement of the axial one is unexpected. Indeed, the forward model considered neglects the pulse-echo waveform as it is assumed to be a Dirac delta function. Thus, there is no reason for the axial resolution, which is mainly governed by the time resolution of the pulse-echo waveform, to be improved.

Both aforementioned issues are linked to the sparsity-promoting regularizer considered in this study. Even if the sparsity averaging in a concatenation of wavelet bases is the best regularizer found to account for speckle patterns, it is clear that it does not allow restoration these patterns. Even worse, it does not preserve proper speckle patterns.

3.6. Conclusion

To conclude, we have derived a generic (approximated) forward model for pulse-echo ultrasound imaging. It offers great flexibility and is ideally suited to ultrafast US imaging. On the computational perspective, it can be reduced to operations similar to that of a DAS operator, enabling efficient parallel implementations of both operators. We showed

3. EFFICIENT MODELING FOR ULTRAFAST ULTRASOUND IMAGING

that it can be used to derive a backprojection-based DAS operator for direct use, but also incorporated in convex optimization tools for reconstructing high-quality images. The proposed sparsity-promoting term in a concatenation of wavelet bases produced visually pleasing images. Unfortunately, speckle patterns are altered in a way that displacement estimation techniques cannot be guaranteed. This represents the major issue of this approach in the context of US imaging as most ultrafast imaging modes heavily rely on these patterns. The main conclusion of our work on finding an appropriate regularizer is that, apparently, conventional image processing regularizers are not well suited to the high dynamic range and statistical properties of RF US images, especially in the presence of speckle patterns.

This observation led us to consider more “powerful” imaging methods based on deep learning techniques, thanks to which amazing performances started emerging in other medical imaging modalities [31], [32]. These CNN-based approaches will be the subject of the remaining chapters on US imaging methods (Chapters 5 to 7). But before we could “embrace” CNN-based methods, we needed data, and a lot of data. At that time, we did not have access to a suitable US research scanner, no large dataset of US images was publicly available, so we opted for simulated data. After a thorough benchmark of many publicly available US simulators, the finding was clear: the generation of a sufficiently large simulated-image dataset to achieve our goal would have required multiple years of computing time on our computers. Therefore, before diving into deep learning, we started by deriving and implementing an accelerated SIR-based US simulator with the hope that it could enable us to generate large datasets of high-quality for the training of CNNs. This simulator is the subject of the next chapter (Chapter 4).

4. A Spline-Based Spatial Impulse Response Simulator

The material presented in this chapter is the result of joint work with F. Martinez, M. Arditi, and J.-Ph. Thiran. This material is currently under preparation for submission.

4.1. Introduction

The need for fast and accurate US simulation tools is as great as ever. Such tools have been and continue to be extensively developed and deployed for the design and characterization of US transducers [71], [72]. For instance, as US transducers are crucial to the quality of the signals acquired, and consequently to the resulting image quality, simulation tools are often deployed to optimize their defining parameters. The use of end-to-end US scanner simulators is also increasing for the training and evaluation of physicians and sonographers [73]–[77], because they eliminate the need for volunteers or patients and can provide on-demand exposure to specific care and diagnosis scenarios. On the US imaging side, simulation tools are crucial for the development, assessment, and validation of image analysis and image reconstruction methods. This includes, for example, the analysis and characterization of speckle patterns [78], [79], the optimization of application-specific acquisition sequences [18], the fine-tuning and benchmarking of image reconstruction parameters and methods [80], or the development of suitable image quality metrics [81]. The recent breakthrough of deep learning-based methods in medical image analysis [82], medical image reconstruction [83], [84], and ultrasound imaging [85] comes with a critical need for large-scale datasets to feed the data-intensive algorithms involved.¹ In this context, the generation of synthetic data is of great interest [86], with the potential to generate (infinitely) large, highly diverse, and unbiased datasets of high quality.

With the myriad of US applications in which simulations can be leveraged, some of which are listed in the previous paragraph, it is no wonder that many US simulators are available.² Among them, and in the context of linear acoustics, Field II [87], [88] is considered a reference, thanks to its ability to calculate acoustic fields of arbitrarily shaped, excited, and apodized transducers, as well as simulating the acquisition of ultrasound pulse-echo signals in the presence of (weak) tissue inhomogeneities [27]. It is based on

¹This is particularly true for supervised learning, which, to date, remains the most used and efficient strategy to train deep neural networks, provided that enough data are available to prevent from overfitting.

²A fairly exhaustive list of available software can be found on the k-Wave website: <http://www.k-wave.org/acousticsoftware.php>.

the SIR method developed by Tupholme [36] and Stepanishen [37], [51], which is an analytic (mesh-free) method dedicated to the evaluation of transient acoustic fields that are of particular interest for pulse-echo US imaging.

The SIR approach, as developed by Tupholme and Stepanishen, provides an elegant way to derive analytic expressions for the evaluation of the velocity potential generated by a transducer aperture that is excited by a uniform surface waveform, at any point in the domain. It is assumed that the medium is homogeneous with a constant speed of sound, that the transducer is planar and embedded in an infinite planar baffle, and that the surface excitation is separable with respect to time and space. The method therefore amounts to solving a surface integral (at each time instant), derived from the Rayleigh-Sommerfeld equations (Section 2.4.3), and a detailed review can be found in [45] in the context of a rigid baffle. The general concept [57] consists of finding analytic expressions for the intersection (i.e., active arc length) between a radiating surface (aperture) and a spherical shell (representing a spherical wave) centered at the evaluation field point with a radius equal to the product of the mean sound speed and the evaluation time. Once the SIR is known, the resulting velocity potential can be obtained by the time convolution between the time derivative of the surface excitation and the SIR. Physical quantities such as pressure or particle velocity can eventually be derived from the velocity potential directly (Section 2.2).

Much attention and efforts were devoted to the derivation of analytic SIR expressions for uniformly excited radiators of specific shapes. The circular piston with a rigid (hard) baffle condition was apparently the first case for which analytic expressions were derived [37], [45], [56], [57], probably because of its simplicity. Other shapes followed, such as the slit [89], [90], the spherically focused radiator [43], [52], [53], and the triangular piston [91]. Of particular interest to array imaging is the rectangular piston, for which analytic expressions were initially derived in the far-field by Freedman [92], [93] and Stepanishen [37], [94]. Analytic expressions for the rectangular piston with a rigid baffle condition for points whose projection lies on a vertex of the rectangle were derived by Lockwood and Willette [44], [95], and generalized to any evaluation field point using the superposition principle by the same authors. Some analytic expressions were further derived by Scarano *et al.* [96] for field points whose projection on the rectangular aperture lie in specific regions. A complete set of analytic expressions were eventually proposed by San Emeterio and Ullate [59] for a uniformly excited rectangular piston with multiple boundary conditions (including hard and soft baffle conditions). Even if the rectangular piston would be suitable to represent flat transducer elements composing 2-D matrix arrays, most 1-D arrays are composed by curved elements. Transducer elements composing linear and phased arrays are cylindrical shells and those composing convex arrays are toroidal shells, and analytic expressions for the SIR of these geometries do not exist. For the cylindrical shell, a 2-D approximation was proposed in [97], and semi-analytic expressions were derived in [98]. Semi-analytic expressions were proposed for the toroidal shell in [99]. It is worth mentioning that non-uniform excitations were also investigated in [100]–[103], but analytic expressions were restricted to specific transducer shapes coupled with specific excitation amplitude distributions.

Unfortunately, analytic expressions only exist for a restricted set of transducer shapes, excitation distributions, and baffle conditions. To cope with this limitation, approximate methods based on discretization were naturally proposed. The main principle consists of representing the transducer surface as a set of characteristic sub-elements for which (simple) analytic expressions of the SIR exist. Using the superposition principle (linear acoustics), the SIR of the transducer is simply obtained by summing all the SIRs of the sub-elements. Provided that the transducer shape is discretized with sufficiently small sub-elements, this also enables approximating surface apodization by appropriately weighting the SIR of each sub-element before summation. Similarly, surface delays can also be accounted for. The most popular characteristic sub-element type is the rectangle, as it enables reasonably good approximations of general surfaces, but more importantly because computationally efficient far-field approximations exist for both rigid and soft baffle conditions. This approach is implemented (among others) in Field II and proved to be an efficient choice compared with using triangles as sub-elements [104]. Another relevant approach proposed by Piwakowski and Delannoy [105] consists of evaluating the surface integral (at each time instant) numerically by discretizing the surface into (infinitesimally small) ideal points. The SIR of the transducer can then be computed by a simple weighted sum of Dirac delta functions. This strategy has the advantage of enabling exact representation of surfaces, at the cost of requiring a much larger number of sub-elements than the strategy deployed in Field II [106]. Interestingly, even though not explicitly noted by the authors, this approach is actually a first-order numerical quadrature to approximate the surface integral, which is probably one of the reasons why an excessive number of “quadrature points” were necessary to achieve high accuracy.

A major numerical difficulty of the SIR method comes from the signal properties of such responses. They are typically of very short duration, approaching a Dirac delta function in limit cases, and are characterized by abrupt slope changes, resulting in very high frequencies [27], [44], [53]. These abrupt changes are induced by the edges and vertices of the transducer aperture. A typical example of a SIR and corresponding frequency spectrum is shown in Figure 4.1 for a rectangular element designed to work at a driving frequency of about 5 MHz, from which it is clear that a sampling rate of few GHz is required. Yet, because the electromechanical impulse responses of conventional transducer elements are bandlimited, the excitation waveform, and most importantly the field signal of interest (e.g., pressure field), are also bandlimited. Ideally, one would want to sample the SIR at the same rate as the excitation waveform before time convolution of the two signals. This is obviously not possible in most cases and sampling at the rate of the SIR (i.e., several additional orders of magnitude) is computationally prohibitive. Different strategies were proposed to tackle this numerical difficulty, mainly based on the integration of the SIR between time samples used for the excitation waveform, either numerically using an adaptive sampling of the SIR (refined around abrupt slope changes) [53] or from analytic expressions [104], [107]. Such strategies enabled for a much more efficient sampling of the SIR to be convolved with the excitation waveform, but still require oversampling ratios of several factors to prevent from detrimental aliasing.

In summary, the most important needs for a computational method relying on the

4. A SPLINE-BASED SPATIAL IMPULSE RESPONSE SIMULATOR

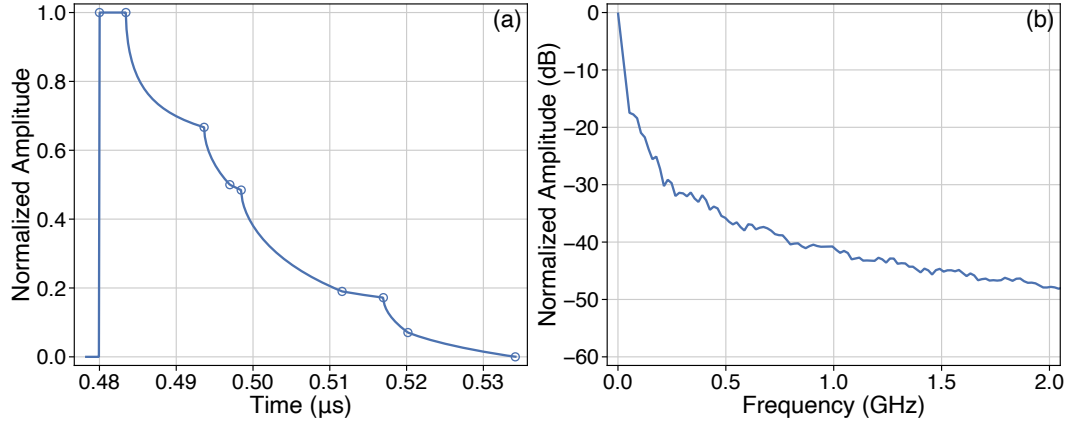


Figure 4.1. Example of (a) a spatial impulse response (SIR) and (b) corresponding frequency spectrum for a rectangular transducer aperture. The rectangle has a width of one wavelength at 5 MHz and a height-to-width ratio of 1.6. The field point at which the SIR was evaluated is positioned relatively to the element center at $(\lambda/5, \lambda/5, 2.5\lambda)$ in the width, height, and depth directions. Characteristic abrupt slope changes are circled in (a). The first one represents the first time of arrival of a spherical shell centered at the position of the evaluation field point. The eight following ones represent time instants at which the spherical shell intersects the edges and vertices of the rectangular surface.

SIR are: an accurate representation of arbitrary transducer shapes; an efficient way of evaluating the surface integral at each time instant; and an efficient sampling strategy such that the SIR can be sampled at rates similar to those required for the excitation waveform. To address these needs, we propose to represent the shape of transducers as NURBS surfaces, to evaluate the surface integral numerically using high-order Gaussian quadrature rules, and to express the SIR in B-spline bases for efficient sampling of the time axis. The use of NURBS enables accurate representation of complex surfaces. In particular, quadric surfaces can be represented exactly. This is a notable feature as most typical transducer shapes are actually quadric surfaces (e.g., cylindrical, spherical, or toroidal shells). Moreover, NURBS representations offer efficient means to evaluate surface derivatives at arbitrary positions [108]. This is typically of interest for the evaluation of Jacobian determinants when deploying Gaussian quadrature rules. Gaussian quadrature rules, such as the Gauss-Legendre quadrature rule, are well-known high-order quadrature rules that require much less quadrature points to achieve a desired accuracy [109, Sec. 25.4]. Finally, relying on numerical integration allowed us to express the calculation of the SIR in B-spline bases [68], [110], [111], and in particular of high-order ones. This enabled the sampling of SIRs at the same rate as the excitation waveform without any loss of accuracy on the (post-convolution) field signal at sampling rates of interest.

4.2. Proposed Approach

Let us recall (Section 2.4.4) that the SIR of a radiating surface $S \subset \mathbb{R}^3$ (assumed to be embedded in an infinite planar baffle) can be expressed at a field point $\mathbf{r} \in V \subset \mathbb{R}^3$ and

at a time instant $t \geq 0$ in a generic form as

$$h_S(\mathbf{r}, t) = \int_S \frac{\xi(\mathbf{r}') \delta\left(t - \frac{\|\mathbf{r} - \mathbf{r}'\|_2}{c_0}\right)}{2\pi \|\mathbf{r} - \mathbf{r}'\|_2} \beta(\tilde{\mathbf{n}}', \mathbf{r} - \mathbf{r}') d\sigma(\mathbf{r}'), \quad (4.1)$$

where $\tilde{\mathbf{n}}'$ is the *inward* surface normal, ξ represents the spatial distribution of surface velocity or pressure amplitudes over the radiating surface, and β is a term depending on the boundary conditions, expressed as

$$\beta(\tilde{\mathbf{n}}', \mathbf{r} - \mathbf{r}') = \begin{cases} 1, & \text{for a hard baffle (Neumann type),} \\ \cos(\tilde{\mathbf{n}}', \mathbf{r} - \mathbf{r}'), & \text{for a soft baffle (Dirichlet type).} \end{cases} \quad (4.2)$$

4.2.1. Numerical Quadrature of the Spatial Impulse Response

At each time instant $t \geq 0$ and field point $\mathbf{r} \in V$, the integrand of (4.1) is a real-valued function that can define as $g : S \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ for the purpose of derivation. Let us assume that the radiating surface S is a *smooth* surface parametrized by a mapping $\mathbf{s} : \hat{S} \subset \mathbb{R}^2 \rightarrow S$ whose Jacobian determinant $\det \mathbf{J}_s(\mathbf{u}) \neq 0, \forall \mathbf{u} \in \hat{S}$, where the Jacobian matrix is defined as $\mathbf{J}_s = (\partial/\partial \mathbf{u})\mathbf{s}$. Thus, the surface integral of g onto S can be rewritten as

$$\int_S g(\mathbf{r}') d\sigma(\mathbf{r}') = \int_{\hat{S}} g(\mathbf{s}(\mathbf{u})) |\det \mathbf{J}_s(\mathbf{u})| d\sigma(\mathbf{u}). \quad (4.3)$$

Assuming that g is a well-behaved function,³ one can approximate (4.3) by means of Gaussian quadrature as

$$\int_S g(\mathbf{r}') d\sigma(\mathbf{r}') \approx \sum_{q=1}^{n_q} g(\mathbf{s}(\mathbf{u}_q)) |\det \mathbf{J}_s(\mathbf{u}_q)| w_q = \sum_{q=1}^{n_q} g(\mathbf{r}_q) j_q w_q, \quad (4.4)$$

where $\{\mathbf{u}_q\}$ are the quadrature coordinates in the parametric space, for which $\{w_q\}$, $\{j_q\}$, and $\{\mathbf{r}_q\}$, are the corresponding quadrature weights, Jacobian determinants, and quadrature points (in the physical space), respectively. Note that Gaussian quadrature rules to obtain quadrature coordinates and corresponding quadrature weights are generally defined in one dimension [109, Sec. 25.4], but can be extended to higher dimensions by means of the tensor product.

³This is generally true as there is no realistic interest for the evaluation of the SIR on the radiating surface, where the integrand of (4.1) is obviously singular.

4. A SPLINE-BASED SPATIAL IMPULSE RESPONSE SIMULATOR

Using (4.4), we can now rewrite (4.1) as the finite sum

$$h_S(\mathbf{r}, t) \approx \sum_{q=1}^{n_q} \frac{\xi(\mathbf{r}_q) \delta\left(t - \frac{\|\mathbf{r} - \mathbf{r}_q\|_2}{c_0}\right)}{2\pi \|\mathbf{r} - \mathbf{r}_q\|_2} \beta(\mathbf{n}_q, \mathbf{r} - \mathbf{r}_q) j_q w_q \quad (4.5)$$

$$= \sum_{q=1}^{n_q} \xi_q j_q w_q \frac{\beta(\mathbf{n}_q, \mathbf{r} - \mathbf{r}_q)}{2\pi \|\mathbf{r} - \mathbf{r}_q\|_2} \delta\left(t - \frac{\|\mathbf{r} - \mathbf{r}_q\|_2}{c_0}\right). \quad (4.6)$$

The (non-unit) normal vector \mathbf{n} at some surface point $\mathbf{r} = \mathbf{s}(\mathbf{u})$ can be computed from the surface parametrization (mapping) as

$$\mathbf{n} = \frac{\partial \mathbf{s}(\mathbf{u})}{\partial u} \times \frac{\partial \mathbf{s}(\mathbf{u})}{\partial v}, \quad (4.7)$$

where (u, v) are the coordinates defining the parametric space of the surface mapping. Note that the term β is defined on the unit normal vector, but, as it can be seen in (4.2), is equal when considering the non-unit normal vector to the surface. The Jacobian determinants, represented as $\{j_q\}$ in (4.6), can then be computed from normal vectors directly as

$$|\det \mathbf{J}_s(\mathbf{u})| = \left\| \frac{\partial \mathbf{s}(\mathbf{u})}{\partial u} \times \frac{\partial \mathbf{s}(\mathbf{u})}{\partial v} \right\|_2. \quad (4.8)$$

Finally, by grouping all spatially dependent weighting terms in (4.6) under a global weighting term α , we can obtain a compact expression for the evaluation of the SIR by means of Gaussian quadrature as

$$h_S(\mathbf{r}, t) \approx \sum_{q=1}^{n_q} \alpha(\mathbf{r}, \mathbf{r}_q) \delta(t - \tau(\mathbf{r}, \mathbf{r}_q)), \quad (4.9)$$

where $\tau(\mathbf{r}, \mathbf{r}_q) = (1/c_0) \|\mathbf{r} - \mathbf{r}_q\|_2$. Thus, provided that such a surface mapping exists and that a suitable Gaussian quadrature rule is deployed with a sufficient amount of quadrature points to achieve a desired accuracy, the calculation of the SIR can be well approximated by a “simple”⁴ sum of shifted and weighted Dirac delta functions.

A typical first-order approximation of such a surface mapping would be to subdivide the radiating surface into small rectangles of identical shape and consider a single quadrature point in the center of each rectangle, in which case the Jacobian determinants would simply be the area of the characteristic rectangular sub-element. This corresponds to the well-know midpoint quadrature rule, which generally requires many more quadrature points to achieve a desired accuracy compared with more elaborated Gaussian quadrature rules such as the Gauss-Legendre one [109, p.887]. Not to mention the potential error in the surface approximation due to the use of rectangular sub-elements that cannot exactly represent non-planar surfaces. Even though not explicitly formulated by Piwakowski and Delannoy [105], their proposed “discrete representation” approach, which was recently used in [112], [113], relies on the midpoint quadrature rule. It is thus no wonder that

⁴Clearly this is not trivial when it comes to sampling.

they observed the need for excessively large number of sub-elements to achieve a desired accuracy [106]. To cope with such limitations, we propose to represent the radiating surface as a NURBS surface and rely on the fact that any NURBS surface can be further decomposed into *smooth* (rational B  zier) patches onto which Gaussian quadrature rules can be applied for the purpose of numerical integration.

4.2.2. Non-Uniform Rational B-Spline Surface Representations

We first review some basic principles of NURBS (surface) representations. For a detailed reference on the topic, the reader is referred to the book by Piegl and Tiller [108]. Let U be a nondecreasing sequence of real numbers representing a nonperiodic, clamped, or open knot vector, with $r + 1$ knots, defined as

$$U = (\underbrace{0, \dots, 0}_{p+1}, u_{p+1}, \dots, u_{r-p-1}, \underbrace{1, \dots, 1}_{p+1}), \quad r = n + p + 1. \quad (4.10)$$

The corresponding (nonnegative) B-spline basis functions $\{b_i^p\}_{i=0}^n$ of degree p (and order $p + 1$) are defined recursively as [114]–[116]

$$b_i^0(u) = \begin{cases} 1, & \text{if } u_i \leq u < u_{i+1}, \\ 0, & \text{otherwise,} \end{cases} \quad (4.11)$$

$$b_i^p(u) = \frac{u - u_i}{u_{i+p} - u_i} b_i^{p-1}(u) + \frac{u_{i+p+1} - u}{u_{i+p+1} - u_{i+1}} b_{i+1}^{p-1}(u), \quad p \geq 1. \quad (4.12)$$

Note that we restrict ourselves to nonperiodic knot vectors onto which B-spline basis functions are interpolating at the endpoints of such knot vectors but are (in general) non-interpolating at interior knots. Also, definitions are not strictly limited to the $[0, 1]$ interval. Yet, it is so common to the NURBS community that it is also adopted in the present work. Typical examples of B-spline basis functions on knot vectors of the form $U = (0, \dots, 0, 1, \dots, 1)$ are shown in Figure 4.2 for different degrees. This type of knot vectors results in B-spline basis functions that are Bernstein polynomials, because B-spline representations are a generalization of B  zier representations [108, Sec. 2.2].

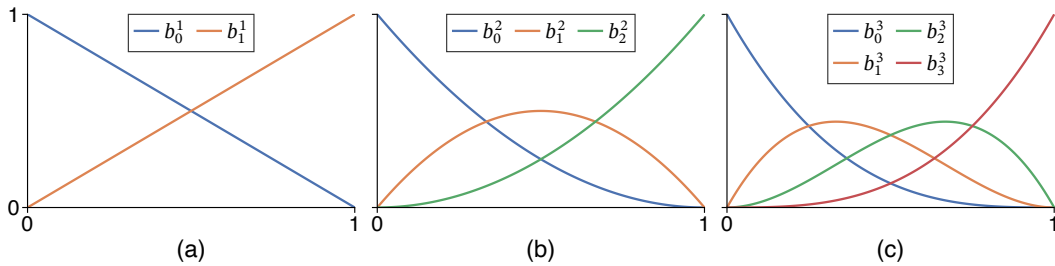


Figure 4.2. Typical examples of B-spline basis functions of different degrees defined on uniform knot vectors with no interior knots: (a) linear basis functions defined on the knot vector $(0, 0, 1, 1)$; (b) quadratic basis functions defined on the knot vector $(0, 0, 0, 1, 1, 1)$; (c) cubic basis functions defined on the knot vector $(0, 0, 0, 0, 1, 1, 1, 1)$.

4. A SPLINE-BASED SPATIAL IMPULSE RESPONSE SIMULATOR

A NURBS surface S of degree (p, q) in directions (u, v) can be represented by a bivariate vector-valued piecewise rational function (mapping) $\mathbf{s} : [0, 1]^2 \rightarrow S \subset \mathbb{R}^3$ defined as [108, Sec. 4.4]

$$\mathbf{s}(u, v) = \frac{\sum_{i=0}^n \sum_{j=0}^m b_i^p(u) b_j^q(v) w_{i,j} \mathbf{p}_{i,j}}{\sum_{i=0}^n \sum_{j=0}^m b_i^p(u) b_j^q(v) w_{i,j}}, \quad (4.13)$$

where $\{\mathbf{p}_{i,j}\}$ are the control points (forming a bidirectional net), $\{w_{i,j}\}$ are the corresponding weights, $\{b_i^p\}$ and $\{b_j^q\}$ are the (nonrational) B-spline basis functions of degrees p and q that are defined on the nondecreasing and nonperiodic (i.e., clamped) knot vectors

$$U = (\underbrace{0, \dots, 0}_{p+1}, u_{p+1}, \dots, u_{r-p-1}, \underbrace{1, \dots, 1}_{p+1}), \quad r = n + p + 1, \quad (4.14)$$

$$V = (\underbrace{0, \dots, 0}_{q+1}, v_{q+1}, \dots, v_{s-q-1}, \underbrace{1, \dots, 1}_{q+1}), \quad s = m + q + 1, \quad (4.15)$$

respectively. Expressions for the derivatives of NURBS surfaces also exist and can be found in [108, Sec. 4.5]. They are essential tools for the calculations of the surface normal vectors and the Jacobian determinants defined in (4.7) and (4.8), respectively. Note that if both U and V are defined with no interior points, such as in the examples shown in Figure 4.2, the NURBS surface is a (smooth) rational Bézier surface.

Equipped with these definitions, we can now assume that the radiating surface S can be represented exactly by a NURBS surface. An important property of NURBS representations is that any NURBS surface can be decomposed into a union of (smooth) rational Bézier surfaces (or patches), that share at most a common edge (and associated vertices) or a common vertex. Such a decomposition can be performed by a simple procedure known as knot refinement [108, Sec. 5.3]. Each (decomposed) rational Bézier patch is actually also a NURBS surface such that it can be readily represented by (4.13). Rational Bézier patches are of particular interest as they are *smooth* surfaces as opposed to “general” NURBS surfaces that may contain breakpoints. This is an essential property to allow the use of Gaussian quadrature rules onto such surfaces (Section 4.2.1).

As most transducer elements composing conventional US transducers are (at most) quadric surfaces, the assumption that the radiating surface can be represented exactly by a NURBS surface is generally valid in US imaging since NURBS surfaces can represent quadric surfaces exactly [108, Chap. 8]. For instance, transducer elements composing linear and phased arrays are cylindrical shells (i.e., elliptic cylinders), those composing convex arrays are toroidal shells (i.e., hyperbolic paraboloids), and those composing 2-D matrix arrays are simple rectangular planes. Spherically focused transducer elements are spherical caps. An example of the latter is shown in Figure 4.3, for which the NURBS surface definition can be obtained by revolving a circular arc (i.e., represented by NURBS curve). The NURBS surface can then be decomposed into four rational biquadratic Bézier smooth patches. Each of these smooth patches is defined by (4.13) onto the characteristic

unit square space $[0, 1]^2$. As they are smooth surfaces, Gaussian quadrature rules can be deployed onto the 2-D parametric space (also referred to as parent element) by means of the tensor product. Resulting quadrature coordinates can then be mapped onto the radiating surface exactly using (4.13). Note that Gaussian quadrature rules are typically defined on a characteristic $[-1, 1]$ interval [109, Sec. 25.4], but can be mapped to arbitrary intervals. This process is also illustrated in Figure 4.3, in which the Gauss-Legendre quadrature rule was considered to obtain the quadrature coordinates of a (5×3) grid.

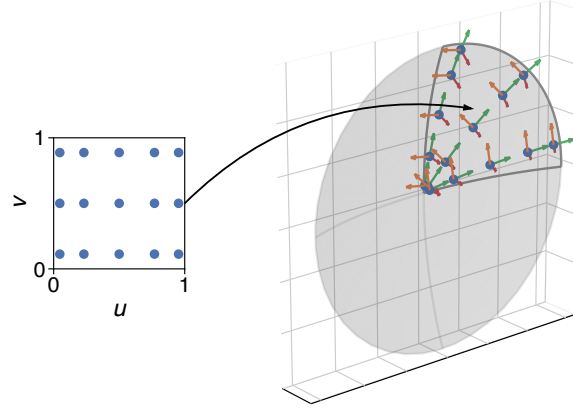


Figure 4.3. Example of a non-uniform rational B-spline (NURBS) surface representing a spherical cap, decomposed into four rational biquadratic Bézier smooth patches. The NURBS mapping from the parametric space to the physical space is illustrated with a parent element comprising (5×3) Gauss-Legendre quadrature points. The two (unit) tangent vectors and the (unit) normal vector at each quadrature point in the physical space are depicted by green, orange and red arrows, respectively.

Thanks to their small (or nonexistent) curvatures,⁵ all other aforementioned transducer element shapes can be represented by NURBS surfaces that are themselves single rational Bézier patches already, and thus do not require any further decomposition. An example of such as case for the cylindrical shell is shown in Figure 4.4.

4.2.3. Spatial Impulse Response in B-Spline Bases

Thanks to the proposed NURBS representation of the radiating surface, we are equipped to evaluate the global weighing term α involved in the sum of (weighted and shifted) Dirac delta functions (4.9) derived in Section 4.2.1 to approximate the SIR h_S .

Before diving into the well-known difficulty of sampling Dirac delta functions, it is important to keep in mind that the SIR is mainly a (very elegant) physical concept as such a quantity cannot be measured in physical conditions because of the electromechanical impulse response of transducer elements that are bandlimited. Thus, signals of interest that will be measured eventually, and are thereby of interest to be simulated, are also

⁵Recall that, strictly speaking, the Rayleigh-Sommerfeld equations Section 2.4.3 used to derive the SIR assume that the radiating surface is embedded in an infinite planar baffle (i.e., no curvature). Yet, it was demonstrated that they remain good approximations in the case of radiating surfaces with curvatures that are small with respect to the characteristic wavelength [43], [52], [54]

4. A SPLINE-BASED SPATIAL IMPULSE RESPONSE SIMULATOR

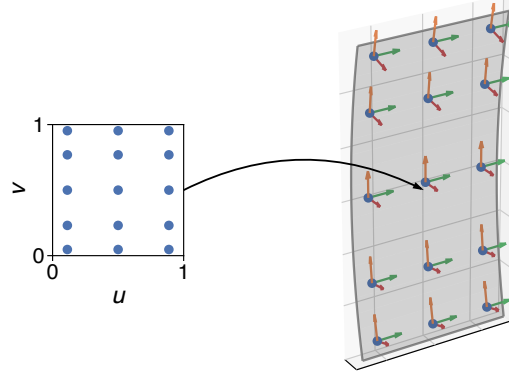


Figure 4.4. Example of a non-uniform rational B-spline (NURBS) surface representing a cylindrical shell that is also a smooth rational Bézier of degree $(1, 2)$, that is, linear in u and quadratic in v directions. The NURBS mapping from the parametric space to the physical space is illustrated with a parent element comprising (3×5) Gauss-Legendre quadrature points. The two (unit) tangent vectors and the (unit) normal vector at each quadrature point in the physical space are depicted by green, orange and red arrows, respectively.

bandlimited. Such a signal can be expressed in a generic form as

$$y(\mathbf{r}, t) = v(t) *_{\mathbf{r}} h_S(\mathbf{r}, t), \quad (4.16)$$

where v represent some bandpass excitation waveform of interest (independent of the field point \mathbf{r}). For instance, recall (Section 2.4.4) that the pressure field for a rigid baffle excited by a velocity waveform \bar{v}_n (including the electromechanical impulse response) can be expressed as $p(\mathbf{r}, t) = \rho_0(\partial/\partial t)\bar{v}_n(t) *_{\mathbf{r}} h_S(\mathbf{r}, t)$. Using (4.9) we obtain the corresponding approximation of such a signal as

$$y(\mathbf{r}, t) \approx v(t) * \sum_{q=1}^{n_q} \alpha(\mathbf{r}, \mathbf{r}_q) \delta(t - \tau(\mathbf{r}, \mathbf{r}_q)). \quad (4.17)$$

Equation (4.17) tells us that the signal y is composed of time-shifted and weighted replicas of the excitation waveform v . The excitation waveform v is typically measured in physical conditions or approximated using a model pulse. Thus, v can be expressed as a sampled signal with uniformly spaced samples $\{v(kT)\}$, $k \in [0, \dots, K-1]$. Equation (4.17) can then be rewritten as the discrete convolution

$$y(\mathbf{r}, t) \approx \sum_{k \in \mathbb{Z}} v(kT) \sum_{q=1}^{n_q} \frac{\alpha(\mathbf{r}, \mathbf{r}_q)}{T} \varphi_{\text{int}}\left(\frac{t - \tau(\mathbf{r}, \mathbf{r}_q)}{T} - k\right) \quad (4.18)$$

$$= \sum_{q=1}^{n_q} \frac{\alpha(\mathbf{r}, \mathbf{r}_q)}{T} \sum_{k \in \mathbb{Z}} v(kT) \varphi_{\text{int}}\left(\frac{t - \tau(\mathbf{r}, \mathbf{r}_q)}{T} - k\right), \quad (4.19)$$

where φ_{int} is some (interpolating) basis function (sometimes referred to as sampling kernel), and where it is assumed that boundary conditions are handled properly (via signal extension of v). Since v is a bandpass signal, a natural and error-free choice for φ_{int} would be the (normalized) sinc function, $\text{sinc}(x) := \sin(\pi x)/(\pi x)$. But because the sinc

function is of infinite support, it needs to be truncated, by multiplying it by some window of finite support. Generally, truncated sinc functions require rather large supports to achieve high accuracy [68]. Other options are standard (finite-support) interpolating basis functions such as those deployed for nearest-neighbor and linear interpolation, which are computationally efficient but of low accuracy.

Instead, as the discrete convolution of (4.19) consists of time shifting and interpolating v , we propose to rely on the concept of generalized interpolation as proposed by Thévenaz, Blu, and Unser [68]. By doing so, we can rewrite (4.19) as

$$y(\mathbf{r}, t) \approx \sum_{q=1}^{n_q} \frac{\alpha(\mathbf{r}, \mathbf{r}_q)}{T} \sum_{k \in \mathbb{Z}} c(kT) \varphi\left(\frac{t - \tau(\mathbf{r}, \mathbf{r}_q)}{T} - k\right), \quad (4.20)$$

where φ is some basis function and $\{c(kT)\}$ are the corresponding basis coefficients defined [68] such that

$$v[k_0] = \sum_{k \in \mathbb{Z}} c[k] \varphi(k_0 - k). \quad (4.21)$$

As a result, one can rely on potentially non-interpolating basis functions, possessing better properties than interpolating ones, with an additional operation that consists of finding the basis coefficients. A particularly elegant and efficient way of finding these coefficients is yet another (pre-filtering) convolution operation with the convolution-inverse⁶ φ^{-1} as

$$c[k_0] = \sum_{k \in \mathbb{Z}} \varphi^{-1}(k_0) v[k_0 - k], \quad (4.22)$$

provided that the convolution-inverse φ^{-1} exists. In the case of interpolating basis functions, the sequence of coefficients $\{c(kT)\}$ is equal to the sequence of samples $\{v(kT)\}$ (i.e., the convolution-inverse of any φ_{int} is logically a Dirac delta function).

We will restrict ourselves to the (extended) family of B-spline basis functions as they possess fantastic properties [68], [110], [111], [117]–[119], in particular that they benefit from maximal approximation orders for a given supports. For the purpose of interpolation, we can express a B-spline basis function of degree n as [120], [121]

$$\beta^n(x) = \sum_{k=0}^{n+1} \frac{(-1)^k (n+1)}{(n+1-k)! k!} \left(\frac{n+1}{2} + x - k\right)_+^n, \quad \forall x \in \mathbb{R}, \quad \forall n \in \mathbb{R}, \quad (4.23)$$

where the one-sided power function $(x)_+^n$ is defined as [68]

$$(x)_+^n = \begin{cases} 0 & n = 0 \text{ and } x < 0, \\ 1/2 & n = 0 \text{ and } x = 0, \\ 1 & n = 0 \text{ and } x > 0, \\ (x)_+^0 x^n & n > 0. \end{cases} \quad (4.24)$$

We are typically interested in B-spline basis functions β^n of degree greater than one,

⁶Not to be confused with the inverse function of φ .

since β^0 is almost identical to the nearest-neighbor basis function and β^1 corresponds to the basis function of linear interpolation. Also, B-spline basis functions of even degrees are usually not computationally interesting because they have the same support as B-spline basis functions of the following (odd) degrees. Since B-spline basis functions are symmetric, the pre-filtering operation defined in (4.22) can be efficiently performed by a series of $m = \lfloor n/2 \rfloor$ consecutive causal and anti-causal infinite impulse response (IIR) filters with poles $\{z_i\}$ and $\{z_i^{-1}\}$, respectively [110], [111], [117]. The m pairs of poles can be derived from the Z-transform of the convolution-inverse $(\beta^n)^{-1}$, many of which are tabulated in [110], [111], [117].

So far we have assumed that the samples $\{v(kT)\}$ and corresponding coefficients $\{c(kT)\}$ were extended properly $\forall k \in \mathbb{Z} \setminus [0, \dots, K-1]$ for the purpose of convolution operations. In many signal processing tasks involving interpolation, it is commonly advised to consider mirror boundary conditions for the purpose of signal extension, preventing from potential ripple artifacts. Also, “standard” interpolation typically consists of approximating signal values at rational coordinates (i.e., between samples). In the approach proposed here and described in (4.19), the sequence of samples $\{v(kT)\}$ are shifted, weighted, and interpolated for each quadrature point to form the signal of interest y . As the corresponding excitation waveform is bandlimited and can be modeled as a windowed sinusoidal RF signal of finite support,⁷ we consider zero boundary conditions. Finally, since we want to keep working with finite support sequences, we impose zero boundary conditions to the coefficients directly, namely $c[k] = 0, \forall k \in \mathbb{Z} \setminus [0, \dots, K-1]$. Coefficients satisfying these boundary conditions can be computed⁸ by the following operations. First, we initialize a sequence of coefficients $\{c_1[k]\}$ by applying a gain to the excitation waveform samples $\{v[k]\}$ as

$$c_0[k] = v[k] \prod_{i=1}^m (1 - z_i)(1 - z_i^{-1}), \quad k \in [0, \dots, K-1]. \quad (4.25)$$

Then the series of causal and anti-causal IIR filters for $i = 1, \dots, m$ can be evaluated as follows: a causal recursion preceded by suitable initialization

$$c_i^+[0] = \frac{1 - z_i^2}{1 - z_i^{2K+2}} \left[c_{i-1}[0] - \frac{z_i^2}{1 - z_i^2} \sum_{k=1}^{K-2} z_i^k (c_{i-1}[k] - z_i^{K+1} c_{i-1}[K-1-k]) - z_i^{K+1} c_{i-1}[K-1] \right], \quad (4.26)$$

$$c_i^+[k] = c_{i-1}[k] + z_i c_i^+[k-1], \quad k \in [1, \dots, K-1], \quad (4.27)$$

⁷Acoustic pulses are theoretically infinite responses with (rapidly) decaying trailing oscillations. Thus, they may be truncated when their trailing pulse envelope falls below some level to achieve a desired accuracy.

⁸Special thanks to Philippe Thévenaz for his invaluable help in deriving these operations.

and an anti-causal recursion preceded by suitable initialization

$$c_i[K-1] = -z_i c_i^+[K-1], \quad (4.28)$$

$$c_i[K-1-k] = z_i(c_i[K-k] - c_i^+[K-1-k]), \quad k \in [1, \dots, K-1]. \quad (4.29)$$

The coefficients $\{c[k]\}$ are then simply equal to $\{c_m[k]\}$. This preliminary step only needs to be performed a single time, with almost no computational cost as the excitation waveform is typically composed of few tens (or at most hundreds) of samples at realistic sample rates.

Once these coefficients are computed, they can be used readily for all quadrature points to evaluate the signal of interest y at any field point \mathbf{r} . The approximated (nonzero) signal samples at some field point, $\{y(\mathbf{r}, t_0 + nT)\}$ with $t_0 = \min_q \tau(\mathbf{r}, \mathbf{r}_q)$, can then be computed by a discrete (full) convolution of the (pre-filtered) coefficients and the summation of n_q basis functions shifted with respect to the corresponding propagation times $\{\tau(\mathbf{r}, \mathbf{r}_q)\}$, namely

$$y(\mathbf{r}, t) \approx \sum_{k \in \mathbb{Z}} c(kT) \sum_{q=1}^{n_q} \frac{\alpha(\mathbf{r}, \mathbf{r}_q)}{T} \varphi\left(\frac{t - \tau(\mathbf{r}, \mathbf{r}_q)}{T} - k\right). \quad (4.30)$$

Note that (4.30) and (4.20) only differ in the (reverted) order of summations. The latter reflects the “physical order,” as in (4.16), which involves the time convolution of the excitation waveform and the SIR of the radiating surface. For the purpose of interpretation (and compactness), let us define the basis SIR as

$$\hat{h}_S(\mathbf{r}, t) = \sum_{q=1}^{n_q} \frac{\alpha(\mathbf{r}, \mathbf{r}_q)}{T} \varphi\left(\frac{t - \tau(\mathbf{r}, \mathbf{r}_q)}{T} - k\right), \quad (4.31)$$

such that (4.30) can be interpreted as the convolution of the (pre-filtered) coefficients and the basis SIR, namely

$$y(\mathbf{r}, t) \approx \sum_{k \in \mathbb{Z}} c(kT) \hat{h}_S(\mathbf{r}, t - kT). \quad (4.32)$$

Even if the SIR h_S is not a quantity that can be measured, and as such is less important to simulate accurately than the signal of interest y , it remains interesting to be able to evaluate it.⁹ Since the basis SIR \hat{h}_S is a linear combination of n_q basis functions $\{\varphi\}$, it is possible to obtain an approximation of h_S from \hat{h}_S by convolution of the latter with the convolution-inverse φ^{-1} , namely

$$h_S(\mathbf{r}, t) \approx \sum_{k \in \mathbb{Z}} \varphi^{-1}(k) \hat{h}_S(\mathbf{r}, t - kT). \quad (4.33)$$

This is similar to expressing a cardinal spline basis function from its corresponding (non-interpolating) B-spline basis function. Thus, the approximation of the SIR obtained from

⁹Note that it is also possible to obtain the field response to a continuous-wave excitation directly from the SIR by evaluating its Fourier transform.

4. A SPLINE-BASED SPATIAL IMPULSE RESPONSE SIMULATOR

(4.33) when considering non-interpolating basis functions will be of infinite support with rapidly decaying oscillations [68].

4.2.4. Implementation Details

An illustration of the convolution involved in the evaluation of (4.30) for each quadrature point is depicted in Figure 4.5 for four different basis functions, namely nearest-neighbor, linear, quadratic¹⁰ Keys [122], and cubic B-spline. As the first three basis functions are interpolating, their corresponding (pre-filtered) coefficients are equal to the excitation waveform samples. The coefficients of the cubic B-spline have larger amplitudes than the excitation waveform to compensate for its non-interpolating property. One can already note the major issue associated with near-neighbor and linear interpolation, namely a shift error and an underestimated response, respectively. The advantage of using a cubic B-spline over a quadratic Keys, both having a support of four samples, can already be noticed. This will be confirmed by the numerical validation of the expected convergence orders (Section 4.3.1).

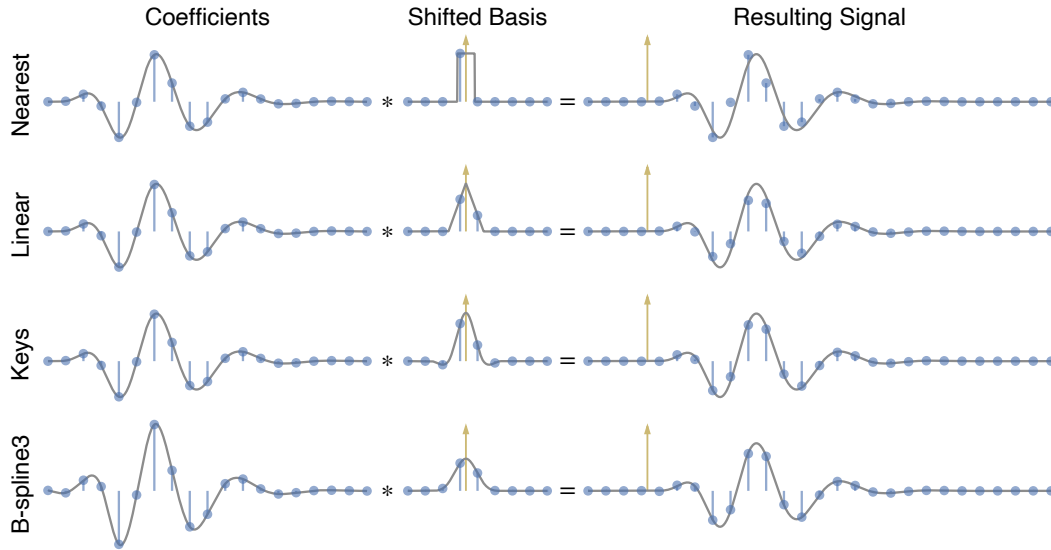


Figure 4.5. Illustration of the generalized interpolation strategy deployed for approximating the signal radiated by a transducer element. This strategy involves the convolution of coefficients, evaluated from the excitation waveform, with shifted basis functions. Three interpolating basis functions are depicted, namely nearest-neighbor, linear, and (quadratic) Keys. One non-interpolating basis function is also shown, namely cubic B-spline (B-spline3).

Figure 4.6 summarizes the complete process of the proposed approach developed in Sections 4.2.1 to 4.2.3 for the approximation of (bandpass) field signals using the SIR of radiating surfaces. The cylindrical shape of the transducer element [Figure 4.6(a)] is represented (exactly) as a NURBS surface using (4.13). As the NURBS surface is also

¹⁰By quadratic, it is implied that the optimal $\alpha = -1/2$ parameter was used for the Keys basis function, resulting in a quadratic interpolation method of cubic order.

a *smooth* Bézier surface, a Gaussian quadrature rule can be defined in the parametric space [Figure 4.6(b)], which in this case is a (2×4) Gauss-Legendre quadrature rule. The corresponding eight quadrature points can be mapped into the physical space using the NURBS representation, as well as the normal vectors (4.7) and Jacobian determinants (4.8). From the distances between the quadrature points and the field point, some basis function (e.g., cubic B-spline) can be evaluated at the corresponding time instants on a sampled time axis of minimum support [Figure 4.6(c)], and weighted accordingly. Their summation (4.31) results in the basis SIR [Figure 4.6(d)]. The (pre-filtered) coefficients [Figure 4.6(f)] are obtained by a series of causal and anti-causal IIR filters applied to the excitation waveform [Figure 4.6(e)], which can be evaluated from (4.25) to (4.29). Finally, the field signal [Figure 4.6(g)] is obtained by the convolution of the basis SIR [Figure 4.6(d)] and the (pre-filtered) coefficients [Figure 4.6(f)]. This process must be repeated for all field points of interest, except for the coefficients that only need to be computed once. An important aspect of the proposed approach is that it only involves simple arithmetic operations¹¹ and that all steps can be performed in parallel, making it a perfect fit for GPU implementations.

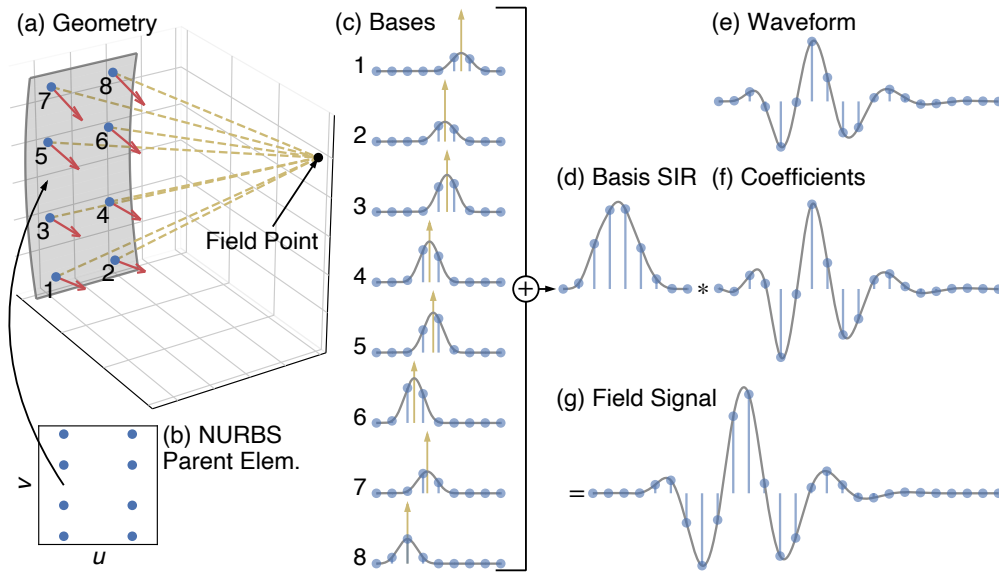


Figure 4.6. Illustration of all steps composing the proposed approach for the approximation of field signals using the spatial impulse response (SIR) of radiating surfaces. (Please refer to the associated text for a summary of each step involved.)

4.2.5. Extension to Arrays

The extension to US transducers composed of multiple transducer elements, typically arranged as arrays, is straightforward. Let us consider a generic definition of a transducer

¹¹Most analytic expressions for the SIR are typically composed of complex hyperbolic and hyperbolic inverse operations.

4. A SPLINE-BASED SPATIAL IMPULSE RESPONSE SIMULATOR

array composed of a set of n_e transducer elements. Each element can be represented by a NURBS surface and decomposed into a set of smooth Bézier patches onto which a Gaussian quadrature rule can be defined. These elements are generally of the same shape, which would result in similar NURBS representations, although this is not required. They can be driven by a set of excitation waveforms $\{v_i(kT)\}_{i=1}^{n_e}$, from which the set of corresponding coefficients $\{c_i(kT)\}_{i=1}^{n_e}$ needs to be evaluated once. Using the principle of superposition (linear acoustics), the field signal at field point \mathbf{r} radiated by such an array can be expressed using (4.32) as

$$y(\mathbf{r}, t) \approx \sum_{i=1}^{n_e} \sum_{k \in \mathbb{Z}} c_i(kT) \hat{h}_i(\mathbf{r}, t - kT), \quad (4.34)$$

where each element basis SIR \hat{h}_i can be computed using (4.31). It is common to define explicitly the delays $\{\tau_i\}_{i=1}^{n_e}$ and apodization weights $\{a_i\}_{i=1}^{n_e}$ applied across the array elements, rather than implicitly including them in the different excitation waveforms. Such delays and apodization weights are typically used to shape the transmit beam (beamforming), for instance to focus at a desired position or to steer an unfocused wavefront. Thus, and without loss of generality, (4.34) can be rewritten as

$$y(\mathbf{r}, t) \approx \sum_{i=1}^{n_e} a_i \sum_{k \in \mathbb{Z}} c_i(kT) \hat{h}_i(\mathbf{r}, t - \tau_i - kT). \quad (4.35)$$

In pulse-echo imaging, it is conventionally assumed that all (identical) transducer elements composing the array have the same electromechanical impulse response. An identical electric excitation is also typically used, such that the characteristic excitation waveform is identical. Only the delays and apodization weights applied to the elements may therefore differ such that (4.35) can be simplified as

$$y(\mathbf{r}, t) \approx \sum_{k \in \mathbb{Z}} c(kT) \sum_{i=1}^{n_e} a_i \hat{h}_i(\mathbf{r}, t - \tau_i - kT). \quad (4.36)$$

Finally, the echo signal scattered by an ideal reflector at \mathbf{r} can be easily obtained by the convolution of the field signal of the array (4.36) with the field signal of the receive element (4.32), multiplied by the scattering amplitude $f_m(\mathbf{r})$. Note that the excitation waveform on transmit is generally different from the one on receive, because the first contains both the electric excitation and the electromechanical impulse response, whereas the second contains the electromechanical impulse response only. Again, all computations can be performed in parallel (with proper handling of the different signal supports), and thereby efficiently implemented on a GPU. It should be noted that the computing time is generally governed by the many convolutions of the different field signals involved. Thus, the sampling rate deployed is critical to the computing time.

4.3. Experiments and Results

To validate the proposed approach, we performed two numerical experiments (Sections 4.3.1 and 4.3.2). The goal of the first one is to validate the core of the proposed approach, as described in (4.32), namely a convolution of (pre-filtered) coefficients and a signal (representing the basis SIR) composed of shifted and weighted basis functions. The second experiment consists of evaluating the accuracy of the proposed approach on field signals radiated by transducer elements with specific shapes allowing analytic expressions for the SIR.

For the two experiments, we considered an analytic pulse model for the excitation waveform such that reference signals are exact. The pulse model considered (Figure 4.7) is an analytic expression of the time derivative of a log-normal-modulated sinusoidal RF pulse. It is a fairly good model for the electromechanical impulse response of transducer elements. Note that the time derivative guarantees a zero direct current (DC) component [Figure 4.7(b)], a physical property of such electromechanical impulse response. For the log-normal distribution $\text{Lognormal}(\mu, \sigma)$, we used the parameters $\mu \approx -14.80$ and $\sigma \approx 0.26$. The frequency of the sinusoidal was set to 4.75 MHz, resulting in a waveform centered at ~ 5.3 MHz with a bandwidth of $\sim 71\%$ at -6 dB [Figure 4.7(b)]. The waveform was truncated at a trailing pulse envelope level of -320 dB (double-precision floating-point format), resulting in a duration of ~ 3.15 μs . The FWHM of the resulting waveform, conventionally used to quantify the time resolution of acoustic pulses, corresponds to approximately 0.23 μs .

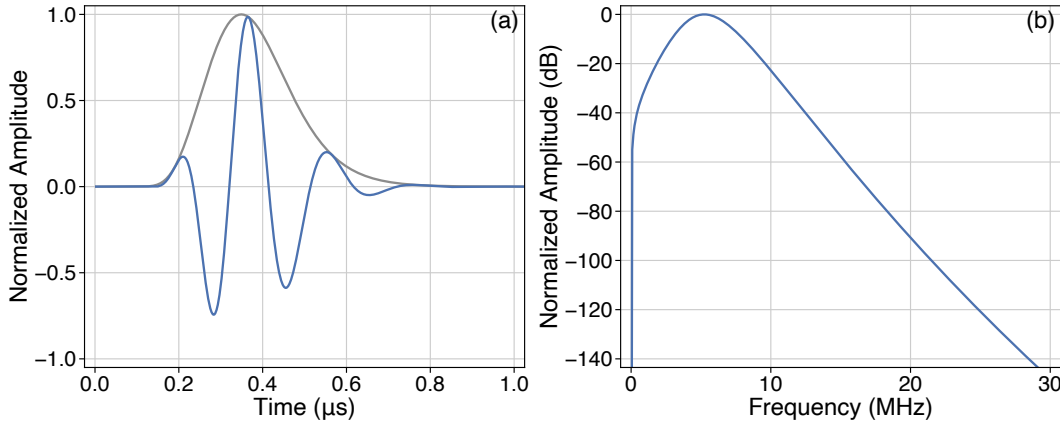


Figure 4.7. Normalized amplitude of (a) the excitation waveform and (b) corresponding frequency spectrum of the pulse model considered for the numerical experiments.

In addition to the B-spline basis functions, we also considered maximal-order-minimal-support (MOMS) functions [123], [124], in particular optimal-maximal-order-minimal-support (O-MOMS) functions [124], that are derived from B-spline basis functions. In general, we compared the classical interpolating basis functions for nearest-neighbor, linear, and quadratic Keys interpolation, as well as B-spline and O-MOMS basis functions of different degrees. For the purpose of quantifying the accuracy in the following experiments, we relied on the relative two-norm error defined between an estimated (RF

signal) vector and its true counterpart $\hat{\mathbf{y}}, \mathbf{y} \in \mathbb{R}^n$ as

$$\varepsilon = \frac{\|\mathbf{y} - \hat{\mathbf{y}}\|_2}{\|\mathbf{y}\|_2}. \quad (4.37)$$

4.3.1. Convergence Order of Various Basis Functions

For a robust validation of the core of the proposed approach, we considered a reference signal consisting of the time convolution between the analytic waveform v considered [Figure 4.7(a)] and a stream of n Dirac delta functions (stream of Diracs) at random-uniform times $\{\tau_i\}_{i=1}^n$ and of random-normal amplitudes $\{a_i\}_{i=1}^n$. The corresponding reference signal can be expressed as

$$y(t) = v(t) * \sum_{i=1}^n a_i \delta(t - \tau_i), \quad (4.38)$$

which can be seen as a generic expression for (4.17), namely signals we want to approximate using the proposed approach (4.32). Note that (4.38) can be evaluated exactly using the analytic expression for the excitation waveform.

We considered a signal duration corresponding to 500 times the time resolution of the waveform considered, that is approximately $114.34 \mu\text{s}$, and populated it with 51 378 random Diracs (i.e., 100 Diracs per time resolution cell, with appropriate pre-padding). We compared three interpolating basis functions, namely nearest-neighbor (degree zero), linear (degree one), and quadratic Keys (degree two), four (non-interpolating) B-spline basis functions of degrees 2, 3, 4, and 5, and the (non-interpolating) O-MOMS basis function of degree three. To validate that the theoretical approximation order of different basis functions is achieved by the proposed method, we computed the relative two-norm error using (4.37) between each approximated signal and the reference one at 15 logarithmically spaced sampling rates ranging from 20 MHz to 1 GHz. For a given basis function of some degree, the theoretical approximation order is defined by its degree plus one.

The resulting convergence curves are depicted in Figure 4.8. One can see that the theoretical approximation orders are accurately validated for all basis functions. For instance, B-spline basis functions of degrees 2, 3, 4, and 5 have approximation orders of 3, 4, 5, and 6, respectively. From these results, it is clear that the use of high-order non-interpolating B-spline basis functions provides a major advantage. For instance, the ratio of approximation error between the B-spline of degree five (B-spline5) and the quadratic Keys is greater than two orders of magnitude at a sampling rate of approximately 50 MHz. It can also be mentioned that the O-MOMS of degree three seems to be “over-performing” at low frequencies, namely from 20 MHz to 50 MHz. This low-frequency range represents a “rough” Nyquist-rate range for the excitation waveform considered [Figure 4.7(b)], especially below 30 MHz.

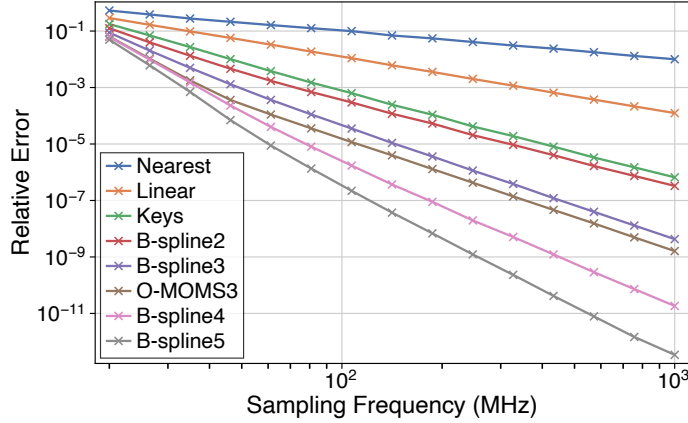


Figure 4.8. Results of the numerical experiment performed to validate the theoretical orders of convergence for different basis functions. Different types of basis functions were considered: three interpolating basis functions, namely nearest-neighbor (degree zero), linear (degree one), and quadratic Keys (degree two); four non-interpolating B-spline basis functions of degree 2, 3, 4, and 5; and the non-interpolating optimal-maximal-order-minimal-support (O-MOMS) basis function of degree three. The theoretical convergence order of a basis function is equal to its degree plus one.

4.3.2. Validation Against Analytic Solutions

The purpose of this experiment is to validate the complete proposed approach, namely including the NURBS surface representation and Gaussian quadrature. To do so, we considered two transducer-element shapes and corresponding baffle conditions for which analytic expressions are available for evaluating the SIR at any field point. The first is the spherically focused element (i.e., spherical cap) with a rigid baffle boundary for which a complete set of analytic expressions were derived in [53]. The second shape considered is the rectangular plane with both rigid and soft baffle conditions for which a complete set of analytic expression were derived in [59]. In all cases, we considered the same excitation waveform as described in Figure 4.7, characterized by a (center) wavelength $\lambda \approx 291 \mu\text{m}$ for the mean sound speed $c_0 = 1540 \text{ m/s}$ considered. In each case, we studied the field signal radiated by the transducer element at three characteristic field points (A, B, and C) positioned relatively to the transducer element center. The three field points were positioned such that their projections onto the surface lie on an axis of symmetry, an edge, and outside the surface projection.

We restricted the comparison to the following basis functions: nearest-neighbor, linear, quadratic Keys, B-spline of degrees three and five, and O-MOMS of degree three. Two different sampling rates of 30 MHz and 80 MHz were considered, namely a rather low-sampling rate for the excitation waveform considered, and a pretty decent one [Figure 4.7(b)]. Both shapes considered were represented exactly by NURBS surfaces. For the Gaussian quadrature rule, we relied on the well-known Gauss-Legendre one. As the regularity of the integrand involved in the evaluation of the SIR (4.1) has not been studied in depth, we relied on a heuristic strategy consisting of selecting the number of quadrature points in each (u, v) direction of the NURBS surface to obtain a spatial sampling rate equivalent to the sampling rate considered for the time dimension. By

doing so, it was observed that the resulting accuracy was not bound by the quadrature rule in the cases studied.

To compute the reference field signals, and because no analytic expressions could be derived for the convolution of the analytic excitation waveform and the analytic SIR, we relied on a very high sampling rate of 20 THz. The evaluation of the analytic SIR and excitation waveform as well as their convolution were performed at that sampling rate. We ensured that all time instants sampled in estimated signals were also present in the reference signals, to avoid interpolating the signals before comparing them in terms of the relative error defined in (4.37). We also compared the SIR obtained using (4.33) from the basis SIR, even if this quantity is less relevant than the field signal.

4.3.2.1. Spherically Focused Element with a Rigid Baffle Condition

The geometry of the spherically focused transducer element considered is defined by an active diameter $D = 20\lambda$ and a spherical radius R defined such that $2R/D = 4.8$, namely a similar ratio to the one studied in [53, Fig. 4]. The NURBS representation and corresponding four *smooth* Bézier patches is identical to the one described in Figure 4.3. Our heuristic strategy to define the number of quadrature points led to (59×91) Gauss-Legendre quadrature points per Bézier patch for the sampling rate of 30 MHz, and (155×243) for the sampling rate of 80 MHz. This corresponds to a total of 21 476 and 150 660 quadrature points, respectively. The three field points (A, B, and C) at which field signals were evaluated all lie in the same plane of revolution at a depth of $D/2 = 10\lambda$. The lateral coordinate of the first one (A) is $x_A = 0$. For the second one (B), the lateral coordinate was computed such that its projection onto the surface lies on an edge, resulting in $x_B \approx 8.1\lambda$. The lateral coordinate of the last one (C) was simply set to $x_C = 2x_B$ such that its projection onto the surface lies outside.

The relative two-norm errors of the field signals at each field point for both sampling rates and all basis functions considered are reported in Table 4.1. These results indicate that the basis function of the highest order performs best (i.e., B-spline5), with a relative error of approximately 10^{-4} and 10^{-7} at a sampling rate of 30 MHz and 80 MHz, respectively. The field signals and SIRs for the B-spline basis function of degree five is depicted in Figure 4.9. Despite the very different SIR at the three field points considered, the relative errors are similar at a given sampling rate. One can note that the approximated SIR typically contains ripple artifacts because the very high frequencies cannot be accounted for at such low sampling rates. This is especially visible for the field point A at a sampling rate of 30 MHz. Yet, the field signals do not suffer from such artifacts. We can also observe that the approach would tend to an accurate approximation of the SIR at much higher rates, should such a quantity be of interest. Field signals obtained with the other basis functions considered can be found in Figures A.1 to A.5.

TABLE 4.1
Relative Two-Norm Errors of Field Signals Radiated by a Spherically Focused
Transducer Element with a Rigid Baffle Condition

Freq.	Point	Nearest	Linear	Keys	B-spline3	O-MOMS3	B-spline5
30 MHz	A	2.89×10^{-1}	9.50×10^{-2}	2.14×10^{-2}	4.38×10^{-3}	1.71×10^{-3}	7.13×10^{-4}
	B	8.91×10^{-2}	9.06×10^{-2}	1.94×10^{-2}	4.64×10^{-3}	1.27×10^{-3}	7.96×10^{-4}
	C	1.41×10^0	1.99×10^{-1}	1.06×10^{-1}	1.57×10^{-2}	6.96×10^{-3}	1.93×10^{-3}
80 MHz	A	1.17×10^{-1}	1.40×10^{-2}	6.02×10^{-4}	6.13×10^{-5}	2.35×10^{-5}	8.62×10^{-7}
	B	1.25×10^{-2}	1.37×10^{-2}	4.51×10^{-4}	6.12×10^{-5}	2.34×10^{-5}	9.50×10^{-7}
	C	1.46×10^{-1}	1.53×10^{-2}	2.01×10^{-3}	1.03×10^{-4}	3.50×10^{-5}	1.55×10^{-6}

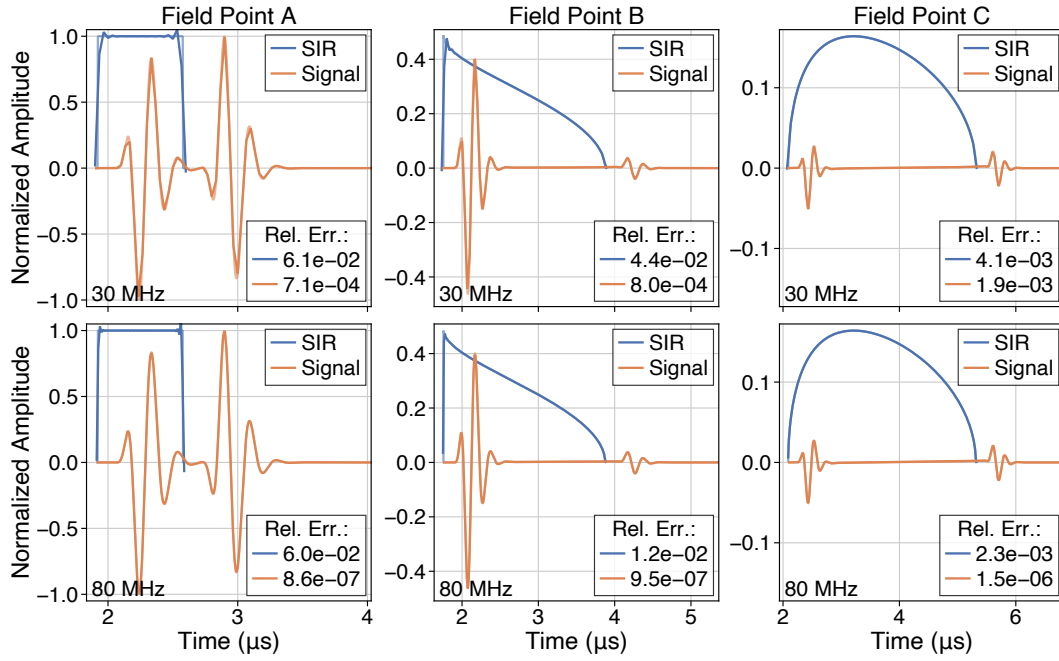


Figure 4.9. Comparison of the spatial impulse responses (SIRs) and field signals radiated at different field points by a spherically focused transducer element with a rigid baffle condition, excited by a windowed-sinusoidal waveform. The excitation waveform is a differentiated log-normal-windowed sine wave, with a characteristic (center) wavelength λ . The geometry of the spherical cap is defined by an active diameter of 20λ and a spherical radius of 240λ . The three field points (A, B, C) lie in the same revolution plane at a depth of 10λ and a lateral coordinate of $0, 8.1\lambda$ (i.e., projection on the edge of the surface), and 16.2λ , respectively. The proposed approach was implemented with a B-spline basis function of degree five, and was evaluated at two sampling rates of (first row) 30 MHz and (second row) 80 MHz. The reference SIRs and field signals were evaluated at a sampling rate of 20 THz. They are depicted with the same colors as the approximated counterparts, with a lower opacity.

4.3.2.2. Rectangular Element with Rigid and Soft Baffle Conditions

The geometry of the rectangular transducer element considered is defined by a width of λ and a height of 10λ , chosen to reflect typical width-to-height ratios of transducer elements composing linear arrays. The NURBS representation is simply a bilinear surface, resulting in a single *smooth* Bézier patch onto which the Gauss-Legendre quadrature rule was deployed. Our heuristic strategy to define the number of quadrature points led to

(7×59) Gauss-Legendre quadrature points for the sampling rate of 30 MHz, and (17×155) for the sampling rate of 80 MHz. The three field points (A, B, and C) at which field signals were evaluated all lie in a plane parallel to the element. They were positioned at a depth of $\lambda/2$ and an elevation of $\lambda/2$, with lateral coordinates of 0, $\lambda/2$, and λ , respectively. Note that this is quite an extreme case that was selected on purpose as this results in SIRs with very high frequencies. Also, the far field approximation made for the derivation of the SIR in the case of a soft baffle condition typically does not hold with a relative position that close to the radiating surface. Yet, since both the analytic and the proposed approaches are based on the same approximation, they can be compared for the purpose of validation.

The relative two-norm errors of the field signals at each field point for both sampling rates and all basis functions considered are reported in Table 4.2 (rigid baffle) and Table 4.3 (soft baffle). All results are consistent for the two different baffle conditions, even at field point C where the SIRs of the rigid and soft baffle conditions differ greatly. As for the spherically focused case, the higher-order basis function performs best. In general, the order of relative errors are also comparable between the two element shapes. For the two baffle conditions, the field signals and SIRs for the B-spline basis function of degree five are depicted in Figures 4.10 and 4.11, respectively. The effect of the soft baffle condition is visible for all field points, and in particular at field point C. Similarly to the spherically focused case, ripples due to the high-frequency components of the SIR can be observed in the approximations. Field signals obtained with the other basis functions considered can be found in Figures A.6 to A.10 for the rigid baffle condition, and in Figures A.11 to A.15 for the soft baffle condition.

TABLE 4.2
Relative Two-Norm Errors of Field Signals Radiated by a Rectangular
Transducer Element with a Rigid Baffle Condition

Freq.	Point	Nearest	Linear	Keys	B-spline3	O-MOMS3	B-spline5
30 MHz	A	2.14×10^{-1}	1.23×10^{-1}	4.34×10^{-2}	7.83×10^{-3}	3.55×10^{-3}	1.31×10^{-3}
	B	3.65×10^{-1}	1.03×10^{-1}	3.82×10^{-2}	6.58×10^{-3}	3.29×10^{-3}	1.14×10^{-3}
	C	3.59×10^{-1}	9.45×10^{-2}	2.37×10^{-2}	4.68×10^{-3}	2.24×10^{-3}	6.55×10^{-4}
80 MHz	A	2.09×10^{-2}	1.39×10^{-2}	4.20×10^{-4}	5.94×10^{-5}	2.90×10^{-5}	8.28×10^{-7}
	B	2.19×10^{-2}	9.50×10^{-3}	3.19×10^{-4}	3.39×10^{-5}	1.70×10^{-5}	7.74×10^{-7}
	C	6.06×10^{-2}	1.08×10^{-2}	4.79×10^{-4}	3.58×10^{-5}	1.76×10^{-5}	3.44×10^{-7}

4.4. Discussion

We demonstrated that field signals radiated by surfaces embedded in both rigid and soft baffles and excited by bandpass RF waveforms can be accurately approximated by the proposed approach. This approach consists of representing the radiating surface by a NURBS surface, decomposing it into *smooth* Bézier patches onto which high-order Gaussian quadrature rules can be deployed. From the resulting quadrature points (and weights), the field signal can then be evaluated by the convolution of the basis SIR, expressed as a

TABLE 4.3
Relative Two-Norm Errors of Field Signals Radiated by a Rectangular
Transducer Element with a Soft Baffle Condition

Freq.	Point	Nearest	Linear	Keys	B-spline3	O-MOMS3	B-spline5
30 MHz	A	1.13×10^{-1}	1.10×10^{-1}	3.05×10^{-2}	6.18×10^{-3}	2.16×10^{-3}	1.20×10^{-3}
	B	2.66×10^{-1}	1.05×10^{-1}	2.89×10^{-2}	5.80×10^{-3}	2.04×10^{-3}	1.09×10^{-3}
	C	2.44×10^{-1}	1.08×10^{-1}	1.90×10^{-2}	5.01×10^{-3}	1.48×10^{-3}	7.12×10^{-4}
80 MHz	A	1.37×10^{-2}	1.42×10^{-2}	4.13×10^{-4}	6.15×10^{-5}	2.99×10^{-5}	8.93×10^{-7}
	B	1.06×10^{-2}	1.08×10^{-2}	3.16×10^{-4}	4.31×10^{-5}	2.12×10^{-5}	9.00×10^{-7}
	C	5.02×10^{-2}	1.15×10^{-2}	4.58×10^{-4}	3.95×10^{-5}	2.07×10^{-5}	3.95×10^{-7}

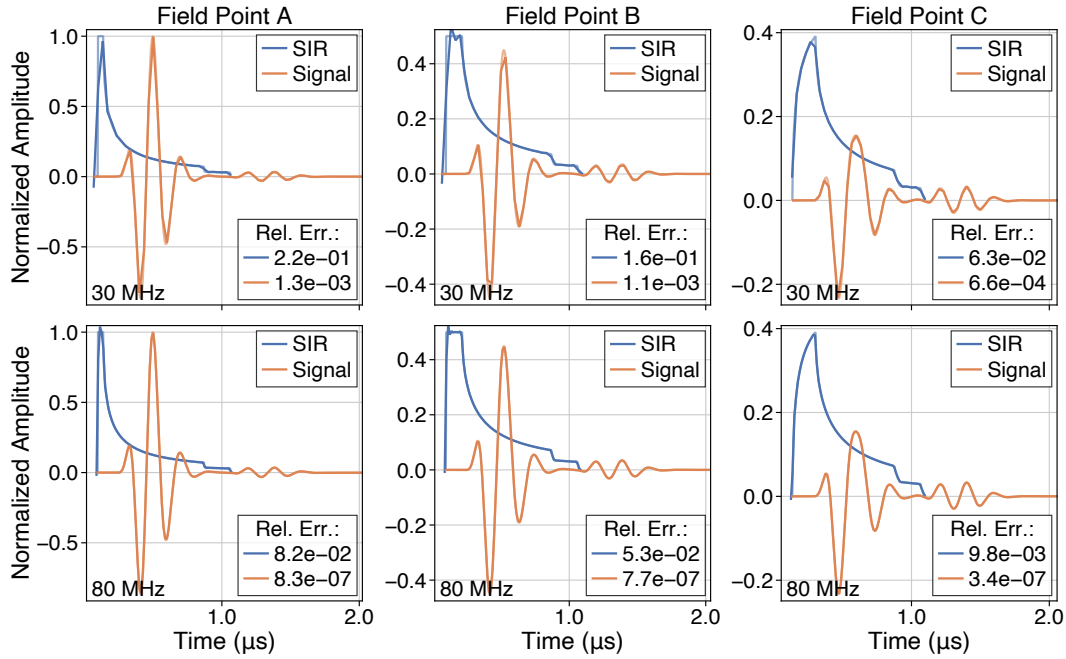


Figure 4.10. Comparison of the spatial impulse responses (SIRs) and field signals radiated at different field points by a rectangular transducer element with a rigid baffle condition, excited by a windowed-sinusoidal waveform. The excitation waveform is a differentiated log-normal-windowed sine wave, with a characteristic (center) wavelength λ . The geometry of the rectangular element is defined by a width of λ and a height of 10λ . The three field points (A, B, C) lie in a plane parallel to the element. They were positioned at a depth of $\lambda/2$ and an elevation of $\lambda/2$, with lateral coordinates of 0, $\lambda/2$, and λ , respectively. The proposed approach was implemented with a B-spline basis function of degree five, and was evaluated at two sampling rates of (first row) 30 MHz and (second row) 80 MHz. The reference SIRs and field signals were evaluated at a sampling rate of 20 THz. They are depicted with the same colors as the approximated counterparts, with a lower opacity.

sum of shifted-and-weighted basis functions, and (pre-filtered) basis coefficients.

The results obtained from the two experiments carried out demonstrate that the proposed approach can attain a high accuracy with respect to analytic reference signals. The comparison of the relative errors obtained in all cases of the second experiment on realistic transducer element shapes (Tables 4.1 to 4.3) with those of the convergence-order study on a random stream of Dirac delta functions (Figure 4.8) shows that the latter provides an upper bound on the relative error. As such, the first experiment provides

4. A SPLINE-BASED SPATIAL IMPULSE RESPONSE SIMULATOR

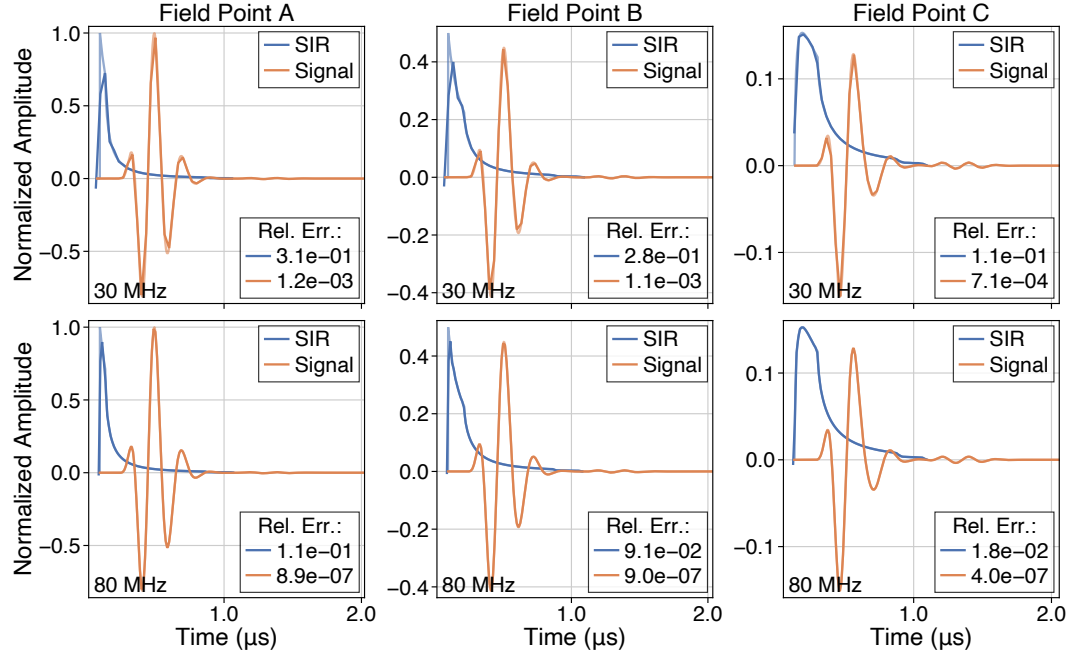


Figure 4.11. Comparison of the spatial impulse responses (SIRs) and field signals radiated at different field points by a rectangular transducer element with a soft baffle condition, excited by a windowed-sinusoidal waveform. The remaining settings relevant to the results depicted are identical to those detailed in Figure 4.10.

a robust way for selecting appropriate basis functions to guarantee a desired accuracy, assuming that suitable numerical quadrature is deployed.

4.4.1. Benefit of High-Order Basis Functions

In pulse-echo US imaging, it is typically acceptable to consider a sampling rate that enable preserving frequencies with a relative spectrum magnitude of approximately -60 dB. In the different experiments conducted, this corresponds to a sampling rate of approximately 35 MHz [Figure 4.7(b)]. One can note (Figure 4.8) that using a B-spline basis function of degree five results in a relative error of approximately -60 dB, and should therefore not induce additional errors in the simulated signals at that sampling rate. On the (complete) other hand, using a nearest-neighbor basis function would require a sampling rate of approximately 10 GHz to achieve the same relative error as the one achieve using a B-spline basis function of degree five at approximately 35 MHz. This is an important observation and a major advantage of the proposed method as large-scale simulations highly benefit from low-sampling rates (because of the many discrete convolutions). Thus, the use of high-order basis functions is of primary interest, even if it implies basis functions with slightly greater supports. Note that we did not investigate basis functions of higher order than the B-spline of degree five (order six), but the proposed approach can be deployed with B-spline (or O-MOMS) basis functions of arbitrarily large orders.

4.4.2. Non-Uniform Rational B-Spline Representation of Surfaces

Our choice of representing radiating surfaces as NURBS surfaces was primarily motivated by the fact that they enable representing quadric surfaces exactly, and that most US transducer elements are at most quadric surfaces. Obviously, other representations could be used in the proposed approach (such as analytic ones), provided that parametrization is *smooth* such that Gaussian quadrature can be deployed. The fact that any NURBS surface can be decomposed into *smooth* Bézier patches guarantees that this requirement is met.

It is also worth mentioning that NURBS are heavily used in computer-aided design (CAD) software for representing surfaces. This means that the design of new transducer element shapes could be evaluated directly from these NURBS definitions without additional processing (such as meshing), similarly to the concept of isogeometric analysis (IGA). In a more automated design strategy, the optimization of transducer element shapes could be performed directly from their NURBS definitions (e.g., control points and weights) with direct evaluation of the field quantities of interest. Also, thanks to IGA, NURBS representations can be used for spectral analyses of objects (e.g., vibration modes). One could therefore compute the relevant field quantities for different vibration modes directly.

4.4.3. Gaussian Quadrature Rules

We only considered the well-known Gauss-Legendre quadrature rule in the present study. Since there is no realistic scenario in which field signals would need to be computed onto the radiating surface (i.e., non-singular integrand), the Gauss-Legendre quadrature rule guarantees high-order accuracies. Even though not reported here, we also evaluated the performance of typical (low-order) quadrature rules such as the midpoint, trapezoidal, and Simpson ones. As expected, they all performed much less efficiently than the Gauss-Legendre one. We also conducted some preliminary evaluations using the Gauss-Legendre-Lobatto rule. This quadrature rule may be promising in the context of radiating surfaces as it includes the endpoints of the integration interval (i.e., the surface edges), at the cost of a slightly reduced accuracy than the Gauss-Legendre rule. However, we could not yet conclude on which one performs generally best. A more in-depth study to assess the regularity of the integrand should be performed. Such a study would probably lead to an optimal selection of the quadrature point number.

4.4.4. Comparison to Other Strategies

It is interesting to note that the proposed approach for expressing the computation of field signals using the SIR can be used to interpret previously proposed discretized approaches. Indeed, the proposed formulation is pretty generic as different surface representations, different quadrature rules, and different basis functions can be used. For instance, Piwakowski and Delannoy [105] proposed an approach to compute the SIR by discretizing

the radiating surface into many ideal points (Huygens principle) and to average the Dirac delta functions between each time samples. This strategy can be considered in the proposed formulation, namely by using a midpoint quadrature rule together with a nearest-neighbor basis function. This would of course result in much worse accuracy than using a high-order Gaussian quadrature rule together with a high-order basis function. It was for instance reported in [105] that an excessive number of points was required to achieve a suitable accuracy. This observation matches the convergence-order study performed in this work (Figure 4.8), in which it is clear that the use of a nearest-neighbor basis function is the least efficient one. Not to mention that a midpoint quadrature rule also requires many more quadrature points than Gaussian quadrature rules. In general, the strategy of averaging SIR values was proposed and used in many approaches [53], [104], [105], [107]. Thus, these approaches could also benefit from the formulation in basis functions proposed here, in particular to benefit from higher-order basis functions.

From a pure computational perspective, it was mentioned several times that the proposed approach is particularly suited for parallel implementations. This property is largely inherited from the SIR model and the linear pulse-echo model for weak scattering derived therefrom (Chapter 2), which imply that all processes are independent of each other. As such, most approaches based on the SIR model could benefit from efficient GPU-based implementations. A critical point when it comes to GPU-based implementations is the complexity of arithmetic operations. Analytic expressions for the SIR of specific radiating surfaces involve complex operations such as hyperbolic and inverse hyperbolic operations, with many cases depending on the relative positioning of field points with respect to radiating surfaces. Such operations would typically not be ideally suited for GPU-based implementations. The proposed formulation relies on many more operations of much lower arithmetic complexity,¹² and is as such better suited for GPU-based implementations. So far, we did not perform an exhaustive benchmark of our current implementation against well-known software such as Field II. Our initial experiments in the context of SA imaging indicated that we could reduce the computing time by approximately two orders magnitude, considering consumer-level central processing units (CPUs) (for Field II) and GPUs (for the proposed approach). Note that the initial goal of developing this approach was not to propose an alternative to Field II, but rather to enable us generating a sufficient amount of data for the purpose of training CNN-based image reconstruction methods (Chapters 5 to 7).

4.4.5. Hypothetical Perspective for Analytic Expressions

Even though research for the derivation of analytic SIR expressions does not seem to be very active anymore, or at least not as active as it used to be, there may be an interest in working with NURBS representations for this task. Analytic expressions for specific shapes may be derived in or from NURBS representations. Indeed, NURBS representations

¹²The most complicated arithmetic operations are the square root and the cosine in the case of soft baffle conditions.

can not only be used to represent a radiating surface exactly, but also to represent the spherical shell centered at a field point. The main principle proposed by Tupholme [36] and Stepanishen [51] for deriving the SIR of a radiating surface consists of expressing the (active) arc length of the curve representing the intersection of the radiating surface and a spherical shell centered at the field (i.e., outgoing spherical wave). The intersection of two such shells represented as NURBS surfaces may be expressed as a NURBS curve. Such NURBS representations may provide useful mathematical tools for the derivation of analytic or (efficient) semi-analytic expressions for the SIR.

As mentioned in the introduction, there exists several shapes of interest to US imaging that have already been studied but for which no complete analytic or semi-analytic expressions for the SIR were found. For instance, the cylindrical shell, namely the shape of transducer elements forming 1-D linear and phased arrays, was partially solved in [98]. The toroidal shell for transducer elements forming 1-D convex arrays was addressed in [99]. Another transducer shape for which only partial expressions were found is the conical shell [125], [126].

4.5. Conclusion

In this chapter, we proposed a spline-based SIR approach for the simulation of field signals radiated by arbitrary shapes embedded in both rigid and soft baffles and excited by RF bandpass waveforms. This approach consists of representing a transducer surface as a NURBS surface and decomposing it as smooth Bézier patches onto which high-order Gaussian quadrature rules can be deployed to approximate the surface integral involved in the computation of the SIR. Using high-order B-spline bases to express the SIR, the basis SIR amounts to a sum of shifted-and-weighted basis functions that depend on the positions and weights of the quadrature points. The resulting field signal is then obtained by the convolution of the basis SIR and the (pre-filtered) coefficients obtained from the excitation waveform. The use of NURBS enables accurate representations of complex surfaces, and common transducer shapes (quadric surfaces) can even be represented exactly. High-order Gaussian quadrature rules enables using fewer quadrature points to attain a desired accuracy than low-order ones commonly used (e.g., midpoint or trapezoidal). High-order B-spline basis functions enables using a simulation sampling rate identical to the one of interest to represent the excitation waveform accurately. Numerical experiments conducted against analytic reference signals demonstrated that the proposed approach can attain an error as low as the sampling error of the excitation waveform. The extension to transducer arrays and pulse-echo settings is straightforward. This approach is also well-suited to parallel implementations. An initial GPU implementation enabled us to reduce the computing time by up to two orders of magnitude compared to the well-known Field II simulator.

Being able to reduce the computing time requirements to simulated pulse-echo acquisitions accurately by such a factor was literally a game changer. Indeed, it enabled generating sufficiently large datasets of high-quality US images in a realistic time frame

4. A SPLINE-BASED SPATIAL IMPULSE RESPONSE SIMULATOR

to develop, benchmark, and train CNNs for the purpose of image reconstruction. These image reconstruction methods based on CNNs are the subject of the following chapters (Chapters 5 to 7), starting with the core CNN-based image reconstruction method for ultrafast US imaging proposed in this thesis.

5. CNN-Based Image Reconstruction Method for Ultrafast Ultrasound Imaging

The material presented in this chapter is the result of joint work with M. Vonlanthen, F. Martinez, M. Arditi, and J.-Ph. Thiran. This material is also presented in [33].

5.1. Introduction

Deep learning entered the medical image analysis field [82], quickly followed by the image reconstruction one [83], [84], with tight links to inverse problems [31], [32], [127], [128]. As US imaging is achieved through a sophisticated signal processing pipeline, deep learning-based components may be introduced at many stages of this process [85]. The first application appears to date back to 1990 with a pre-beamforming aberration correction method based on a multi-layer fully connected neural network [129]. In recent years, different strategies relying on CNNs have been proposed for post-beamforming speckle reduction [130]–[132] or to mimic the post-processing of clinical scanners [133]. Fully connected neural networks operating on beamformed signals were proposed to remove off-axis artifacts [134] or to learn the apodization weights of an adaptive beamformer [135]. An end-to-end CNN-based method was proposed for segmenting anechoic cysts from raw element data directly [136]. Restoration techniques using CNNs were proposed to enhance low-quality images [137], [138], to learn a compounding operation from a reduced number of insonifications [139], [140], or for super-resolution in the context of US localization microscopy [141].

In this chapter, we explore the reconstruction of high-quality US images from single insonifications using a CNN-based approach. The main objective is to provide images with a minimum amount of diffraction artifacts at the highest possible frame rate, crucial for tracking rapid phenomena in the human body [1]. Inspired by regularized regression methods and [31], we propose a “two-step” image reconstruction approach, consisting of a backprojection operation to obtain a low-quality image estimate, followed by applying a CNN trained to restore a high-quality image. The backprojection operator is derived from linear acoustics and far-field assumptions (Section 3.2), resulting in an operation similar to DAS beamforming, and is further improved with a re-weighting operation. The CNN architecture is based on [31], [32], with notable improvements over our preliminary work [137]. To account for the HDR property of US images while preserving their RF nature, we introduce the mean signed logarithmic absolute error (MSLAE) as training loss function, inspired by both the conventional log-compression applied to visualize

US images and audio-coding companding algorithms. Experiments were conducted on a linear transducer array using a single PW with normal incidence to reconstruct low-quality input images. Reference images were reconstructed from the complete set of SA acquisitions. The training of the CNN was performed using a simulated-image dataset with relevant statistical properties, in particular spanning a wide dynamic range. Leveraging numerical simulations, reference images were pushed “beyond physics” by considering an optimal, spatially oversampled, virtual version of the transducer array to prevent GL artifacts. A special attention was given to speckle patterns as they are essential to most displacement estimation techniques deployed in ultrafast US applications [1], [5], [14], [15], [65]. Extensive quantitative evaluations were performed on a numerical test phantom inspired by [80], and robustness to experimental data was evaluated in both *in vitro* and *in vivo* settings. Current limitations and directions for future improvements are also discussed.

5.2. Background

5.2.1. Pulse-Echo Modeling and Imaging

In this section, we briefly summarize notions presented and derived in Chapters 2 and 3 that are necessary to the present chapter. In Chapter 2, we showed that under the first-order Born approximation, assuming (longitudinal) linear acoustics, and neglecting dispersive attenuation, the signal received (e.g., by a transducer element) from a weakly scattering medium Ω embedded in a homogeneous medium and insonified by a transmitter (e.g., wavefront) can be (compactly) expressed as

$$y_{i,j}(t) = v_{\text{pe}}(t) *_t \int_{\mathbf{r} \in \Omega} \left[h_i^{\text{tx}}(\mathbf{r}, t) *_t h_j^{\text{rx}}(\mathbf{r}, t) \right] x(\mathbf{r}) d\mathbf{r}, \quad (5.1)$$

where $*_t$ denotes time convolution. The terms h_i^{tx} and h_j^{rx} represent the SIRs of the transmitter and receiver, respectively. The pulse-echo waveform¹ v_{pe} includes both electro-acoustic (transmit) and acousto-electric (receive) impulse responses as well as the electric excitation waveform (assumed identical for all transducer elements). Local fluctuations in density and propagation velocity, which induce the scattered echo signal, are represented by x .

Due to the high complexity of the SIRs involved in (5.1), and in order to obtain a computationally tractable measurement model for use in image reconstruction methods, further assumptions need to be made (Section 3.2.1). In Chapter 3, we showed that assuming far-field approximation both for the transmitter (e.g., ideal wavefront) and for the receiver (e.g., narrow transducer element), and assuming v_{pe} to be a Dirac delta

¹Originally called pulse-echo wavelet in [27].

function δ , (5.1) can be approximated as

$$y_{i,j}(t) = \int_{\mathbf{r} \in \Omega} \bar{h}_i^{\text{tx}}(\mathbf{r}) \bar{h}_j^{\text{rx}}(\mathbf{r}) \delta(t - \tau_i^{\text{tx}}(\mathbf{r}) - \tau_j^{\text{rx}}(\mathbf{r})) x(\mathbf{r}) d\mathbf{r}, \quad (5.2)$$

where \bar{h}_i^{tx} and \bar{h}_j^{rx} are scalar functions representing the (far-field) acoustic diffraction effects of the transmitter and the receiver to and from a field point \mathbf{r} , respectively. The terms τ_i^{tx} and τ_j^{rx} represent the acoustic wave propagation times from the transmitter and the receiver to a field point \mathbf{r} , respectively.

Ultrasound transducers typically comprise a set of n_r receivers (i.e., transducer elements) arranged in a regular array and signals (measurements) are sampled at discrete time intervals. As images are composed of discrete (pixel) values, the (approximated) linear physical measurement model defined in (5.2) can be conveniently expressed as a discretized operation for all receivers as $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$, where $\mathbf{H}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is the forward (matrix) operator, $\mathbf{x} \in \mathbb{R}^n$ is the (vectorized) image we seek to recover, $\mathbf{y} \in \mathbb{R}^m$ are the (vectorized) transducer elements measurements (raw data), and $\mathbf{n} \in \mathbb{R}^m$ is the measurement noise. (Note that the transmitter index i has been omitted in the matrix expressions to lighten notations.) Recovering \mathbf{x} from \mathbf{y} is a well-known inverse problem in medical imaging and may be addressed using various strategies (e.g., see [32] for CT imaging).

Classical backprojection algorithms, which exploit the adjoint operator $\mathbf{H}^*: \mathbb{R}^m \rightarrow \mathbb{R}^n$, may be used to address such an inverse problem. Using (5.2), one can express the adjoint operation for all receivers in the continuous domain as (Section 3.2.2)

$$\tilde{x}_i(\mathbf{r}) = \bar{h}_i^{\text{tx}}(\mathbf{r}) \sum_{j=1}^{n_r} \bar{h}_j^{\text{rx}}(\mathbf{r}) y_{i,j}(\tau_i^{\text{tx}}(\mathbf{r}) + \tau_j^{\text{rx}}(\mathbf{r})), \quad (5.3)$$

which is an instance of the well-known DAS algorithm, where \bar{h}_i^{tx} and \bar{h}_j^{rx} can be interpreted as weighting functions. As the resulting RF image \tilde{x}_i is obtained from a single insonification (transmitter), it usually suffers from strong diffraction artifacts. To improve the resulting image quality, it is common to average coherently multiple (low-quality) RF images reconstructed using DAS-based algorithms from measurements acquired by insonifying the medium of interest differently, for instance using steered PWs or DWs in a process called coherent compounding [65], or using SA imaging [19].

As an alternative to using multiple insonifications to improve the image quality, regularization techniques may be used by solving

$$\hat{\mathbf{x}} = \underset{\mathbf{x} \in \mathbb{R}^n}{\operatorname{argmin}} \{ \mathcal{D}(\mathbf{H}\mathbf{x}, \mathbf{y}) + \kappa \mathcal{R}(\mathbf{x}) \}, \quad (5.4)$$

where $\mathcal{D}: \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}_+$ is a data fidelity term (e.g., the ℓ_2 -norm), and $\mathcal{R}: \mathbb{R}^n \rightarrow \mathbb{R}_+$ is a regularizer used to infer prior knowledge on the expected image. The parameter $\kappa \in \mathbb{R}_+$ controls the weighting of the regularization and is typically adjusted manually. One solution to (5.4), when using the ℓ_2 -norm as data fidelity, may be found using the

well-known proximal gradient descent iteration [66]

$$\mathbf{x}^{(k+1)} = \text{prox}_{\gamma\mathcal{R}}(\mathbf{x}^{(k)} - \gamma\mathbf{H}^*\mathbf{H}\mathbf{x}^{(k)} + \mathbf{H}^*\mathbf{y}), \quad (5.5)$$

where \mathbf{H}^* is the adjoint of \mathbf{H} , $\gamma \in \mathbb{R}_+$ is the gradient step size (that may also be iteration-dependent), and the proximity operator $\text{prox}_{\gamma\mathcal{R}} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is defined as

$$\text{prox}_{\mu\mathcal{R}}(\mathbf{z}) = \underset{\mathbf{x} \in \mathbb{R}^n}{\text{argmin}} \frac{1}{2}\|\mathbf{x} - \mathbf{z}\|_2^2 + \mu\mathcal{R}(\mathbf{x}), \quad (5.6)$$

for some parameter $\mu \in \mathbb{R}_+$. The proximity operator in (5.5) acts as a projection to “denoise” each estimate based on some prior knowledge of \mathbf{x} . Such methods remain seldom used in US imaging given the difficulty of finding a regularizer suitable for the very specific and broad statistical properties of US images (especially considering speckle patterns). Also, their inherent iterative process makes it difficult to deploy them in real-time imaging.

One can note that if (5.5) is initialized to zero, namely $\mathbf{x}^{(0)} = \mathbf{0}$, the first estimate is obtained as $\mathbf{x}^{(1)} = \text{prox}_{\gamma\mathcal{R}}(\mathbf{H}^*\mathbf{y})$, which corresponds to a backprojection followed by some “denoising” projection that depends on the regularizer \mathcal{R} . This observation is the basis of the proposed “two-step” approach.

5.2.2. Statistical Considerations of Fully Developed Speckle

Ultrasound speckle is characteristic of images produced by conventional DAS-based pulse-echo imaging systems; it arises from the coherent interferences of echo-components reflected by sub-resolved diffusive scatterers. It is said to be fully developed when scatterers are present in sufficient numbers within resolution cells, and resulting backscattered signals follow a (circular symmetric) complex normal distribution $\mathcal{CN}(0, 2\sigma^2)$, where σ^2 represents the variance of each (independent) component [78]. Following envelope detection, the signal amplitude of these interferences follow a Rayleigh distribution [142], denoted as $\text{Rayleigh}(\sigma)$, with a parameter $\sigma > 0$ related to the underlying (circular symmetric) complex distribution. The corresponding probability density function (PDF) and cumulative distribution function (CDF) are defined as

$$f(x; \sigma) = \frac{x}{\sigma^2} e^{-x^2/(2\sigma^2)}, \quad (5.7)$$

$$F(x; \sigma) = 1 - e^{-x^2/(2\sigma^2)}, \quad (5.8)$$

for $x \geq 0$, respectively. The first moment (i.e., mean or expected value), the second

moment, and the variance of a Rayleigh random variable X are given by

$$\mu_1 = \mathbb{E}[X] = \sqrt{\frac{\pi}{2}}\sigma, \quad (5.9)$$

$$\mu_2 = \mathbb{E}[X^2] = 2\sigma^2, \quad (5.10)$$

$$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{4 - \pi}{2}\sigma^2. \quad (5.11)$$

Even though speckle patterns are sometimes interpreted as noise, they contain positional information about the underlying physical phenomenon, as they result from deterministic interferences, and are therefore extensively exploited in motion analysis [6].

5.2.2.1. First Order Statistics

The analysis of first order statistics provides useful tools to characterize envelope signals regardless of the acquisition system geometry and are therefore extensively used in image quality metrics and tissue characterization. A widely used measure of first order statistics in US imaging is the ratio of mean to standard deviation of a signal (i.e., the reciprocal of the coefficient of variation), often referred to as SNR [142]. In the case of a signal following a Rayleigh distribution, it is given by

$$\text{SNR} = \frac{\mathbb{E}[X]}{\sqrt{\text{Var}(X)}} = \sqrt{\frac{\pi}{4 - \pi}} \approx 1.91. \quad (5.12)$$

This ratio is ideally estimated at a single point in an image by conducting multiple independent realizations and estimating $\mathbb{E}[X]$ and $\text{Var}(X)$ using the sample mean and sample variance. Assuming a zone of a physical domain composed of a large amount of random diffuse scatterers with constant mean amplitude imaged with a system characterized by a slowly varying PSF within such a zone, the resulting speckle patterns will inherit quasi-constant statistical properties. Therefore, one can assume a wide-sense stationary (WSS) process within that image zone and estimate the SNR directly from the samples obtained. It should be noted that, in an attempt to reduce speckle “noise,” one may want to improve the SNR defined in (5.12) by reducing $\text{Var}(X)$. Yet, in scenarios where accurate speckle patterns are required (e.g., motion estimation), the goal is to preserve such patterns or to restore them as they may have been altered by imaging artifacts or thermal noise. The SNR can thus serve to verify that speckle patterns follow the expected (first order) statistics.

5.2.2.2. Second Order Statistics

To study the spatial characteristics of speckle patterns, which depend on the PSF of the imaging system, the evaluation of second order statistics is required. The normalized autocovariance function (ACF), also referred to as Pearson correlation coefficient (PCC),

is commonly used for this purpose [78], [79]. Assuming a WSS process, it is defined as

$$\rho_{XX}(\Delta \mathbf{r}) = \frac{\mathbb{E}[(X(\mathbf{r}_1) - \mathbb{E}[X])(X(\mathbf{r}_2) - \mathbb{E}[X])^*]}{\text{Var}(X)}, \quad (5.13)$$

where \mathbf{r}_1 and \mathbf{r}_2 are two positions (in the image), $\Delta \mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$, and z^* denotes the complex conjugate of z . Note that the numerator of (5.13) is simply the autocorrelation of the image amplitudes from which the mean was subtracted, and may therefore be efficiently estimated using two fast Fourier transforms (FFTs). To characterize the resolution of speckle patterns, also referred to as speckle spot size, the FWHM of the ACF is typically evaluated in all image dimensions.

5.2.2.3. Contrast

A commonly used definition of contrast between two US signals (or image zones) X_1 and X_2 is given by their ratio of mean amplitude, expressed in decibels as [143]

$$C = 20 \log_{10} \left(\frac{\mathbb{E}[X_1]}{\mathbb{E}[X_2]} \right). \quad (5.14)$$

Considering a Rayleigh random variable X , one can note that, from (5.9) and (5.10), $\mathbb{E}[X] = (\pi \mathbb{E}[X^2]/4)^{1/2}$. Hence, if $X_1 \sim \text{Rayleigh}(\sigma_1)$ and $X_2 \sim \text{Rayleigh}(\sigma_2)$, (5.14) can be equivalently expressed, in decibels, on the signal intensity (power) as

$$C = 10 \log_{10} \left(\frac{\mathbb{E}[X_1^2]}{\mathbb{E}[X_2^2]} \right). \quad (5.15)$$

When considering fully developed speckle signals that follow a Rayleigh distribution, the analysis of the signal intensity, which follows an exponential distribution, may be of interest as it is linearly proportional to the concentration of scatterers [143]. However, in general, the signal amplitude is the quantity of interest in US imaging as US systems sense RF signals that are linearly proportional to the amplitude of scatterers [78], irrespectively of their statistical properties. Hence, (5.15) should not be used on signals deviating from Rayleigh statistics as it could result in unrealistic values.

5.2.2.4. Confidence Interval

As the CDF of a Rayleigh random variable, defined in (5.8), is continuous and strictly monotonically increasing, its quantile function $Q = F^{-1}$, and can be expressed as

$$Q(y; \sigma) = \sigma \sqrt{-2 \ln(1 - y)}, \quad (5.16)$$

for $y \in [0, 1)$. Considering a symmetric confidence level $\beta \in [0, 1)$, the lower and upper

confidence bounds are obtained directly from (5.16), and expressed as

$$\left(\sigma \sqrt{-2 \ln \left(\frac{1 + \beta}{2} \right)}, \sigma \sqrt{-2 \ln \left(\frac{1 - \beta}{2} \right)} \right). \quad (5.17)$$

Figure 5.1 shows the confidence bounds of a Rayleigh random variable normalized by its expected value, namely $X/\mu_1 \sim \text{Rayleigh}(\sqrt{2/\pi})$. A 90 % confidence level is therefore achieved when accounting values ranging from approximately -12 dB to $+6$ dB with respect to its expected value (i.e., mean).

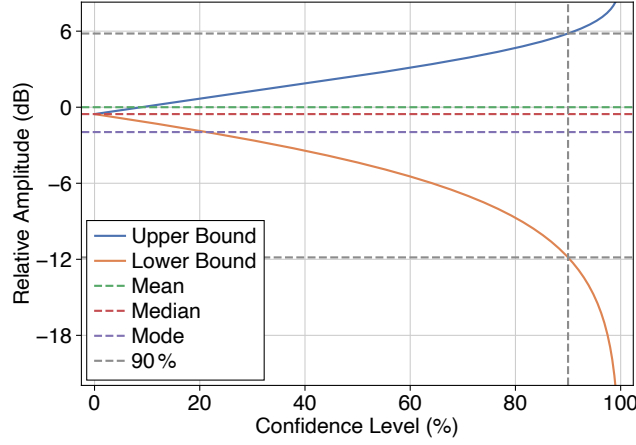


Figure 5.1. Lower and upper confidence bounds (in decibels) with respect to the confidence level (in percent) of a Rayleigh random variable normalized by its expected value (i.e., mean).

5.3. Methods

5.3.1. Proposed Image Reconstruction Method

The proposed method consists of first computing a low-quality estimate of \mathbf{x} , denoted $\tilde{\mathbf{x}}$, from measurements \mathbf{y} acquired with a single insonification, by means of a re-weighted backprojection-based DAS algorithm $\mathbf{D}: \mathbb{R}^m \rightarrow \mathbb{R}^n$. We define $\mathbf{D} := \mathbf{W}\mathbf{H}^*$, where $\mathbf{W}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a “pixel-wise” re-weighting operator (diagonal matrix) defined, for the i -th transmitter, as

$$w_i(\mathbf{r}) = \left(\bar{h}_i^{\text{tx}}(\mathbf{r}) \sum_{j=1}^{n_r} \bar{h}_j^{\text{rx}}(\mathbf{r}) \right)^{-1}. \quad (5.18)$$

It has been designed to compensate for the amplitude-related effects of (far-field) diffraction and can be interpreted as post-DAS image equalization. In a second step, the resulting approximation $\tilde{\mathbf{x}} = \mathbf{D}\mathbf{y}$ is fed to a CNN $\mathbf{f}_{\theta}: \mathbb{R}^n \rightarrow \mathbb{R}^n$, with parameters θ , trained to recover a high-quality estimate of \mathbf{x} as $\hat{\mathbf{x}} = \mathbf{f}_{\theta}(\tilde{\mathbf{x}})$. As opposed to end-to-end approaches

seeking to map a measurement space to an image space directly, we suggest using a CNN for a task in which they are well-known to excel, namely restoration tasks [127], [144].

Put formally, we seek to train a mapping \mathbf{f}_θ between a subspace of low-quality images $W \subset \mathbb{R}^n$ to a subspace of “ground-truth” images $V \subset \mathbb{R}^n$. In order to define these subspaces more precisely, let us consider an US transducer with a given aperture, composed of an array of transducer elements, with given geometry, center frequency, and bandwidth properties, and designed to operate at a given transmit frequency.

We define W as the subspace of US images reconstructed using \mathbf{D} from measurements acquired by a single insonification using the entire aperture. These images are typically contaminated by high SL and EW artifacts (mostly due to the single insonification), as well as potential GL artifacts, if the spatial sampling of the aperture is sub-optimal (e.g., as it is the case in linear-array designs).

For V , we propose to use a transducer similar to the one used for W , namely spanning the same aperture and composed of transducer elements with the same physical properties, but with a spatial sampling ensuring the absence of GLs. In order to produce reference images from this array, we reconstruct them from the complete set of SA measurements obtained by sequentially insonifying the medium using a single element of the array and recording the scattered signals with all elements. Each insonification is reconstructed using the corresponding \mathbf{D} , and they are averaged to produce the final image. As SA makes it possible to virtually focus both on transmit and receive in every point of the image [19], it produces images with a high resolution (tightened main lobe) while minimizing the level of SL and EW artifacts (increased contrast). As a consequence, SA is often considered as the gold standard in US imaging.

Furthermore, in comparison to using more “exotic” ground-truth images (e.g., CT-like), SA images retain high-quality speckle patterns. By ensuring that both subspaces contain speckle patterns arising from sub-resolution scattering interferences, we enforce the CNN to preserve their information content. We emphasize the fact that, due to the assumptions considered in the physical measurement process (5.2) and resulting backprojection operator (5.3) used to define both W and V , the resulting trained CNN is not expected to correct artifacts arising from neglected physical phenomena, such as multiple scattering, speed-of-sound distortions, dispersive attenuation, or non-linear acoustic effects. The focus is on reducing diffraction artifacts and preserving speckle with an increased resolution.

5.3.2. Convolutional Neural Network Architecture

The proposed CNN architecture [Figure 5.2(a)] is derived from our previous work [137], adapted from the popular U-Net architecture [145], as well as the architecture used in [31], [32]. It is a residual CNN, expressed as $\mathbf{f}_\theta(\mathbf{x}) = \mathbf{x} + \mathbf{r}_\theta(\mathbf{x})$, designed to predict the negative noise to be applied to some input \mathbf{x} . It is composed of a series of multi-channel convolutional layers (CLs) and rectified linear units (ReLUs), arranged in a downsampling (left arm) and upsampling (right arm) paths, with intrinsic skip connections to mitigate information losses. The multi-scale structure confers a large receptive field to the CNN,

particularly adapted to the non-stationary restoration mapping to be learned.

The input image first undergoes a channel expansion (leftmost chamoisee arrow) up to N_c channels. It is then fed to a series of convolutional blocks (red arrows) followed by downsampling layers (blue arrows) which reduce the spatial dimension while augmenting the channel number. The upsampling path is performed symmetrically using intrinsic skip connections (yellow arrows), convolutional blocks, and upsampling layers (violet arrows). Eventually, the channel number N_c is contracted back to its initial state (rightmost chamoisee arrow) and the output is then summed to the input image (residual skip connection).

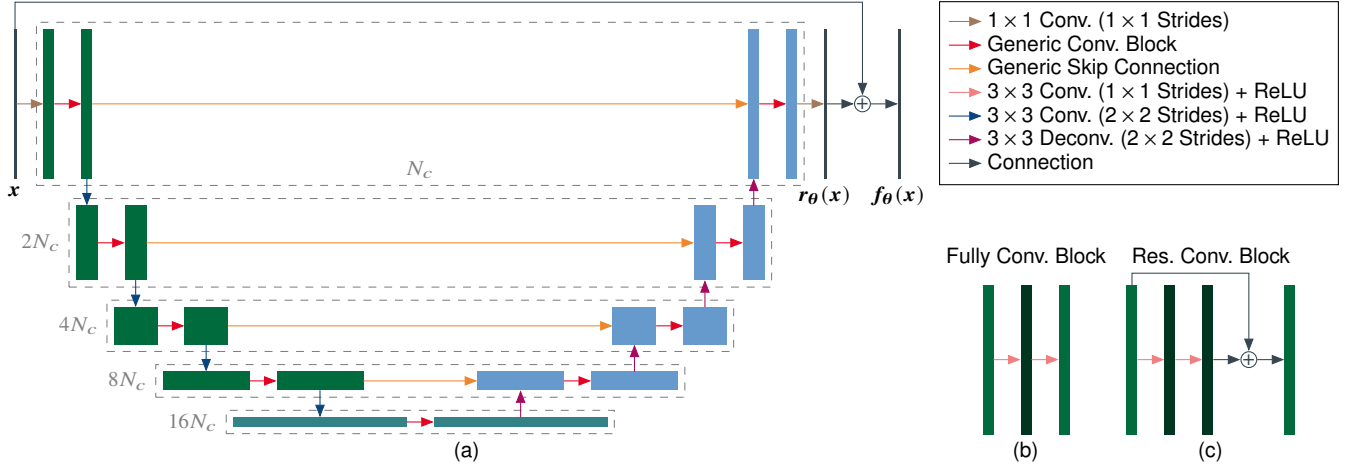


Figure 5.2. Proposed residual convolutional neural network (CNN) architecture, adapted from U-Net [145], and from [31], [137]: (a) generic overall CNN architecture; convolutional blocks considered, namely (b) conventional fully convolutional block (FCB) and (c) proposed residual convolutional block (RCB). Connections and tensor operations (i.e., layers) are represented as straight, colored arrows (legend in the upper right corner). The 3-D tensors are represented as colored rectangles. (Their width and height are depicted in proportion to the number of channels and image dimension, respectively. One image dimension is not represented for readability reasons.)

The main differences with respect to our initial adaptation [137] are as follows. Instead of a max-pooling layer, which seems inadequate for restoration tasks, a 2×2 strided CL was used within each downsampling layer (blue arrow) as a symmetric counterpart to the 2×2 strided “transposed” CL used within each upsampling layer (violet arrow). We used simple additive intrinsic skip connections instead of concatenated ones conventionally used in U-Net-like architectures, resulting in a symmetric amount of trainable parameters in both arms. A residual convolutional block (RCB) [Figure 5.2(c)] is proposed to supersede the fully convolutional block (FCB) [Figure 5.2(b)] used previously. Note that such RCBs would not have been possible to be deployed with concatenated skip connections directly.

5.3.3. Training Loss for High-Dynamic-Range Radio-Frequency Data

The trainable model parameters θ are optimized in a supervised manner by minimizing the empirical risk

$$R(\theta) = \frac{1}{l} \sum_{i=1}^l \mathcal{L}(\mathbf{x}^{(i)}, f_{\theta}(\tilde{\mathbf{x}}^{(i)})), \quad (5.19)$$

where $\mathcal{L}(\mathbf{x}, \hat{\mathbf{x}})$ is a non-negative real-valued (training) loss function, which measures the distance between a prediction $\hat{\mathbf{x}} = f_{\theta}(\tilde{\mathbf{x}})$ and its true value \mathbf{x} . The training set $\{\{\mathbf{x}^{(1)}, \tilde{\mathbf{x}}^{(1)}\}, \dots, \{\mathbf{x}^{(l)}, \tilde{\mathbf{x}}^{(l)}\}\}$ is composed of l image pairs. Common loss functions include the MSE loss and the mean absolute error (MAE) loss expressed as

$$\mathcal{L}_{\text{MSE}}(\mathbf{x}, \hat{\mathbf{x}}) = \frac{1}{n} \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2, \quad (5.20)$$

$$\mathcal{L}_{\text{MAE}}(\mathbf{x}, \hat{\mathbf{x}}) = \frac{1}{n} \|\mathbf{x} - \hat{\mathbf{x}}\|_1, \quad (5.21)$$

respectively.

Due to the inherent HDR property of US images, they are commonly compressed (after envelope detection) before being displayed for interpretation. To account for the HDR property of US images while preserving their RF nature, we introduce the MSLAE, inspired by both the log-compression applied to visualize US images and audio-coding companding algorithms (pulse code modulation). The associated loss is expressed as

$$\mathcal{L}_{\text{MSLAE}}(\mathbf{x}, \hat{\mathbf{x}}) = \frac{1}{n} \|g_{\alpha}(\mathbf{x}) - g_{\alpha}(\hat{\mathbf{x}})\|_1, \quad (5.22)$$

where $g_{\alpha} : \mathbb{R} \rightarrow \mathbb{R}$ is a signed (clipped-and-scaled) logarithmic transform defined element-wise as

$$g_{\alpha}(x_k) = \text{sign}(x_k) \log_{\alpha} \left(\frac{\alpha}{\max(\alpha, |x_k|)} \right), \quad (5.23)$$

where $\alpha \in (0, 1)$ and x_k is an element of \mathbf{x} (e.g., a pixel value). It should be noted that $g_{\alpha}(x_k) = 0 \forall |x_k| < \alpha$. Thus, α can be interpreted as a threshold parameter below which a (pixel) value is assumed “negligible.”

To anticipate the effect of the MSLAE loss, let us define a predicted value $\hat{x} = \epsilon x$ for any true value $x \in \mathbb{R}$ and error ratio $\epsilon \in \mathbb{R}$. (Note that the component index k has been

dropped to lighten notation.) The resulting loss function can be expressed as

$$\begin{aligned} \mathcal{L}_{\text{MSLAE}}(x, \varepsilon x) &= |g_\alpha(x) - g_\alpha(\varepsilon x)| \\ &= \begin{cases} |\log_\alpha(\varepsilon)| & \text{for } |x| > \alpha, |\varepsilon x| > \alpha, \varepsilon > 0, \\ |\log_\alpha(-\alpha^2/(\varepsilon x^2))| & \text{for } |x| > \alpha, |\varepsilon x| > \alpha, \varepsilon < 0, \\ |\log_\alpha(\alpha/|x|)| & \text{for } |x| > \alpha, |\varepsilon x| \leq \alpha, \\ |\log_\alpha(\alpha/|\varepsilon x|)| & \text{for } |x| \leq \alpha, |\varepsilon x| > \alpha, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \quad (5.24)$$

For comparison purposes, and as it served as inspiration for the proposed MSLAE loss, let us also define the mean μ -law absolute error (MMUAE) loss function as

$$\mathcal{L}_{\text{MMUAE}}(\mathbf{x}, \hat{\mathbf{x}}) = \|g_\mu(\mathbf{x}) - g_\mu(\hat{\mathbf{x}})\|_1, \quad (5.25)$$

where $g_\mu: \mathbb{R} \rightarrow \mathbb{R}$ is the μ -law transform (commonly used in audio companding algorithms) defined element-wise as

$$g_\mu(x_k) = \text{sign}(x_k) \ln\left(\frac{1 + \mu|x_k|}{1 + \mu}\right), \quad (5.26)$$

where $\mu \in \mathbb{R}_+$ defines the extent of dynamic range compression. Note that to obtain a dynamic range compression similar to that of (5.23), μ must be set to α^{-1} . Proceeding in the same way as for the derivation of (5.24), and using $\nu = 1 + \mu$, one can express the resulting loss for MMUAE as

$$\begin{aligned} \mathcal{L}_{\text{MMUAE}}(x, \varepsilon x) &= |g_\mu(x) - g_\mu(\varepsilon x)| \\ &= \begin{cases} |\log_\nu((1 + \mu\varepsilon|x|)/(1 + \mu|x|))| & \text{for } \varepsilon > 0, \\ |\log_\nu((1 - \mu\varepsilon|x|)(1 + \mu|x|))| & \text{for } \varepsilon < 0. \end{cases} \end{aligned} \quad (5.27)$$

From (5.24) and (5.27), one can note that both losses are not differentiable for $\varepsilon = 1$, namely for $\hat{x} = x$ (similarly to MAE). They both penalize sign errors, which means that they can preserve the RF property of US images. The main advantage of MSLAE over MMUAE resides in the fact that, for any true x and $\varepsilon > 0$ such that $|x|, |\varepsilon x| > \alpha$, the loss is a positive constant value (i.e., independent of x). Consequently, a specific error ratio between a predicted value and its true counterpart is penalized equally, regardless of the true value magnitude.² This is a highly desirable feature when working on HDR data, as is the case for (RF) US images. Due to the “threshold” parameter α , MSLAE is also not differentiable in a few other cases, namely for $|x| = \alpha$ and/or $|\varepsilon x| = \alpha$. Also, note that in cases where both $|x|, |\varepsilon x| < \alpha$, the penalty is zero. Therefore, the parameter α must be selected carefully based on the statistics of the dataset considered (Section 5.4.4).

²Note that MMUAE approximates such a behavior.

Note that the same derivation can be applied to both MSE and MAE losses, resulting in

$$\mathcal{L}_{\text{MSE}}(x, \varepsilon x) = (1 - \varepsilon)^2 x^2, \quad (5.28)$$

$$\mathcal{L}_{\text{MAE}}(x, \varepsilon x) = |(1 - \varepsilon)x|. \quad (5.29)$$

From (5.28) and (5.29), it is clear that MSE and MAE are not optimal in the context of HDR data as the resulting loss value is proportional to the true value x (i.e., quadratically for MSE and linearly for MAE).

Other log-compressed loss functions were proposed in the context of deep learning and US. In [132], conventional loss functions were computed on log-compressed images (i.e., limited to non-negative data). In [135], the signed-mean-squared-logarithmic error (SMSLE) was introduced as a loss function. Because the SMSLE operates separately on the positive and negative parts of RF signals, it is limited to inputs and predictions that oscillate identically (i.e., cannot account for sign errors, making it unusable in the present study). Also, both losses have a singularity at zero and can become highly unstable as (pixel) values tend to zero.

5.4. Experimental Setup

5.4.1. Imaging Configurations

The imaging configurations considered in this study (Table 5.1) are based on the 9L-D transducer (GE Healthcare, Chicago, Illinois, USA) and the Vantage 256 system (Verasonics, Kirkland, WA, USA) specifications. The 9L-D is a linear array, composed of 192 transducer elements with a center frequency of 5.3 MHz and a bandwidth of 75 % (at -6 dB). The transmit excitation is a single-cycle tri-state waveform of 67 % duty cycle centered at 5.208 MHz, with leading and trailing equalization pulses of quarter-cycle durations and opposite polarities. The received echo signals are sampled at 20.833 MHz (200 % bandwidth sampling).

We first introduce two “natural” imaging configurations, namely low-quality (LQ) and high-quality (HQ), defined by the properties of the 9L-D. A single PW with normal incidence and without apodization is transmitted in the LQ configuration. The complete set of 192 SA measurements are used for HQ. Assuming a typical speed of sound in soft tissue of 1540 m/s, one can note that the element spacing (i.e., pitch) in LQ and HQ configurations is $\sim 0.78\lambda$, namely greater than half the effective wavelength. Hence, images reconstructed by conventional DAS-based algorithms will inevitably be contaminated by GL artifacts.

Furthermore, we introduce the ultra-high-quality (UQ) configuration as an idealized configuration. It is based on the HQ one, but takes advantage of a spatially oversampled aperture with a halved pitch ($\sim 0.39\lambda$), resulting in a virtual 383-element array, guaranteeing GL-free images. In order to preserve the same speckle patterns obtained using the HQ configuration while removing GL artifacts, the same aperture and geometric properties

TABLE 5.1
Specifications of the Imaging Configurations Considered

Parameter	LQ	HQ	UQ ^a
Center frequency	5.3 MHz	5.3 MHz	5.3 MHz
Bandwidth	75 %	75 %	75 %
Aperture	43.93 mm	43.93 mm	43.93 mm
Element number	192	192	383
Pitch	230 μm	230 μm	115 μm ^a
Element width ^b	207 μm	207 μm	207 μm ^a
Element height	6 mm	6 mm	6 mm
Elevation focus	28 mm	28 mm	28 mm
Transmit frequency	5.208 MHz	5.208 MHz	5.208 MHz
Excitation cycles ^c	1	1	1
Transmit-receive scheme	1 PW	192 SA	383 SA
Sampling frequency	20.833 MHz	20.833 MHz	20.833 MHz

^aUQ is not physically possible and can only be simulated.

^bGuessed (no official data available).

^cSingle excitation cycle with equalization pulses.

of the elements were kept. The configuration can thus only be realized in a simulation environment.

For each imaging configuration considered, the images were reconstructed using the corresponding backprojection-based DAS operator \mathbf{D} , defined in Sections 5.2.1 and 5.3.1, for which the scalar weighting functions \bar{h}_i^{tx} and \bar{h}_j^{rx} and the delay functions τ_i^{tx} and τ_j^{rx} need to be specified. In case of PW acquisitions (i.e., LQ), an idealized wavefront was assumed as transmitter, namely $\bar{h}_i^{\text{tx}}(\mathbf{r}) = 1$. In case of SA acquisitions (i.e., HQ and UQ), each transmission was performed with a different transducer element. The diffraction effect of a narrow element evaluated at a field point \mathbf{r} can be derived from a 2-D far-field assumption considering a soft baffle boundary condition as [49], [61]

$$\bar{h}_i^{\text{tx}}(\mathbf{r}) = \frac{d \operatorname{sinc}(d/\lambda \sin(\theta))}{\sqrt{2\pi} \|\mathbf{r} - \mathbf{r}_i\|_2^{1/2}} \cos(\theta), \quad (5.30)$$

where \mathbf{r}_i is the position of the transducer element, θ is the angle between the element normal and the vector $\mathbf{r} - \mathbf{r}_i$, d is the width of the element, and $\operatorname{sinc}(x) := \sin(\pi x)/(\pi x)$. The transducer elements are also the receivers for all imaging configurations and (5.30) was also used to evaluate $\bar{h}_j^{\text{rx}}(\mathbf{r})$, with \mathbf{r}_j the position of the receiving element. The time delay functions τ_i^{tx} and τ_j^{rx} were computed from the distance traveled by the wavefront from the transmitter to a field point \mathbf{r} and from a field point \mathbf{r} to the receiver, respectively, divided by the mean sound speed (assumed 1540 m/s), as in Section 3.3.

It is worth highlighting a slight variation with respect to the weighting function deployed in Section 3.3. The difference comes from the 2-D assumption made here contrary to the 3-D one defined in (3.16). This choice was motivated because our initial 3-D assumption does not hold in the case of 2-D imaging using conventional linear or phased transducer arrays that are composed of (focused) transducer elements with a height-to-width ratio of more than an order of magnitude. The use of (5.30) instead of (3.16)

resulted in images better equalized. Also note that the physical units do not match in (5.30) because of the non-physical 2-D approximation.

The interpolation of element raw-data values (before summation) was performed using a B-spline approximation of degree three [68]. Analytic (complex) images, often called in-phase quadrature (IQ) images, were reconstructed from the analytic raw-data signals, enabling us to have direct access to the RF (real part) and envelope (modulus) image representations. The process was implemented with PyUS,³ a GPU-accelerated Python package for US imaging developed in our laboratory.

The images were reconstructed with a width spanning the 9L-D aperture (Table 5.1) and a depth from 1 mm to 60 mm. A $\lambda/4 \times \lambda/8$ (Cartesian) grid was used to guarantee Nyquist sampling of RF images in both dimensions, resulting in images of 596×1600 pixels. Signals were sampled from 0 μ s to 96.58 μ s to account for all echoes contributing to the final image limits.

As the PSF of DAS-based pulse-echo US imaging systems is spatially varying, especially when considering ultrafast acquisitions, a generic analysis is a complicated task. Yet, the PSF varies slowly over the image domain and its visualization at some locations in the image provides meaningful information about its spread and enables comparing different imaging configurations. Figure 5.3 shows simulated PSFs, evaluated in three distinct positions, for the LQ, HQ, and UQ imaging configurations. One can note that GL artifacts are drastically reduced between LQ and HQ, and are completely removed for UQ. Artifacts caused by SLs are easily identifiable as they develop from main lobes in “cross”-like artifacts. The spread and amplitude of these artifacts are drastically reduced for HQ and UQ with respect to LQ. Artifacts caused by EWs, which are only present in the LQ configuration [Figure 5.3(b)], are the most spatially varying ones and appear as two “defocused” duplicates below each main lobe (except in the center of the lateral dimension where they interfere coherently). The deeper the position in the image, the closer EW artifacts are to the main lobe, and the more they resemble the combination of a main lobe and associated SLs.

From these observations, it is clear that all three imaging configuration considered are characterized by spatially varying PSFs, and that this spatially varying property is most pronounced for the LQ configuration. Therefore, as we seek to learn a restoration mapping (using a CNN) to recover high-quality estimates from low-quality ones, such a mapping needs to be non-stationary as well. Moreover, as the PSF of the LQ configuration spreads over a large portion of the image because of diffraction artifacts, the restoration mapping needs a large receptive field to be effective. This observation was critical to the design of the proposed CNN architecture (Section 5.3.2). In particular, its multi-scale structure results in a large receptive field even when using convolutional kernels of small supports (3×3 in our case).

³<https://gitlab.com/pyus/pyus>

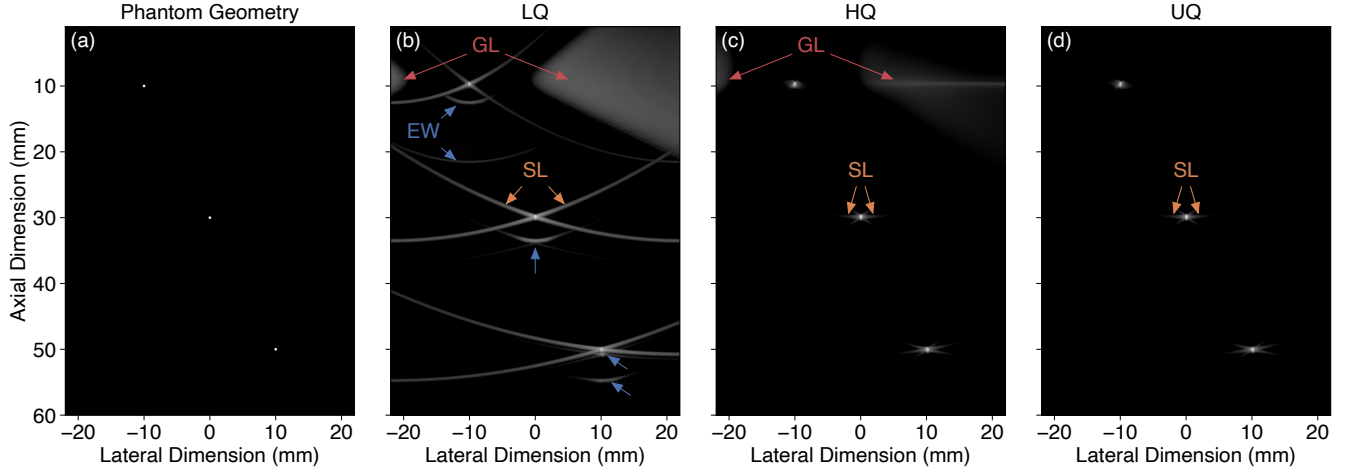


Figure 5.3. B-mode image representations (98-dB range) of simulated point spread function (PSF) examples: (a) point reflector positions in which the PSFs were evaluated; images reconstructed using each imaging configuration considered (Table 5.1), namely (b) low-quality (LQ) configuration, (c) high-quality (HQ) configuration (i.e., gold-standard image for the physical transducer array), and (d) ultra-high-quality (UQ) configuration (i.e., gold-standard image for the spatially oversampled virtual version of the transducer array, considered as ground-truth). Some zones dominated by grating lobe (GL), side lobe (SL), and edge wave (EW) artifacts are highlighted by colorized arrows and associated annotations.

5.4.2. Element Raw-Data Generation

For the generation of the dataset image pairs used to train f_{θ} , we relied on computer simulations. This enables the generation of a training set in a fully controlled environment. In particular, it can overcome the difficulties encountered in SA imaging. Indeed, tissue displacements, attenuation effects, and thermal noise can make such acquisitions rather challenging in experimental conditions (especially at great depth). An arbitrarily diverse dataset can be crafted using simulation, which may be difficult to achieve on experimental set-ups. Furthermore, transducers with optimal spatial sampling and an arbitrary number of elements, even beyond what is physically feasible, are possible by means of simulations. While many commercial US scanners do not allow raw RF data to be collected [133], simulations enable generating realistic RF data provided that governing transducer parameters are available (or possible to estimate).

In order to generate realistic element raw-data, we relied on the exact SIR model described in (5.1), as opposed to the approximated measurement model expression given in (5.2), which was used to derive a computationally tractable backprojection operator. This means that neither the far-field approximation of the SIRs nor the Dirac approximation of the pulse-echo waveform was assumed for simulating the element raw-data involved in the generation of the simulated-image dataset. Equation (5.1) can be accurately evaluated using the well-known Field II simulator [87], [88]. Its only issue lies in the computing time requirements, especially for simulating SA measurements in conditions such as to produce fully developed speckle.

To generate a sufficiently large and realistic dataset in a reasonable time-frame, we implemented an in-house GPU-accelerated simulator for evaluating (5.1). The simulator is described in great detail in Chapter 4, with a brief summary of the key features

described hereafter. The main differences compared with Field II are the spline-based representations used for both time and element surface domains enabling it to exploit fully the high parallelizability offered by the SIR model. NURBS representations and Gauss-Legendre quadrature are used for surface integrations, enabling high accuracy with few integration points. The time domain is represented in a B-spline basis, thus reducing the sampling frequency requirements. It is a full 3-D implementation, including element directivity and elevation effects. The implementation has been validated against Field II and enabled us to obtain an overall speed-up of about 200 \times .

The transducer elements were exactly represented by cylindrical NURBS surfaces of degree (1, 2), and 3×87 quadrature points were used for the surface integral involved in evaluating the SIR terms of (5.1). Their electromechanical impulse response was approximated by a differentiated log-normal-windowed sine wave. A soft boundary condition was assumed and a constant speed of sound of 1540 m/s was set. To minimize the simulation time, we used a B-spline approximation of degree five [68] for the time domain representation, enabling us to obtain a sufficient accuracy (i.e., >60 dB) with a sampling frequency of 31.25 MHz.

5.4.3. Simulation Phantoms

Each numerical phantom is represented by a set of ideal point reflectors defined by their position and echogenicity (amplitude). Figure 5.4 depicts the geometric considerations, in the transducer plane, relevant to the design and generation of the synthetic phantoms. The domains $\Omega_i \subset \Omega_r \subset \Omega_l \subset \Omega_e \subset \mathbb{R}^3$ are defined by the transducer aperture and the considered imaging configurations (Section 5.4.1). All possible scattered contributions to the final image domain Ω_i arise from Ω_l , which is bounded by the Cartesian domain Ω_e . To mitigate the computing time and enable the simulation of a sufficiently large dataset, we only considered contributions arising from Ω_r , where $\theta_r = 29.71^\circ$ is the angle from which the element sensitivity falls below -6 dB.

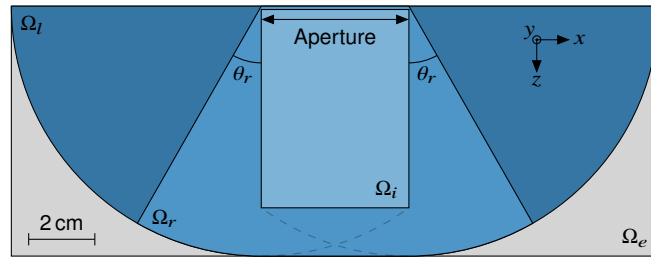


Figure 5.4. Representation (in the transducer plane) of the spatial domains used for generating the simulation phantoms for the imaging configurations considered (Table 5.1).

As we are particularly interested in generating speckle patterns, the positions of scatterers were drawn from a uniform distribution over Ω_e and their amplitudes were drawn from a normal distribution, for each synthetic phantom, providing a uniform speckle reference background (defined as 0 dB). We computed the most restrictive (3-D) resolution cell, defined by the FWHM in all dimensions, for the UQ imaging configuration, namely

$0.71\lambda \times 3.23\lambda \times 1.10\lambda$. To mitigate the computing time while obtaining fully developed speckle patterns throughout the resulting images, an average of ten scatterers per resolution cell was used, namely the lower bound prescribed in [146, Sec. 8.4.4]. To mitigate the computing time further, only a single resolution cell was considered in elevation. This resulted in ~ 153.35 scatterers/mm³ and a total of $\sim 900\,000$ effective scatterers in Ω_r .

Two hundred ellipsoidal zones were incorporated at random positions and with random orientations. The size of each semi-axis was uniformly selected between 0.71λ and 71λ (i.e., between 1 and 100 times the smallest resolution-cell dimension). These choices guaranteed an almost complete filling of Ω_e . Within each inclusion, a mean echogenicity drawn uniformly randomly between -50 dB and $+30$ dB with respect to the background, was set by scaling the initial random amplitude of scatterers contained within it accordingly. Inclusions were processed in descending order of surface area to maximize the resulting diversity of shapes.

5.4.4. Simulated Dataset

A total of 31 000 simulation phantoms were generated as described in Section 5.4.3, each defined by a unique set of scatterers. Element raw-data signals were simulated (Section 5.4.2) from each simulation phantom and for each imaging configuration (Table 5.1). Corresponding images were reconstructed using the associated backprojection operators (Section 5.4.1). For each imaging configuration, a normalization factor was determined on independent realizations of a reference simulation phantom composed of scatterers resulting in a constant mean echogenicity of 0 dB (background reference).

Assuming a fully developed speckle zone of constant mean echogenicity, image envelope values follow a Rayleigh distribution [78]. Thus, for a Rayleigh distributed speckle with a 0 dB mean echogenicity, an interval of -12 dB to $+6$ dB covering 90 % of the envelope values can be determined (Section 5.2.2.4). Since the simulated phantoms contain inclusions of constant mean echogenicities ranging from -50 dB to $+30$ dB, an interval of -62 dB to $+36$ dB (i.e., 98-dB range) guarantees an absolute minimum of 90 % envelope value coverage.

Representations of a simulated dataset sample are shown in Figure 5.5 on a 98-dB range. One can appreciate the significant difference in artifact levels between LQ, HQ, and UQ configurations [Figures 5.5(b) to 5.5(d)]. In particular, the image obtained from the UQ configuration, which exploits a spatially oversampled optimal version of the array, is indeed free from GL artifacts, while the HQ image still suffers from them (although significantly reduced compared with LQ). Both HQ and UQ configurations result in images free from EW artifacts and with SL artifacts substantially reduced compared with LQ. Due to the spatial dependency of the imaging configuration PSFs (Section 5.4.1), the spread of these artifacts is also spatially dependent.

The total computing time required to generate a single sample for all imaging configurations, which was mainly governed by the element raw-data simulation, was ~ 1500 s on a single NVIDIA GeForce GTX 1080 Ti GPU. The complete dataset was simulated on multiple GPUs for about six weeks.

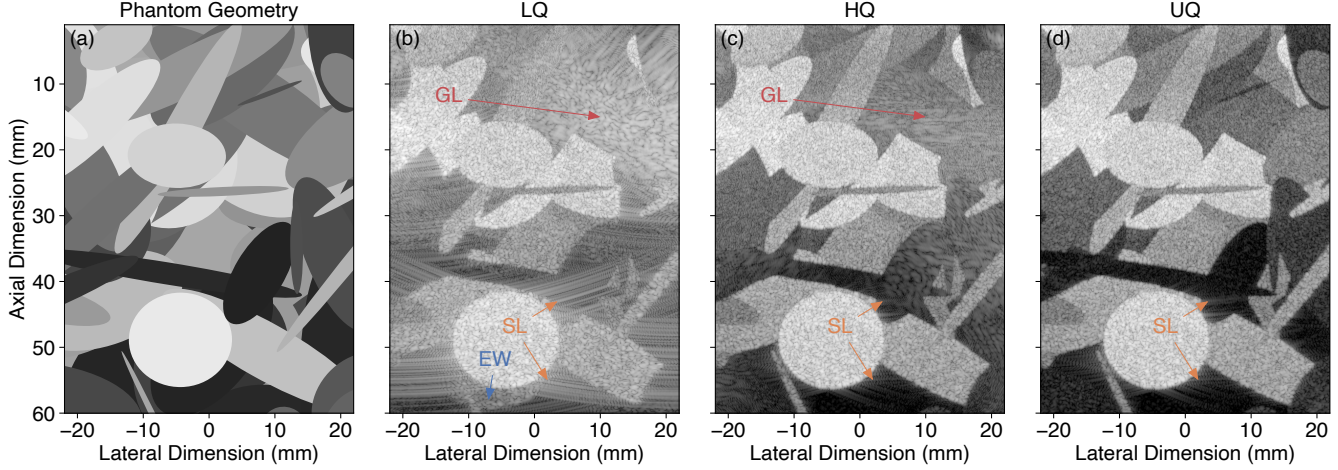


Figure 5.5. B-mode image representations (98-dB range) of a simulated dataset sample: (a) the phantom mask composed of elliptical inclusions; images reconstructed using each imaging configuration considered (Table 5.1), namely (b) low-quality (LQ) configuration, (c) high-quality (HQ) configuration (i.e., gold-standard image for the physical transducer array), and (d) ultra-high-quality (UQ) configuration (i.e., gold-standard image for the spatially oversampled virtual version of the transducer array, considered as ground-truth). Some zones dominated by grating lobe (GL), side lobe (SL), and edge wave (EW) artifacts are highlighted by colorized arrows and associated annotations.

5.5. Training and Hyperparameter Search

5.5.1. Training and Validation Setup

Many training experiments were performed for hyperparameter search (Section 5.5.2). For each training experiment, kernel weights were initialized using the well-known Glorot initialization [147], also referred to as Xavier initialization, implemented with a uniform distribution, and biases were initialized to zero. Model parameters were optimized using the Adam optimizer [148] with a learning rate of 5×10^{-5} , exponential decay rates for both the 1st and the 2nd moment estimates $\beta_1 = 0.9$ and $\beta_2 = 0.999$, as well as a numerical stability constant $\hat{\epsilon} = 10^{-7}$. Mini-batch learning was deployed with a batch size of two (due to memory constraints). The training set consisted of 30 000 image pairs (i.e., low-quality input and corresponding high-quality reference). The size of the training set was motivated by a dedicated study to prevent training experiments from overfitting (Section 5.5.2.6). A total of 500 000 iterations were performed (stopping criterion), corresponding to ~ 33 epochs. Complete random shuffling of the training set was performed after each epoch. Neither training regularization (e.g., dropout or weight regularization) nor data augmentation was used. To fulfill the downsampling restrictions imposed by the proposed CNN architecture (Figure 5.2), namely four 2×2 strided convolutions in the image domain, input images were zero padded symmetrically to reach the closest supported image shape (i.e., 608×1600), and cropped to their original size after inference.

To monitor and evaluate the performance of each training experiment, we used a validation set of 500 image pairs (extracted from the simulated dataset). Both the peak signal-to-noise ratio (PSNR) and the structural similarity (SSIM) index [149] were computed at each validation step (i.e., every 1000 iterations). These metrics were evaluated on

B-mode representations between -62 dB and $+36$ dB (confidence interval discussed in Section 5.4.4), and were averaged over the entire validation set. Even though not optimal, the B-mode SSIM correlated particularly well with visual assessments for evaluating the overall quality of recovered images, and was used to select the best performing CNN instance among the 500 validation steps of each training experiment. For comparison purposes, a fixed random seed was used for initializing kernel weights, identical training set shufflings were performed, and the same validation set was used for each training experiment.

The implementation was carried out using TensorFlow⁴ (v1.14), and the trainings were performed on NVIDIA Tesla V100 GPUs.

5.5.2. Hyperparameter Search

Hundreds of training experiments were carried out heuristically to select the hyperparameters involved in the proposed approach. The selected training experiments presented in this section are the ones that guided the selection of the trained CNNs evaluated in Section 5.6. Each experiment was conducted using the global setup parameters as well as the training and validation strategy described in Section 5.5.1.

5.5.2.1. Image Representations

Ultrasound images can be expressed, analyzed, and displayed in different representations, namely RF, IQ, envelope, and B-mode (log-compressed envelope). We thus compared the impact of training on these different image representations using the proposed residual CNN (Figure 5.2) deployed with 16 initial expansion channels, RCBs, and additive intrinsic skip connections. All instances were trained using the MSLAE as loss function and UQ images as references, except when trained on B-mode representations in which case the MAE was used, as this image representation is already log-compressed.

Even though it may seem intuitive to train on image representations that we actually look at (i.e., B-mode), it is clear (Figure 5.6) that trainings performed on both envelope and B-mode representations are worse than those performed on RF and IQ ones. This presumably comes from the fact that both envelope and B-mode representations do not contain the RF property of US images anymore (due to the envelope detection process), a property carrying additional information that can be exploited by the learning process. The images from CNNs trained on B-mode and envelope representations are characterized by blurred speckle patterns [Figures 5.7(e) and 5.7(f)].

Trainings performed on RF and IQ representations resulted in similar performances. This was expected as RF and IQ images contain the same information. Hence, both are valid choices. Yet, training (and inference) on IQ representations is more demanding than on RF ones as IQ images are composed of “two channels” (i.e., real and imaginary parts), but this only affects the first and last CLs. (i.e., initial channel expansion and final

⁴<https://www.tensorflow.org>

channel contraction, Figure 5.2). On the other hand, the use of IQ images simplifies the following envelope detection step compared with RF ones, namely a simple element-wise modulus compared with a Hilbert transform (followed by an element-wise modulus). This is the reason why IQ was preferred.

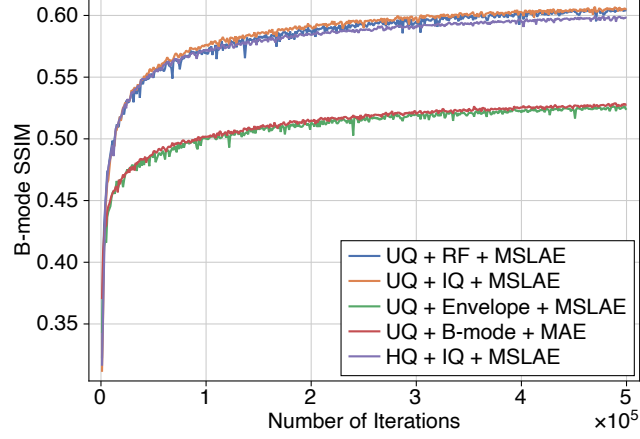


Figure 5.6. Validation metric curves (SSIM evaluated on B-mode representations) of training experiments performed using different image representations (i.e., RF, IQ, envelope, and B-mode) and different reference images (i.e., HQ and UQ). All training experiments were performed on identical instances of the proposed residual CNN (Figure 5.2) with 16 initial expansion channels, RCBs, and additive intrinsic skip connections.

5.5.2.2. Reference Image Configurations

In our preliminary work [137], we observed that training on reference images in which GL artifacts were still present (i.e., HQ) resulted in predicted images with a surprising reduction of said artifacts. This observation inspired us to develop (and simulate) reference images free from these artifacts (i.e., UQ). In this experiment, we evaluated the effect of using UQ images, obtained from the optimal (and virtual) UQ imaging configuration, as reference images during training, compared to using HQ images, obtained from the HQ imaging configuration (Section 5.4.1). In both cases, UQ images were used as references for computing validation metrics. As for Section 5.5.2.1, we used a CNN with 16 initial expansion channels, RCBs, and additive intrinsic skip connections. Each instance was trained on IQ representations using the MSLAE as loss function.

Figure 5.6 demonstrates the benefit of training on UQ rather than HQ reference images in terms of B-mode SSIM. As expected, we observed that imaging artifacts, in particular those caused by GLs, were better reduced when trainings were performed using UQ images as references [Figure 5.7(h)], than when using HQ ones [Figure 5.7(g)]. Details initially hidden by GL artifacts were also better recovered. Yet, and as observed in our preliminary work [137], the training on HQ images as references resulted in images with far less GL artifacts compared with the corresponding HQ images [Figures 5.7(c) and 5.7(g)]. This effect remains unexplained and could therefore be unpredictable. Nonetheless, training CNNs on the newly designed UQ reference images resulted in a more consistent GL

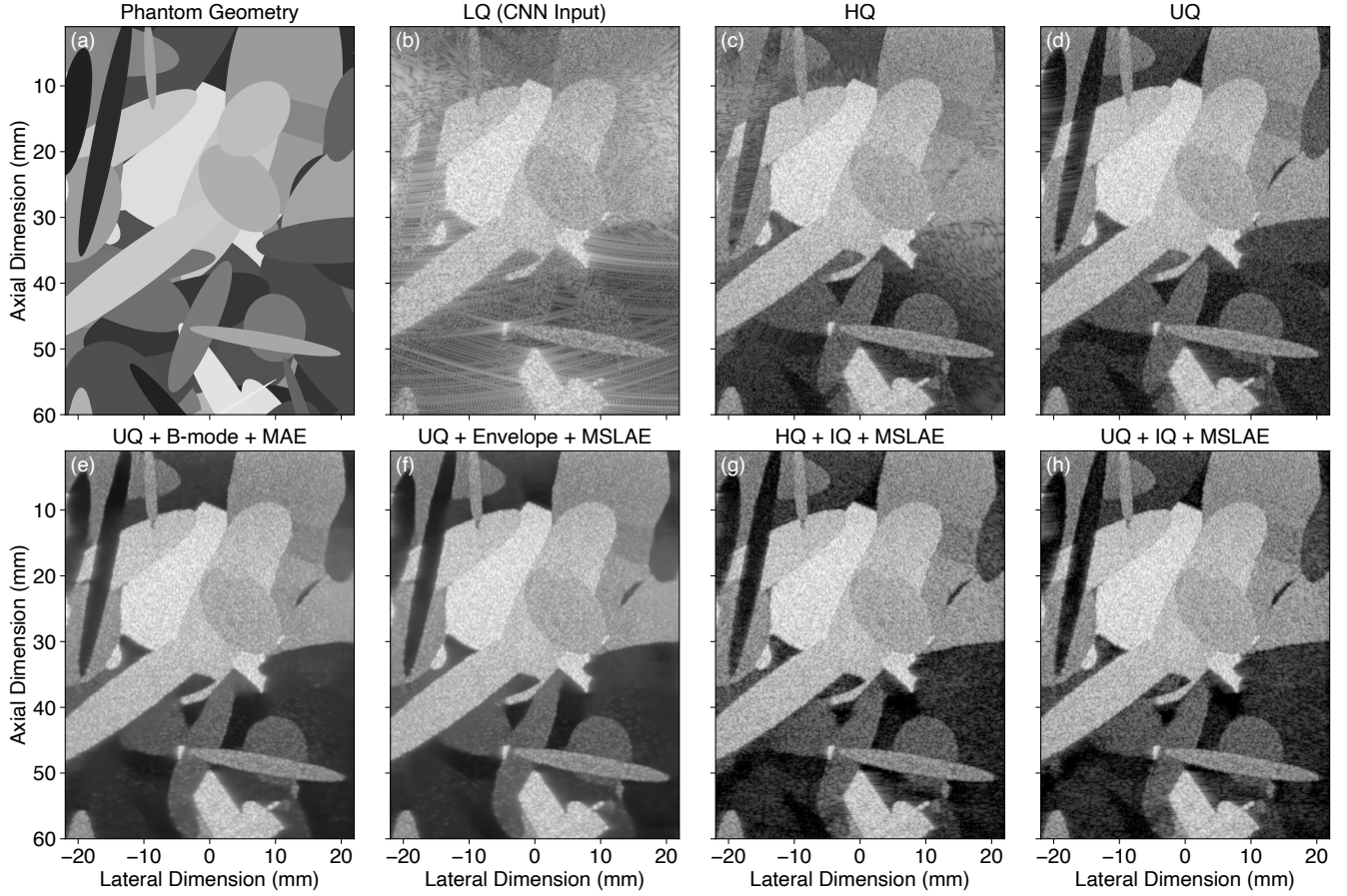


Figure 5.7. B-mode image representations (98-dB range) of a numerical test phantom sample (extracted from the simulated dataset): (a) the phantom mask; images reconstructed using each imaging configuration considered (Table 5.1), namely (b) low-quality (LQ), (c) high-quality (HQ) (i.e., gold-standard image for the physical transducer array), and (d) ultra-high-quality (UQ) (i.e., reference image); images recovered from the LQ input image using the proposed approach with different convolutional neural networks (CNNs), deployed with 16 initial expansion channels, residual convolutional blocks (RCBs), and additive intrinsic skip connections, trained on different image representations and image references, namely (e) UQ + B-mode + mean absolute error (MAE), (f) UQ + envelope + mean signed logarithmic absolute error (MSLAE), (g) HQ + in-phase quadrature (IQ) + MSLAE, and (h) UQ + IQ + MSLAE.

reduction as UQ reference images did not contain such artifacts. Therefore, we opted for using UQ images as reference images during training.

5.5.2.3. Training Losses

We compared the effect of using different training losses, namely the MSE, the MAE, and the proposed MSLAE (implemented with a “threshold” parameter α corresponding to -62 dB). For this comparison, we used a CNN with 16 initial expansion channels, RCBs, and additive intrinsic skip connections. Trainings were performed on IQ representations using UQ images as references.

Figure 5.8 clearly shows that, despite being the standard loss in regression problems, and the loss we used in our preliminary work [137], the MSE is the least effective one to

address the restoration problem involved in the proposed approach. Indeed, the HDR property of RF US images makes the use of the MSE suboptimal, as too much emphasis is put on image samples with large values (i.e., highly echogenic). The use of the MAE as loss function, which has been increasingly reported in similar regression problems such as image super-resolution and MRI [150]–[152], performed better than using the MSE. As expected, the fact that MAE is less sensitive to “outliers” makes it more robust to HDR contents. A substantial increase in performance with respect to MAE was observed when using the proposed MSLAE as loss function, confirming its superiority over the other two losses compared for learning a restoration mapping on HDR RF (or IQ) data.

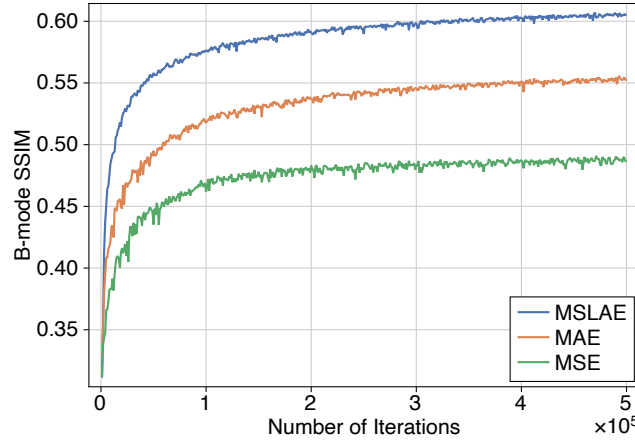


Figure 5.8. Validation metric curves (SSIM evaluated on B-mode representations) of training experiments performed using different training losses, namely mean squared error (MSE), mean absolute error (MAE), and mean signed logarithmic absolute error (MSLAE). All training experiments were performed on identical instances of the proposed residual CNN (Figure 5.2) with 16 initial expansion channels, RCBs, and additive intrinsic skip connections.

One can also note that, the more effective the loss, the more stable the training, appearing as smoother validation curves. It was also observed that the use of a more effective loss resulted in trainings less prone to overfitting, thus less demanding in terms of data quantity, as it maximized the usage of the available information content. The flattening of the validation curve observed when using the MSE as loss function is an indication that overfitting would most probably appear earlier than when using the other losses. Further analysis and discussions on losses can be found in Sections 5.6 and 5.7.

5.5.2.4. Convolutional Blocks and Skip Connections

These experiments were conducted to evaluate the effects of the proposed CNN architecture improvements (Section 5.3.2). All trainings were performed on IQ representations using MSLAE as loss function and UQ images as references. All experiments were carried out on CNN instances with 16 initial expansion channels. Two types of intrinsic skip connections, namely additive and concatenated as originally proposed in [145], were compared on CNN instances with traditional FCBs. We also compared the use of the proposed RCBs instead of FCBs on CNN instances with additive intrinsic skip connections,

as concatenated ones cannot be used with RCBs directly.

The comparison of concatenated and additive intrinsic skip connections implemented with the jointly compatible FCBs shows (Figure 5.9) that the use of concatenated ones results in slightly better performances than additive ones. This was somehow expected as the use of concatenated intrinsic skip connections increases the total number of trainable parameters (i.e., increased capacity) by approximately 7 % in the “decoding” arm only [Figure 5.2(a)]. As a result it also significantly increases both training and inference times, due to augmented convolution operations which are the most costly ones. (Especially the last intrinsic skip connection which results in the most computationally intense convolutional operation of the CNN architecture.) The comparison of RCBs and FCBs implemented with additive intrinsic skip connections showed that the use of RCBs performs best at same CNN capacity with (almost) no effect on the inference time. It also outperformed the greater-capacity CNN instance with FCBs and concatenated intrinsic skip connections both in terms of validation metric (B-mode SSIM) and training stability (smoother validation curve). The use of RCBs together with additive intrinsic skip connections was therefore selected.

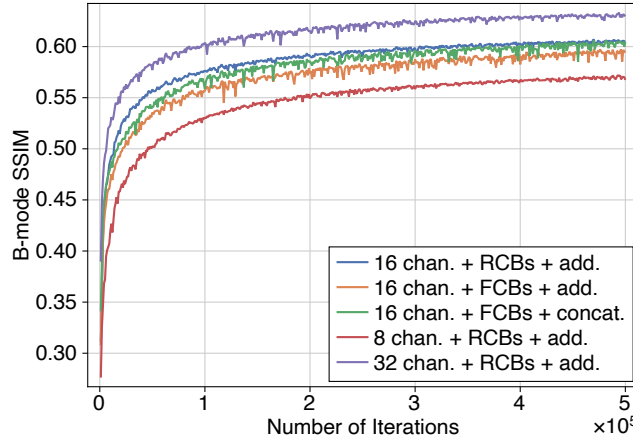


Figure 5.9. Validation metric curves (SSIM evaluated on B-mode representations) of training experiments performed using different combinations of initial channel expansion numbers (i.e., 8, 16, and 32), convolutional blocks (i.e., RCBs and FCBs), and intrinsic skip connections (i.e., additive and concatenated). Each training experiment was performed using the MSLAE as loss function.

5.5.2.5. Initial Channel Expansion Numbers

Since the initial channel expansion number affects the entire architecture, this parameter has a major impact on the overall CNN capacity, the training time, and the inference time. Three CNN instances (RCBs, additive intrinsic skip connections) with initial channel expansion numbers of 8, 16, and 32 were trained on IQ representations using MSLAE as loss function and UQ images as references. In these settings, the total number of trainable parameters were 687 704, 2 748 592, and 10 989 920, respectively (i.e., approximately quadrupled when the initial channel expansion number is doubled).

As US imaging is, in essence, a real-time imaging modality, inference speed tests were

also performed on these three configurations. To quantify the impact on the inference time of using IQ images rather than RF images (Section 5.5.2.2), inference speed tests were also performed on the same configurations but trained on RF images. We computed the averaged inference time, over 5000 runs, on images of size 596×1600 , with appropriate zero-padding, for each configuration using both TensorFlow⁵ (v1.14) and TensorRT⁶ (v5.1.5), an inference optimizer. Different GPU models were compared, namely the NVIDIA GeForce MX 150 (laptop, 384 cores, Pascal arch., entry-level), the NVIDIA GeForce GTX 1080 Ti (desktop, 3584 cores, Pascal arch.), and the NVIDIA TITAN V (desktop, 5120 cores, Volta arch.).

As expected, the more initial expansion channels the better the validation metric (Figure 5.9), provided that enough data is available to avoid overfitting. Inference speed tests (Table 5.2) demonstrated that, depending on code optimization and GPU model, real-time imaging is feasible using the proposed approach and a 16-channel version. Since we are using simulations and can theoretically generate a dataset of infinite size preventing from any overfitting, the architecture optimization really comes down to speed vs. quality in scenarios where real-time imaging is a necessity. One can also note that the increase in inference time of using IQ images rather than RF images was of about 5 % to 10 %, and did not result in loosing real-time capabilities.

TABLE 5.2
Average Inference Time for Different Image Representations
and Initial Channel Expansion Numbers

Image Repr.	Channel Number	MX150		1080 Ti		TITAN V	
		TF ^a	TRT ^b	TF	TRT	TF	TRT
RF	8	130 ms	83 ms	21 ms	10 ms	18 ms	8 ms
	16	249 ms	167 ms	36 ms	21 ms	29 ms	14 ms
	32	× ^c	× ^c	74 ms	52 ms	52 ms	37 ms
IQ	8	136 ms	86 ms	24 ms	12 ms	21 ms	9 ms
	16	256 ms	172 ms	39 ms	22 ms	32 ms	15 ms
	32	× ^c	× ^c	77 ms	53 ms	54 ms	39 ms

^aTensorFlow ^bTensorRT ^cNot enough memory

5.5.2.6. Training Set Sizes

This experiment was performed to evaluate the impact of the training set size, and most importantly, to guarantee that the selected configuration is not prone to overfitting. We considered the proposed residual CNN (Figure 5.2) deployed with 16 initial expansion channels, RCBs, and additive intrinsic skip connections. Each instance was trained on IQ representations using MSLAE as loss function and UQ images as references. Different training set sizes (spanning a logarithmic range) were compared, namely 200, 409, 837,

⁵<https://www.tensorflow.org>

⁶<https://developer.nvidia.com/tensorrt>

1713, 3504, 7168, 14 664, and 30 000.

From Figure 5.10, it is evident that the training of the analyzed, comparatively small CNN with only 16 initial expansion channels, suffers from obvious overfitting up to $\sim 7k$ training image pairs. In these settings, it seems like the use of a training set composed of $\sim 10k$ image pairs would be sufficient to avoid overfitting. Yet, the training of CNNs with more capacity, such as with 32 initial expansion channels, necessarily requires larger training sets. Thus, we chose to use 30 000 image pairs for the reported experiments. One can also note (magnified inset of Figure 5.10) that the larger the training set, the better the resulting validation loss, even after “obvious” overfitting cases.

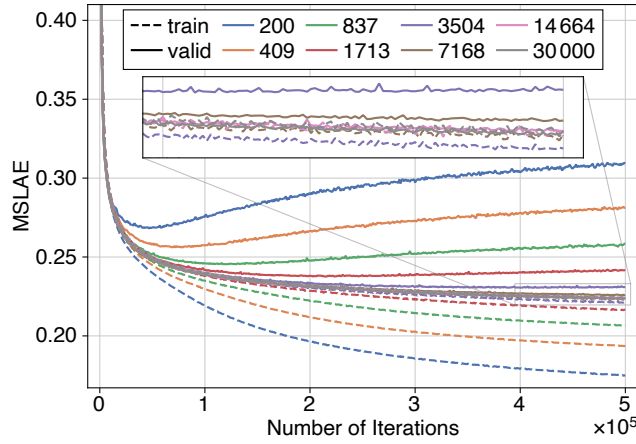


Figure 5.10. Training and validation loss curves of training experiments performed using different training set sizes. All training experiments were performed on identical instances of the proposed residual CNN (Figure 5.2) with 16 initial expansion channels, RCBS, and additive intrinsic skip connections, using the MSLAE as loss function.

5.5.2.7. Kernel Initializers

We confirmed our choice of using Glorot uniform as kernel initializer by comparing the performances of differently initialized CNNs instances with 16 initial expansion channels, RCBS, and additive intrinsic skip connections. As the proposed architecture is composed of CLs and ReLU activations, we were particularly interested in evaluating the He initializer proposed in [153] to maintain the variance through such layers and activations. We compared both Glorot [147] and He initializers implemented with uniform and normal distributions. All kernels were initialized by the initializers considered, except for the initial channel expansion layer and the final channel contraction layer which were always initialized using the Glorot (uniform) initializer, as they are not followed by a ReLU activation.

Interestingly, both implementations of the Glorot initializer (i.e., uniform and normal) performed similarly better than both implementations of the He initializer (Figure 5.11). This may be explained by the many residual connections (i.e., all intrinsic ones and the outer one) and/or the multi-scale property of the proposed architecture, for which the benefit of He initializer does not seem to be confirmed.

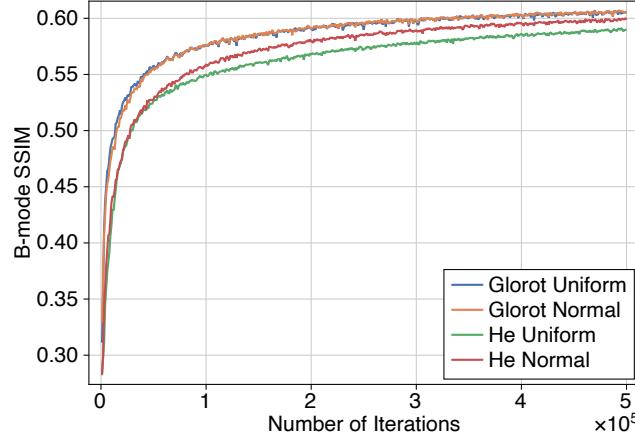


Figure 5.11. Validation metric curves (SSIM evaluated on B-mode representations) of training experiments performed using different kernel initializers (i.e., Glorot uniform, Glorot normal, He uniform, and He normal). Each training experiment was performed on the proposed residual CNN (Figure 5.2) with 16 initial expansion channels, RCBs, and additive intrinsic skip connections, using the MSLAE as loss function.

5.5.2.8. Learning Rates

We also compared different learning rates of 1×10^{-5} , 5×10^{-5} , 1×10^{-4} , 5×10^{-4} , and 1×10^{-3} . Identical instances of the proposed residual CNN (Figure 5.2) deployed with 16 initial expansion channels, RCBs, and additive intrinsic skip connections were trained using the Adam optimizer [148] with each learning rate. Trainings were performed using the MSLAE as loss function and UQ images as references.

From Figure 5.12, it is clear that a learning rate of 1×10^{-5} is too small and that a learning rate of 1×10^{-3} is too large. The other three, namely 5×10^{-5} , 1×10^{-4} , and 5×10^{-4} , resulted in fairly similar performances. Even though a learning rate of 5×10^{-5} was the least performing among these three, we decided to select this one as it resulted in the most stable validation curve and adapted best to all other experiments carried out for hyperparameter search (in particular when training CNNs with larger capacities).

5.5.3. Summary

All proposed improvements to the neural network architecture (Section 5.3.2) resulted in increased performances. The use of optimal (virtual) UQ images as reference images for the training process was successful. It provided better results than using HQ images, with controlled GL artifacts removal (Section 5.5.2.2). The image representation onto which the training is performed is crucial. Trainings performed on B-mode and envelope representations resulted in much worse image restoration capabilities than RF and IQ ones, as the envelope detection process removes high-frequency content that can be exploited by the CNN. Even though both IQ and RF trainings performed similarly, as the information contained in both representations is identical, we opted for the IQ ones as it allows for a simplified subsequent envelope detection process (Section 5.5.2.1). The loss choice was observed as the most impactful parameter (Section 5.5.2.3) and was therefore

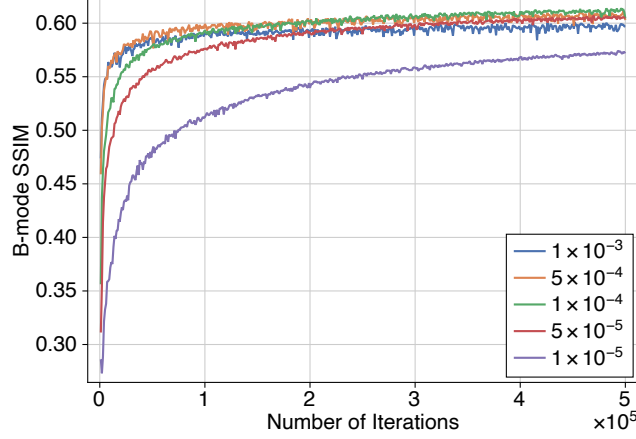


Figure 5.12. Validation metric curves (SSIM evaluated on B-mode representations) of training experiments performed using different learning rates (i.e., 1×10^{-5} , 5×10^{-5} , 1×10^{-4} , 5×10^{-4} , and 1×10^{-3}). Each training experiment was performed on the proposed residual CNN (Figure 5.2) with 16 initial expansion channels, RCBs, and additive intrinsic skip connections, using the MSLAE as loss function.

further evaluated in an US-specific test environment (Section 5.6.1).

5.6. Experiments and Results

To evaluate thoroughly the performance of the proposed CNN-based image reconstruction method in realistic settings, we conducted two types of experiments. A dedicated numerical test phantom (Section 5.6.1) with US-specific metrics was designed and deployed for in-depth evaluation of the proposed method in ideal conditions, namely, on images obtained using the same simulation settings as the ones used to generate the training dataset. The robustness to experimental acquisitions was evaluated in both *in vitro* and *in vivo* settings (Section 5.6.2).

5.6.1. Numerical Test Phantom

5.6.1.1. Setup

Standard image-processing metrics such as PSNR and SSIM may be used to compare different experiments globally (as used for the hyperparameter search, Section 5.5.2), but they remain suboptimal for assessing US image quality. Traditionally, US image quality is evaluated using metrics that (ideally) reflect lesion detectability while guaranteeing that images accurately represent the echogenicity of the underlying tissue. The most commonly used metrics are the contrast, the CNR, and the imaging system resolution. Yet, as demonstrated in [80], both contrast and CNR measures, which are often estimated between image regions with and without signal, can be affected by dynamic range alterations (DRAs), common to most image-dependent adaptive beamformers (e.g., minimum variance or coherence-based) or resulting from simple non-linear transformations of im-

age amplitudes (e.g., low-amplitude thresholding). These DRAs may result in improved contrast or CNR measures without actually improving image quality or lesion detectability, and can even potentially conceal information relevant to clinical diagnosis, in the form of distorted image amplitudes. Inspired by [80], we designed a “challenging” numerical test phantom for extensive evaluations of images reconstructed from measurements acquired with the 9L-D transducer. It is composed of tissue-mimicking echogenic zones embedded in an anechoic background. The test phantom zones are depicted in Figure 5.13(a) and are described in detail in the following paragraphs, in conjunction with their associated metrics.

A block with a square section of $20 \text{ mm} \times 20 \text{ mm}$ is centered at $(-5 \text{ mm}, 20 \text{ mm})$. A low-echogenic cylindrical inclusion with a diameter of 8.5 mm is embedded at its center. The contrast between the two was set to -36 dB . This value was selected such that the diffraction artifacts arising from the high-echogenic block and covering the low-echogenic inclusion are significantly higher ($\sim 8 \text{ dB}$) than the inclusion level, in images obtained using the LQ configuration. The restoration quality of the low-echogenic inclusion was assessed by computing the obtained contrast, expressed in decibels as

$$C = 20 \log_{10} \left(\frac{\mathbb{E}[\mathbf{s}_I]}{\mathbb{E}[\mathbf{s}_B]} \right), \quad (5.31)$$

where \mathbf{s}_B and \mathbf{s}_I are the envelope-detected image amplitude values in Ω_B and Ω_I [Figure 5.13(a)], respectively, and $\mathbb{E}[\cdot]$ is the expected value, evaluated as the sample mean. In the case of fully developed speckle zones following a Rayleigh distribution, the contrast can be equivalently expressed (in decibels) on the image intensity (Section 5.2.2.3). Note that, by considering an inclusion with a prescribed contrast, reconstruction errors such as DRAs would most likely result in erroneous contrast estimates.

Another block with a rectangular section of $43.93 \text{ mm} \times 10 \text{ mm}$ (i.e., spanning the probe aperture) is positioned at a depth of 50 mm , characterized by a lateral log-linear echogenicity ranging from $+30 \text{ dB}$ to -50 dB (i.e., spanning a range of 80 dB with a gradient of $\sim 1.82 \text{ dB/mm}$). The capacity of the proposed method to preserve prescribed linearity (while removing artifacts) was assessed by averaging the obtained image amplitudes within Ω_{LG} along the depth axis, and the accuracy was visually assessed by comparing it with the prescribed one. Again, any potential DRAs would result in (highly) distorted amplitude gradients (as demonstrated in [80]).

Four ideal bright reflectors (p_0 , p_1 , p_2 , and p_3) are arranged at a lateral position of 12.5 mm and at different depths, namely at 10 mm , 20 mm , 30 mm , and 40 mm . Both axial and lateral FWHM measures were evaluated on the image amplitude using a 2-D spline-based interpolation and a sub-pixel peak finder within $2\lambda \times 2\lambda$ regions centered at the position of each bright reflector [Figure 5.13(a)].

The level of diffraction artifacts was quantified by averaging the image amplitudes within different anechoic rectangular regions. These regions were selected on LQ images [Figure 5.13(b)] to be dominated by significant diffraction artifact levels primarily caused by GLs (Ω_{GL}), SLs (Ω_{SL}), and EWs (Ω_{EW}).

The speckle was assessed from within a square region (Ω_S) of size $10\lambda \times 10\lambda$, centered

at (0 mm, 27 mm) [Figure 5.13(a)]. First order statistics was assessed by computing the ratio between the mean and the standard deviation of image amplitudes, often referred to as SNR, expressed as,

$$\text{SNR} = \frac{\mathbb{E}[\mathbf{s}_S]}{\sqrt{\text{Var}(\mathbf{s}_S)}}. \quad (5.32)$$

In the case of samples following a Rayleigh distribution, this ratio would be equal to 1.91 [142] (Section 5.2.2.1). Second order statistics was evaluated by computing the FWHM of the 2-D ACF [78], [79] (Section 5.2.2.2). This metric represents a statistical measure of the “speckle resolution,” in both axial and lateral dimensions, which is of great importance to many post-processings (e.g., displacement tracking algorithms).

Global PSNR and SSIM metrics were also computed on B-mode images between -62 dB and $+36$ dB against UQ reference images.

Three hundred statistically independent realizations (i.e., random scatterers) were simulated for each imaging configuration, identically to the simulated dataset (Sections 5.4.1 to 5.4.3). Images were reconstructed identically, using the same normalization factors (Section 5.4.4). An additional normalization factor was evaluated on the average of all UQ test images such that the reconstructed gradient would fit (on average) the prescribed one. This factor was applied to all images of each imaging configurations (i.e., also before inference). No renormalization was applied after inference. All metrics were evaluated separately for each of the 300 realizations. Both the mean value and standard deviation were computed for each metric through all independent realizations.

From the extensive hyperparameter search carried out (Section 5.5.2), four trained CNNs were selected for further evaluations using the numerical test phantom detailed above. To evaluate the effect of the training loss function, we considered three instances of the proposed residual CNN, deployed with RCBs, additive intrinsic skip connections, and 16 initial expansion channels (Figure 5.2), trained using MSE (MSE-16), MAE (MAE-16), and MSLAE (MSLAE-16) as loss functions. A 32-channel instance, trained using MSLAE as loss function (MSLAE-32), was also selected to evaluate the effect of increasing network capacity. The proposed MSLAE, defined in (5.23), was implemented with a “threshold” parameter α corresponding to -62 dB. The reasoning for such a value comes from the confidence interval on the dataset values discussed in Section 5.4.4.

5.6.1.2. Results

Visual assessment of the test phantom images (Figure 5.13) and the metrics obtained (Table 5.3) confirm that the proposed image reconstruction method significantly improves the image quality compared with LQ using any of the trained CNNs. A global comparison of CNNs with identical capacities (i.e., MSE-16, MAE-16, and MSLAE-16) demonstrates the superiority of the proposed HDR-sensitive MSLAE loss. Both MSLAE trainings achieved higher global quality metrics (i.e., PSNR and SSIM) than HQ, mainly due to the impressive reduction of GL artifacts. The added capacity of MSLAE-32 resulted in improved overall performances compared with MSLAE-16.

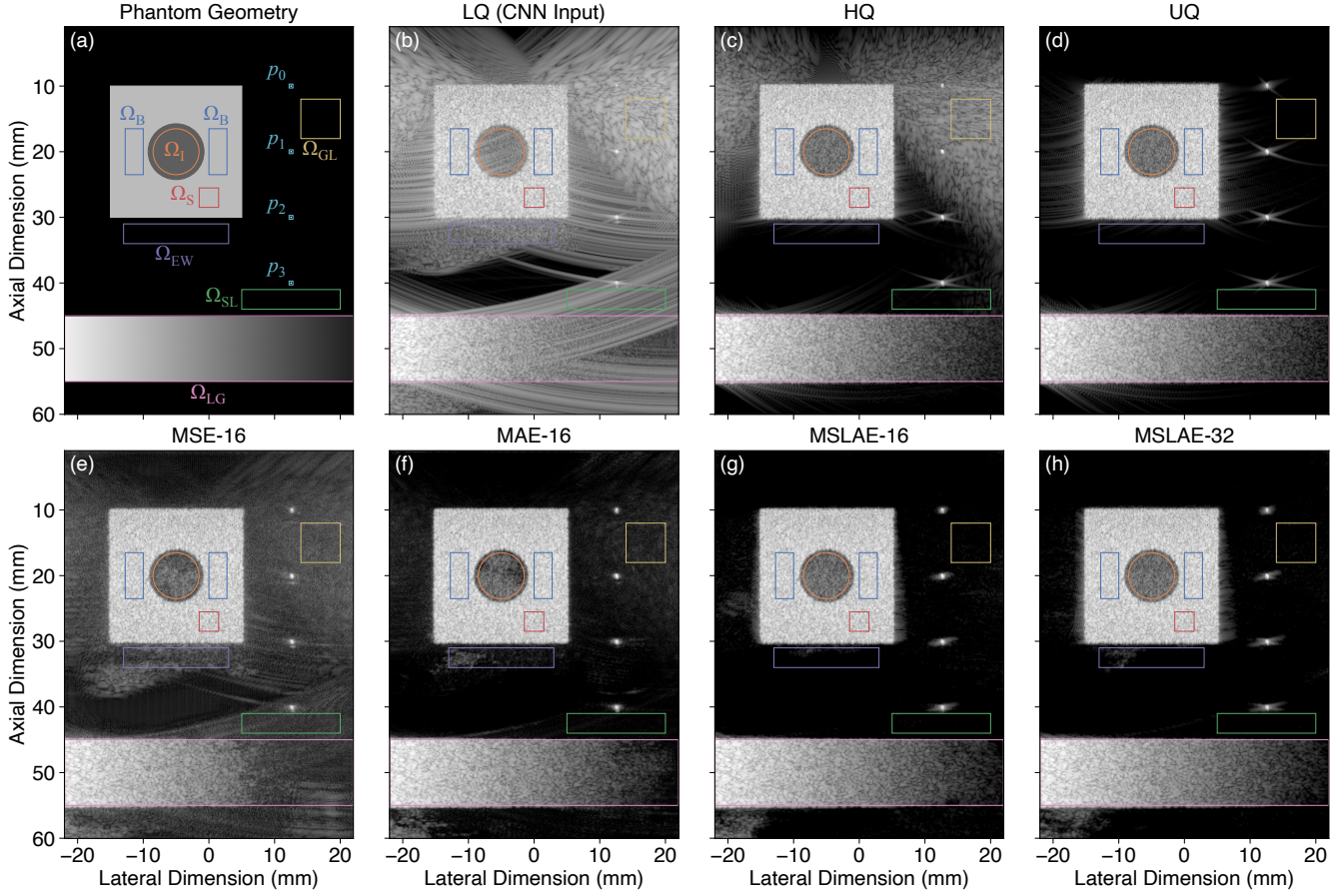


Figure 5.13. B-mode image representations (98-dB range) of a numerical test phantom sample: (a) the phantom mask and annotated zones in which the metrics were evaluated; images reconstructed using each imaging configuration considered (Table 5.1), namely (b) low-quality (LQ) configuration, (c) high-quality (HQ) configuration (i.e., gold-standard image for the physical transducer array), and (d) ultra-high-quality (UQ) configuration (i.e., reference image); images recovered from the low-quality (LQ) input image using the proposed approach with each of the trained convolutional neural networks (CNNs) considered (Section 5.6.1.1), namely (e) MSE-16, (f) MAE-16, (g) MSLAE-16, and (h) MSLAE-32.

The restoration of the low-echogenic inclusion (prescribed contrast of -36 dB) and the resulting contrast obtained were improved drastically compared with LQ. Both MSE-16 and MAE-16 suffer from important “dark region artifacts” [154] in the low-echogenic inclusion [Ω_i in Figures 5.13(f) and 5.13(g)], whereas MSLAE-16 and MSLAE-32 are more robust to them, resulting in a more accurate restoration of the inclusion. This is confirmed by the contrast obtained which are tending to the reference one (i.e., UQ) for MSLAE-16 and MSLAE-32 (Table 5.3).

All trained CNNs resulted in diffraction artifact levels drastically reduced compared with LQ ones. Remaining GL artifacts were far below HQ ones (>18 dB). Artifacts caused by EWs appeared to be the most complex artifact to deal with [e.g., lower-left corner in Figures 5.13(e) to 5.13(h)]. It can also be observed that the restoration of the SL artifacts present in the UQ reference image was handled more accurately with MSLAE-32.

The SNR obtained within the speckle zone Ω_s for DAS-based methods (i.e., LQ, HQ, and UQ) did not reach the theoretical value of 1.91 for fully developed speckle. This was

TABLE 5.3
Numerical Test Phantom Metrics

Metric ^a	LQ	HQ	UQ	MSE-16	MAE-16	MSLAE-16	MSLAE-32
C (dB) ^b	−28.33 (0.60)	−36.06 (0.32)	−36.06 (0.32)	−39.41 (0.83)	−39.49 (1.03)	−37.74 (0.70)	−37.40 (0.51)
GL (dB)	+6.39 (0.45)	−10.27 (0.52)	−66.62 (0.46)	−27.99 (0.32)	−45.34 (0.35)	−61.56 (0.36)	−62.24 (0.32)
SL (dB)	−14.48 (1.08)	−58.16 (0.55)	−68.64 (0.38)	−38.89 (0.54)	−60.49 (0.53)	−67.51 (0.52)	−67.91 (0.64)
EW (dB)	−13.64 (0.30)	−59.61 (0.67)	−59.96 (0.70)	−31.23 (0.78)	−46.74 (3.04)	−60.74 (4.05)	−55.17 (6.55)
SNR	1.84 (0.09)	1.80 (0.09)	1.80 (0.09)	1.80 (0.09)	1.80 (0.09)	1.79 (0.09)	1.81 (0.09)
ACF lat. (μm)	262.1 (19.5)	219.4 (15.2)	219.6 (15.3)	245.6 (17.5)	246.2 (17.6)	251.1 (18.4)	246.0 (17.4)
ACF ax. (μm)	293.5 (21.5)	302.6 (21.4)	302.7 (21.4)	301.5 (21.6)	302.9 (21.9)	301.9 (21.6)	301.8 (21.6)
FWHM lat.	p_0 (μm)	276.6 (24.4)	202.1 (1.8)	202.2 (0.0)	226.5 (8.5)	211.5 (6.9)	232.5 (7.9)
	p_1 (μm)	336.2 (5.6)	242.5 (0.9)	242.7 (0.0)	255.2 (3.2)	243.1 (2.6)	270.1 (6.3)
	p_2 (μm)	388.6 (1.5)	280.0 (0.0)	280.5 (0.0)	286.3 (2.0)	293.7 (1.6)	301.1 (1.6)
	p_3 (μm)	446.6 (4.9)	321.9 (0.0)	322.4 (0.0)	345.9 (6.1)	340.5 (4.9)	359.4 (2.2)
FWHM ax.	p_0 (μm)	264.8 (8.3)	266.5 (0.8)	266.6 (0.0)	265.0 (2.9)	256.3 (2.3)	286.2 (5.7)
	p_1 (μm)	316.7 (2.5)	314.6 (0.3)	314.6 (0.0)	313.0 (2.4)	310.8 (1.6)	308.5 (3.3)
	p_2 (μm)	317.5 (0.9)	318.3 (0.0)	318.3 (0.0)	314.7 (0.6)	311.7 (0.5)	303.5 (1.3)
	p_3 (μm)	320.7 (1.6)	324.0 (0.0)	324.0 (0.0)	324.8 (2.2)	328.4 (1.5)	319.3 (1.6)
PSNR (dB)	8.60 (0.04)	14.23 (0.04)	× ^c	14.71 (0.05)	21.96 (0.12)	24.18 (0.25)	25.14 (0.32)
SSIM	0.31 (0.00)	0.73 (0.00)	× ^c	0.39 (0.00)	0.58 (0.00)	0.75 (0.00)	0.78 (0.00)

^a Metrics were averaged over 300 independent realizations. The standard deviation is given in parentheses.

^b Prescribed contrast of −36 dB.

^c PSNR and SSIM metrics were computed against UQ.

expected since ten scatterers per resolution cell were used for numerical simulations (i.e., lower bound to obtain fully developed speckle). All trained CNNs improved the SNR compared with LQ (i.e., closer to the UQ one). The lateral resolution of speckle patterns (Table 5.3, ACF lat.) was only slightly improved, without reaching the one of HQ and UQ. On the other hand, the lateral resolution evaluated on bright reflectors was improved significantly.

The restoration results of log-linear gradient are depicted in Figure 5.14. Almost perfect restoration was achieved from +30 dB to −30 dB, with a slight but increasing deviation for lower echogenicity values, by all trained CNNs except for MSE-16, which only preserved linearity from +30 dB to −15 dB. Note that the LQ response is slightly overestimated (offset). This effect results from the ideal PW assumption used to derive the backprojection (DAS) operator (Section 5.4.1). It is easily resolved by all trained CNNs.

Since we generated a simulated test set obtained from 300 statistically independent realizations (i.e., random scatterers) of the same numerical test phantom (Section 5.6.1.1), we also analyzed the incoherent average (performed after envelope detection). As independent realizations of scatterers with identical statistical properties result in similar images with uncorrelated speckle patterns, the incoherent averaging of a large amount of such images provides us with an interesting visualization of stationary structures; the underlying phantom mask and the image zones suffering from imaging artifacts are fully revealed. The visual assessment of such a representation (Figure 5.15) leads to the

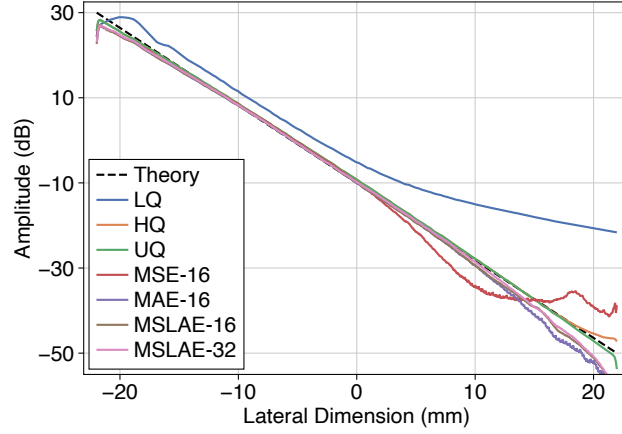


Figure 5.14. Mean amplitude responses (averaged along the axial dimension) of the horizontal gradient zone in the numerical phantom [Ω_{LG} in Figure 5.13(a)], averaged over 300 independent realizations.

same conclusions, some of which deserve to be re-emphasized. The comparison of the averaged restoration of the low-echogenic inclusion is of particular interest and shows again the benefit of using the proposed MSLAE as loss function over the conventional MSE and MAE losses [Figures 5.15(e) to 5.15(g)]. By comparing the results obtained with MSLAE-16 and MSLAE-32, one can note that the greater the CNN capacity, the closer the recovered image to the corresponding UQ reference. This increase in performance is especially visible on the remaining SL artifacts, which more closely resemble those of the UQ reference. This visualization makes it very clear that EW artifacts are the most complex to deal with. It also reveals a remaining EW artifact arising from the deepest bright reflector and located within the log-linear gradient that was indistinguishable in the test phantom sample displayed in Figure 5.13.

5.6.2. Experimental Evaluations

5.6.2.1. Setup

Experimental data were acquired using a 9L-D transducer on a Vantage 256 system using the imaging configurations defined in Section 5.4.1 (Table 5.1), except for the UQ case (only implementable in a simulation environment). Compounded acquisitions were performed at maximum pulse repetition frequency (PRF) (~ 9.5 kHz) to minimize the effect of potential inter-acquisition motion. The single-PW insonification (LQ) was performed first in the ultrafast sequence, directly followed by 192 SA acquisitions (HQ), performed in an alternated manner from central to outer elements. A peak-to-peak voltage of 50 V was used for the transmit excitation. Time gain compensation (TGC) was implemented to compensate for a mean tissue attenuation of -0.5 dB/(cm \cdot MHz).

In vitro frames were acquired on the CIRS model 054GS general purpose ultrasound phantom (CIRS, Norfolk, VA, USA). The transducer was clamped on a stand during acquisitions and its face was immersed in water for acoustic coupling. A normalization

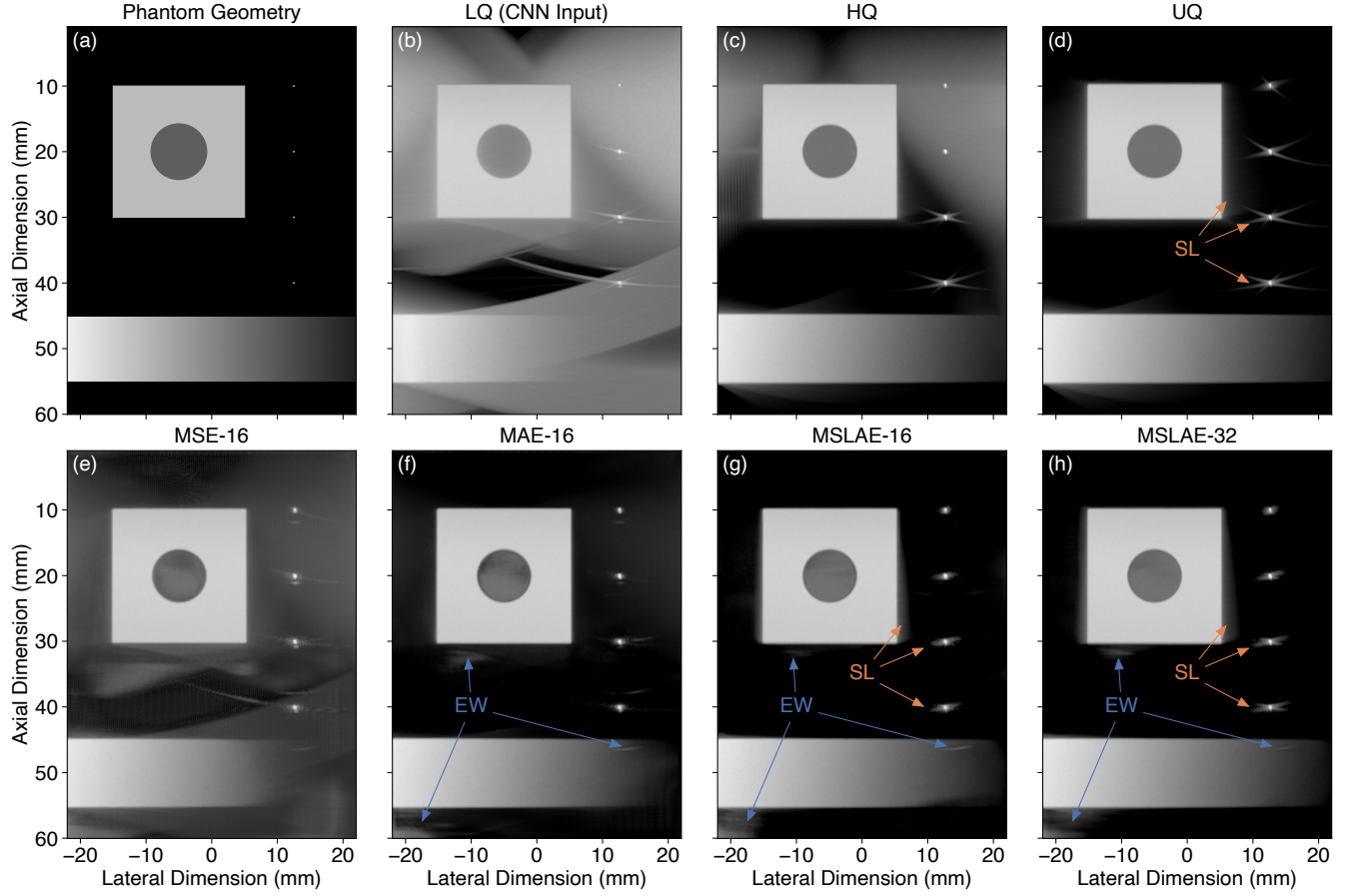


Figure 5.15. B-mode image representations (98-dB range) of the incoherent average (performed after envelope detection) of all images reconstructed from the 300 independent realizations (random scatterers) of the numerical test phantom: (a) the phantom mask; images reconstructed using each imaging configuration considered (Table 5.1), namely (b) low-quality (LQ) configuration, (c) high-quality (HQ) configuration (i.e., gold-standard image for the physical transducer array), and (d) ultra-high-quality (UQ) configuration (i.e., reference image); images recovered from the low-quality (LQ) input image using the proposed approach with each of the trained convolutional neural networks (CNNs) considered (Section 5.6.1.1), namely (e) MSE-16, (f) MAE-16, (g) MSLAE-16, and (h) MSLAE-32. Some remaining side lobe (SL) and edge wave (EW) artifacts are highlighted by colored arrows and associated annotations.

factor was determined, in the same manner as described in Section 5.4.4, for both LQ and HQ imaging configurations on fully developed speckle zones of the *in vitro* phantom. These normalization factors were applied to all images reconstructed from experimental acquisitions (including before inference). Quantitative metrics were evaluated on a frame acquired with the transducer positioned to image three low-echogenic circular inclusions with a radius of 4 mm and centered at a depth of 40 mm [Figures 5.16(a) to 5.16(c)], simultaneously. That is an anechoic inclusion (Ω_A) and two low-echogenic inclusions of -6 dB (Ω_B) and -3 dB (Ω_C) with respect to their surrounding. For each inclusion, the contrast was computed against a background zone (Ω_D). The speckle was assessed using first and second order statistics (Section 5.6.1.1) within a square region (Ω_S) of $10\lambda \times 10\lambda$ centered at (0 mm, 27 mm).

An *in vivo* sequence of 60 frames was acquired at a frame rate of 30 Hz on the carotid

of a volunteer. The transducer was positioned on the neck of the volunteer to image a longitudinal view of the right carotid. Acoustic coupling was achieved by applying a layer of conventional US coupling gel between the face of the transducer and the skin of the volunteer. All images within the *in vivo* sequence were reconstructed identically, in particular with the normalization factors evaluated on the *in vitro* phantom.

Experimental acquisitions were evaluated on images obtained with the proposed approach using the trained MSLAE-16 CNN, and compared with images obtained from the LQ imaging configuration (also used as inputs to the CNN) and the HQ imaging configuration (considered as references). We opted for a CNN deployed with 16 initial expansion channels because of its real-time inference capabilities (Table 5.2).

5.6.2.2. Results

Figure 5.16 depicts the experimental results of an example image for both *in vitro* (top row) and *in vivo* (bottom row) acquisitions. Overall, it can be observed that, despite using only simulated dataset for training, key effects of the proposed approach translated well to experimental settings.

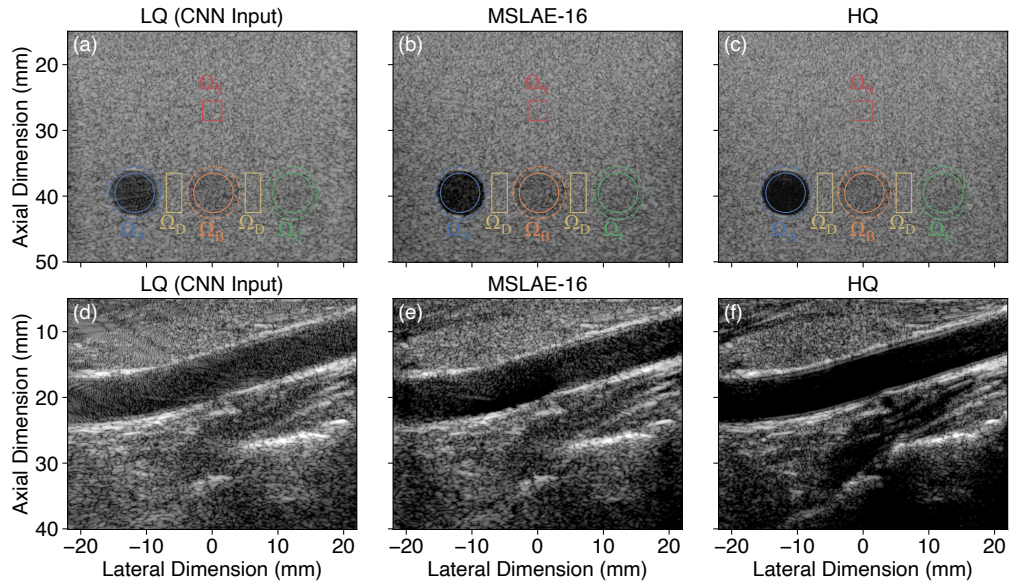


Figure 5.16. B-mode image representations of an *in vitro* test phantom (top row, 78-dB range) and an *in vivo* carotid sample (bottom row, 50-dB range): (a) and (d) single plane wave (PW) low-quality (LQ) images; (b) and (e) images recovered from LQ using the proposed convolutional neural network (CNN)-based image reconstruction method with the selected trained CNN (i.e., MSLAE-16); (c) and (f) reference high-quality (HQ) images reconstructed from the complete set of synthetic aperture (SA) acquisitions.

The visual assessment of *in vitro* results shows that both SL and GL artifacts (clearly visible in the anechoic inclusion) were strongly reduced. (Note that EW artifacts are harder to identify as they result in patterns that resemble speckle.) This effect was confirmed quantitatively by the contrast measured in each inclusion of the *in vivo* phantom

(Table 5.4). The contrast in the anechoic inclusion was largely improved compared with the LQ case. However, the proposed approach seemed to slightly “overshoot” in the other two inclusions compared with the HQ reference. Speckle patterns were generally well-preserved. Yet, almost no improvement in the lateral resolution of speckle patterns was observed and measured (Table 5.4).

TABLE 5.4
Experimental Test Phantom Metrics

Metric	LQ	HQ	MSLAE-16
C_A (dB)	-19.77	-29.55	-25.93
C_B (dB)	-6.18	-6.30	-6.71
C_C (dB)	-2.77	-3.41	-4.04
SNR	1.93	1.91	1.92
ACF lat. (μm)	282.3	220.6	281.0
ACF ax. (μm)	284.2	279.7	291.0

The *in vivo* experiments cover the full complexity of US imaging, namely highly diverse scattering processes, a wide range of echogenicities, and all physical effects neglected in the simulated dataset used for training. Yet, diffraction artifacts were strongly reduced, especially visible in zones where SL and GL artifacts aggregate [e.g., upper left of Figure 5.16(d)]. Structures initially shadowed by such artifacts were well restored, up to some degree of residual artifacts. Image quality improvements were less visible in deeper regions, partially due to the fact that diffraction artifacts do not seem to be dominant there. Very fine and low-echogenic details, such as the carotid intima, were not accurately restored. An overall remaining clutter noise was observed, for instance within the carotid or in the anechoic regions below it (also present in the LQ case). The complete image sequence from the longitudinal carotid acquisitions is presented in video format as supplementary material to [33].

5.7. Discussion

5.7.1. Performance in Ideal Conditions

The potential of the proposed CNN-based image reconstruction method for ultrafast US imaging was demonstrated through the results obtained in numerical experiments, in which the physical assumptions of the SIR model, defined in (5.1) and used to simulate the training dataset, were fully satisfied. Indeed, the proposed “two-step” image reconstruction method, which consists first of computing a low-quality estimate of the image using a re-weighted backprojection operator (akin to DAS), and then using a CNN trained to recover a high-quality image from that estimate, includes the learning of a complex image-to-image mapping, capable of strongly reducing (non-stationary) diffraction artifacts, mainly caused by GLs, SLs, and EWs, while preserving speckle patterns that result from main lobes.

We demonstrated that the proposed approach is not only capable of strongly reducing diffraction artifacts, but also to recover accurately zones initially hidden by them, on a dynamic range exceeding 60 dB (Section 5.6.1.2). This means that the detectability of lesions potentially hidden by such artifacts would be largely improved. The use of simulated reference images obtained from an optimal version of the linear transducer considered (i.e., UQ configuration), together with the HDR-sensitive MSLAE training loss, enabled us to reconstruct images from single-PW acquisitions with a quality similar to that of (gold-standard) SA imaging. This represents a more than 100-fold reduction in acquisition requirements, such as acquisition time, power requirements, or subsequent data transfer bandwidth requirements.

Artifacts caused by EWs were the most difficult to tackle, most likely due to their close resemblance to speckle patterns. This issue may be addressed using a CNN with greater capacity. These EW artifacts could also be reduced significantly by a suitable transmit apodization [14], [18], at the cost of a lower insonification energy on the sides of the field of view (probably restorable if accounted for in the training set), and an increased transmitter complexity. Due to the complexity of the task, the lateral resolution of speckle patterns was only slightly improved compared with LQ. This improvement was more pronounced for MSLAE-32 than MSLAE-16, suggesting that using a CNN with greater capacity could improve further the tightening of speckle patterns.

It should be noted that many elements composing the numerical test phantom, were not present in the training dataset. As they were formed by random ellipses of constant mean echogenicity filling out the entire image domain, no rectilinear boundaries, isolated bright reflectors, anechoic zones, or echogenicity gradients were explicitly present in training samples. The resulting trained CNNs adapted very well to these features, which is a strong sign of robustness, suggesting that the complex restoration mapping involved in the proposed method can be learned accurately. In particular, the robustness observed on bright reflectors strongly suggests that the learned mapping is not limited to fully developed speckle zones that composed the simulated-image dataset.

It was also observed that when using a CNN with increased capacity (i.e., MSLAE-16 vs. MSLAE-32), not only the metrics were improved, but also the restoration of remaining diffraction artifacts present in the UQ reference images (i.e., SLs). This confirms that the learning of the restoration mapping (from LQ to UQ images) involved in the proposed image reconstruction is effective, and may be achieved exactly with a CNN of even greater capacity than MSLAE-32, provided that a sufficiently large training dataset is available to avoid overfitting.

5.7.2. Performance in Experimental Conditions

In vitro experiments showed that images were improved significantly over conventional single PW (LQ) images. A reduction of diffraction artifacts was also observed on *in vivo* acquisitions, in particular at shallow depths. Yet, performance drops were observed compared with numerical evaluations performed in ideal conditions. Such performance drops were expected because CNNs were trained exclusively on simulated data. They are

likely to be caused by all physical phenomena not considered in the simulated training set, for which we relied on the SIR model defined in (5.1). These discrepancies come from several factors discussed hereafter.

A first set of potential differences come from the transducer array itself. Indeed, not all parameters were known or possible to be measure accurately. In particular, the electromechanical impulse response and exact geometry of each piezoelectric forming the transducer array could only be approximated. These parameters have an influence on the resulting (spatially dependent) PSF of the imaging system, namely the main lobe and all diffraction artifacts.

Another important set of differences come from the physical assumptions inherent to the SIR model considered for both simulating the dataset and deriving the backprojection operator (Sections 5.2.1 and 5.3.1). The SIR model only accounts for diffusive scattering in the medium, hence neither specular nor diffractive scattering regimes were taken into account [146, Sec. 8.2]. Such scattering regimes result in image statistics deviating from purely diffusive (Rayleigh) ones [78], [155] and may therefore disrupt trained models not accounting for such statistical features. Speed of sound is also assumed constant throughout the medium in the SIR model, hence deviations in mean speed of sound and/or local fluctuations may alter speckle patterns (i.e., image statistics). Dispersive attenuation was compensated using a standard TGC, which only corrects for a constant and frequency-independent attenuation throughout the medium. This may represent a limitation, especially at great depths, because the frequency-dependent attenuation continuously distorts the acoustic pulse as it travels through the medium. Thermal noise and quantization noise also have an increased impact with depth, as the backscattered signal amplitude decreases (mainly due to attenuation). While less likely to have a significant impact in the imaging schemes considered, thanks to the low mechanical index of unfocused transmit wavefronts, non-linear effects could also result in discrepancies.

Even though 3-D simulations were performed, only extruded 2-D phantoms were considered on a layer of 3.23λ , namely approximately one-sixth of the transducer height. This choice was made for computational reasons, but this means that potential out-of-plane artifacts were not exactly accounted for in the training set. It should be noted that such artifacts would also have impacted reference images.

5.7.3. Potential Improvements

As the training dataset is crucial to the learning of the complex restoration mapping involved in the proposed image reconstruction method, it is also a great opportunity for improvements. Indeed, as observed when analyzing the performance in ideal conditions (Section 5.7.1), the deployed CNN architecture trained using the proposed MSLAE as loss function is even capable of partially restoring remaining artifacts of UQ reference images (i.e., SLs). One may thus consider using reference images with a quality even higher than UQ in the training phase, for instance obtained with ideal PSFs without SLs. As the main reason for performance drops in experimental conditions seems to be related to the physical phenomena neglected for simulating the training dataset, the use

of more sophisticated simulations and/or experimental training datasets could lead to improved results in experimental conditions. The use of an experimental dataset is of particular interest as its acquisition would be faster than its simulation and all physical phenomena would be taken into account. However, undesirable physical phenomena such as frequency-dependent attenuation would also impact reference images. Moreover, the acquisition of such a dataset, free of motion artifacts and with a high diversity of images, remains a challenging task. The use of a more sophisticated simulated dataset remains extremely appealing as one could generate (low-quality) input images suffering not only from diffraction artifacts, but also from other physical phenomena (e.g., frequency-dependent attenuation), while generating (high-quality) reference images free of these.

As detailed in Sections 5.2.1 and 5.3.1, the proposed CNN-based image reconstruction method relies on a backprojection operator. This operator, defined in (5.3), is a DAS algorithm with weighting (apodization) functions that result from the (far-field) physical assumptions made to obtain a computationally tractable measurement model. While deviating from the theoretical derivations, the proposed method could also be implemented with common apodization functions (e.g., a Hamming window), conventionally designed to reduce diffraction artifacts at the cost of a lower lateral image resolution (i.e., broader main lobe). One should keep in mind that the trained CNNs were more efficient at reducing diffraction artifacts than improving the lateral resolution of speckle patterns. Also, note that the backprojection-based DAS operator results in a PSF with a tighter main lobe and higher diffraction artifacts than more “restrictive” apodization functions commonly used.

The fact that the restoration mapping is learned (using a CNN) on a specific imaging configuration (i.e., array geometry, impulse response, transmit wavefront, etc.) theoretically limits its use to said configuration. While this limits the approach, it maximizes its potential as the entire CNN capacity is used to learn an already complex non-stationary restoration problem. Not to mention that it is pretty common in US imaging systems to have finely tuned image reconstruction settings for each imaging configuration. It would thus be reasonable to consider dedicated pre-trained models for different imaging configurations. More elaborate datasets, accounting for variations in some imaging configuration parameters, could also be considered. This would probably result in improved generalization at the cost of degraded specialized performance.

It can also be mentioned that the proposed method does not contain an explicit data fidelity feedback mechanism, as opposed to convex optimization techniques. Data fidelity is “only” inferred implicitly by the overall training strategy (supervised learning). Subtle combinations of optimization algorithms and learned projections, such as proposed in [32], could be considered, at the risk of losing real-time imaging capabilities, which, in some cases, would not be detrimental.

Among all hyperparameter searches carried out (Section 5.5.2) the use of the proposed HDR-sensitive and RF-compatible MSLAE as loss function provided the largest increase in performance. Other architectural and/or optimization parameters could of course be optimized. An in-depth study of the activation function would be of particular interest in the context of RF signals. Indeed, due to its asymmetric (positive) nature, ReLU

does not strike as the most adapted activation function for bi-polar (RF) signals. Yet, preliminary studies conducted on this aspect using anti-rectifier-like activations did not yield satisfactory results so far.

5.7.4. Application Perspectives

The proposed approach may provide a viable solution to ultrafast US imaging modes in challenging environments, such as shear-wave elastography or high-velocity flow measurements (e.g., echocardiography), where only few acquisitions are possible to track the underlying (fast-evolving) physical phenomena accurately, and where artifacts (including diffraction ones) can severely degrade the accuracy of these imaging modes [6]. It should be noted that such modes heavily rely on the time-coherence between consecutive frames, and that current static image metrics used throughout this work cannot guarantee such a coherence. From the visual assessment of the *in vivo* carotid sequences, provided in video format as supplementary material to [33], the proposed approach also seems to preserve the time-coherence of moving structures (smooth movement of speckle patterns). A preliminary study on the time coherence between consecutive frames reconstructed using a CNN-based approach was carried out recently with positive outcomes [156]. This essential aspect is the subject of Chapter 6.

Portable systems could also benefit from the proposed approach to reduce the number of transmit-receive events required per frame, enabling, for instance, to reach more efficient power-down states of some electronic components [157]. Not to mention the potential simplifications in the transmitter as beamforming can be avoided when transmitting an unfocused wavefront, for which neither delay nor apodization are required. Sparse-array imaging is another area of research that could benefit from this approach, as it is even more prone to diffraction artifacts. Indeed, sub-sampling the transducer aperture inevitably entails substantial increases in GL, SL, and EW artifacts. An application to ultrafast imaging with sparse linear arrays is presented in Chapter 7.

It is also interesting to note that the use of a backprojection-based operation (akin to DAS in its computational complexity) followed by an inference is readily compatible with real-time imaging; and this using consumer-level GPUs (Table 5.2). This preserves one of the main advantage of US imaging compared with other medical imaging modalities.

5.8. Conclusion

We proposed a CNN-based image reconstruction method for high-quality ultrafast US imaging. A low-quality estimate is obtained by means of a backprojection operation, akin to conventional DAS beamforming, from which a high-quality image is then restored using a CNN trained specifically to remove diffraction artifacts, inherent to ultrafast US imaging. Trainings were performed on a simulated dataset using a loss function designed to account for both the HDR and the RF property of US images. Through extensive numerical experiments, we demonstrated that the proposed method can effectively reconstruct

images from single-PW insonifications with a quality comparable to that of gold-standard SA imaging. This represents a more than 100-fold reduction in acquisition requirements, which could unlock ultrafast image modalities where only a single insonification is possible, but also to reduce power requirements for portable systems. *In vitro* and *in vivo* experiments confirmed that trainings performed on simulated images translated well to experimental settings. Yet, trainings could also be performed on a more precise dataset, such as one using more accurate simulations or acquired experimentally. This could enable the CNN to learn an even more complex restoration mapping to remove other type of artifacts, currently neglected by the simulated dataset crafted for this work. The proposed method is readily compatible with real-time imaging, and could also benefit to other acquisition and imaging systems relying on antenna arrays.

As already mentioned, ultrafast US imaging modes heavily rely on the coherent displacement of speckle patterns between consecutive frames. Because the proposed CNN-based image reconstruction method was evaluated on static images only, and despite extensive evaluation of speckle patterns, it is thus not guaranteed that our method can be applied to ultrafast US imaging modes. As such, we also investigated an approach to ultrafast displacement estimation that exploits consecutive frames reconstructed by our CNN-based imaging approach. This will be the topic of the following chapter (Chapter 6).

Acknowledgment

We would like to thank warmly the many contributors to the work presented in this chapter and in [33]; Adrien Besson for the initial discussions and contributions; Olivier Bernard for the discussions on ultrasound-specific image metrics and speckle quality; Arthur Père and Jérémie Gringet for their invaluable help in implementing and benchmarking hundreds of models, and for the many suggestions for improvements; Malo Grisard and Philippe Rossinelli for their help in the early stages of this work; Verasonics for the technical support on the Vantage 256 system.

6. CNN-Based Ultrasound Image Reconstruction for Ultrafast Displacement Tracking

The material presented in this chapter is the result of joint work with M. Vonlanthen, F. Martinez, M. Arditi, and J.-Ph. Thiran. This material is also presented in [34].

6.1. Introduction

As discussed in Section 1.1, the ability of ultrafast US to image large tissue regions at frame rates of multiple kilohertz unlocked several breakthrough US imaging modes such as shear-wave elastography [5], high-frame-rate vector flow imaging [14], ultrasensitive flow imaging [13], and functional US neuroimaging [15]. All of these imaging modes rely on displacement estimation between consecutive frames acquired at ultrafast rates. The accuracy achievable with such displacement estimation techniques is strongly contingent upon two contradictory requirements: a high quality of consecutive frames and a high frame rate. Indeed, the image quality of consecutive frames can usually be improved by increasing the number of steered ultrafast acquisitions, but at the expense of a reduced frame rate and possible motion artifacts. Consequently, there exists a great need for a robust displacement estimation technique that does not rely on multiple acquisitions to reconstruct consecutive frames. This is of particular interest when analyzing rapidly evolving physical phenomena, such as the propagation of naturally occurring or externally induced shear waves [5]–[9], in highly heterogeneous tissue prone to strong diffraction artifacts.

In Chapter 5, we introduced a method for reconstructing high-quality US images from single unfocused acquisitions. It consists of a backprojection-based DAS operation followed by the application of a CNN, specifically trained to reduce the diffraction artifacts inherent to the deployed ultrafast US imaging setup. Strong artifact reduction was demonstrated in simulated, *in vitro*, and *in vivo* environments. The CNN-based image reconstruction method works strictly on a frame-by-frame basis and relies on the spatial information of each image only. Hence, it is completely agnostic to displacements that may occur between consecutive frames, making it a perfect fit for combination with state-of-the-art image-based displacement estimation techniques. In a preliminary work [156] we showed that a CNN-based image reconstruction method may preserve the time-coherence of speckle patterns between consecutive frames, which is essential to any image-based displacement estimation technique.

In this work, we propose an approach for estimating 2-D inter-frame displacements

at maximum frame rates, by combining our CNN-based image reconstruction method (Chapter 5) with a state-of-the-art 2-D speckle tracking algorithm. Although estimating the axial displacement (only) remains the standard in US imaging, 2-D displacement estimation is increasingly gaining attention in both flow and tissue motion applications [12], [158], [159], as it enables the analysis of more complex motion patterns. In elastography, 2-D displacement maps may be of interest to increase the quality and robustness of the estimated elasticity maps [65]. Also, 2-D speckle tracking represents an optimal fit for high-frame-rate displacement estimation since, unlike vector Doppler techniques, it does not rely on multi-angle acquisitions. Moreover, displacement estimation can be performed accurately from two consecutive frames only, whereas Doppler-based techniques usually require multiple consecutive frames to estimate the phase accurately.

Since our aim is to tackle displacement estimation at maximum frame rates, the proposed approach relies only on single unfocused acquisitions to reconstruct consecutive frames and on two consecutive frames only to obtain 2-D displacement estimates. The primary goal of this work is to assess whether the diffraction artifact reduction and speckle restoration capabilities of our CNN-based image reconstruction method (Chapter 5) can enable accurate estimation of displacements in zones initially shadowed by GL, SL, and EW artifacts. This work was conducted in the context of PW imaging with a linear transducer array (Section 6.2). The accuracy of the proposed approach was evaluated both in numerical and in *in vivo* experiments, and was compared with a state-of-the-art coherent plane wave compounding (CPWC)-based displacement estimation approach (Section 6.3). The results obtained demonstrate that the proposed approach is capable of estimating displacements in zones initially shadowed by SL and GL artifacts accurately. However, only slight improvements were observed in zones initially shadowed by EW artifacts, which still prevent accurate displacement estimates. In-depth results, implications, and limitations of the experiments carried out are analyzed and discussed in Sections 6.3 and 6.4, respectively. Concluding remarks are given in Section 6.5.

6.2. Materials and Methods

6.2.1. Imaging Configurations

We considered an US acquisition system composed of a 9L-D transducer (GE Healthcare, Chicago, Illinois, USA) and a Vantage 256 system (Verasonics, Kirkland, WA, USA), identical to the one considered in Section 5.4.1. Relevant imaging configuration parameters are summarized in Table 6.1. The 9L-D is a 192-element linear transducer array with a center frequency of 5.3 MHz and a bandwidth of 75 % (at -6 dB), and is commonly used for vascular imaging. All pulse-echo acquisitions were carried out by transmitting a single-cycle tri-state waveform of 67 % duty cycle centered at 5.208 MHz, with leading and trailing equalization pulses of quarter-cycle durations and opposite polarities. The received echo signals were sampled at 20.833 MHz, guaranteeing a Nyquist sampling rate up to a bandwidth of 200 %. To reconstruct images up to a depth of 60 mm, we considered

a maximum PRF of 9 kHz.

TABLE 6.1
Specifications of the Imaging Configurations Considered

Parameter	Value
Center frequency	5.3 MHz
Bandwidth	75 %
Aperture	43.93 mm
Element number	192
Pitch	230 μm
Element width ^a	207 μm
Element height	6 mm
Elevation focus	28 mm
Transmit frequency	5.208 MHz
Excitation cycles ^b	1
Sampling frequency	20.833 MHz

^aGuessed (no official data available).

^bSingle excitation cycle with equalization pulses.

All image reconstruction methods considered in this study rely on PW acquisitions performed without transmit apodization. Single PW acquisitions with normal incidence were used for the proposed CNN-based image reconstruction method (Section 6.2.2), and steered PW acquisitions were used for CPWC-based comparison methods (Section 6.2.3). For each transmit-receive event, echo signals were recorded on all transducer elements (i.e., full aperture). A typical speed of sound in soft tissue of 1540 m/s was assumed, resulting in an element spacing (i.e., pitch) of $\sim 0.78\lambda$ at the transmit frequency. As a result, images reconstructed with this transducer in the context of ultrafast imaging by conventional DAS algorithms will inevitably be contaminated by GL artifacts. As discussed in [18], most linear transducer arrays available commercially were optimized for line-by-line scanning, and are thus suboptimal when used in the context of ultrafast imaging. Nonetheless, these transducer arrays remain commonly used in ultrafast imaging [1], [6], [18], thanks to their wide aperture and resulting high lateral resolution.

6.2.2. CNN-Based Image Reconstruction Method

To obtain high-quality images from single unfocused acquisitions, we relied on our CNN-based image reconstruction method presented in Chapter 5, and briefly summarized hereafter.

The method consists of first reconstructing a (vectorized) low-quality estimate $\tilde{\mathbf{x}} \in \mathbb{R}^n$ from the (vectorized) transducer elements measurements $\mathbf{y} \in \mathbb{R}^m$, obtained from a single unfocused insonification, by means of a backprojection-based DAS operator $\mathbf{D} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ as $\tilde{\mathbf{x}} = \mathbf{D}\mathbf{y}$. The operator \mathbf{D} is composed of the adjoint of a linear measurement model (backprojection) and a pixel-wise reweighing operator (image equalization). The measurement model is based on linear acoustics and is derived from the SIR model [27], assuming far-field approximation both for the transmitter (e.g., ideal wavefront) and

the receiver (e.g., narrow transducer element), an ideal Dirac pulse-echo waveform, and neglecting tissue attenuation. Before summation, measurement values were interpolated using a B-spline approximation of degree three [68]. Analytic (complex) images, also called IQ images, were reconstructed on a $\lambda/4 \times \lambda/8$ (Cartesian) grid, with a width spanning the 9L-D aperture (Table 6.1) and a depth from 1 mm to 60 mm. The image grid resolution was chosen to guarantee Nyquist sampling of RF content of US images in both dimensions, resulting in images of 596×1600 pixels. The process was implemented with PyUS,¹ a GPU-accelerated Python package for US imaging developed in our laboratory.

In a second step, the low-quality estimate $\tilde{\mathbf{x}}$ is fed to a CNN $\mathbf{f}_{\theta}: \mathbb{R}^n \rightarrow \mathbb{R}^n$, with parameters θ , trained to recover a high-quality estimate as $\hat{\mathbf{x}} = \mathbf{f}_{\theta}(\tilde{\mathbf{x}})$, with strongly reduced diffraction artifacts and well-preserved speckle patterns. The CNN architecture is based on the popular U-Net [145] and on [31], with several improvements such as the use of RCBs and additive intrinsic skip connections (Section 5.3.2). It is a residual CNN with multi-scale and multi-channel filtering properties, composed of 2-D CLs and ReLUs arranged in symmetric downsampling and upsampling paths. As real-time displacement estimation was not a primary goal of this work, we used the best-performing CNN architecture analyzed in Section 5.6, with 32 initial expansion channels. The CNN was trained precisely as detailed in Section 5.5.1, namely in a supervised manner using a dataset composed of 30 000 simulated image pairs (i.e., input and ground-truth). The well-known Adam optimizer [148] was used to minimize the MSLAE loss, introduced in Section 5.3.3 to account for both the HDR and the RF property of US images. A total of 500 000 iterations were performed with a batch size of two and a learning rate of 5×10^{-5} . The same training dataset of simulated images was used. It is composed of low-quality input images reconstructed from single PW acquisitions with normal incidence. High-quality reference images were reconstructed from the complete set of SA acquisitions using a spatially oversampled version of the transducer array to ensure the absence of GL artifacts (only possible in a simulation environment). To reconstruct both input and reference images, element raw-data were simulated using an in-house 3-D SIR simulator, validated against the well-known Field II simulator [87]. Each numerical phantom was composed of random scatterers with a density that ensured fully developed speckle patterns throughout the resulting images. The simulated images composing the training dataset are characterized by overlapping ellipsoidal zones of random size, position, and orientation, with mean echogenicities spanning an 80-dB range.

6.2.3. Comparative Image Reconstruction Methods

For the CPWC-based comparison methods, acquisitions to reconstruct consecutive frames consisted of sequential transmit-receive events of N_a differently steered PWs, fired at maximum PRF. The PW steering angle spacing was evaluated as [6], [17]

$$\Delta\beta = \arcsin\left(\frac{\lambda}{L}\right) \approx 0.38^\circ, \quad (6.1)$$

¹<https://gitlab.com/pyus/pyus>

where λ is the wavelength of the transmit excitation and L is the transducer aperture. We restricted ourselves to odd acquisition numbers, thus the linearly increasing sequence of steering angles can be expressed as

$$\beta_n = n\Delta\beta, \quad n = -M, -M + 1, \dots, 0, \dots, M - 1, M, \quad (6.2)$$

where $M = (N_a - 1)/2$. We deployed an alternate steering angle sequence $(-\beta_M, \beta_M, -\beta_{M-1}, \beta_{M-1}, \dots, -\beta_1, \beta_1, 0)$, as proposed in [17].

In particular, we considered single-PW acquisitions with normal incidence, used both with the proposed CNN-based image reconstruction method and with DAS beamforming, as well as sequences of 3, 9, 15, and 87 steered PW acquisitions used with DAS beamforming. Comparison DAS-based methods are denoted CPWC-1, CPWC-3, CPWC-9, CPWC-15, and CPWC-87. The parameters for each imaging acquisition sequence considered are summarized in Table 6.2; the corresponding maximum achievable frame rates, given the deployed PRF of 9 kHz, are also provided. A sketch of the imaging acquisition schemes is depicted in Figure 6.1.

The CPWC-87 was used for reference purposes only and exclusively in settings where motion artifacts were negligible. This reference number of acquisitions was computed following [6] as

$$N_a^{\text{ref}} = \frac{L}{\lambda F_{\#}} \approx 87, \quad (6.3)$$

with an F-number $F_{\#} = 1.75$. The other comparison methods, namely CPWC-1 to CPWC-15, were selected to obtain a range of maximum achievable frame rates, namely from 9 kHz to 0.6 kHz, spanning typical values necessary for analyzing rapid events occurring in the human body.

TABLE 6.2
Plane Wave Imaging Acquisition Sequences Considered

Method	Sequence Parameters				Maximum Frame Rate	
	N_a	$\Delta\beta$	β_M	Type	PRF	
CNN	1	\times^a	\times^a	\times^a	\times^a	9 kHz
CPWC-1	1	\times^a	\times^a	\times^a	\times^a	9 kHz
CPWC-3	3	0.38°	0.38°	Alternate	9 kHz	3 kHz
CPWC-9	9	0.38°	1.52°	Alternate	9 kHz	1 kHz
CPWC-15	15	0.38°	2.66°	Alternate	9 kHz	0.6 kHz
CPWC-87	87	0.38°	16.34°	Alternate	9 kHz	0.1 kHz

^aSingle PW with normal incidence.

Each PW acquisition was reconstructed using the DAS algorithm detailed in Section 6.2.2. Coherent compounding of images reconstructed from steered acquisitions was realized by simple pixel-wise averaging. Note that as CPWC-1 only relies on single-PW acquisitions, it is not a compounding method. Its designation was adopted to simplify the naming convention. Also, images obtained from CPWC-1 are identical to input images

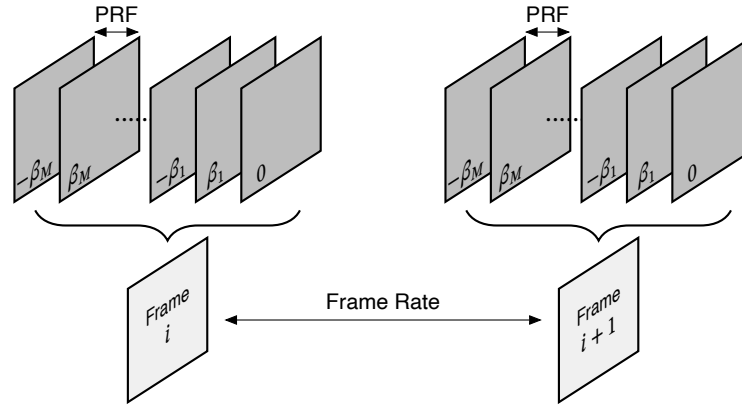


Figure 6.1. Sketch of the acquisition schemes deployed for the different plane wave (PW) imaging configurations considered. To form a single frame, a sequence of echo-signals from differently steered PWs is acquired at a pulse repetition frequency (PRF) of 9 kHz. The number of PWs composing each acquisition sequence depends on the imaging configuration (Table 6.2). The resulting frame rate is determined by the time interval between consecutive acquisition sequences, and is limited by the duration of a single acquisition sequence.

of the CNN-based image reconstruction (Section 6.2.2), as the same DAS algorithm was deployed in both cases.

6.2.4. Speckle Tracking Algorithm

The proposed speckle tracking algorithm is a block-matching algorithm based on normalized cross-correlation. It is heavily inspired by both the speckle tracking method described in [160], which won the challenge on synthetic aperture vector flow imaging (SA-VFI) organized during the 2018 IEEE International Ultrasonic Symposium (IUS) [161], and the PIVlab toolbox [162], a popular software for particle image velocimetry (PIV). Speckle tracking is fundamentally linked to PIV. However, instead of tracking particles to visualize flows, speckle tracking estimates displacements by tracking speckle patterns arising from interferences by scatterers separated by sub-resolution distances, assuming that these patterns are highly correlated between consecutive frames.

To estimate the 2-D displacement field between two consecutive frames S_1 and S_2 , both frames were identically subdivided into overlapping interrogation windows. The most probable displacement that occurred between a pair of interrogation windows was obtained by finding the maximum value (peak) of the (2-D) zero-normalized cross-correlation (ZNCC). To achieve sub-pixel precision, we applied a 2-D Gaussian regression around the ZNCC peak, as proposed in [163]. In order to analyze complex displacements, including shear and rotation, this process was deployed in a coarse-to-fine multi-pass algorithm [162]. Between each pass, S_2 was deformed (B-spline interpolation) using the estimated displacements to resemble S_1 more closely. For the next pass, the displacements between S_1 and the deformed S_2 were estimated in a similar way. The remaining displacement estimates of each pass were accumulated, resulting in more accurate estimates after a few passes. After each pass, statistical outliers of the estimates were smoothed using

the unsupervised smoothing algorithm described in [164].

Speckle tracking was performed on envelope images, obtained by computing the (pixel-wise) modulus of IQ images. Envelope images were downsampled by a factor of two in the axial dimension, in a uniformly spaced spatial grid of $\lambda/4 \times \lambda/4$ (i.e., 596×800 pixels). While applying normalized cross-correlation-based speckle tracking directly to RF signals may lead to a higher precision than using envelope signals [165], especially when analyzing very small displacements close to the Cramér-Rao lower bound [166], it is also much more prone to faulty displacement estimation because of speckle decorrelation [167, Sec. 14.2.1]. Speckle decorrelation increases when analyzing larger displacements, more complex displacement patterns with strong gradients (e.g., rotation), and tissue deformation [168], [169]. As our method is designed to be a robust displacement estimator over a wide range of displacements and flow patterns, envelope images were preferred for the purpose of speckle tracking. However, it is easily adapted to work with RF images if the potential increase in precision for small displacements is of interest.

For adapting the speckle tracking parameters to the imaging configurations and displacement ranges considered, we cross-validated a wide range of different interrogation window sizes, number of passes, and window overlaps using a dedicated numerical test phantom, namely a rotating cylinder centered at the elevation focus of the transducer, similar to the ones deployed in the numerical experiment (Section 6.3.1). Two different angular velocities were considered, resulting in the same inter-frame displacements considered in this work. Consecutive frames were generated by simulating high-quality images using CPWC-87 without rotating the cylinder between successive steered PW acquisitions (only achievable in a simulation environment). This strategy of “pausing” motion during a complete compounded acquisition sequence was exclusively deployed for the purpose of finding optimal speckle tracking parameters, to avoid being biased by potential motion artifacts. Inter-acquisition motion was considered in the following numerical experiment (Section 6.3.1).

Interestingly, the speckle tracking parameters yielding best overall displacement estimates in our settings were identical to the ones deployed in [160]. Thus, for all experiments conducted in this work, irrespectively of the displacement range and frame rate under consideration, we deployed the proposed speckle algorithm with four passes, square interrogation windows of 4 mm, 2.5 mm, 2 mm, and 1.5 mm, and a window overlap of 65%.

6.2.5. Metrics

To evaluate the accuracy of displacement estimates throughout the experiments, we relied on the relative endpoint error (REPE), a normalized version of the well-known endpoint error (EPE), commonly used in flow estimation techniques [170], [171]. Considering a displacement estimate vector $\hat{\mathbf{u}} \in \mathbb{R}^2$ and its true counterpart $\mathbf{u} \in \mathbb{R}^2$, the REPE can be

expressed as

$$\text{REPE} = \frac{\|\mathbf{u} - \hat{\mathbf{u}}\|_2}{\|\mathbf{u}\|_2}, \quad (6.4)$$

where $\|\cdot\|_2$ represents the Euclidean norm. The main advantage of REPE over EPE comes from its relative nature, enabling a more reliable comparison over a wide range of displacements. On the other hand, REPE becomes unstable as the reference displacement tends to zero. Such cases should therefore be analyzed with care.

We also relied on the mean relative endpoint error (MREPE) as a global metric to assess a set of l displacement estimates and true counterparts $\{\{\hat{\mathbf{u}}_1, \mathbf{u}_1\}, \dots, \{\hat{\mathbf{u}}_l, \mathbf{u}_l\}\}$ (e.g., extracted from a region of interest), by simply computing the sample mean of all REPE values over the set.

6.3. Experiments and Results

We conducted two experiments (numerical and *in vivo*) to assess the performance of the proposed 2-D displacement estimation approach, which combines our CNN-based image reconstruction method (Section 6.2.2) to reconstruct consecutive frames from single-PW acquisitions and the deployed speckle tracking algorithm (Section 6.2.4). In both experiments, we compared the proposed CNN-based displacement estimation method to CPWC-based tracking, which consists of applying the same speckle tracking algorithm to consecutive frames reconstructed using conventional CPWC (Section 6.2.3). For CPWC, a larger number of compounded acquisitions results, if motion artifacts are negligible, in better image quality and consequently in improved displacement estimation, at the cost of a reduced achievable frame rate. Thus, by studying different numbers of compounded acquisitions (Table 6.2) we compared the proposed approach to multiple levels of displacement estimation accuracy.

6.3.1. Numerical Experiment

6.3.1.1. Setup

For the first experiment, we used computer simulations to control the motion pattern, the relative echogenicities of tissue-mimicking structures, and the diffraction artifact levels precisely. The goal is to evaluate the quality of displacement tracking that can be achieved using the proposed method in rapidly moving, highly heterogeneous tissue, where strong diffraction artifacts hinder proper motion analysis with conventional CPWC-based tracking. All simulations were conducted using the same SIR simulator used to generate the training dataset (Section 6.2.2).

We designed a dynamic numerical test phantom composed of scatterers randomly positioned within four cylinders [A, B, C, and D in Figure 6.2(a)], embedded in an anechoic background. Each cylinder has a radius of 6.86 mm and a height of 1.0 mm, the

latter corresponding to the resolution cell size in elevation evaluated for the imaging configuration considered (Section 5.4.3). Within each of the four zones, an average of ten scatterers per resolution cell was used to ensure fully developed speckle patterns in the resulting images [146, Sec. 8.4.4]. The cylinders were centered such that cylinder A spawns distinct and spatially separable diffraction artifacts onto cylinders B, C, and D. Cylinders B, C, and D were positioned such that they are maximally covered by EW, SL, and GL artifacts, respectively [Figure 6.2(b)]. The mean amplitudes of scatterers located within cylinders B, C, and D were chosen to blend in with the amplitude of EW, SL, and GL artifacts arising from cylinder A [Figure 6.2(b)]. Specifically, the mean amplitudes in cylinders A, B, C, and D were set to 20 dB, -20 dB, -20 dB, and 0 dB with respect to an arbitrary 0 dB reference, respectively. Between successive simulated transmit-receive events (i.e., steered PWs), the scatterers were rotated with a constant counter-clockwise angular velocity around the center of the cylinder within which they are positioned. The same angular velocity was used for all cylinders.

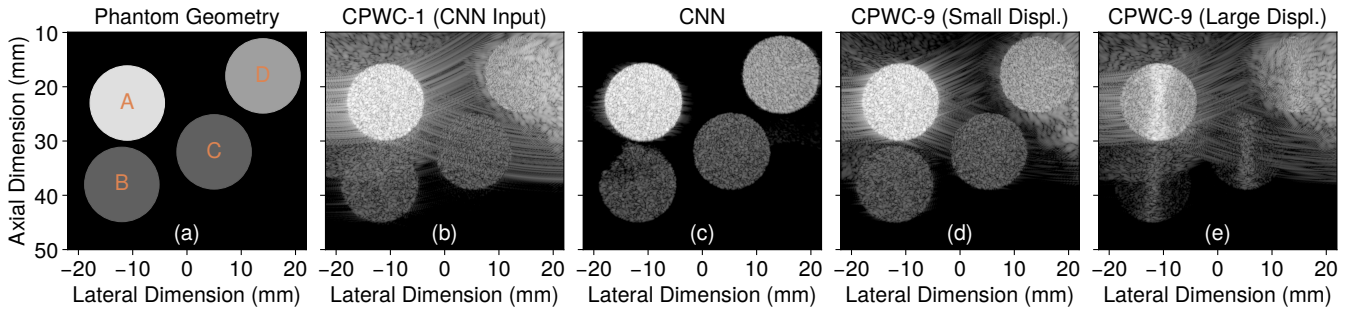


Figure 6.2. B-mode image representations (80-dB range) of a numerical test phantom sample: (a) the 2-D geometry of the deployed numerical phantoms, composed of four cylinders (A, B, C, and D) filled with dense point-scatterers rotating at constant angular velocity around their respective cylinder center; (b) image reconstructed by delay-and-sum (DAS) beamforming a single plane-wave (PW) acquisition (CPWC-1), simultaneously representing the convolutional neural network (CNN) input image for the proposed method; (c) image reconstructed using CNN-based reconstruction; images reconstructed by coherent plane wave compounding (CPWC) using nine steered PW acquisitions (CPWC-9): (d) small displacement range and (e) large displacement range. The frame rate and displacement range for each image reconstruction method considered are given in Table 6.3.

This experiment is designed to evaluate the accuracy of displacement estimates, obtained using the same speckle tracking algorithm on consecutive frames reconstructed with the different image reconstruction methods considered. For this purpose, displacements were estimated using the proposed CNN-based approach, as well as CPWC-1, CPWC-3, CPWC-9, and CPWC-15. Each method was deployed at its maximum achievable frame rate (Table 6.2), while always considering the same range of inter-frame displacements for comparison purposes. Inter-frame displacements ranging from $3.3 \mu\text{m}$ to $600 \mu\text{m}$ (i.e., approximately from $\lambda/10$ to 2λ) were analyzed, covering a range from the small displacements that typically occur in shear-wave elastography [6] or acoustic radiation force imaging [172], up to the large displacements that typically occur in external compression-based elastography [172]. It can be noted that these inter-frame displacement ranges correspond to velocities up to 5.4 m/s for the two methods capable of achieving a frame rate of 9 kHz in these settings (i.e., CPWC-1 and CNN). Such velocities

are close to peak velocities inside the cardiovascular system [173].

Two different sets of numerical phantoms were simulated for each image reconstruction method considered and associated frame rate, covering two inter-frame displacement ranges, namely $3.3\text{ }\mu\text{m}$ to $60\text{ }\mu\text{m}$ (small displacement range) and $33\text{ }\mu\text{m}$ to $600\text{ }\mu\text{m}$ (large displacement range). The respective angular velocities were determined such that the maximum inter-frame displacement occurs at a radius of 6.5 mm . The remaining border of 0.36 mm was ignored to avoid speckle tracking border effects in the quality evaluation. It corresponds to the approximate average resolution cell size in the transducer plane. A similar zone was ignored in the center of each cylinder. Displacement ranges are made explicit in Table 6.3 for each image reconstruction method considered, and the corresponding cross-radial velocity ranges are also provided as additional information. It can be noted that the large-displacement case involves displacements greater than half the deployed wavelength. As a result, motion artifacts are expected for CPWC-based methods [17].

TABLE 6.3
Displacement and Velocity Ranges Considered
for the Numerical Experiment

Method	Frame Rate	Large Ranges		Small Ranges	
		D. (μm)	V. (cm/s)	D. (μm)	V. (cm/s)
CNN	9 kHz	33–600	29.7–540	3.3–60	2.97–54
CPWC-1	9 kHz	33–600	29.7–540	3.3–60	2.97–54
CPWC-3	3 kHz	33–600	9.9–180	3.3–60	0.99–18
CPWC-9	1 kHz	33–600	3.3–60	3.3–60	0.33–6
CPWC-15	0.6 kHz	33–600	2.0–36	3.3–60	0.20–3.6

For all test configurations considered (i.e., method and displacement range), 50 statistically independent scatterer realizations were simulated, resulting in 50 inter-frame displacement estimate maps for each configuration. The accuracy of each method was measured locally in terms of REPE, by computing (6.4) for each displacement estimate (grid point) and corresponding true (analytical) value. The average local REPE was also computed over the 50 independent realizations (in each displacement estimate grid point).

6.3.1.2. Results

Figure 6.3 displays local REPE values, averaged over the 50 independent realizations performed in each configuration considered. To support the analysis, we also provide two global metrics computed for each zone, method, and displacement range considered (Table 6.4), namely the MREPE and the ratio of valid estimates (RVE). For the RVE, a local REPE value (averaged over the 50 independent realizations) exceeding 100 % was deemed invalid. It is thus directly related to the amount of saturated REPE values depicted in Figure 6.3, and provides a global metric less sensitive than MREPE to potentially huge-but-scarce local REPE values.

Zone A was designed such that it did not suffer from diffraction artifacts and could be

6.3. EXPERIMENTS AND RESULTS

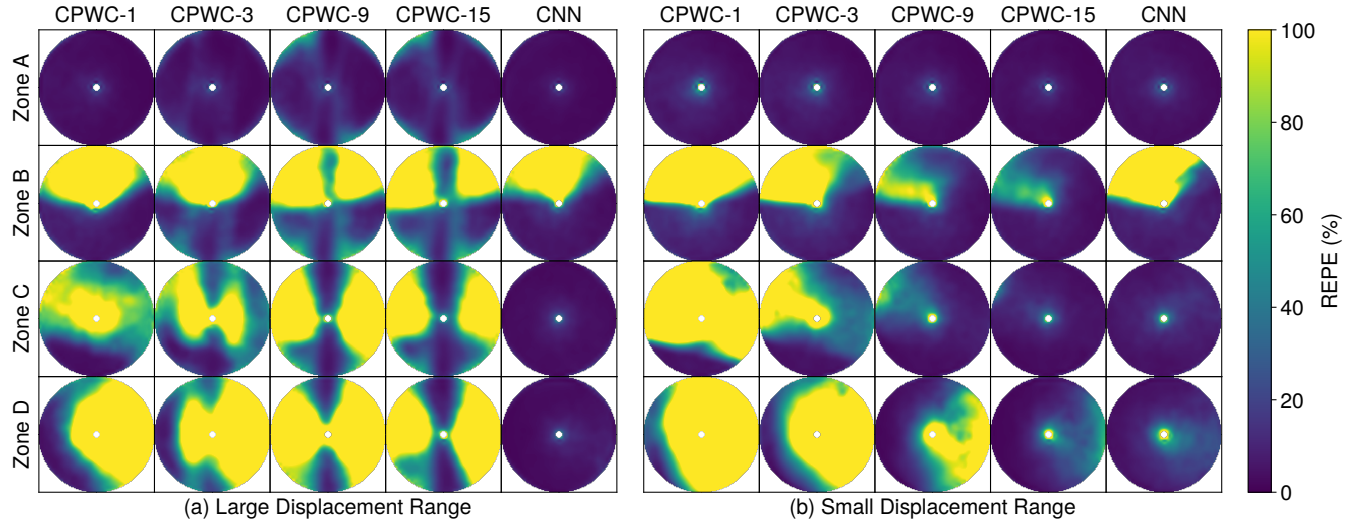


Figure 6.3. Local relative endpoint error (REPE), averaged over 50 independent realizations, of the 2-D displacement estimates inside each of the numerical phantom zones [A, B, C, and D in Figure 6.2(a)], obtained by applying the deployed 2-D speckle tracking algorithm (Section 6.2.4) on two consecutive frames for the two inter-frame displacement ranges considered: (a) large displacement range (from $33\text{ }\mu\text{m}$ to $600\text{ }\mu\text{m}$); (b) small displacement range (from $3.3\text{ }\mu\text{m}$ to $60\text{ }\mu\text{m}$). Consecutive frames were reconstructed either by coherent plane wave compounding (CPWC) from 1, 3, 9, and 15 differently steered PWs, or using the proposed convolutional neural network (CNN)-based image reconstruction method from single PWs. The frame rate and displacement range for each image reconstruction method considered are given in Table 6.3. The displayed REPE range is limited to 100 %. Local REPE values were interpolated onto a fine grid for display purposes.

TABLE 6.4
Global Evaluation Metrics of the Numerical Experiment

Zone	Metric	Large Displacement Range					Small Displacement Range				
		CPWC-1	CPWC-3	CPWC-9	CPWC-15	CNN	CPWC-1	CPWC-3	CPWC-9	CPWC-15	CNN
A	RVE (%)	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	MREPE (%)	4.45	7.24	12.99	12.84	3.62	7.34	6.91	5.25	4.36	5.81
B	RVE (%)	63.10	69.48	63.10	68.86	74.61	57.76	73.79	99.38	99.69	67.42
	MREPE (%)	78.58	61.32	82.72	72.65	48.36	143.91	64.30	26.66	19.37	95.44
C	RVE (%)	85.27	77.80	51.56	65.25	100.00	29.25	81.64	100.00	100.00	100.00
	MREPE (%)	67.50	66.30	91.49	71.03	4.98	192.57	54.37	17.64	8.29	9.61
D	RVE (%)	44.59	44.18	34.81	49.43	100.00	22.14	42.02	82.29	99.69	99.59
	MREPE (%)	135.18	120.69	123.74	100.19	5.51	504.38	159.15	52.90	17.83	15.67

used to assess displacement estimation in pure speckle zones. In the large-displacement case [Figure 6.3(a)], CPWC-based tracking suffered from increasing motion artifacts with the number of compounded acquisitions when tracking identical inter-frame displacements (i.e., at decreasing frame rates), reaching a stable motion artifact level after nine compounded acquisitions. The proposed method performed best and improved over CPWC-1 both in terms of local and global metrics. In the small-displacement case [Figure 6.3(b)], motion artifacts are negligible and all methods performed efficiently. A typical comparison of CPWC with and without motion artifacts is shown in Figures 6.2(d)

and 6.2(e) for CPWC-9.

Zone B was designed to suffer from EW artifacts. The proposed method was not capable of restoring speckle patterns shadowed by EW artifacts accurately, resulting in performance metrics only slightly improved compared with CPWC-1. Inaccurate restoration of speckle patterns plagued by EW artifacts can be observed in Figure 6.2(c) (e.g., clipped values). These artifacts could only be progressively resolved in the small displacement case [Figure 6.3(b)] with the increase in compounded acquisitions, because motion artifacts are negligible in that case.

Zone C was designed to suffer from SL artifacts. In the large-displacement case [Figure 6.3(a)], the reduction in SL artifacts achieved by compounding several acquisitions was counteracted by the induced motion artifacts, except in zones of pure lateral movement, making proper tracking impossible using CPWC-based tracking. The proposed method was capable of properly estimating displacements, with a quality only slightly worse than in artifact-free zone A. In the small-displacement case [Figure 6.3(b)], CPWC-based tracking was improved with the increase in compounded acquisitions, thanks to a more efficient SL reduction than with motion artifacts. The proposed method achieved a quality slightly worse than CPWC-15 but significantly better than CPWC-9.

Zone D was designed to suffer from GL artifacts, that increase in strength towards the right edge of the image. In the large-displacement case [Figure 6.3(a)], compounding multiple acquisitions reduced GL artifacts. Yet, motion artifacts prevented accurate displacement estimation except in zones of pure lateral movement. The proposed method significantly improved the displacement estimation quality over CPWC-1 and was the only method to enable tracking displacements in this case. In the small-displacement case [Figure 6.3(b)], the increase in compounded acquisitions enabled CPWC-based tracking to reduce the effect of GLs and restore the underlying speckle patterns, progressively resulting in an increased RVE and lower MREPE. The proposed method performed slightly better than CPWC-15.

6.3.2. *In Vivo* Experiment

6.3.2.1. Setup

For the second experiment, we applied the proposed approach to *in vivo* acquisitions, to analyze the natural tissue motion around the carotid artery. The goal of this experiment is to evaluate the robustness and translatability of the results obtained in the numerical experiment to the full complexity of *in vivo* imaging. As the natural tissue motion induced by cardiac pulsations in the vicinity of the carotid artery is small compared with the one considered in the numerical experiment, similar inter-frame displacements could be studied at a much lower frame rate, enabling the use of CPWC-87 for obtaining high-quality reference displacement estimates.

We analyzed the slow-moving tissue between the skin and the carotid artery of a healthy volunteer. In particular, motion within a specific tissue region of size 5 mm × 5 mm (Figure 6.4) was analyzed from consecutive frames acquired at a frame rate of 10 Hz. This

resulted in inter-frame displacements similar to those studied in the numerical experiment (Table 6.3), namely ranging from $5\text{ }\mu\text{m}$ to $125\text{ }\mu\text{m}$ approximately [Figure 6.4(e)]. Therefore, identical speckle tracking settings were used (Section 6.2.4). Speckle tracking was performed on full images, but we restricted our analysis to a specific zone characterized by fully developed speckle patterns, plagued by diffraction artifacts mainly originating from the highly echogenic carotid walls when imaged using CPWC-1 [Figure 6.4(a)]. The mean echogenicity of the analyzed speckle zone was approximately 20 dB lower than the echogenicity of the carotid walls, thus similar to the relative echogenicity between cylinders A and D studied in the numerical experiment.

We compared displacement estimates obtained using the proposed CNN-based approach, CPWC-1 (i.e., the CNN input), and CPWC-15 with respect to reference displacement estimates obtained with CPWC-87 (Table 6.2). As compounded acquisition sequences were performed at a PRF of 9 kHz, motion artifacts were negligible. More specifically, the maximum mean displacement estimated during a complete compounded acquisition sequence for CPWC-87 was approximately of $12\text{ }\mu\text{m}$. This amounts to approximately $\lambda/25$ and motion artifacts can therefore be neglected [17]. For each method being compared, consecutive frames were reconstructed using the relevant subset of steered PW(s) acquired for the reference CPWC-87 method (Section 6.2.3).

A total of 30 frames were obtained at a frame rate of 10 Hz, resulting in 29 inter-frame displacement estimate maps. For each inter-frame displacement estimate map, the accuracy of each method was measured locally in terms of REPE, by computing (6.4) for each displacement estimate (grid point) and corresponding reference value (CPWC-87). The quality of the displacement estimates for each pair of frames was assessed by computing the MREPE obtained within the region of interest.

6.3.2.2. Results

From the example images and corresponding displacement estimates [Figures 6.4(a) to 6.4(d)], one can observe that CPWC-1 suffered from diffraction artifacts (mainly caused by GLs and SLs arising from the carotid walls), disturbing both the speckle patterns and the resulting displacement estimates. These artifacts were strongly reduced using CPWC-15, leading to speckle patterns similar to the reference ones (CPWC-87), resulting in accurate displacement estimates. The proposed CNN-based imaging approach also reduced these artifacts, restoring the underlying speckle patterns accurately. This resulted in local displacement estimates with a quality similar to that obtained with CPWC-15.

The analysis of the MREPE values over time [Figure 6.4(f)] shows that, while CPWC-1 was generally unable to estimate inter-frame motion properly, the proposed method resulted in high and stable displacement estimation quality, similar to (though slightly worse than) CPWC-15. This observation matches the results of the numerical experiments on small displacements (Section 6.3.1.2). At 2.2 s, significant deviations in the MREPE values for all methods compared can be observed [Figure 6.4(f)]. As the estimated reference tissue displacement at this time instant is very small ($\sim 5\text{ }\mu\text{m}$) [Figure 6.4(e)], local REPE values, and as a consequence MREPE values, can be very sensitive to small absolute errors.

6. CNN-BASED ULTRASOUND IMAGE RECONSTRUCTION FOR ULTRAFAST DISPLACEMENT TRACKING

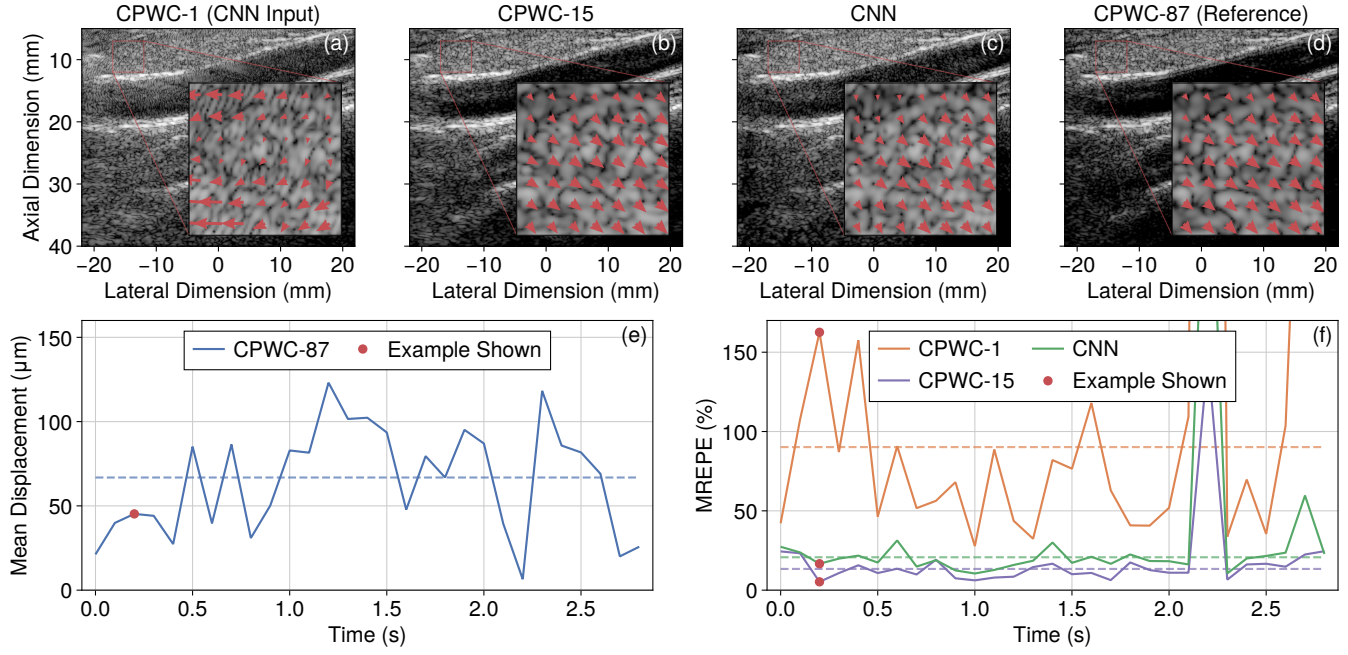


Figure 6.4. Examples of displacement estimates, mean reference displacement magnitude, and mean relative endpoint error (MREPE), obtained using the displacement estimation methods considered, in a fully developed speckle zone above the carotid artery: images of a longitudinal view of the carotid artery, are shown for (a) CPWC-1 (also CNN input), (b) CPWC-15, (c) CNN, and (d) CPWC-87 (reference); the bottom row shows (e) the mean reference displacement magnitude and (f) the MREPE along the entire *in vivo* sequence for each method considered. In each B-mode image of the top row, the square region of interest is highlighted and the corresponding magnified inset displays the 2-D displacement estimates. B-mode images are displayed using a dynamic range of 50 dB. The mean value (through time) of each quantity represented by a colorized solid line in (e) and (f) is represented by a horizontal dashed line of the same color. These mean values were computed ignoring samples at 2.2 s due to the resulting extreme MREPE values. An animation of the figure and the corresponding slideshow are provided as supplementary material to [34].

Moreover, it is likely that such small displacements are close to the minimum achievable displacement estimation error (Cramér-Rao lower bound), thus over amplifying the inherent sensitivity of the REPE to very small reference displacements. This behavior can also be observed in the numerical experiment on small displacements towards the center of rotation of zones C and D [Figure 6.3(b)]. Therefore, all values estimated at 2.2 s were ignored in the computation of the following global metrics. As global metrics, we computed the mean value through time (ignoring said time instant) of each estimated quantity [represented as dashed lines in Figures 6.4(e) and 6.4(f)]. The estimated mean inter-frame displacement is $66.83 \mu\text{m}$. The MREPE values are 90.18 %, 13.38 %, and 20.73 % for CPWC-1, CPWC-15, and the proposed CNN-based method, respectively.

6.4. Discussion

In this work, we proposed a 2-D motion estimation approach based on single unfocused acquisitions to reconstruct consecutive frames and on pairs of consecutive frames to estimate local displacements. This approach relies on our CNN-based image reconstruction

method (Chapter 5) to reconstruct full-view US frames from single unfocused acquisitions. It consists of first reconstructing low-quality images using a backprojection-inspired DAS algorithm and then feeding them to a CNN, specifically trained to reduce diffraction artifacts inherent to ultrafast US imaging. Inter-frame displacements are estimated by applying a state-of-the-art 2-D speckle tracking algorithm on consecutive frame pairs only.

6.4.1. Performance in Numerical Conditions

An important observation is that the proposed approach could not estimate displacements accurately in zones dominated by EW artifacts (Figure 6.3, zone B). This is directly related to the fact that the CNN deployed is not capable of restoring the underlying speckle patterns accurately [Figure 6.2(c)]. Slight improvements were observed compared with conventional single PW imaging (CPWC-1), but far less striking than in zones dominated by SL and GL artifacts (Figure 6.3, zones C and D). In Section 5.7.1, we already observed that EW artifacts were the most difficult artifacts to deal with, but also that the restoration quality improved with the increase of the CNN capacity. The latter implies that the reduction of these artifacts might be further improved using a more efficient CNN architecture or training process.

As expected, we observed in the large-displacement case that compounding multiple acquisitions in an attempt to improve the obtained image quality induces strong motion artifacts, mainly due to destructive interferences caused by axial motion. In the presence of motion artifacts, conventional CPWC-based speckle tracking was generally incapable of providing valid displacement estimation, in particular in zones plagued by strong diffraction artifacts. Consequently, compounding multiple acquisitions decreased the displacement estimation quality compared with single-PW acquisitions (CPWC-1). While motion compensation techniques have been proposed to tackle this issue [22], it remains unclear if motion-compensated coherent compounding can be deployed in zones plagued by diffraction artifacts (as it is based on inter-acquisition motion estimation), and if it actually improves displacement estimation quality in artifact-free zones compared with single unfocused acquisitions. We demonstrated that the proposed single PW CNN-based approach is capable of providing high-quality displacement estimates in artifact-free zones, as well as in zones plagued by SL and GL artifacts.

In the case of small displacements, increasing the number of compounded acquisitions using CPWC-based tracking progressively increased, as expected, the accuracy of displacement estimation. The proposed CNN-based approach achieves a displacement estimation quality comparable to CPWC-15 in zones suffering from SL and GL artifacts and comparable to CPWC-9 in artifact-free zones. It can be noted that the relative estimation precision achieved by the proposed approach was generally worse when analyzing small inter-frame displacements than in larger displacement cases. This was also observed for conventional CPWC-based tracking in artifact-free zones [e.g., compare CPWC-1, zone A in Figures 6.3(a) and 6.3(b)]. This mainly comes from the fact that the minimum estimation error of correlation-based tracking converges to a minimum value (Cramér-Rao

lower bound), which, relatively speaking, becomes more significant for smaller displacements [172]. For quantifying very small displacements, applying speckle tracking to RF data instead of envelope data may improve precision [165], [167, Sec. 14.2.1], at the expense of a reduced robustness to speckle decorrelation.

6.4.2. Performance in Physical Conditions

We demonstrated that the proposed CNN-based approach, which relies on single-PW acquisitions, significantly improved over conventional single PW imaging (CPWC-1). It also achieved an accuracy of inter-frame displacement estimation similar to that of 15 compounded acquisitions (CPWC-15), in conditions where motion artifacts were negligible and thus did not limit the performance of the comparative CPWC-15 method.

Overall, the quantitative evaluations performed in the *in vivo* experiment are comparable to those of the numerical one. This does not only show that the proposed method can be applied to *in vivo* data successfully, even though the CNN used for image reconstruction was trained on simulated data only, it also suggests that the results of the numerical experiments are robust and translatable (to some extent) to experimental conditions. More specifically, as motion artifacts were negligible in the *in vivo* experiment, the results obtained are best compared with the ones obtained in the numerical experiment on small displacements [Figure 6.3(b)]. It can be noted that the artifacts initially shadowing the zone in which displacement estimates were analyzed seem to be a combination of GL and SL artifacts spawned by the highly echogenic carotid walls [Figure 6.4(a)]. Thus, zones C and D of the numerical experiment are of interest for comparison purposes, as they contain SL and GL artifacts, respectively [Figure 6.2(b)]. While the quantitative metrics are similar, it is important to note that this presumptive combination of GL and SL artifacts was not present in the numerical experiment, and that the “signal-to-artifact” ratio was probably more favorable in the *in vivo* experiment than in the numerical one. One can observe that CPWC-15 performed better than the proposed method in the *in vivo* experiment [Figure 6.4(f)], whereas both methods performed similarly well in zones C and D of the numerical experiment on small displacements (Table 6.4). A performance drop of the proposed approach from numerical to physical conditions was expected since the deployed CNN was trained on simulated images only. This performance drop was already observed in Section 5.7.2, in which a detailed discussion on the discrepancies between the numerical and physical conditions can be found.

It should be noted that the *in vivo* experiment was intentionally carried out on a slow moving tissue zone. This enabled us to obtain reference displacement estimates for quantitative evaluation purposes, and to select a frame rate, identical for all methods considered, resulting in inter-frame displacements within ranges of interest. However, as speckle tracking is agnostic to the underlying frame rate, the results are fully translatable to fast motion cases with similar inter-frame displacement ranges, provided that the required frame rate is achievable by the method deployed.

6.4.3. Potential, Perspectives, and Limitations

The proposed approach is overall able to provide high-quality estimates for a wide range of 2-D inter-frame displacements, even in tissue regions dominated by SL and GL artifacts. As it only relies on single unfocused acquisitions to reconstruct consecutive frames, it is immune to motion artifacts. Moreover, it is limited only by the propagation time of acoustic waves, making it especially interesting for the analysis of rapidly changing events at very high frames rate, such as the propagation of shear waves in tissue or complex flow patterns within the cardiovascular system, where displacement estimation techniques based on multi-acquisition image reconstruction methods may not be deployable.

The major limitation is that the current implementation of the proposed approach was not able to provide accurate displacement estimates in regions dominated by EW artifacts. This is most probably due to the fact that the patterns resulting from EW artifacts resemble speckle patterns much more closely than the ones resulting from SL and GL artifacts [Figure 6.2(b)]. Since CNNs are, in essence, based on pattern recognition, the close resemblance of two patterns, one sought to be removed, the other to be preserved, represents a greater difficulty compared with a situation in which the two patterns are very distinctive. Both the EW behavior and the general performance of the approach might be further improved by augmenting the performance of the CNN used for image reconstruction. For instance, the use of a higher-capacity CNN or a more efficient training process may improve the restoration of tissue structures hidden by EW artifacts. Another way to tackle this limitation would be to use transmit apodization [18]. This technique can significantly reduce EW artifacts, at the cost of limited energy towards the image borders. However, its effectiveness is limited by the apodization capability of US systems, in particular by the transmitter complexity. If the method is not used at maximum achievable frame rate, and in the presence of sufficiently stationary motion, the robustness and precision of the displacement estimation could be improved, for instance, by averaging multiple displacement estimates or by using ensemble correlation [160].

This study was limited to tracking fully developed speckle patterns, hence no insights about tracking tissue structures arising from specular or diffractive scattering should be drawn from it directly. Yet, carotid wall movement was observed to be similar to that of conventional methods (see animation of Figure 6.4 provided as supplementary material to [34]). The training set was also limited to simulated images of fully developed speckle zones resulting from diffusive scattering. In Section 5.7.2, we observed that, while reconstructing other tissue structures is generally possible, the performance may be less potent than in fully developed speckle zones. Using a versatile training set may be considered to widen the applicability of both the reconstruction approach and the displacement tracking method proposed here.

6.5. Conclusion

In this work, we proposed an approach for estimating 2-D inter-frame displacements in the context of ultrafast US imaging. The approach consists of a CNN trained to restore high-quality images from single unfocused acquisitions and a speckle tracking algorithm to estimate inter-frame displacements from two consecutive frames only. Compared with conventional multi-acquisition strategies, this approach is immune to motion artifacts and enables accurate motion estimation at maximum frame rates, even in highly heterogeneous tissues prone to strong diffraction artifacts. Numerical and *in vivo* results demonstrated that the proposed approach is capable of estimating displacement vector fields from single-PW acquisitions accurately, including in zones initially hidden by SL and GL artifacts. The proposed approach may thus unlock the full potential of ultrafast US, with direct applications to imaging modes that depend on accurate motion estimation at maximum frame rates, such as shear-wave elastography or ultrasensitive echocardiography.

On a more general perspective, this work further validates the potency of the CNN-based image reconstruction method presented in Chapter 5. Indeed, this method not only provides high-quality images from single unfocused acquisitions, but also preserves the information of underlying physical phenomena that can be further exploited for estimating inter-frame displacements accurately.

Acknowledgment

We would like to thank Quentin Ligier for his important contribution to the implementation of the speckle tracking algorithm deployed in the work reported in this chapter. We would also like to thank the anonymous reviewers of [34] for their comments, which contributed significantly to the improvement of the work reported in this chapter and in [34].

7. CNN-Based Ultrafast Ultrasound Imaging with Sparse Linear Arrays

The material presented in this chapter is the result of joint work with M. Vonlanthen, F. Martinez, M. Arditi, and J.-Ph. Thiran. It is a direct application of the CNN-based image reconstruction method proposed in Chapter 5 to the problem of sparse-array imaging. This material is also presented in [35].

7.1. Introduction

Sparse array design and image formation have been studied extensively during the past decades, mainly to reduce costs and complexity of large-channel-count 3-D US scanners [174], and more recently for low-power portable ones [175]. Real-time 3-D US imaging relies on 2-D transducer arrays with thousands of elements, rendering both the fabrication process and element control involved very challenging. The use of sparse 2-D arrays has proven to be a promising solution to reduce the complexity of 3-D US systems and is still heavily studied [176]. On the other hand, the recent advent of low-power, portable US systems [157] has led to increased interest in 2-D imaging using sparse 1-D arrays [175], not only to reduce the costs and complexity, but also to keep the energy consumption and data transfers minimal.

Many sparse array design strategies have been studied with the goal of maintaining a high image quality, while reducing the number of array elements deployed as low as possible. While “naively” undersampling a transducer array uniformly keeps both the main lobe shape and SL levels largely intact, the interferences between element channels can lead to devastating GL artifacts, dominating a large portion of the resulting image. Thus, most design strategies focus on mitigating GL artifacts, while making compromises on the resulting resolution, SL levels, and image uniformity. Some consider random element positioning (or subsampling) to break the periodicity of the array [177], others leverage probabilistic optimization techniques (e.g., simulated annealing) under some arbitrary cost function [176] or adopt different transmit and receive element arrangements with the aim of cancelling out their respective GLs [175], [178].

In Chapter 5, we proposed reconstructing US images from single unfocused acquisitions, by first computing a low-quality image estimate with a backprojection-based DAS algorithm and then restoring a high-quality image using a CNN, specifically trained to reduce diffraction artifacts. The CNN has proven a strong ability to remove GL artifacts while properly restoring previously shadowed tissue structures of interest.

In this chapter, we investigate, in a numerical environment, the applicability of our CNN-based image reconstruction method (Chapter 5) to the case of ultrafast PW imaging using undersampled arrays. To ensure GL-free transmit wavefronts, we considered unapodized single-PW insonifications using the full array on transmit and uniformly undersampled arrays on receive. As this simple transmit scheme could be implemented in the front-end, receive-only sparsity is sufficient for the sought-after data-rate reduction. In contrast to most sparse array design approaches, we considered uniform undersampling for its simplicity and its ability to maintain proper main-lobe and SL characteristics, and relied on a trained CNN to reduce the resulting GL artifacts significantly.

7.2. Materials and Methods

7.2.1. Imaging Configurations

As in Section 5.4.1, we considered an US acquisition system composed of the 9L-D transducer (GE Healthcare, Chicago, Illinois, USA) and the Vantage 256 system (Verasonics, Kirkland, WA, USA). Relevant system specifications are summarized in Table 7.1. The 9L-D is a linear array transducer containing 192 elements with a center frequency of 5.3 MHz and a bandwidth of 75 % (at -6 dB). An excitation composed of a single-cycle tri-state waveform of 67 % duty cycle centered at 5.208 MHz, with leading and trailing equalization pulses of quarter-cycle durations and opposite polarities, was used for all pulse-echo measurements. The received echo signals were sampled at 20.833 MHz, guaranteeing a Nyquist sampling rate up to a bandwidth of 200 %. We assumed a speed of sound of 1540 m/s (typical for soft tissue), resulting in an element spacing (i.e., pitch) of more than half a wavelength (i.e., $\sim 0.78\lambda$) at the effective frequency. Thus, even when using all transducer elements on receive, images reconstructed from unfocused acquisitions with a DAS algorithm will inevitably be contaminated by GL artifacts, especially at shallow depths.

Based on this acquisition system, we considered full-aperture unapodized single-PW insonifications with normal incidence and different uniformly undersampled receive configurations, namely by a factor of 2, 3, 4, and 6 (U2–U6). For comparison purposes, we also considered single-PW insonifications with full-aperture receive, named LQ to keep the same terminology as in Chapter 5. As it will serve as reference for CNN trainings (Section 7.2.2), we also introduce here the UQ configuration, which relies on SA imaging from a spatially oversampled version of the transducer array with 383 elements and a halved pitch of $\sim 0.39\lambda$ (i.e., simulation exclusive), guaranteeing GL-free images of gold-standard quality. All configurations considered are summarized in Table 7.2.

From the acquisitions obtained with each configuration considered, full-view images were reconstructed using the backprojection-based DAS algorithm used in Chapter 5. Analytic IQ images were reconstructed on a $\lambda/4 \times \lambda/8$ (Cartesian) grid, with a width spanning the 9L-D aperture (Table 7.1) and a depth ranging from 1 mm to 60 mm. The image grid resolution guaranteed Nyquist sampling for the RF content of US images in

TABLE 7.1
Acquisition System Specifications

Parameter	Value
Center frequency	5.3 MHz
Bandwidth	75 %
Aperture	43.93 mm
Element number	192
Pitch	230 μm
Element width ^a	207 μm
Element height	6 mm
Elevation focus	28 mm
Transmit frequency	5.208 MHz
Excitation cycles ^b	1
Sampling frequency	20.833 MHz

^aGuessed (no official data available).

^bSingle excitation cycle with equalization pulses.

TABLE 7.2
Transmit and Receive Array Configurations

Parameter	UQ ^a	LQ	U2	U3	U4	U6
Tx Element Number	383	192	192	192	192	192
Tx Pitch	0.39 λ	0.78 λ	0.78 λ	0.78 λ	0.78 λ	0.78 λ
Rx Element Number	383	192	96	64	48	32
Rx Pitch	0.39 λ	0.78 λ	1.56 λ	2.34 λ	3.12 λ	4.68 λ
Tx-Rx scheme	383 SA	1 PW	1 PW	1 PW	1 PW	1 PW

^aSimulation exclusive.

both lateral and axial dimensions, and resulted in images of size 596×1600 pixels.

7.2.2. CNN-Based Image Reconstruction Method

To obtain higher quality images from single unfocused acquisitions using sparse array configurations, we deployed the two-step CNN-based image reconstruction method proposed in Section 5.3.1, briefly summarized hereafter. In a first step, a (vectorized) low-quality estimate $\tilde{\mathbf{x}} \in \mathbb{R}^n$ is computed from the (vectorized) transducer elements measurements $\mathbf{y} \in \mathbb{R}^m$, obtained from a single unfocused insonification, by means of a backprojection-based DAS operator $\mathbf{D} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ as $\tilde{\mathbf{x}} = \mathbf{D}\mathbf{y}$. In this study, the low-quality estimates corresponded to either of the U2 to U6 images described in Section 7.2.1. The low-quality estimate $\tilde{\mathbf{x}}$ is then fed to a CNN $\mathbf{f}_{\theta} : \mathbb{R}^n \rightarrow \mathbb{R}^n$, with parameters θ , trained to reduce diffraction artifacts while restoring the underlying image content of interest and preserving proper speckle patterns, thereby recovering a high-quality estimate as $\hat{\mathbf{x}} = \mathbf{f}_{\theta}(\tilde{\mathbf{x}})$.

The deployed network architecture is a residual CNN with multi-scale and multi-channel filtering properties, adapted from the popular U-Net. It is composed of 2-D CLs and ReLUs arranged in symmetric downsampling and upsampling paths and includes several adaptations such as the use of RCBs and additive intrinsic skip connections

(Section 5.3.2). For this study, we deployed the best-performing CNN architecture analyzed in Section 5.6 that still maintains real-time capabilities, namely using 16 initial expansion channels.

Trainings were performed precisely as detailed in Section 5.5.1, namely in a supervised manner using a dataset composed of 30 000 simulated image pairs (i.e., input and reference). The MSLAE loss (Section 5.3.3), which accounts for both the HDR and the RF property of US images, was minimized using the well-known Adam optimizer over 500 000 iterations, using a learning rate of 5×10^{-5} and a batch size of two. For each input image type, a dedicated CNN was trained using a specific simulated-images training dataset composed of low-quality input images (U2–U6) and corresponding reference images (UQ). An in-house 3-D SIR simulator, validated against the well-known Field II simulator [87], was used to generate the element raw-data for both input and reference images. The numerical phantoms were composed of random scatterers with a density that ensured fully developed speckle patterns throughout the resulting images. The resulting training images are characterized by overlapping ellipsoidal zones of random size, position, and orientation, with mean echogenicities spanning an 80-dB range.

7.2.3. Numerical Experiment

To evaluate the image quality achieved using CNN-based reconstruction (Section 7.2.2) on single-PW acquisitions with the different uniformly undersampled receive configurations considered (U2–U6, Section 7.2.1), we relied on the numerical test phantom deployed in Section 5.4.3, previously adapted from [80] to the US acquisition system under study. The test phantom is composed of tissue-mimicking echogenic zones embedded in an anechoic background [Figure 7.1(f)], and allows to compute a wide range of image quality metrics. It comprises a cylindrical inclusion embedded in a block of square section with a prescribed contrast of 36 dB, a rectangular region with a log-linear lateral echogenicity gradient ranging from 30 dB to -50 dB, and four bright reflectors p_0 , p_1 , p_2 , and p_3 at increasing depths. Similar to the numerical phantoms used for training purposes (Section 7.2.2), echogenic zones were filled with densely populated random scatterers to guarantee fully developed speckle patterns.

A total of 300 statistically independent realizations of the test phantom were simulated for each undersampled configuration considered with and without applying the corresponding trained CNN, as well as for the comparative methods LQ and UQ. Image quality was assessed using a subset of the quality metrics from the numerical test phantom, namely using the SSIM against UQ reference images as a global metric (evaluated on log-compressed B-mode representations), the contrast achieved between the block and the inclusion, the lateral FWHM at each bright reflector position, and the linearity preservation. All metrics were computed precisely as described in Section 5.4.3 and averaged over the 300 statistically independent realizations.

7.3. Results and Discussion

Figure 7.1 shows example images of a test phantom sample (from the 300 realizations) for each imaging configuration considered (Table 7.2), and all resulting image quality metrics, computed as described in Section 7.2.3, are provided in Table 7.3. Note that the contrast metrics reported in Table 7.3 differ slightly from those reported in [35] because the background zone [Figure 5.13(a), Ω_B] was modified from its initial version [33]. The resulting contrast measured did not contradict any conclusion drawn in [35].

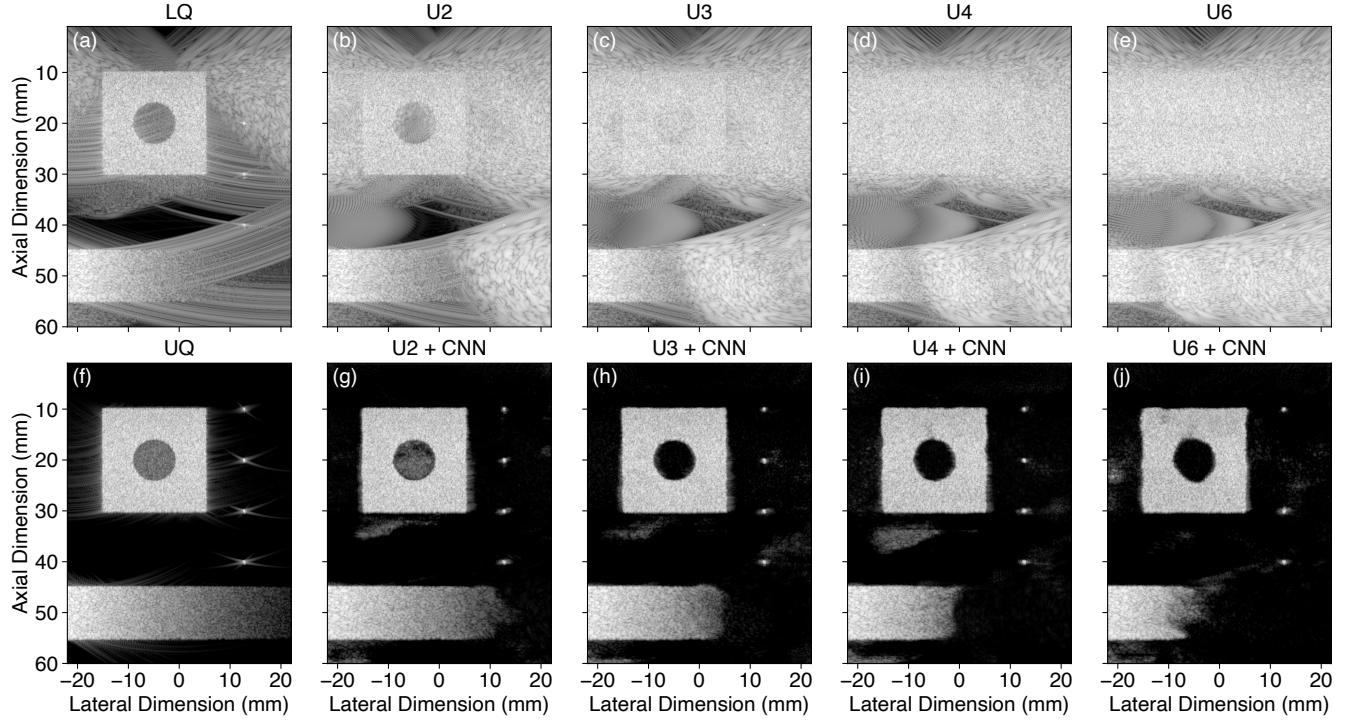


Figure 7.1. B-mode image representations (98-dB range) of a numerical test phantom sample, reconstructed using each imaging configuration considered (Table 7.2): (a) low-quality (LQ) configuration (i.e., full array on receive) and (f) ultra-high-quality (UQ) configuration (i.e., reference image), providing image quality comparisons; (b) to (e) U2, U3, U4, and U6 uniformly undersampled configurations (by a factor of 2, 3, 4, and 6 on receive only); (g) to (j) images recovered from U2, U3, U4, and U6 using the proposed approach with the corresponding trained convolutional neural network (CNN).

The impact of increased uniform undersampling can be analyzed in Figures 7.1(a) to 7.1(e), where both a significant increase in the area covered by GL artifacts and a decrease of the signal-to-artifact ratio within zones contaminated by these artifacts can be observed with the increase of the undersampling factor. Furthermore, comparing images obtained from the undersampled configurations, in zones free of GL artifacts (e.g., highly echogenic portion of the rectangular gradient), [Figures 7.1(b) to 7.1(e)] to the LQ image which uses the full array [Figure 7.1(a)], one can note that both main-lobe and SL characteristics are not effected by the uniform undersampling.

Overall, the CNN-based image reconstruction demonstrated a high efficiency at recovering many structures previously hidden by the significant GL artifacts arising from the undersampled configurations. For instance, one can appreciate the recovery performance

7. CNN-BASED ULTRAFAST ULTRASOUND IMAGING WITH SPARSE LINEAR ARRAYS

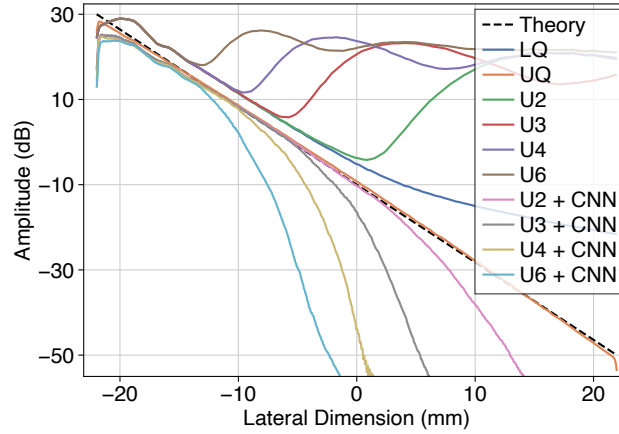


Figure 7.2. Mean amplitude responses (averaged along the axial dimension) of the lateral gradient zone in the numerical phantom [Figure 7.1(f)], averaged over 300 independent realizations.

TABLE 7.3
Numerical Test Phantom Metrics

Metric ^a	LQ	UQ	U2	U2 + CNN	U3	U3 + CNN	U4	U4 + CNN	U6	U6 + CNN	
SSIM	0.31	× ^b	0.12	0.68	0.07	0.62	0.03	0.56	0.00	0.56	
C (dB) ^c	−28.33	−36.06	−16.51	−43.61	−3.47	−71.42	−0.58	−75.66	+0.05	−75.90	
FWHM lat.	p_0 (μm)	276.6	202.2	× ^d	265.9	× ^d	257.0	× ^d	269.0	× ^d	298.8
	p_1 (μm)	336.2	242.7	339.5	273.7	× ^d	304.4	× ^d	× ^d	× ^d	× ^d
	p_2 (μm)	388.6	280.5	× ^d	299.7	× ^d	316.1	× ^d	306.4	× ^d	× ^d
	p_3 (μm)	446.6	322.4	× ^d	359.7	× ^d	376.9	× ^d	410.8	× ^d	357.7

^a Metrics were averaged over 300 independent realizations.

^b SSIM metric was computed against UQ.

^c Prescribed contrast of -36 dB.

^d At least one image sample resulted in an undetectable reflector.

of the bright reflectors which were almost undetectable [e.g., compare Figure 7.1(d) and Figure 7.1(i)], or the recovery of the square block which progressively disappears below GL artifacts as the undersampling factor increases. This is confirmed by the increase in SSIM, for instance improved from 0.07 to 0.62 in the U3 case, even surpassing the SSIM obtained for LQ significantly.

However, these recovery performances were limited to a certain signal-to-artifact ratio of the input images, below which the CNNs generally removed both the GL artifacts and the underlying signal of interest. Because of decreasing signal-to-artifact ratios with increased undersampling, a progressively larger portion of the image could not be recovered accurately and the image quality achieved thus decreased steadily (Table 7.3). This behavior can be well observed when comparing the recovery of the circular inclusion from the U2 and U3 images [Figures 7.1(g) and 7.1(h)]. While the CNN was able to (partially) recover the structures within the circular inclusion in the U2 case, the inclusion was fully thresholded by the CNN from an undersampling factor of three (U3–U6).

The structures that could be recovered by the CNN, namely in zones of the input images with a signal-to-artifact ratio sufficiently high before restoration, were recovered

with high-quality speckle patterns, preserved log-linearity (see Figure 7.2) and, in the case of the point scatterers, with a lateral resolution comparable or better than the one achieved for LQ (FWHM in Table 7.3). The analysis of the mean amplitude responses within the lateral gradient zone (Figure 7.2) can be used to estimate the signal-to-artifact ratio below which the CNN is not capable to recover the underlying signal of interest with the prescribed log-linear amplitude. For instance, in the U3 case, the restored mean amplitude (U3 + CNN) starts to deviate from UQ at approximately -5 mm, where the input signal-to-artifact ratio (between LQ and U3) is of approximately -3 dB.

7.4. Conclusion

We showed that the proposed CNN-based image reconstruction method can increase the image quality achieved from single-PW acquisitions using uniformly undersampled linear arrays on receive. Indeed, the overall quality of the restored images is competitive to images reconstructed by a DAS algorithm from single-PW acquisitions on the full array up to an undersampling factor of three, demonstrating a high potential in the context of sparse array imaging. As observed in Section 5.7.3, the performances can be further improved by optimizing the CNN architecture, its training process, or simply increasing its capacity. To evaluate the potency of the studied approach further, additional studies, including *in vitro* and *in vivo* data should be conducted.

8. Concluding Remarks

As most chapters contain in-depth discussions on the results obtained, the limitations observed, and the perspectives envisaged, only prospective notes of particular interest for future research are discussed hereafter.

The CNN-based image reconstruction presented in Chapter 5 proved to be a promising approach for reconstructing high-quality US images from single unfocused acquisitions; most importantly, images that can benefit displacement estimation techniques (Chapter 6) such that the accuracy of ultrafast US imaging modes could be further improved. Thus, it is not unrealistic to believe that such approaches may be useful in clinical settings. In an effort to improving these approaches and push them further, few topics are discussed in this section.

Concerning the structure of the CNN deployed, we relied on a “U-Net-like” architecture, primarily because it was successfully applied to other medical image reconstruction methods [31], [32]. Some modifications of the architecture were proposed, mainly inspired by multi-level wavelet decomposition, the sparsity-promoting term deployed in our initial SR approach. All these modifications resulted in performance increases, thus it is expected that further optimization and hyperparameter searches on the architecture itself should further increase the resulting performances. Nevertheless, compared with the improvements obtained by optimizing the training datasets and adapting the training loss function, these improvements can be considered “minor.”

As for any deep learning method, the quality of the dataset used to train some artificial neural network is obviously crucial. This is one of the reasons why we actually developed a dedicated simulator that allowed the generation of relatively large datasets in realistic time frames. In total three complete training datasets of about 30 000 were simulated before obtaining the one used in the different CNN-based contributions of this thesis (Chapters 5 to 7). The two most important features of the datasets designed are: a high diversity of structures resulting in fully developed speckle patterns; and echogenicities spanning at least the expected (high) dynamic range of signals (including a proper confidence interval). I personally find the use of simulated-image datasets particularly adapted to research because the environment is fully controlled and testing can be performed in identical settings. It also allows using reference images that go “beyond physics” (as we did for the UQ reference image configuration). Yet, experimental acquisitions could be highly beneficial to more applied scenarios. A good experimental setup is certainly complicated to design, but, once available, datasets can be obtained much faster than simulated ones. Also, I have the feeling that they would result in drastic performance

8. CONCLUDING REMARKS

increases in experimental conditions.

Adapting the loss to the signal properties was probably the second most important improvement obtained. The use of the proposed MSLAE as loss function, designed to be compatible with both the HDR and RF properties of US images, resulted in drastic performance improvements compared with less adapted losses such as the MSE and MAE. Yet, MSLAE remains a “pixel-based” function, which makes it very efficient on a computational perspective, but also means that only first-order statistics can be accounted for by such a loss function. And, because second-order statistics is crucial to many post-processing algorithms such as 2-D speckle tracking, a suitable (distance) function accounting for the expected second-order statistics should result in notable improvements (though not trivial).

Another component of major importance to deep learning is the activation function. Because RF US images are inherently “symmetric” signals, the use of ReLU as activation function is most probably suboptimal. Yet, from the experiments carried out using “off-the-shelf” symmetric rectifiers, none outperformed ReLU. Preliminary works on examining feature maps of CNNs trained on RF images using ReLU as activation function confirmed that the “positive” and “negative” components of RF content is somehow processed by the CNN separately (i.e., symmetric feature maps). It is thus expected that a CNN trained with a suitable bi-polar activation function should be twice as efficient as the same CNN trained with ReLU. We investigated the use of trainable and non-trainable soft-thresholding activation functions, here again with disappointing results so far. There is however a good chance that the initialization of the weights was not optimal. Also, note that the recent work by Aziznejad *et al.* [179] on trainable activations could be of great interest in the case of RF signals.

As already discussed in Section 5.7, and pointed out previously, the use of simulated datasets offers many interesting perspectives. In this thesis, we exploited a spatially oversampled version of the transducer array to guarantee GL-free images. Another US application could benefit from simulated datasets, namely speckle reduction (also referred to as despeckling). Even though speckle patterns are crucial to displacement estimation, they are also known to reduce both the contrast and the resolution of images. Thus, removing these patterns in applications where tracking is not required could be of great interest. Having access to both experimental acquisitions and underlying echogenicity maps is certainly not easy (if not impossible). Simulated datasets could be a good solution to assess whether a CNN-based speckle reduction technique could increase both contrast and resolution of US images.

Other physical phenomena have deterministic impacts on speckle patterns (especially second-order statistics), which could thus benefit from CNN-based approaches. In particular, local fluctuations of mean sound speed with respect to the “global” mean sound speed assumed for DAS beamforming results in visible alterations of speckle patterns, such as shifted and defocused patterns. This is typically the case when an aberration layer is present between the tissue of interest and the US transducer, such as a fatty layer or a skull bone (in zones and at frequencies for which the porous layer does not induce detrimental scattering of the acoustic waves). Also, some tissues have a mean sound

speed that deviate sufficiently from the one in soft tissue such that images reconstructed from insonifications with different angles result in shifted speckle patterns. This effect is currently exploited by speed-of-sound imaging techniques, which are also based on the correlation of speckle patterns. I believe that these approaches could highly benefit from CNN-based approaches, in particular with simulated datasets. The main (temporary) problem I see is that the simulation of such complex physical phenomena would require even more computational resources.

A. Supplementary Material to Chapter 4

This appendix presents all additional results to Section 4.3.2.

Spherically Focused Element with a Rigid Baffle Condition

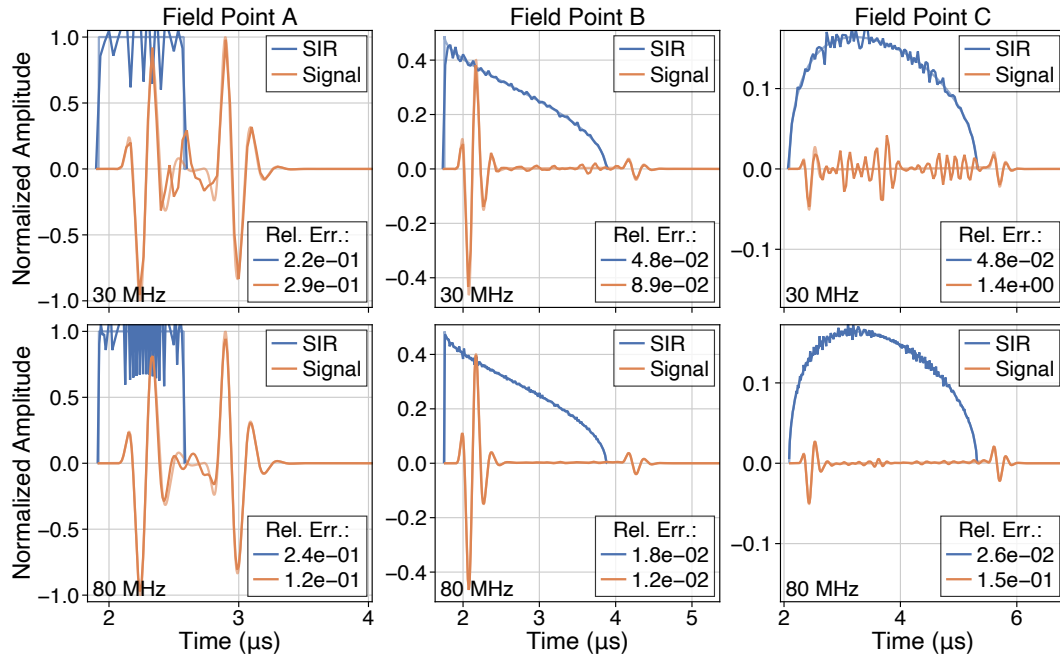


Figure A.1. Comparison of the spatial impulse responses (SIRs) and field signals radiated at different field points by a spherically focused transducer element with a rigid baffle condition, excited by a windowed-sinusoidal waveform. The excitation waveform is a differentiated log-normal-windowed sine wave, with a characteristic (center) wavelength λ . The geometry of the spherical cap is defined by an active diameter of 20λ and a spherical radius of 240λ . The three field points (A, B, C) lie in the same revolution plane at a depth of 10λ and a lateral coordinate of $0, 8.1\lambda$ (i.e., projection on the edge of the surface), and 16.2λ , respectively. The proposed approach was implemented with a nearest-neighbor basis function (degree zero), and was evaluated at two sampling rates of (first row) 30 MHz and (second row) 80 MHz. The reference SIRs and field signals were evaluated at a sampling rate of 20 THz. They are depicted with the same colors as the approximated counterparts, with a lower opacity.

A. SUPPLEMENTARY MATERIAL TO CHAPTER 4

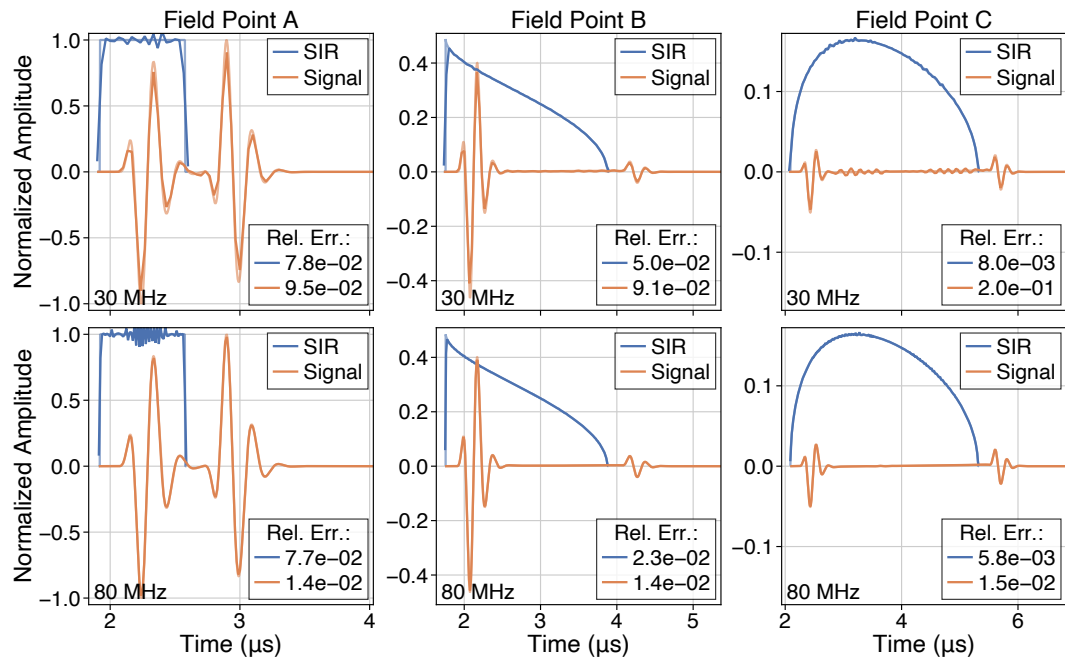


Figure A.2. *Ibid.* for a linear basis function (degree one).

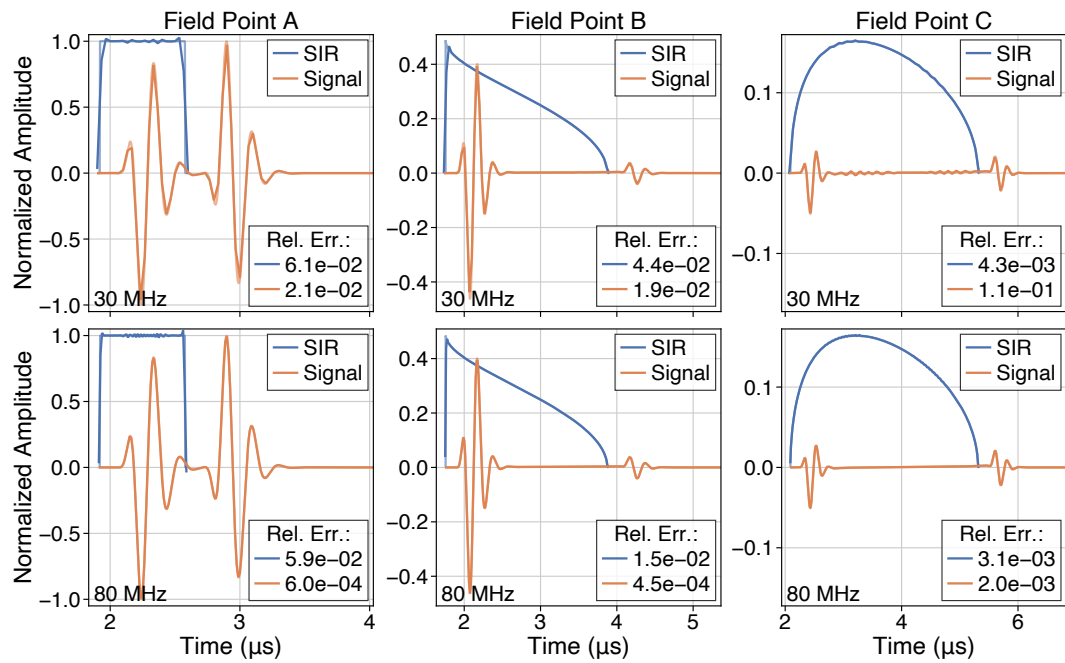


Figure A.3. *Ibid.* for a quadratic Keys basis function.

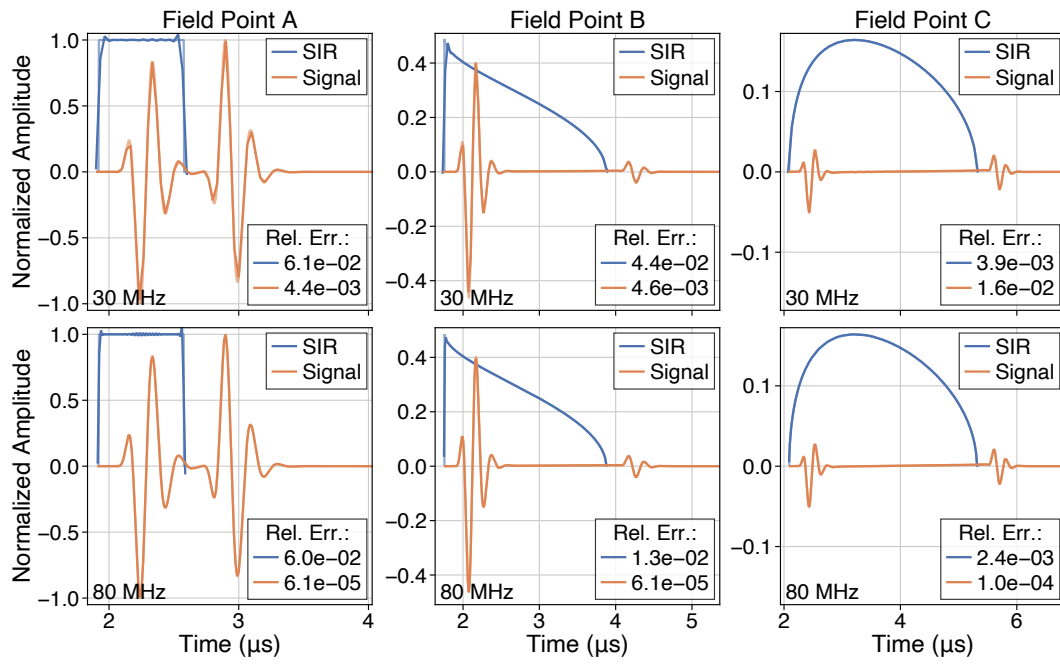


Figure A.4. *Ibid.* for a cubic B-spline basis function.

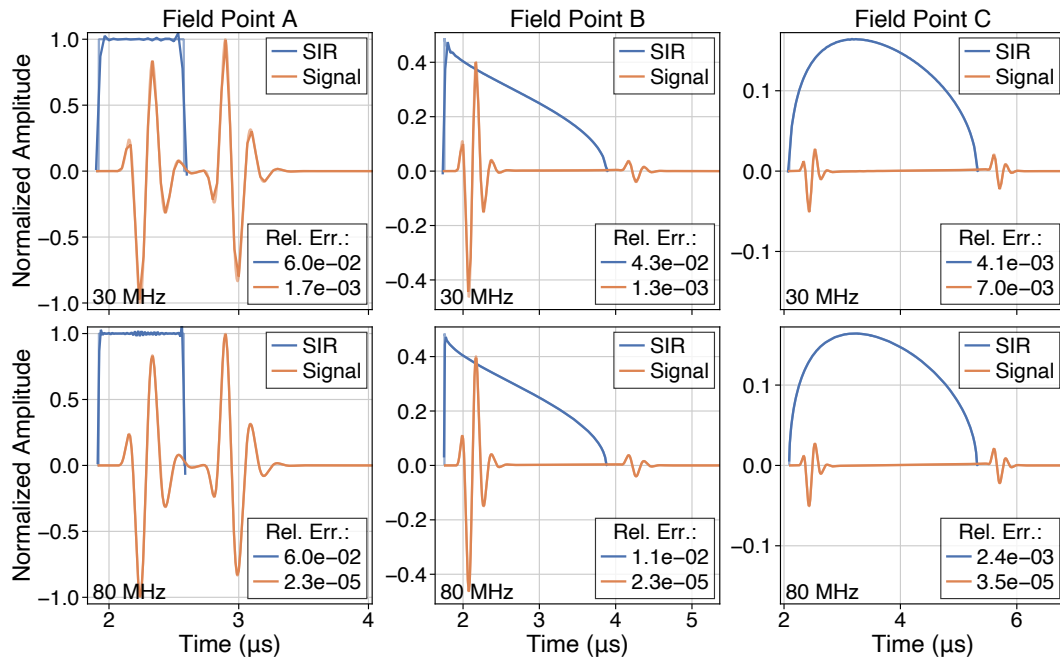


Figure A.5. *Ibid.* for a cubic O-MOMS basis function.

Rectangular Element with Rigid and Soft Baffle Conditions

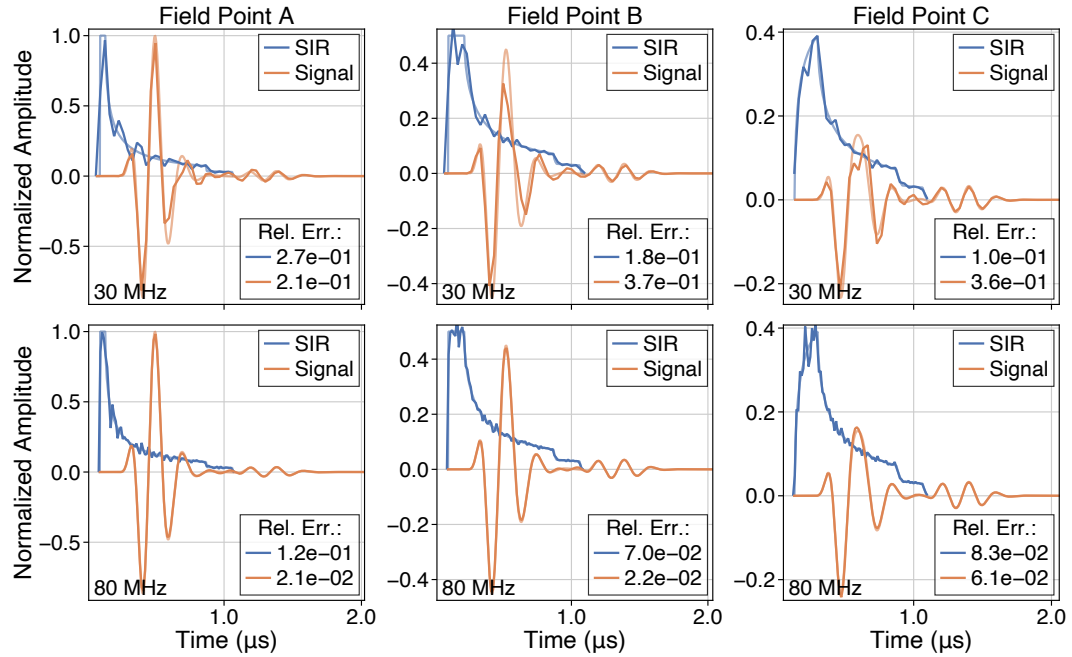


Figure A.6. Comparison of the spatial impulse responses (SIRs) and field signals radiated at different field points by a rectangular transducer element with a rigid baffle condition, excited by a windowed-sinusoidal waveform. The excitation waveform is a differentiated log-normal-windowed sine wave, with a characteristic (center) wavelength λ . The geometry of the rectangular element is defined by a width of λ and a height of 10λ . The three field points (A, B, C) lie in a plane parallel to the element. They were positioned at a depth of $\lambda/2$ and an elevation of $\lambda/2$, with lateral coordinates of 0, $\lambda/2$, and λ , respectively. The proposed approach was implemented with a nearest-neighbor basis function (degree zero), and was evaluated at two sampling rates of (first row) 30 MHz and (second row) 80 MHz. The reference SIRs and field signals were evaluated at a sampling rate of 20 THz. They are depicted with the same colors as the approximated counterparts, with a lower opacity.

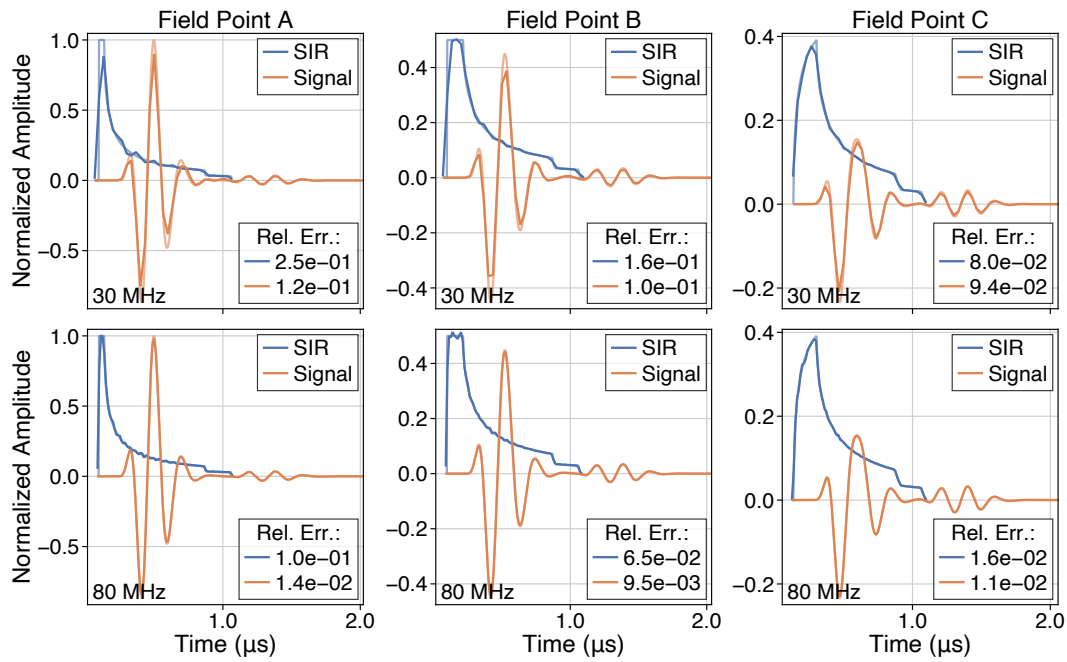


Figure A.7. *Ibid.* for a linear basis function (degree one).

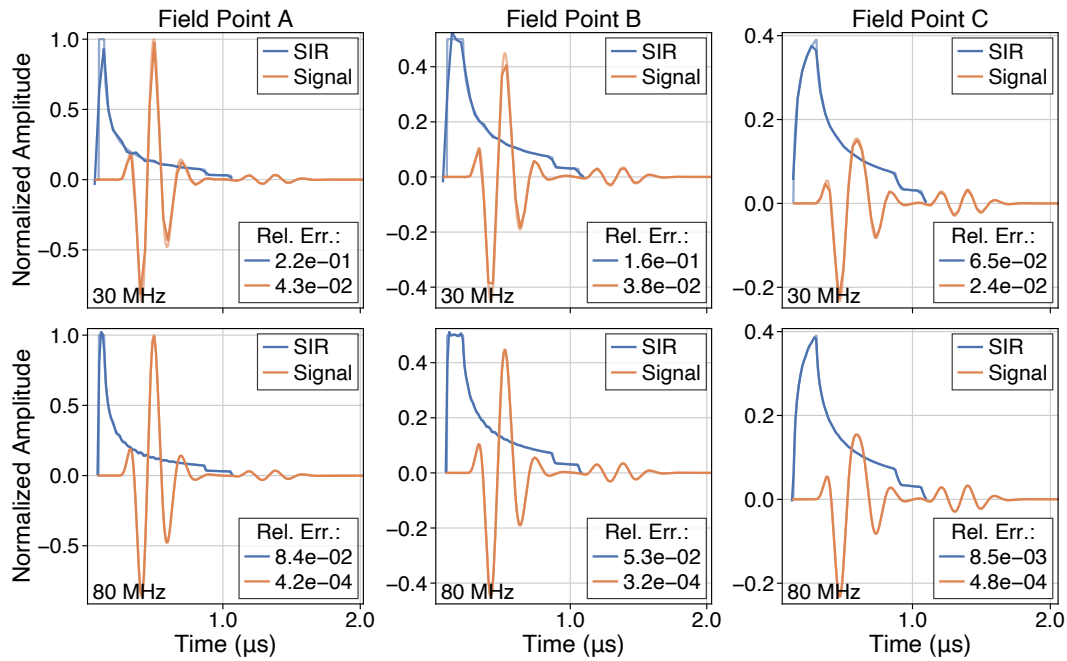


Figure A.8. *Ibid.* for a quadratic Keys basis function.

A. SUPPLEMENTARY MATERIAL TO CHAPTER 4

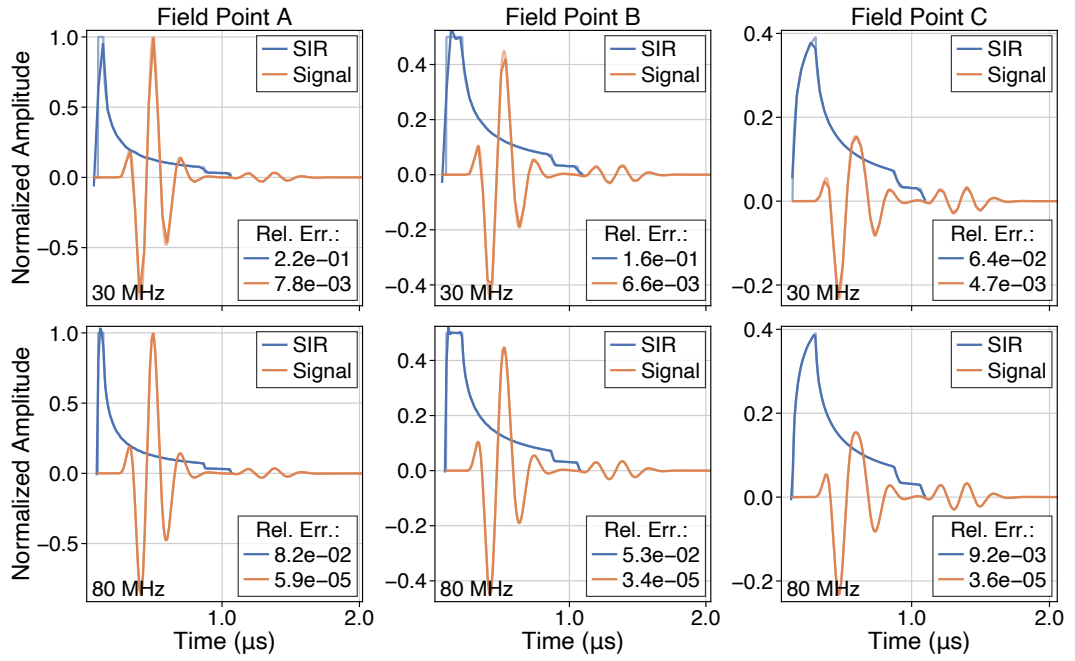


Figure A.9. *Ibid.* for a cubic B-spline basis function.

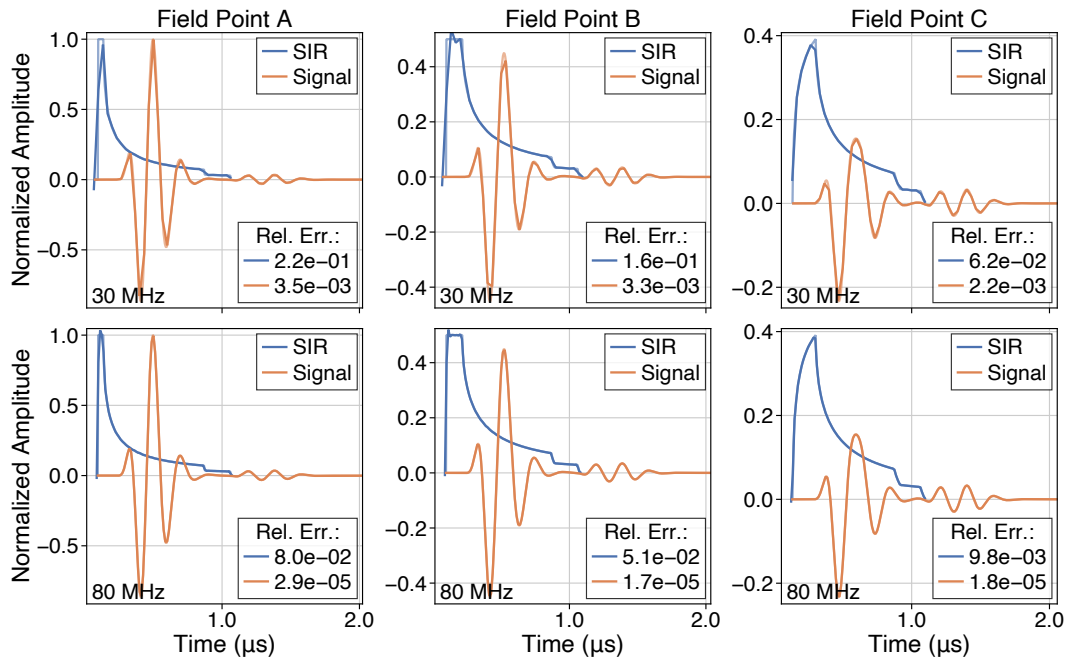


Figure A.10. *Ibid.* for a cubic O-MOMS basis function.

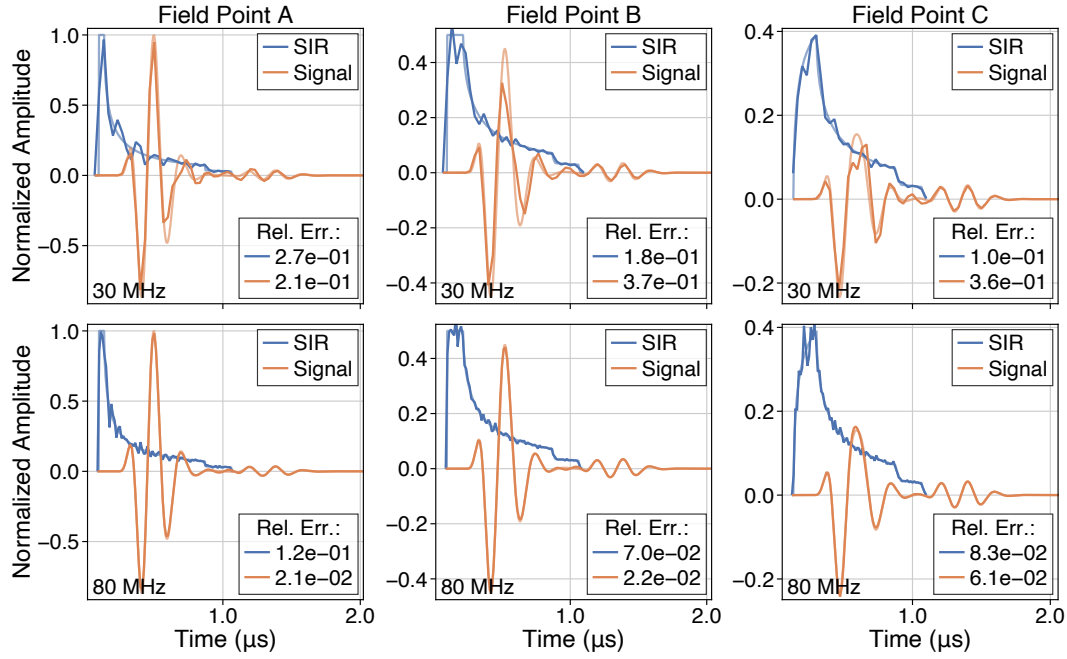


Figure A.11. Comparison of the spatial impulse responses (SIRs) and field signals radiated at different field points by a rectangular transducer element with a soft baffle condition, excited by a windowed-sinusoidal waveform. The remaining settings relevant to the results depicted are identical to those detailed in Figure A.6. The proposed approach was implemented with a nearest-neighbor basis function (degree zero), and was evaluated at two sampling rates of (first row) 30 MHz and (second row) 80 MHz. The reference SIRs and field signals were evaluated at a sampling rate of 20 THz. They are depicted with the same colors as the approximated counterparts, with a lower opacity.

A. SUPPLEMENTARY MATERIAL TO CHAPTER 4

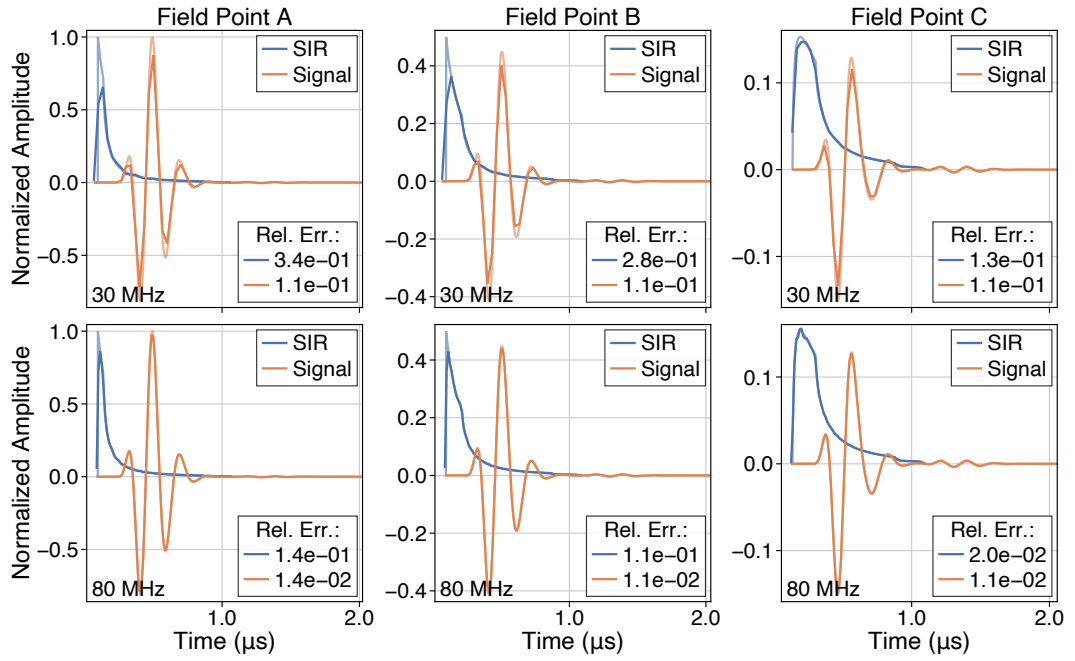


Figure A.12. *Ibid.* for a linear basis function (degree one).

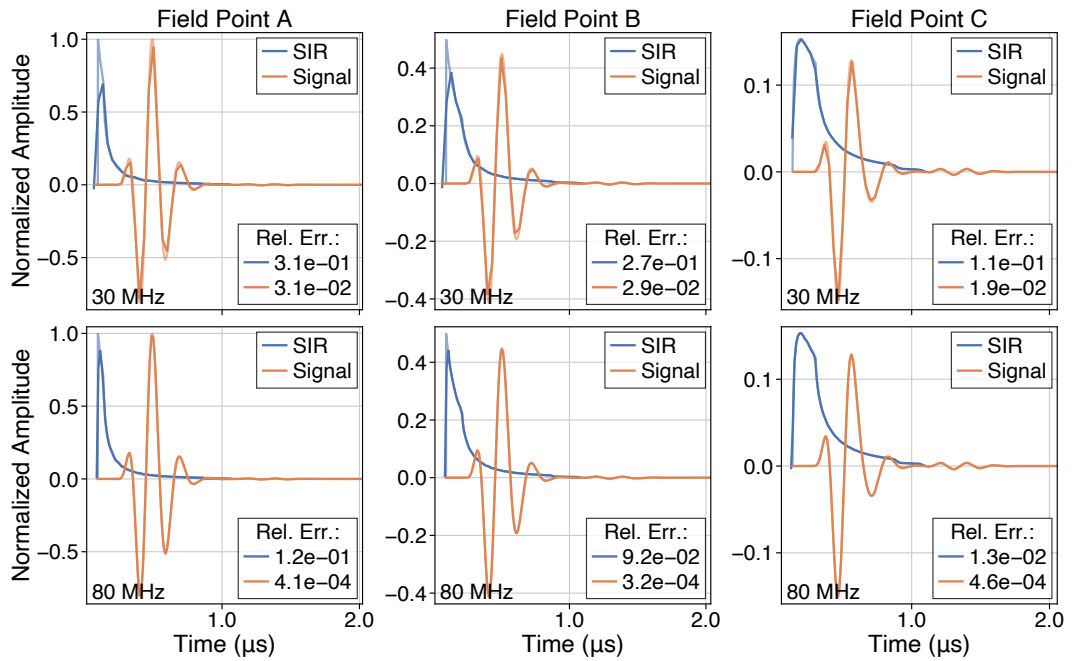


Figure A.13. *Ibid.* for a quadratic Keys basis function.

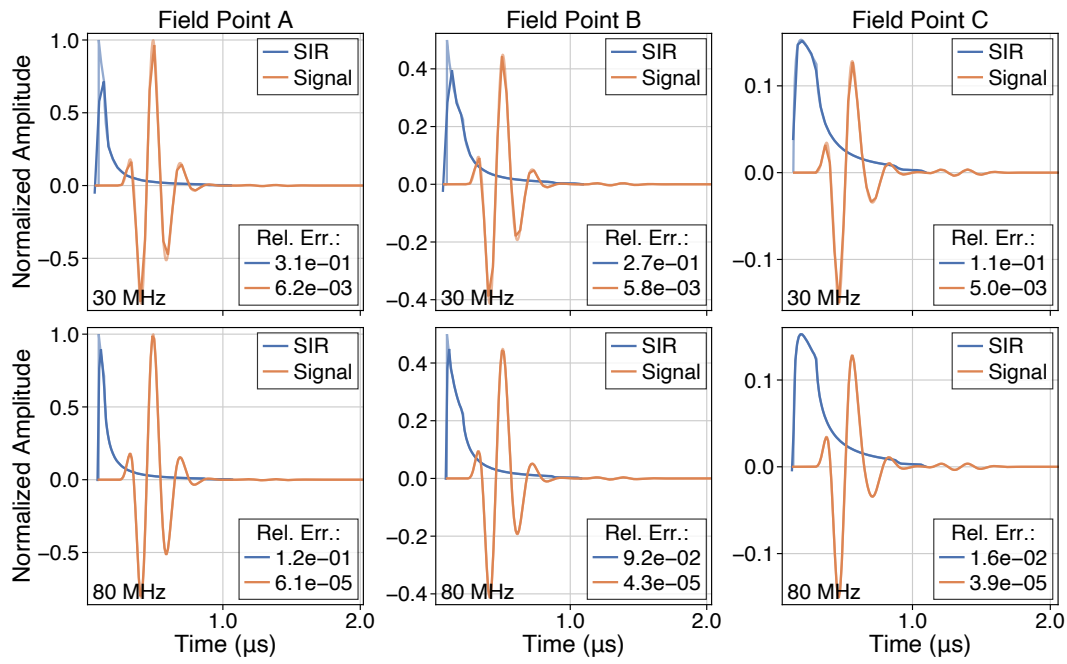


Figure A.14. *Ibid.* for a cubic B-spline basis function.

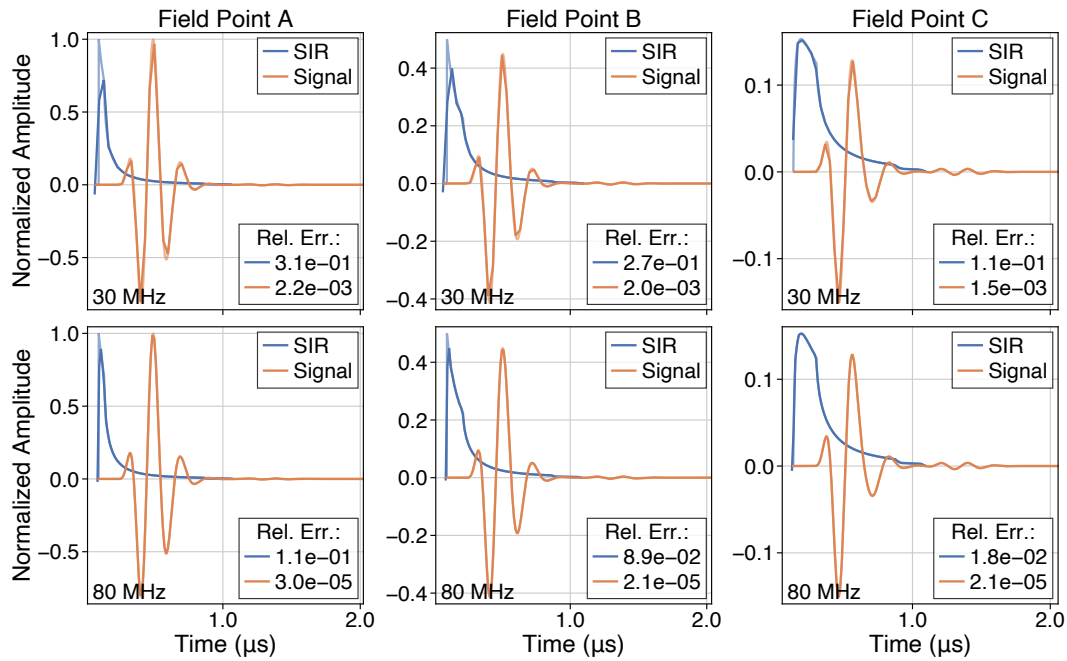


Figure A.15. *Ibid.* for a cubic O-MOMS basis function.

References

- [1] M. Tanter and M. Fink, "Ultrafast imaging in biomedical ultrasound," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 61, no. 1, pp. 102–119, Jan. 2014. DOI: 10.1109/TUFFC.2014.6689779.
- [2] C. Bruneel, R. Torguet, K. M. Rouvaen, E. Bridoux, and B. Nongaillard, "Ultrafast echotomographic system using optical processing of ultrasonic signals," *Applied Physics Letters*, vol. 30, no. 8, pp. 371–373, Apr. 1977. DOI: 10.1063/1.89436.
- [3] B. Delannoy, R. Torguet, C. Bruneel, E. Bridoux, J. M. Rouvaen, and H. Lasota, "Acoustical image reconstruction in parallel-processing analog electronic systems," *Journal of Applied Physics*, vol. 50, no. 5, pp. 3153–3159, May 1979. DOI: 10.1063/1.326397.
- [4] D. P. Shattuck, M. D. Weinshenker, S. W. Smith, and O. T. von Ramm, "Explososcan: A parallel processing technique for high speed ultrasound imaging with linear phased arrays," *The Journal of the Acoustical Society of America*, vol. 75, no. 4, pp. 1273–1282, Apr. 1984. DOI: 10.1121/1.390734.
- [5] J. Bercoff, M. Tanter, and M. Fink, "Supersonic shear imaging: A new technique for soft tissue elasticity mapping," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 51, no. 4, pp. 396–409, Apr. 2004. DOI: 10.1109/TUFFC.2004.1295425.
- [6] G. Montaldo, M. Tanter, J. Bercoff, N. Benech, and M. Fink, "Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 56, no. 3, pp. 489–506, Mar. 2009. DOI: 10.1109/TUFFC.2009.1067.
- [7] P. Santos, A. M. Petrescu, J. Pedrosa, M. Orlowska, V. Komini, J.-U. Voigt, *et al.*, "Natural shear wave imaging in the human heart: Normal values, feasibility, and reproducibility," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 66, no. 3, pp. 442–452, Mar. 2019. DOI: 10.1109/TUFFC.2018.2881493.
- [8] C. Papadacci, M. Pernot, M. Couade, M. Fink, and M. Tanter, "High-contrast ultrafast imaging of the heart," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 61, no. 2, pp. 288–301, Feb. 2014. DOI: 10.1109/TUFFC.2014.6722614.
- [9] M. Pernot, M. Couade, P. Mateo, B. Crozatier, R. Fischmeister, and M. Tanter, "Real-time assessment of myocardial contractility using shear wave imaging,"

REFERENCES

- Journal of the American College of Cardiology*, vol. 58, no. 1, pp. 65–72, Jun. 2011. DOI: 10.1016/j.jacc.2011.02.042.
- [10] H. Geyer, G. Caracciolo, H. Abe, S. Wilansky, S. Carerj, F. Gentile, *et al.*, “Assessment of myocardial mechanics using speckle tracking echocardiography: Fundamentals and clinical applications,” *Journal of the American Society of Echocardiography*, vol. 23, no. 4, pp. 351–369, Apr. 2010. DOI: 10.1016/j.echo.2010.02.015.
- [11] M. Cikes, L. Tong, G. R. Sutherland, and J. D’hooge, “Ultrafast cardiac ultrasound imaging,” *JACC: Cardiovascular Imaging*, vol. 7, no. 8, pp. 812–823, Aug. 2014. DOI: 10.1016/j.jcmg.2014.06.004.
- [12] J.-U. Voigt, G. Pedrizzetti, P. Lysyansky, T. H. Marwick, H. Houle, R. Baumann, *et al.*, “Definitions for a common standard for 2D speckle tracking echocardiography: Consensus document of the EACVI/ASE/Industry task force to standardize deformation imaging,” *European Heart Journal - Cardiovascular Imaging*, vol. 16, no. 1, pp. 1–11, Jan. 2015. DOI: 10.1093/ehjci/jeu184.
- [13] J. Bercoff, G. Montaldo, T. Loupas, D. Savary, F. Mézière, M. Fink, *et al.*, “Ultrafast compound doppler imaging: Providing full blood flow characterization,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 58, no. 1, pp. 134–147, Jan. 2011. DOI: 10.1109/TUFFC.2011.1780.
- [14] J. Udesen, F. Gran, K. Hansen, J. A. Jensen, C. Thomsen, and M. B. Nielsen, “High frame-rate blood vector velocity imaging using plane waves: Simulations and preliminary experiments,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 55, no. 8, pp. 1729–1743, Aug. 2008. DOI: 10.1109/TUFFC.2008.858.
- [15] E. Macé, G. Montaldo, I. Cohen, M. Baulac, M. Fink, and M. Tanter, “Functional ultrasound imaging of the brain,” *Nature Methods*, vol. 8, no. 8, pp. 662–664, Aug. 2011. DOI: 10.1038/nmeth.1641.
- [16] J. Cheng and J.-y. Lu, “Extended high-frame rate imaging method with limited-diffraction beams,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 53, no. 5, pp. 880–899, May 2006. DOI: 10.1109/TUFFC.2006.1632680.
- [17] B. Denarie, T. A. Tangen, I. K. Ekroll, N. Rolim, H. Torp, T. Bjastad, *et al.*, “Coherent plane wave compounding for very high frame rate ultrasonography of rapidly moving targets,” *IEEE Transactions on Medical Imaging*, vol. 32, no. 7, pp. 1265–1276, Jul. 2013. DOI: 10.1109/TMI.2013.2255310.
- [18] J. Jensen, M. B. Stuart, and J. A. Jensen, “Optimized plane wave imaging for fast and high-quality ultrasound imaging,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 63, no. 11, pp. 1922–1934, Nov. 2016. DOI: 10.1109/TUFFC.2016.2591980.
- [19] J. A. Jensen, S. I. Nikolov, K. L. Gammelmark, and M. H. Pedersen, “Synthetic aperture ultrasound imaging,” *Ultrasonics*, vol. 44, e5–e15, Dec. 2006. DOI: 10.1016/j.ultras.2006.07.017.
- [20] J. Porée, D. Posada, A. Hodzic, F. Tournoux, G. Cloutier, and D. Garcia, “High-frame-rate echocardiography using coherent compounding with doppler-based

- motion-compensation,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 7, pp. 1647–1657, Jul. 2016. DOI: 10.1109/TMI.2016.2523346.
- [21] J. Porée, D. Garcia, B. Chayer, J. Ohayon, and G. Cloutier, “Noninvasive vascular elastography with plane strain incompressibility assumption using ultrafast coherent compound plane wave imaging,” *IEEE Transactions on Medical Imaging*, vol. 34, no. 12, pp. 2618–2631, Dec. 2015. DOI: 10.1109/TMI.2015.2450992.
- [22] P. Joos, J. Poree, H. Liebgott, D. Vray, M. Baudet, J. Faurie, *et al.*, “High-frame-rate speckle-tracking echocardiography,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 65, no. 5, pp. 720–728, May 2018. DOI: 10.1109/TUFFC.2018.2809553.
- [23] H. Liebgott, A. Rodriguez-Molares, F. Cervenansky, J. A. Jensen, and O. Bernard, “Plane-wave imaging challenge in medical ultrasound,” in *2016 IEEE International Ultrasonics Symposium (IUS)*, Sep. 2016, pp. 1–4. DOI: 10.1109/ULTSYM.2016.7728908.
- [24] B. Byram, K. Dei, J. Tierney, and D. Dumont, “A model and regularization scheme for ultrasonic beamforming clutter reduction,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 62, no. 11, pp. 1913–1927, Nov. 2015. DOI: 10.1109/TUFFC.2015.007004.
- [25] A. Besson, D. Perdios, F. Martinez, Z. Chen, R. E. Carrillo, M. Arditi, *et al.*, “Ultrafast ultrasound imaging as an inverse problem: Matrix-free sparse image reconstruction,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 65, no. 3, pp. 339–355, Mar. 2018. DOI: 10.1109/TUFFC.2017.2768583.
- [26] E. Ozkan, V. Vishnevsky, and O. Goksel, “Inverse problem of ultrasound beamforming with sparsity constraints and regularization,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 65, no. 3, pp. 356–365, Mar. 2018. DOI: 10.1109/TUFFC.2017.2757880.
- [27] J. A. Jensen, “A model for the propagation and scattering of ultrasound in tissue,” *The Journal of the Acoustical Society of America*, vol. 89, no. 1, pp. 182–190, Jan. 1991. DOI: 10.1121/1.400497.
- [28] A. Besson, D. Perdios, F. Martinez, M. Arditi, Y. Wiaux, and J.-P. Thiran, “Ussr: An ultrasound sparse regularization framework,” in *2017 IEEE International Ultrasonics Symposium (IUS)*, Sep. 2017, pp. 1–4. DOI: 10.1109/ULTSYM.2017.8092360.
- [29] A. Besson, L. Roquette, D. Perdios, M. Simeoni, M. Arditi, P. Hurley, *et al.*, “A physical model of nonstationary blur in ultrasound imaging,” *IEEE Transactions on Computational Imaging*, vol. 5, no. 3, pp. 381–394, Sep. 2019. DOI: 10.1109/TCI.2019.2897951.
- [30] D. Perdios, A. Besson, F. Martinez, M. Vonlanthen, M. Arditi, and J.-P. Thiran, “On problem formulation, efficient modeling and deep neural networks for high-quality ultrasound imaging: Invited presentation,” in *2019 53rd Annual Conference on Information Sciences and Systems (CISS)*, Mar. 2019, pp. 1–4. DOI: 10.1109/CISS.2019.8692870.

REFERENCES

- [31] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, “Deep convolutional neural network for inverse problems in imaging,” *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4509–4522, Sep. 2017. DOI: 10.1109/TIP.2017.2713099.
- [32] H. Gupta, K. H. Jin, H. Q. Nguyen, M. T. McCann, and M. Unser, “CNN-based projected gradient descent for consistent CT image reconstruction,” *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1440–1453, Jun. 2018. DOI: 10.1109/TMI.2018.2832656.
- [33] D. Perdios, M. Vonlanthen, F. Martinez, M. Arditi, and J.-P. Thiran, “CNN-based image reconstruction method for ultrafast ultrasound imaging,” Aug. 2020. arXiv: 2008.12750. [Online]. Available: <https://arxiv.org/abs/2008.12750>.
- [34] —, “CNN-based ultrasound image reconstruction for ultrafast displacement tracking,” Sep. 2020. arXiv: 2009.01816. [Online]. Available: <https://arxiv.org/abs/2009.01816>.
- [35] —, “Single-shot CNN-based ultrasound imaging with sparse linear arrays,” in *2020 IEEE International Ultrasonics Symposium (IUS)*, Sep. 2020, pp. 1–4. DOI: 10.1109/IUS46767.2020.9251442.
- [36] G. E. Tupholme, “Generation of acoustic pulses by baffled plane pistons,” *Mathematika*, vol. 16, no. 2, pp. 209–224, Dec. 1969. DOI: 10.1112/S0025579300008184.
- [37] P. R. Stepanishen, “Transient radiation from pistons in an infinite planar baffle,” *The Journal of the Acoustical Society of America*, vol. 49, no. 5B, pp. 1629–1638, May 1971. DOI: 10.1121/1.1912541.
- [38] P. M. Morse and H. Feshbach, *Methods of theoretical physics: Part I*, ser. International series in pure and applied physics. McGraw-Hill, 1953.
- [39] P. M. Morse and K. U. Ingard, *Theoretical Acoustics*, ser. International series in pure and applied physics. New York: McGraw-Hill, 1986.
- [40] L. A. Chernov, *Wave propagation in a random medium*. New York: McGraw-Hill, 1960.
- [41] J. C. Gore and S. Leeman, “Ultrasonic backscattering from human tissue: A realistic model,” *Physics in Medicine and Biology*, vol. 22, no. 2, p. 011, Mar. 1977. DOI: 10.1088/0031-9155/22/2/011.
- [42] R. S. C. Cobbold, *Foundations of Biomedical Ultrasound*. New York: Oxford University Press, 2007.
- [43] H. T. O’Neil, “Theory of focusing radiators,” *The Journal of the Acoustical Society of America*, vol. 21, no. 5, pp. 516–526, Sep. 1949. DOI: 10.1121/1.1906542.
- [44] J. C. Lockwood and J. G. Willette, “High-speed method for computing the exact solution for the pressure variations in the nearfield of a baffled piston,” *The Journal of the Acoustical Society of America*, vol. 53, no. 3, pp. 735–741, Mar. 1973. DOI: 10.1121/1.1913385.
- [45] G. R. Harris, “Review of transient field theory for a baffled planar piston,” *The Journal of the Acoustical Society of America*, vol. 70, no. 1, pp. 10–20, Jul. 1981. DOI: 10.1121/1.386687.
- [46] Lord Rayleigh, “XXXVII. On the passage of waves through apertures in plane screens, and allied problems,” *The London, Edinburgh, and Dublin Philosophical*

- Magazine and Journal of Science*, vol. 43, no. 263, pp. 259–272, Apr. 1897. DOI: 10.1080/14786449708620990.
- [47] A. Sommerfeld, *Optics: Lectures on Theoretical Physics, Vol. IV*. New York: Academic Press, 1954.
 - [48] J. W. S. Rayleigh, *The Theory of Sound*, Second ed. rev. and enlarged, [reprint]. New York: Dover Publications, 1945.
 - [49] B. Delannoy, H. Lasota, C. Bruneel, R. Torguet, and E. Bridoux, “The infinite planar baffles problem in acoustic radiation and its experimental verification,” *Journal of Applied Physics*, vol. 50, no. 8, pp. 5189–5195, Aug. 1979. DOI: 10.1063/1.326656.
 - [50] D. Guyomar and J. Powers, “Boundary effects on transient radiation fields from vibrating surfaces,” *The Journal of the Acoustical Society of America*, vol. 77, no. 3, pp. 907–915, Mar. 1985. DOI: 10.1121/1.392058.
 - [51] P. R. Stepanishen, “The time-dependent force and radiation impedance on a piston in a rigid infinite planar baffle,” *The Journal of the Acoustical Society of America*, vol. 49, no. 3B, pp. 841–849, Mar. 1971. DOI: 10.1121/1.1912424.
 - [52] A. Penttinen and M. Luukkala, “The impulse response and pressure nearfield of a curved ultrasonic radiator,” *Journal of Physics D: Applied Physics*, vol. 9, no. 10, pp. 1547–1557, Jul. 1976. DOI: 10.1088/0022-3727/9/10/020.
 - [53] M. Arditi, F. S. Foster, and J. W. Hunt, “Transient fields of concave annular arrays,” *Ultrasonic Imaging*, vol. 3, no. 1, pp. 37–61, Jan. 1981. DOI: 10.1016/0161-7346(81)90081-X.
 - [54] A. Freedman, “Sound field of plane or gently curved pulsed radiators,” *The Journal of the Acoustical Society of America*, vol. 48, no. 1B, pp. 221–227, Jul. 1970. DOI: 10.1121/1.1912119.
 - [55] S. Leeman, V. C. Roberts, P. E. Chandler, and L. A. Ferrari, “Inverse imaging with strong multiple scattering,” in *Mathematics and Computer Science in Medical Imaging*, ser. NATO ASI Series (Series F: Computer and Systems Sciences), M. A. Viergever and A. Todd-Pokropek, Eds., vol. 39, Berlin, Heidelberg: Springer, 1988, pp. 279–289. DOI: 10.1007/978-3-642-83306-9_12.
 - [56] F. Oberhettinger, “On transient solutions of the ‘baffled piston’ problem,” *Journal of Research of the National Bureau of Standards, Section B: Mathematics and Mathematical Physics*, vol. 65B, no. 1, pp. 1–6, Jan. 1961. DOI: 10.6028/jres.065B.001.
 - [57] D. Robinson, S. Lees, and L. Bess, “Near field transient radiation patterns for circular pistons,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 22, no. 6, pp. 395–403, Dec. 1974. DOI: 10.1109/TASSP.1974.1162612.
 - [58] T. L. Szabo, B. Ü. Karbeyaz, R. O. Cleveland, and E. L. Miller, “Determining the pulse-echo electromechanical characteristic of a transducer using flat plates and point targets,” *The Journal of the Acoustical Society of America*, vol. 116, no. 1, pp. 90–96, Jul. 2004. DOI: 10.1121/1.1756893.
 - [59] J. L. San Emeterio and L. G. Ullate, “Diffraction impulse response of rectangular transducers,” *The Journal of the Acoustical Society of America*, vol. 92, no. 2, pp. 651–662, Aug. 1992. DOI: 10.1121/1.403990.

REFERENCES

- [60] S. W. Smith, O. T. von Ramm, M. E. Haran, and F. L. Thurstone, "Angular response of piezoelectric elements in phased array ultrasound scanners," *IEEE Transactions on Sonics and Ultrasonics*, vol. 26, no. 3, pp. 185–190, May 1979. DOI: 10.1109/T-SU.1979.31085.
- [61] A. R. Selfridge, G. S. Kino, and B. T. Khuri-Yakub, "A theory for the radiation pattern of a narrow-strip acoustic transducer," *Applied Physics Letters*, vol. 37, no. 1, pp. 35–36, Jul. 1980. DOI: 10.1063/1.91692.
- [62] J. Sato, H. Fukukita, M. Kawabuchi, and A. Fukumoto, "Farfield angular radiation pattern generated from arrayed piezoelectric transducers," *The Journal of the Acoustical Society of America*, vol. 67, no. 1, pp. 333–335, Jan. 1980. DOI: 10.1121/1.384466.
- [63] P. Pesque and M. Fink, "Effect of the planar baffle impedance in acoustic radiation of a phased array element theory and experimentation," in *IEEE 1984 Ultrasonics Symposium*, Nov. 1984, pp. 1034–1038. DOI: 10.1109/ULTSYM.1984.198462.
- [64] A. Besson, "Imaging from Echoes: On Inverse Problems in Ultrasound," PhD thesis, École polytechnique fédérale de Lausanne (EPFL), Lausanne, 2019. DOI: 10.5075/epfl-thesis-9310.
- [65] M. Tanter, J. Bercoff, L. Sandrin, and M. Fink, "Ultrafast compound imaging for 2-D motion vector estimation: Application to transient elastography," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 49, no. 10, pp. 1363–1374, Oct. 2002. DOI: 10.1109/TUFFC.2002.1041078.
- [66] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, ser. Springer Optimization and Its Applications, H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, Eds., New York, NY: Springer New York, 2011, pp. 185–212. DOI: 10.1007/978-1-4419-9569-8_10.
- [67] R. E. Carrillo, J. D. McEwen, D. Van De Ville, J.-P. Thiran, and Y. Wiaux, "Sparsity averaging for compressive imaging," *IEEE Signal Processing Letters*, vol. 20, no. 6, pp. 591–594, Jun. 2013. DOI: 10.1109/LSP.2013.2259813.
- [68] P. Thévenaz, T. Blu, and M. Unser, "Interpolation revisited," *IEEE Transactions on Medical Imaging*, vol. 19, no. 7, pp. 739–758, Jul. 2000. DOI: 10.1109/42.875199.
- [69] D. H. Turnbull and F. S. Foster, "Fabrication and characterization of transducer elements in two-dimensional arrays for medical ultrasound imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 39, no. 4, pp. 464–475, Jul. 1992. DOI: 10.1109/58.148536.
- [70] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, Jan. 2009. DOI: 10.1137/080716542.
- [71] G. S. Kino, *Acoustic Waves: Devices, Imaging, and Analog Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
- [72] J. W. Hunt, M. Arditi, and F. S. Foster, "Ultrasound transducers for pulse-echo medical imaging," *IEEE Transactions on Biomedical Engineering*, vol. BME-30, no. 8, pp. 453–481, Aug. 1983. DOI: 10.1109/TBME.1983.325150.

- [73] A. Ziv, P. R. Wolpe, S. D. Small, and S. Glick, "Simulation-based medical education: An ethical imperative," *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, vol. 1, no. 4, pp. 252–256, 2006. DOI: 10.1097/01.SIH.0000242724.08501.63.
- [74] T. Blum, A. Rieger, N. Navab, H. Friess, and M. Martignoni, "A review of computer-based simulators for ultrasound training," *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, vol. 8, no. 2, pp. 98–108, Apr. 2013. DOI: 10.1097/SIH.0b013e31827ac273.
- [75] R. E. Lewiss, B. Hoffmann, Y. Beaulieu, and M. B. Phelan, "Point-of-care ultrasound education," *Journal of Ultrasound in Medicine*, vol. 33, no. 1, pp. 27–32, Jan. 2014. DOI: 10.7863/ultra.33.1.27.
- [76] G. Gunabushanam, F. Nautsch, I. Mills, and L. M. Scoutt, "Accessible personal ultrasound training simulator," *Journal of Ultrasound in Medicine*, vol. 38, no. 6, pp. 1425–1432, Jun. 2019. DOI: 10.1002/jum.14820.
- [77] S. Hani, G. Chalouhi, Z. Lakissian, and R. Sharara-Chami, "Introduction of ultrasound simulation in medical education: Exploratory study," *JMIR Medical Education*, vol. 5, no. 2, e13568, Sep. 2019. DOI: 10.2196/13568.
- [78] R. F. Wagner, S. W. Smith, J. M. Sandrik, and H. Lopez, "Statistics of speckle in ultrasound B-scans," *IEEE Transactions on Sonics and Ultrasonics*, vol. 30, no. 3, pp. 156–163, May 1983. DOI: 10.1109/T-SU.1983.31404.
- [79] D. R. Foster, M. Arditi, F. S. Foster, M. S. Patterson, and J. W. Hunt, "Computer simulations of speckle in B-scan images," *Ultrasonic Imaging*, vol. 5, no. 4, pp. 308–330, Oct. 1983. DOI: 10.1177/016173468300500403.
- [80] O. M. H. Rindal, A. Austeng, A. Fatemi, and A. Rodriguez-Molares, "The effect of dynamic range alterations in the estimation of contrast," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 66, no. 7, pp. 1198–1208, Jul. 2019. DOI: 10.1109/TUFFC.2019.2911267.
- [81] A. Rodriguez-Molares, O. M. H. Rindal, J. D'hooge, S.-E. Masoy, A. Austeng, M. A. Lediju Bell, *et al.*, "The generalized contrast-to-noise ratio: A formal definition for lesion detectability," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 67, no. 4, pp. 745–759, Apr. 2020. DOI: 10.1109/TUFFC.2019.2956855.
- [82] H. Greenspan, B. van Ginneken, and R. M. Summers, "Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1153–1159, May 2016. DOI: 10.1109/TMI.2016.2553401.
- [83] G. Wang, "A perspective on deep imaging," *IEEE Access*, vol. 4, pp. 8914–8924, 2016. DOI: 10.1109/ACCESS.2016.2624938.
- [84] G. Wang, J. C. Ye, K. Mueller, and J. A. Fessler, "Image reconstruction is a new frontier of machine learning," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1289–1296, Jun. 2018. DOI: 10.1109/TMI.2018.2833635.

REFERENCES

- [85] R. J. G. van Sloun, R. Cohen, and Y. C. Eldar, "Deep learning in ultrasound imaging," *Proceedings of the IEEE*, vol. 108, no. 1, pp. 11–29, Jan. 2020. DOI: 10.1109/JPROC.2019.2932116.
- [86] A. F. Frangi, S. A. Tsaftaris, and J. L. Prince, "Simulation and synthesis in medical imaging," *IEEE Transactions on Medical Imaging*, vol. 37, no. 3, pp. 673–679, Mar. 2018. DOI: 10.1109/TMI.2018.2800298.
- [87] J. A. Jensen, "FIELD: A program for simulating ultrasound systems," in *10th Nordic-Baltic Conference on Biomedical Imaging*, vol. 4, 1996, pp. 351–353.
- [88] J. A. Jensen and N. B. Svendsen, "Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 39, no. 2, pp. 262–267, Mar. 1992. DOI: 10.1109/58.139123.
- [89] H. Lasota and R. Salamon, "Application of time-space impulse responses to calculations of acoustic fields in imaging systems," in *Acoustical Imaging: Proceedings of the International Symposium*, P. Alais and A. F. Metherell, Eds., vol. 10, Springer, Boston, MA, 1982, pp. 493–512. DOI: 10.1007/978-1-4684-3944-1_32.
- [90] H. Lasota, R. Salamon, and B. Delannoy, "Acoustic diffraction analysis by the impulse response method: A line impulse response approach," *The Journal of the Acoustical Society of America*, vol. 76, no. 1, pp. 280–290, Jul. 1984. DOI: 10.1121/1.391115.
- [91] J. A. Jensen, "Ultrasound fields from triangular apertures," *The Journal of the Acoustical Society of America*, vol. 100, no. 4, pp. 2049–2056, Oct. 1996. DOI: 10.1121/1.417914.
- [92] A. Freedman, "Farfield of pulsed rectangular acoustic radiator," *The Journal of the Acoustical Society of America*, vol. 49, no. 3B, pp. 738–748, Mar. 1971. DOI: 10.1121/1.1912411.
- [93] —, "Reply to 'Comments on 'Farfield of Pulsed Rectangular Acoustic Radiator' [P. R. Stepanishen, J. Acoust. Soc. Amer. 52 , 434–436(L) (1972)]," *The Journal of the Acoustical Society of America*, vol. 52, no. 1B, pp. 437–438, Jul. 1972. DOI: 10.1121/1.1913110.
- [94] P. R. Stepanishen, "Comments on 'Farfield of Pulsed Rectangular Acoustic Radiator' [A. Freedman, J. Acoust. Soc. Amer. 49 , 738–748 (1971)]," *The Journal of the Acoustical Society of America*, vol. 52, no. 1B, pp. 434–436, Jul. 1972. DOI: 10.1121/1.1913109.
- [95] J. C. Lockwood and J. G. Willette, "Erratum: 'High-speed method for computing the exact solution for the pressure variations in the nearfield of a baffled piston' [J. C. Lockwood and J. G. Willette, J. Acoust. Soc. Am. 53 , 735–74 (1973)]," *The Journal of the Acoustical Society of America*, vol. 54, no. 6, pp. 1762–1762, Dec. 1973. DOI: 10.1121/1.1914483.
- [96] G. Scarano, N. Denisenko, M. Matteucci, and M. Pappalardo, "A new approach to the derivation of the impulse response of a rectangular piston," *The Journal of the Acoustical Society of America*, vol. 78, no. 3, pp. 1109–1113, Sep. 1985. DOI: 10.1121/1.393030.

- [97] F. S. Foster, M. Arditi, and J. W. Hunt, "Cylindrical transducer scatter scanner," *The Journal of the Acoustical Society of America*, vol. 68, no. 1, pp. 85–92, Jul. 1980. DOI: 10.1121/1.384508.
- [98] J.-F. Theumann, M. Arditi, J.-J. Meister, and E. Jaques, "Acoustic fields of concave cylindrical transducers," *The Journal of the Acoustical Society of America*, vol. 88, no. 2, pp. 1160–1169, Aug. 1990. DOI: 10.1121/1.399812.
- [99] D. B. Bæk, J. A. Jensen, and M. Willatzen, "Spatial impulse response of a rectangular double curved transducer," *The Journal of the Acoustical Society of America*, vol. 131, no. 4, pp. 2730–2741, Apr. 2012. DOI: 10.1121/1.3693659.
- [100] G. R. Harris, "Transient field of a baffled planar piston having an arbitrary vibration amplitude distribution," *The Journal of the Acoustical Society of America*, vol. 70, no. 1, pp. 186–204, Jul. 1981. DOI: 10.1121/1.386671.
- [101] P. R. Stepanishen, "Acoustic transients from planar axisymmetric vibrators using the impulse response approach," *The Journal of the Acoustical Society of America*, vol. 70, no. 4, pp. 1176–1181, Oct. 1981. DOI: 10.1121/1.386949.
- [102] J. N. Tjøtta and S. Tjøtta, "Nearfield and farfield of pulsed acoustic radiators," *The Journal of the Acoustical Society of America*, vol. 71, no. 4, pp. 824–834, Apr. 1982. DOI: 10.1121/1.387582.
- [103] W. A. Verhoef, M. J. T. M. Cloostermans, and J. M. Thijssen, "The impulse response of a focused source with an arbitrary axisymmetric surface velocity distribution," *The Journal of the Acoustical Society of America*, vol. 75, no. 6, pp. 1716–1721, Jun. 1984. DOI: 10.1121/1.390973.
- [104] J. A. Jensen, "Speed-accuracy trade-offs in computing spatial impulse responses for simulating medical ultrasound imaging," *Journal of Computational Acoustics*, vol. 09, no. 03, pp. 731–744, Sep. 2001. DOI: 10.1142/S0218396X01001248.
- [105] B. Piwakowski and B. Delannoy, "Method for computing spatial pulse response: Time-domain approach," *The Journal of the Acoustical Society of America*, vol. 86, no. 6, pp. 2422–2432, Dec. 1989. DOI: 10.1121/1.398449.
- [106] B. Piwakowski and K. Sbai, "A new approach to calculate the field radiated from arbitrarily structured transducer arrays," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 46, no. 2, pp. 422–440, Mar. 1999. DOI: 10.1109/58.753032.
- [107] J. D'hooge, J. Nuyts, B. Bijmens, B. De Man, P. Suetens, J. Thoen, *et al.*, "The calculation of the transient near and far field of a baffled piston using low sampling frequencies," *The Journal of the Acoustical Society of America*, vol. 102, no. 1, pp. 78–86, Jul. 1997. DOI: 10.1121/1.419715.
- [108] L. Piegl and W. Tiller, *The NURBS Book*, 2nd ed., ser. Monographs in Visual Communication. Berlin, Heidelberg: Springer, 1997. DOI: 10.1007/978-3-642-59223-2.
- [109] M. Abramowitz and I. A. Stegun, Eds., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, ser. Applied Mathematics Series 55. Washington, D.C.: US Government Printing Office, 1964, Tenth Printing, December 1972, with corrections.

REFERENCES

- [110] M. Unser, A. Aldroubi, and M. Eden, “B-spline signal processing: Part I—Theory,” *IEEE Transactions on Signal Processing*, vol. 41, no. 2, pp. 821–833, Feb. 1993. DOI: 10.1109/78.193220.
- [111] ———, “B-spline signal processing: Part II—Efficiency design and applications,” *IEEE Transactions on Signal Processing*, vol. 41, no. 2, pp. 834–848, Feb. 1993. DOI: 10.1109/78.193221.
- [112] L. Aguilar, R. S. C. Cobbold, and D. Steinman, “Fast and mechanistic ultrasound simulation using a point source/receiver approach,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 60, no. 11, pp. 2335–2346, Nov. 2013. DOI: 10.1109/TUFFC.2013.6644737.
- [113] L. Aguilar, J. Wong, D. A. Steinman, and R. S. C. Cobbold, “FAMUS II: A fast and mechanistic ultrasound simulator using an impulse response approach,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 64, no. 2, pp. 362–373, Feb. 2017. DOI: 10.1109/TUFFC.2016.2632706.
- [114] M. G. Cox, “The numerical evaluation of B-splines,” *IMA Journal of Applied Mathematics*, vol. 10, no. 2, pp. 134–149, Oct. 1972. DOI: 10.1093/imat/10.2.134.
- [115] C. de Boor, “On calculating with B-splines,” *Journal of Approximation Theory*, vol. 6, no. 1, pp. 50–62, Jul. 1972. DOI: 10.1016/0021-9045(72)90080-9.
- [116] ———, *A Practical Guide to Splines*, 4th ed., ser. Applied Mathematical Sciences. New York: Springer-Verlag, 1987.
- [117] M. Unser, A. Aldroubi, and M. Eden, “Fast B-spline transforms for continuous image representation and interpolation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 3, pp. 277–285, Mar. 1991. DOI: 10.1109/34.75515.
- [118] M. A. Unser, “Ten good reasons for using spline wavelets,” in *Wavelet Applications in Signal and Image Processing V*, A. Aldroubi, A. F. Laine, and M. A. Unser, Eds., vol. 3169, SPIE, Oct. 1997, pp. 422–431. DOI: 10.1117/12.292801.
- [119] M. Unser, “Splines: A perfect fit for signal and image processing,” *IEEE Signal Processing Magazine*, vol. 16, no. 6, pp. 22–38, 1999. DOI: 10.1109/79.799930.
- [120] I. J. Schoenberg, “Contributions to the problem of approximation of equidistant data by analytic functions. Part A. On the problem of smoothing or graduation. A first class of analytic approximation formulae,” *Quarterly of Applied Mathematics*, vol. 4, no. 1, pp. 45–99, Apr. 1946. DOI: 10.1090/qam/15914.
- [121] ———, “Contributions to the problem of approximation of equidistant data by analytic functions. Part B. On the problem of osculatory interpolation. A second class of analytic approximation formulae,” *Quarterly of Applied Mathematics*, vol. 4, no. 2, pp. 112–141, Jul. 1946. DOI: 10.1090/qam/16705.
- [122] R. Keys, “Cubic convolution interpolation for digital image processing,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 6, pp. 1153–1160, Dec. 1981. DOI: 10.1109/TASSP.1981.1163711.
- [123] A. Ron, “Factorization theorems for univariate splines on regular grids,” *Israel Journal of Mathematics*, vol. 70, no. 1, pp. 48–68, Feb. 1990. DOI: 10.1007/BF02807218.

- [124] T. Blu, P. Thévenaz, and M. Unser, “MOMS: Maximal-order interpolation of minimal support,” *IEEE Transactions on Image Processing*, vol. 10, no. 7, pp. 1069–1080, Jul. 2001. DOI: 10.1109/83.931101.
- [125] F. S. Foster, M. Arditi, M. S. Patterson, and J. W. Hunt, “The conical scanner: A two transducer ultrasound scatter imaging technique,” *Ultrasonic Imaging*, vol. 3, no. 1, pp. 62–82, Jan. 1981. DOI: 10.1016/0161-7346(81)90082-1.
- [126] M. S. Patterson and F. S. Foster, “Acoustic fields of conical radiators,” *IEEE Transactions on Sonics and Ultrasonics*, vol. 29, no. 2, pp. 83–91, Mar. 1982. DOI: 10.1109/T-SU.1982.31312.
- [127] M. T. McCann, K. H. Jin, and M. Unser, “Convolutional neural networks for inverse problems in imaging: A review,” *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 85–95, Nov. 2017. DOI: 10.1109/MSP.2017.2739299.
- [128] A. Lucas, M. Iliadis, R. Molina, and A. K. Katsaggelos, “Using deep neural networks for inverse problems in imaging: Beyond analytical methods,” *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 20–36, Jan. 2018. DOI: 10.1109/MSP.2017.2760358.
- [129] M. Nikoonahad and D. C. Liu, “Medical ultrasound imaging using neural networks,” *Electronics Letters*, vol. 26, no. 8, p. 545, Apr. 1990. DOI: 10.1049/el:19900354.
- [130] S. Vedula, O. Senouf, A. M. Bronstein, O. V. Michailovich, and M. Zibulevsky, “Towards CT-quality ultrasound imaging using deep learning,” pp. 1–4, Oct. 2017. arXiv: 1710.06304. [Online]. Available: <https://arxiv.org/abs/1710.06304>.
- [131] F. Dietrichson, E. Smistad, A. Ostvik, and L. Lovstakken, “Ultrasound speckle reduction using generative adversarial networks,” in *2018 IEEE International Ultrasonics Symposium (IUS)*, Oct. 2018, pp. 1–4. DOI: 10.1109/ULTSYM.2018.8579764.
- [132] D. Hyun, L. L. Brickson, K. T. Looby, and J. J. Dahl, “Beamforming and speckle reduction using neural networks,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 66, no. 5, pp. 898–910, May 2019. DOI: 10.1109/TUFFC.2019.2903795.
- [133] O. Huang, W. Long, N. Bottenus, M. Lerendegui, G. E. Trahey, S. Farsiu, *et al.*, “MimickNet, mimicking clinical image post- processing under black-box constraints,” *IEEE Transactions on Medical Imaging*, vol. 39, no. 6, pp. 2277–2286, Jun. 2020. DOI: 10.1109/TMI.2020.2970867.
- [134] A. C. Luchies and B. C. Byram, “Deep neural networks for ultrasound beamforming,” *IEEE Transactions on Medical Imaging*, vol. 37, no. 9, pp. 2010–2021, Sep. 2018. DOI: 10.1109/TMI.2018.2809641.
- [135] B. Luijten, R. Cohen, F. J. de Bruijn, H. A. W. Schmeitz, M. Mischi, Y. C. Eldar, *et al.*, “Adaptive ultrasound beamforming using deep learning,” *IEEE Transactions on Medical Imaging*, vol. 39, no. 12, pp. 3967–3978, Dec. 2020. DOI: 10.1109/TMI.2020.3008537.
- [136] A. A. Nair, K. N. Washington, T. D. Tran, A. Reiter, and M. A. Lediju Bell, “Deep learning to obtain simultaneous image and segmentation outputs from a single input of raw ultrasound channel data,” *IEEE Transactions on Ultrasonics, Ferro-*

REFERENCES

- electrics, and Frequency Control*, vol. 67, no. 12, pp. 2493–2509, Dec. 2020. DOI: 10.1109/TUFFC.2020.2993779.
- [137] D. Perdios, M. Vonlanthen, A. Besson, F. Martinez, M. Arditi, and J.-P. Thiran, “Deep convolutional neural network for ultrasound image enhancement,” in *2018 IEEE International Ultrasonics Symposium (IUS)*, Oct. 2018, pp. 1–4. DOI: 10.1109/ULTSYM.2018.8580183.
- [138] D. Mishra, S. Chaudhury, M. Sarkar, and A. S. Soin, “Ultrasound image enhancement using structure oriented adversarial network,” *IEEE Signal Processing Letters*, vol. 25, no. 9, pp. 1349–1353, Sep. 2018. DOI: 10.1109/LSP.2018.2858147.
- [139] M. Gasse, F. Millioz, E. Roux, D. Garcia, H. Liebgott, and D. Friboulet, “High-quality plane wave compounding using convolutional neural networks,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 64, no. 10, pp. 1637–1639, Oct. 2017. DOI: 10.1109/TUFFC.2017.2736890.
- [140] Z. Zhou, Y. Wang, J. Yu, Y. Guo, W. Guo, and Y. Qi, “High spatial-temporal resolution reconstruction of plane-wave ultrasound images with a multichannel multiscale convolutional neural network,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 65, no. 11, pp. 1983–1996, Nov. 2018. DOI: 10.1109/TUFFC.2018.2865504.
- [141] R. J. G. van Sloun, O. Solomon, M. Bruce, Z. Z. Khaing, Y. C. Eldar, and M. Mischi, “Deep learning for super-resolution vascular ultrasound imaging,” in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019, pp. 1055–1059. DOI: 10.1109/ICASSP.2019.8683813.
- [142] C. B. Burckhardt, “Speckle in ultrasound B-mode scans,” *IEEE Transactions on Sonics and Ultrasonics*, vol. 25, no. 1, pp. 1–6, Jan. 1978. DOI: 10.1109/T-SU.1978.30978.
- [143] S. W. Smith, H. Lopez, and W. J. Bodine, “Frequency independent ultrasound contrast-detail analysis,” *Ultrasound in Medicine & Biology*, vol. 11, no. 3, pp. 467–477, May 1985. DOI: 10.1016/0301-5629(85)90158-9.
- [144] M. Weigert, U. Schmidt, T. Boothe, A. Müller, A. Dibrov, A. Jain, *et al.*, “Content-aware image restoration: Pushing the limits of fluorescence microscopy,” *Nature Methods*, vol. 15, no. 12, pp. 1090–1097, Dec. 2018. DOI: 10.1038/s41592-018-0216-7.
- [145] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds., Cham: Springer International Publishing, Oct. 2015, pp. 234–241. DOI: 10.1007/978-3-319-24574-4_28.
- [146] T. L. Szabo, *Diagnostic Ultrasound Imaging: Inside Out*, 2nd ed., ser. Biomedical Engineering. Academic Press, 2014.
- [147] X. Glorot and Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks,” in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, Y. W. Teh and M. Titterton, Eds., ser. Pro-

- ceedings of Machine Learning Research, vol. 9, Chia Laguna Resort, Sardinia, Italy: PMLR, 2010, pp. 249–256.
- [148] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” pp. 1–15, Dec. 2014. DOI: <http://doi.acm.org.ezproxy.lib.ucf.edu/10.1145/1830483.1830503>. arXiv: 1412.6980. [Online]. Available: <https://arxiv.org/abs/1412.6980>.
 - [149] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, Apr. 2004. DOI: 10.1109/TIP.2003.819861.
 - [150] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, “Enhanced deep residual networks for single image super-resolution,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Jul. 2017, pp. 1132–1140. DOI: 10.1109/CVPRW.2017.151.
 - [151] X. Zhao, Y. Zhang, T. Zhang, and X. Zou, “Channel splitting network for single MR image super-resolution,” *IEEE Transactions on Image Processing*, vol. 28, no. 11, pp. 5649–5662, Nov. 2019. DOI: 10.1109/TIP.2019.2921882.
 - [152] X. Hu, H. Mu, X. Zhang, Z. Wang, T. Tan, and J. Sun, “Meta-SR: A magnification-arbitrary network for super-resolution,” in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2019, pp. 1575–1584. DOI: 10.1109/CVPR.2019.00167.
 - [153] K. He, X. Zhang, S. Ren, and J. Sun, “Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification,” in *2015 IEEE International Conference on Computer Vision (ICCV)*, Dec. 2015, pp. 1026–1034. DOI: 10.1109/ICCV.2015.123.
 - [154] O. M. H. Rindal, A. Rodriguez-Molares, and A. Austeng, “The dark region artifact in adaptive ultrasound beamforming,” in *2017 IEEE International Ultrasonics Symposium (IUS)*, Sep. 2017, pp. 1–4. DOI: 10.1109/ULTSYM.2017.8092255.
 - [155] T. A. Tuthill, R. H. Sperry, and K. J. Parker, “Deviations from rayleigh statistics in ultrasonic speckle,” *Ultrasonic Imaging*, vol. 10, no. 2, pp. 81–89, Apr. 1988. DOI: 10.1016/0161-7346(88)90051-X.
 - [156] D. Perdios, M. Vonlanthen, F. Martinez, M. Arditi, and J.-P. Thiran, “Deep learning based ultrasound image reconstruction method: A time coherence study,” in *2019 IEEE International Ultrasonics Symposium (IUS)*, Oct. 2019, pp. 448–451. DOI: 10.1109/ULTSYM.2019.8925595.
 - [157] P. A. Hager and L. Benini, “LightProbe: A digital ultrasound probe for software-defined ultrafast imaging,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 66, no. 4, pp. 747–760, Apr. 2019. DOI: 10.1109/TUFFC.2019.2898007.
 - [158] J. A. Jensen, S. I. Nikolov, A. C. H. Yu, and D. Garcia, “Ultrasound vector flow imaging—Part II: Parallel systems,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 63, no. 11, pp. 1722–1732, Nov. 2016. DOI: 10.1109/TUFFC.2016.2598180.

REFERENCES

- [159] S. Fadnes, S. A. Nytnes, H. Torp, and L. Lovstakken, “Shunt flow evaluation in congenital heart disease based on two-dimensional speckle tracking,” *Ultrasound in Medicine & Biology*, vol. 40, no. 10, pp. 2379–2391, Oct. 2014. DOI: 10.1016/j.ultrasmedbio.2014.03.029.
- [160] V. Perrot and D. Garcia, “Back to basics in ultrasound velocimetry: Tracking speckles by using a standard PIV algorithm,” in *2018 IEEE International Ultrasonics Symposium (IUS)*, Oct. 2018, pp. 206–212. DOI: 10.1109/ULTSYM.2018.8579665.
- [161] J. A. Jensen, H. Liebgott, F. Cervenansky, and C. A. Villagomez Hoyos, “SA-VFI: The IEEE IUS challenge on synthetic aperture vector flow imaging,” in *2018 IEEE International Ultrasonics Symposium (IUS)*, Oct. 2018, pp. 1–5. DOI: 10.1109/ULTSYM.2018.8580208.
- [162] W. Thielicke and E. J. Stamhuis, “PIVlab—Towards user-friendly, affordable and accurate digital particle image velocimetry in matlab,” *Journal of Open Research Software*, vol. 2, pp. 1–10, Oct. 2014. DOI: 10.5334/jors.bl.
- [163] H. Nobach and M. Honkanen, “Two-dimensional gaussian regression for sub-pixel displacement estimation in particle image velocimetry or particle position estimation in particle tracking velocimetry,” *Experiments in Fluids*, vol. 38, no. 4, pp. 511–515, Apr. 2005. DOI: 10.1007/s00348-005-0942-3.
- [164] D. Garcia, “Robust smoothing of gridded data in one and higher dimensions with missing values,” *Computational Statistics & Data Analysis*, vol. 54, no. 4, pp. 1167–1178, Apr. 2010. DOI: 10.1016/j.csda.2009.09.020.
- [165] W. F. Walker and G. E. Trahey, “A fundamental limit on the performance of correlation based phase correction and flow estimation techniques,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 41, no. 5, pp. 644–654, Sep. 1994. DOI: 10.1109/58.308499.
- [166] —, “A fundamental limit on delay estimation using partially correlated speckle signals,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 42, no. 2, pp. 301–308, Mar. 1995. DOI: 10.1109/58.365243.
- [167] C. P. Loizou, C. S. Pattichis, and J. D’hooge, Eds., *Handbook of Speckle Filtering and Tracking in Cardiovascular Ultrasound Imaging and Video*, ser. Healthcare Technologies. Institution of Engineering and Technology, Jan. 2018. DOI: 10.1049/PBHE013E.
- [168] L. N. Bohs, B. J. Geiman, M. E. Anderson, S. C. Gebhart, and G. E. Trahey, “Speckle tracking for multi-dimensional flow estimation,” *Ultrasonics*, vol. 38, no. 1-8, pp. 369–375, Mar. 2000. DOI: 10.1016/S0041-624X(99)00182-1.
- [169] J. Meunier and M. Bertrand, “Ultrasonic texture motion analysis: Theory and simulation,” *IEEE Transactions on Medical Imaging*, vol. 14, no. 2, pp. 293–300, Jun. 1995. DOI: 10.1109/42.387711.
- [170] M. Otte and H. -H. Nagel, “Optical flow estimation: Advances and comparisons,” in *Computer Vision — ECCV ’94*, J.-O. Eklundh, Ed., Springer Berlin Heidelberg, 1994, pp. 49–60. DOI: 10.1007/3-540-57956-7_5.

- [171] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black, and R. Szeliski, "A database and evaluation methodology for optical flow," *International Journal of Computer Vision*, vol. 92, no. 1, pp. 1–31, Mar. 2011. DOI: 10.1007/s11263-010-0390-2.
- [172] G. F. Pinton, J. J. Dahl, and G. E. Trahey, "Rapid tracking of small displacements with ultrasound," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 53, no. 6, pp. 1103–1117, Jun. 2006. DOI: 10.1109/TUFFC.2006.1642509.
- [173] H. F. Routh, "Doppler ultrasound," *IEEE Engineering in Medicine and Biology Magazine*, vol. 15, no. 6, pp. 31–40, 1996. DOI: 10.1109/51.544510.
- [174] K. E. Thomenius, "Evolution of ultrasound beamformers," in *1996 IEEE Ultrasonics Symposium. Proceedings*, vol. 2, 1996, pp. 1615–1622. DOI: 10.1109/ULTSYM.1996.584398.
- [175] J. H. Song, J. Lee, S. Yeo, G.-D. Kim, and T.-K. Song, "An analytical approach to designing optimal sparse 1-D phased arrays for handheld ultrasound imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 67, no. 7, pp. 1354–1365, Jul. 2020. DOI: 10.1109/TUFFC.2020.2973419.
- [176] E. Roux, F. Varray, L. Petrusca, C. Cachard, P. Tortoli, and H. Liebgott, "Experimental 3-D ultrasound imaging with 2-D sparse arrays using focused and diverging waves," *Scientific Reports*, vol. 8, no. 1, p. 9108, Dec. 2018. DOI: 10.1038/s41598-018-27490-2.
- [177] R. E. Davidsen, J. A. Jensen, and S. W. Smith, "Two-dimensional random arrays for real time volumetric imaging," *Ultrasonic Imaging*, vol. 16, no. 3, pp. 143–163, Jul. 1994. DOI: 10.1177/016173469401600301.
- [178] S. S. Brunke and G. R. Lockwood, "Broad-bandwidth radiation patterns of sparse two-dimensional vernier arrays," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 44, no. 5, pp. 1101–1109, Sep. 1997. DOI: 10.1109/58.655635.
- [179] S. Aziznejad, H. Gupta, J. Campos, and M. Unser, "Deep neural networks with trainable activations and controlled lipschitz constant," *IEEE Transactions on Signal Processing*, vol. 68, pp. 4688–4699, 2020. DOI: 10.1109/TSP.2020.3014611.

DIMITRIS PERDIOS

Chemin de Chandolin 5, 1005 Lausanne, Switzerland • +41 (0)79 509 18 60 • dimitris.perdios@epfl.ch

PROFILE

M.Sc. in Mechanical Engineering, I currently pursue the Ph.D. degree in Electrical Engineering with the Signal Processing Laboratory 5 (LTS5), École polytechnique fédérale de Lausanne (EPFL), under the supervision of Prof. Jean-Philippe Thiran. My research focuses on the development of image reconstruction algorithms in the field of medical ultrasound imaging, with an emphasis on inverse problem formulations, measurement process and physical modeling, numerical simulations, and deep learning. Innovative, conceptual, analytical, and pragmatic, I am a team player with proven leadership skills, keen to conceive and build projects and computer programs from the ground up.



EDUCATION

2016 – Present **EPFL – École polytechnique fédérale de Lausanne** Lausanne

Ph.D. candidate in Electrical Engineering.

- **Ph.D. Thesis:** “Ultrasound Imaging: From Physical Modeling to Deep Learning”
- **Research focus:** Medical ultrasound imaging, image reconstruction, image restoration, inverse problems, physical modeling, B-spline approximation, computer simulations, machine learning, deep learning.
- Supervisor: Prof. Jean-Philippe Thiran, Signal Processing Laboratory 5 (LTS5)

2012 – 2014 **EPFL – École polytechnique fédérale de Lausanne** Lausanne

M.Sc. in Mechanical Engineering (GPA of 5.31 / 6.0). Specialization in Aero- and Hydrodynamic. Minor in Computational Science and Engineering.

- **M.Sc. thesis:** “Development of an Unsteady Three-Dimensional Potential Flow Solver Including Free-Surface”
Keywords: Potential flows, panel method, boundary element method (BEM), free-surface waves.
Supervisors: Dr. Marc-Antoine Habisreutinger, Prof. François Gallaire
Note: Member of the EPFL team for the 2014 HYDROcontest (2nd place).
- **Semester project:** “Optimization of Guide Vanes Geometry” at LMH
Keywords: Geometric and mesh parameterization, Splines, CFD, objective optimization.
Supervisor: Prof. François Avellan, Laboratory for Hydraulic Machines (LMH)

2011 – 2012 **TUB – Technische Universität Berlin** Berlin

Exchange term (Erasmus) as part of M.Sc. studies (GPA of 1.3 / 1.0).

- **Semester project:** “Transient Investigations of Rayleigh-Bénard Instability” at ISTA
Keywords: Compact finite difference schemes, solver development, linearized stability.
- Supervisor: Prof. Jörn Sesterhenn, CFD Group, ISTA

2007 – 2011 **EPFL – École polytechnique fédérale de Lausanne** Lausanne

B.Sc. in Mechanical Engineering (GPA of 5.18 / 6.0, top 15%).

- **B.Sc. thesis:** “Experimental Tests of a Micro-Turbine”
Keywords: Testbed, sensor calibration, data acquisition, processing, and analysis.
Supervisor: Prof. François Avellan, Laboratory for Hydraulic Machines (LMH)

PROFESSIONAL EXPERIENCE

2016 – Present **EPFL – LTS5** Lausanne

Doctoral Assistant, Signal Processing Laboratory 5 (LTS5), EPFL, Lausanne

Doctoral assistant in the field of medical ultrasound. Scientific production: 6 journal publications in the main international journals of the specialty, 2 international filed patents, and 18 conferences. Developed a real-time demonstrator to showcase our image reconstruction methods at 10+ national and international events. Supervised 15+ M.Sc. student projects. Participated actively in 4 funding proposals (SNSF), with ~CHF 1.5M granted. Specification, public bidding follow-up, and selection of a unique high-end 3-D ultrasound research scanner.

2014 – 2016 **EPFL – LTS5** Lausanne

Scientific Assistant, Signal Processing Laboratory 5 (LTS5), EPFL, Lausanne

Fulfilled a proof-of-concept project on transcranial focusing (time reversal) using a consumer ultrasound probe. Developed a multi-domain boundary element method for simulation pressure fields and predicting delay corrections. Conducted experimental validations.

DIMITRIS PERDIOS

Chemin de Chandolin 5, 1005 Lausanne, Switzerland • +41 (0)79 509 18 60 • dimitris.perdios@epfl.ch

2013

ALMATECH

Lausanne

Swiss-based technology company active in naval, space, and various “non-space” markets.

3-month M.Sc. internship in naval engineering (fluid-structure interaction). Developed a tool predicting flutter instabilities of custom hydrofoil designs. Implemented an interpolation and smoothing method (Thin Plate Spline) between solid and fluid domains. Reorganized former source code and simplified input files. Wrote complete documentation.

2011 – 2013

EPFL – LMAF

Lausanne

Teaching Assistant, Laboratory of applied mechanics and reliability analysis (LMAF), EPFL, Lausanne

Courses: Continuum Mechanics and Mechanics of Solids and Structures. Exercise sessions and lecture slides.

SELECTED PUBLICATIONS

- D. Perdios, M. Vonlanthen, F. Martinez, M. Ardit, and J.-P. Thiran, “CNN-based ultrasound image reconstruction for ultrafast displacement tracking,” *IEEE Transactions on Medical Imaging*, DOI: 10.1109/TMI.2020.3046700. [Online]. Available: <https://arxiv.org/abs/2009.01816>, forthcoming.
- D. Perdios, M. Vonlanthen, F. Martinez, M. Ardit, and J.-P. Thiran, “CNN-based image reconstruction method for ultrafast ultrasound imaging,” Aug. 2020. arXiv: 2008.12750. [Online]. Available: <https://arxiv.org/abs/2008.12750>, under review.
- A. Besson, L. Roquette, D. Perdios, M. Simeoni, M. Ardit, P. Hurley, Y. Wiaux, and J.-P. Thiran, “A physical model of nonstationary blur in ultrasound imaging,” *IEEE Transactions on Computational Imaging*, vol. 5, no. 3, pp. 381–394, Sep. 2019. DOI: 10.1109/TCI.2019.2897951.
- A. Besson, D. Perdios, F. Martinez, Z. Chen, R. E. Carrillo, M. Ardit, Y. Wiaux, and J.-P. Thiran, “Ultrafast ultrasound imaging as an inverse problem: Matrix-free sparse image reconstruction,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 65, no. 3, pp. 339–355, Mar. 2018. DOI: 10.1109/TUFFC.2017.2768583.

VOLUNTEER EXPERIENCE (NON-PROFIT / ASSOCIATION)

2012 – 2015, 2019

Electrosanne Festival

Berlin & Lausanne

International electronic music festival organized each September in Lausanne.

Artistic programming and coordination manager. Established schedule (transports, hostels, catering, shows, etc.) for 50+ international artists. Trained and managed up to 30+ volunteers yearly.

2009 – 2018

Traffic Association

Lausanne

A non-profit, cultural organization focused on promoting electronic-based music culture and technologies.

Co-founder, vice president, and treasurer. Artistic director and coordination manager during events. Managed budgets and accounts with a yearly turnover of CHF 30'000.

LANGUAGES

- French (mother tongue)
- English (C1)
- German (B1 – B2, one-year stay in Berlin, Germany)
- Greek (A2, oral comprehension)

PROGRAMMING AND IT SKILLS

- Python (skilled, incl. TensorFlow), C/C++ (average), Bash (Unix shell), Git, Matlab
- Linux, macOS, SSH, LaTeX, MS Word, MS Excel, MS PowerPoint

PERSONAL INFORMATION

- Born February 21, 1989
- Nationality: Swiss, French, and Greek
- Reviewer for the IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control
- Other interests: Vinyl enthusiast, typography, graphic design, tennis, computer simulations

Journals

- [1] D. Perdios, F. Martinez, M. Arditi, and J.-P. Thiran, "A spline-based spatial impulse response ultrasound simulator," in preparation.
- [2] D. Perdios, M. Vonlanthen, F. Martinez, M. Arditi, and J.-P. Thiran, "CNN-based ultrasound image reconstruction for ultrafast displacement tracking," *IEEE Transactions on Medical Imaging*, DOI: 10.1109/TMI.2020.3046700. [Online]. Available: <https://arxiv.org/abs/2009.01816>, forthcoming.
- [3] D. Perdios, M. Vonlanthen, F. Martinez, M. Arditi, and J.-P. Thiran, "CNN-based image reconstruction method for ultrafast ultrasound imaging," Aug. 2020. arXiv: 2008.12750. [Online]. Available: <https://arxiv.org/abs/2008.12750>, submitted.
- [4] A. Besson, L. Roquette, D. Perdios, M. Simeoni, M. Arditi, P. Hurley, Y. Wiaux, and J.-P. Thiran, "A physical model of nonstationary blur in ultrasound imaging," *IEEE Transactions on Computational Imaging*, vol. 5, no. 3, pp. 381–394, Sep. 2019. DOI: 10.1109/TCI.2019.2897951.
- [5] A. Besson, D. Perdios, Y. Wiaux, and J.-P. Thiran, "Joint sparsity with partially known support and application to ultrasound imaging," *IEEE Signal Processing Letters*, vol. 26, no. 1, pp. 84–88, Jan. 2019. DOI: 10.1109/LSP.2018.2880571.
- [6] A. Besson, D. Perdios, F. Martinez, Z. Chen, R. E. Carrillo, M. Arditi, Y. Wiaux, and J.-P. Thiran, "Ultrafast ultrasound imaging as an inverse problem: Matrix-free sparse image reconstruction," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 65, no. 3, pp. 339–355, Mar. 2018. DOI: 10.1109/TUFFC.2017.2768583.

Conferences

- [1] D. Perdios, M. Vonlanthen, F. Martinez, M. Arditi, and J.-P. Thiran, "High-dynamic-range-sensitive training loss for CNN-based ultrasound image reconstruction," presented at the 2020 IEEE International Ultrasonics Symposium (IUS).
- [2] D. Perdios, M. Vonlanthen, F. Martinez, M. Arditi, and J.-P. Thiran, "CNN-based image reconstruction for tracking displacements under extreme conditions," presented at the 2020 IEEE International Ultrasonics Symposium (IUS).
- [3] D. Perdios, M. Vonlanthen, F. Martinez, M. Arditi, and J.-P. Thiran, "Single-shot CNN-based ultrasound imaging with sparse linear arrays," in *IEEE International Ultrasonics Symposium (IUS)*, Sep. 2020, pp. 1–4. DOI: 10.1109/IUS46767.2020.9251442.
- [4] S. Beuret, D. Perdios, and J.-P. Thiran, "Refraction-aware integral operator for speed-of-sound pulse-echo imaging," in *IEEE International Ultrasonics Symposium (IUS)*, Sep. 2020, pp. 1–4. DOI: 10.1109/IUS46767.2020.9251601.
- [5] D. Perdios, M. Vonlanthen, F. Martinez, M. Arditi, and J.-P. Thiran, "Deep learning based ultrasound image reconstruction method: A time coherence study," in *IEEE International Ultrasonics Symposium (IUS)*, Oct. 2019, pp. 448–451. DOI: 10.1109/ULTSYM.2019.8925595.
- [6] D. Perdios, A. Besson, F. Martinez, M. Vonlanthen, M. Arditi, and J.-P. Thiran, "On problem formulation, efficient modeling and deep neural networks for high-quality ultrasound imaging: Invited presentation," in *53rd Annual Conference on Information Sciences and Systems (CISS)*, Mar. 2019, pp. 1–4. DOI: 10.1109/CISS.2019.8692870.
- [7] D. Perdios, M. Vonlanthen, A. Besson, F. Martinez, M. Arditi, and J.-P. Thiran, "Deep convolutional neural network for ultrasound image enhancement," in *IEEE International Ultrasonics Symposium (IUS)*, Oct. 2018, pp. 1–4. DOI: 10.1109/ULTSYM.2018.8580183.
- [8] E. Bezzam, A. Besson, H. Pan, D. Perdios, J.-P. Thiran, and M. Vetterli, "Sparse recovery of strong reflectors with an application to non-destructive evaluation," in *IEEE International Ultrasonics Symposium (IUS)*, Oct. 2018, pp. 1–4. DOI: 10.1109/ULTSYM.2018.8580184.
- [9] A. Besson, D. Perdios, M. Arditi, Y. Wiaux, and J.-P. Thiran, "Compressive multiplexing of ultrasound signals," in *IEEE International Ultrasonics Symposium (IUS)*, Oct. 2018, pp. 1–4. DOI: 10.1109/ULTSYM.2018.8580067.
- [10] O. Bernard, D. Bradway, H. H. Hansen, P. Kruijzinga, A. Nair, D. Perdios, S. Ricci, O. M. H. Rindal, A. Rodriguez-Molares, M. B. Stuart, and P. F. Viseu Dos Santos, "The ultrasound file format (UFF)—First draft," in *IEEE International Ultrasonics Symposium (IUS)*, Oct. 2018, pp. 1–4. DOI: 10.1109/ULTSYM.2018.8579642.
- [11] D. Perdios, F. Martinez, A. Besson, M. Arditi, and J.-P. Thiran, "A spline-based spatial impulse response ultrasound simulator with GPU implementation," presented at the 2018 IEEE International Ultrasonics Symposium (IUS).
- [12] A. Besson, D. Perdios, Y. Wiaux, and J.-P. Thiran, "Pulse-stream models in time-of-flight imaging," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2018, pp. 3389–3393. DOI: 10.1109/ICASSP.2018.8461767.

- [13] D. Perdios, A. Besson, P. Rossinelli, and J.-P. Thiran, "Learning the weight matrix for sparsity averaging in compressive imaging," in *IEEE International Conference on Image Processing (ICIP)*, Sep. 2017, pp. 3056–3060. DOI: 10.1109/ICIP.2017.8296844.
- [14] D. Perdios, A. Besson, M. Arditi, and J.-P. Thiran, "A deep learning approach to ultrasound image recovery," in *IEEE International Ultrasonics Symposium (IUS)*, Sep. 2017, pp. 1–4. DOI: 10.1109/ULTSYM.2017.8092746.
- [15] A. Besson, D. Perdios, F. Martinez, M. Arditi, Y. Wiaux, and J.-P. Thiran, "USSR: An ultrasound sparse regularization framework," in *IEEE International Ultrasonics Symposium (IUS)*, Sep. 2017, pp. 1–4. DOI: 10.1109/ULTSYM.2017.8092360.
- [16] A. Besson, R. E. Carrillo, D. Perdios, M. Arditi, Y. Wiaux, and J.-P. Thiran, "A compressed-sensing approach for ultrasound imaging," in *Signal Processing with Adaptive Sparse Structured Representations (SPARS) Workshop*, Jun. 2017, pp. 1–2.
- [17] A. Besson, R. E. Carrillo, D. Perdios, E. F. Bezzam, M. Arditi, Y. Wiaux, and J.-P. Thiran, "Morphological component analysis for sparse regularization in plane wave imaging," in *IEEE International Ultrasonics Symposium (IUS)*, Sep. 2016, pp. 1–4. DOI: 10.1109/ULTSYM.2016.7728896.
- [18] A. Besson, R. E. Carrillo, D. Perdios, M. Arditi, O. Bernard, Y. Wiaux, and J.-P. Thiran, "A compressed beamforming framework for ultrafast ultrasound imaging," in *IEEE International Ultrasonics Symposium (IUS)*, Sep. 2016, pp. 1–4. DOI: 10.1109/ULTSYM.2016.7728743.

Patents

- [1] D. Perdios, A. Besson, F. Martinez, M. Arditi, and J.-P. Thiran, "Image reconstruction method based on a trained non-linear mapping," PCT/EP2019/073152, 2019.
- [2] A. Besson, D. Perdios, M. Arditi, and J.-P. Thiran, "Model-based image reconstruction method," PCT/EP2018/072579, 2018.