The background features a dark blue field with several overlapping, semi-transparent spheres in shades of red, orange, and blue. These spheres are interconnected by a network of thin, semi-transparent lines in the same color palette, creating a complex, web-like structure that suggests a network or a system of relationships.

Data-driven
reconstruction
methods

for
photoacoustic
tomography:

learning
structures

*Yoeri
Boink*

by
structured
learning

DATA-DRIVEN RECONSTRUCTION METHODS FOR PHOTOACOUSTIC TOMOGRAPHY

Learning structures by structured learning

Yoei Ewald Boink

Committee members:

Chairman:

prof. dr. J. L. Herek University of Twente

Supervisors:

prof. dr. C. Brune University of Twente

prof. dr. S. Manohar University of Twente

Other members:

prof. dr. S.R. Arridge University College London

prof. dr. M. Haltmeier University of Innsbruck

prof. dr. I. Išgum Academic Medical Center – University of Amsterdam

prof. dr. W. Steenbergen University of Twente

prof. dr. J.J.W. van der Vegt University of Twente

prof. dr. M. Versluis University of Twente

The work in this thesis was carried out at the chairs:

- Applied Analysis, Faculty of Electrical Engineering, Mathematics and Computer Science,
- Biomedical Photonic Imaging, Faculty of Science and Technology,
- Multi-Modality Medical Imaging, Faculty of Science and Technology,
University of Twente, Drienerlolaan 5, 7522NB, The Netherlands.

This work was co-funded by the PPP allowance made available by Health~Holland,
Top Sector Life Sciences & Health, to stimulate public private partnerships.

Cover design: Twan Eshuis – Studio Antwan

Lay-out: Yoei Ewald Boink

Printed by: Ipskamp Printing

ISBN: 978-90-365-5087-1

DOI: 10.3990/1.9789036550871

Copyright © 2020 by Yoei Ewald Boink, Enschede, The Netherlands.

All rights reserved. No part of this thesis may be reproduced, stored in a retrieval system or transmitted in any form or by any means without permission of the author.

DATA-DRIVEN RECONSTRUCTION METHODS FOR PHOTOACOUSTIC TOMOGRAPHY

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof. dr. A. Veldkamp,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op vrijdag 5 februari 2021 om 14.45 uur

door

Yoen Ewald Boink

geboren op 7 oktober 1990
te Enschede, Nederland

Dit proefschrift is goedgekeurd door de promotoren:

prof. dr. C. Brune

en

prof. dr. S. Manohar

Contents

1	Introduction	1
1.1	Main challenges	3
1.2	Thesis outline	4
	References	6
2	Technical background	13
2.1	Photoacoustic process	13
2.2	Inverse problems in imaging	18
2.3	Artificial neural networks	27
	References	38
3	A reconstruction framework for photoacoustic tomography	47
3.1	Introduction	48
3.2	Variational methods using regularisation	50
3.3	Numerical framework and algorithm	53
3.4	Experimental setup and forward model	59
3.5	Test objects	61
3.6	Synthetic results	65
3.7	Experimental results	68
3.8	Summary and outlook	71
	References	72
4	Partially learned joint reconstruction and segmentation	77
4.1	Introduction	78
4.2	Theory of photoacoustic tomography	80
4.3	From non-learned to learned primal-dual methods	82
4.4	Methodology	84
4.5	Results	90
4.6	Conclusion and outlook	97
	References	98
	Appendices	102

5	Data-consistent networks for nonlinear inverse problems	107
5.1	Introduction	108
5.2	Convergence of regularised data-consistent networks	112
5.3	Convergence rates	115
5.4	Examples and numerical setup	117
5.5	Numerical results	123
5.6	Conclusion	128
	References	129
	Appendices	131
6	Learned SVD for solving inverse problems	137
6.1	Introduction	138
6.2	Motivation: SVD and inversion methods	141
6.3	The learned singular value decomposition	142
6.4	Analysis	144
6.5	Research context	151
6.6	Experiments and implementation	154
6.7	Numerical results	157
6.8	Conclusion and outlook	164
	References	165
	Appendices	169
7	Conclusion and outlook	177
	References	182
	Summary	185
	Samenvatting	187
	List of publications	189
	Acknowledgements	193

Chapter 1

Introduction

Photoacoustic imaging has received an increasing interest over the past 25 years. This soft tissue imaging modality achieves high contrast and fine resolution by the intrinsically hybrid combination of optical absorption and acoustic wave propagation. Moreover, compared to purely optical modalities, imaging depth is relatively large while attaining high resolution. The combination of contrast, resolution and depth has led to the use of photoacoustics in the form of microscopy [1, 2], needle-based image guidance during surgery [3, 4], imaging by handheld devices [5, 6] and in a tomography setup [6, 7, 8, 9].

This thesis focuses on photoacoustic tomography (PAT), that has found applications in various fields of biomedicine [10], such as small animal imaging [8, 11, 12], rheumatoid arthritis research [13, 14, 15, 16] and breast imaging [9, 17, 18, 19]. In many of these applications, optical contrast is obtained from vascular structures containing haemoglobin. The use of non-ionising light makes it possible to easily change wavelength, which paves the way for extracting biochemical information from the tissue. This may aid in more robust detection or diagnosis of various diseases that manifest as changes in vasculature. While useful applications have provided a boost in research, they are only a first reason for the interest in PAT.

A second reason is the affordability of some photoacoustic devices. Nowadays, optical energy can be delivered by pulsed laser diodes (PLD) [20, 21] or light-emitting diodes (LED) [12, 21, 22]. The detection of ultrasound can be done by commercially available linear transducer arrays, even in a tomography setup [12]. By combining these systems with open-source reconstruction toolboxes like k-Wave [23] it is relatively easy to create an acoustic reconstruction of some initial pressure.

A third reason is that modifications and extensions to both hardware and software can provide substantial improvement to the imaging results, which makes them worth to be investigated. For instance, multiple laser fibres can be

used, as well as larger detector arrays. Both can be placed in many desired configurations. Moreover, multiple wavelengths provide the possibility for spectroscopic imaging [24, 25, 26]. Reconstruction software can be improved by changing from direct methods such as filtered backprojection and time-reversal, to iterative methods, often in a variational framework [27, 28, 29, 30]. An advantage of this variational framework is that it can be used to jointly perform a second task, such as segmentation or motion estimation [31]. Very recently, research has focused on combining classical reconstruction methods with artificial neural networks to solve the inverse problem generated by PAT [32, 33, 34, 35, 36]. The hybrid nature of photoacoustic tomography allows to focus either on just the acoustic or optical inverse problem, or on the combined inverse problem [26, 37, 38, 39, 40, 41]. The latter is a challenge in terms of implementation and computation, especially when the three-dimensional (3D) inverse problem is considered. This almost ‘modular’ choice of the photoacoustic system and reconstruction method enables some researchers to focus on one of its aspects from a fundamental point-of-view, while it enables more applied researchers to focus on the biomedical application of PAT.

The large-scale nature of two coupled processes, acoustic and optical, makes PAT reconstruction a challenging nonlinear inverse problem to be investigated. This is not unique for PAT, since coupled processes also play an important role in other nonlinear hybrid inverse problems [42, 43]. Large-scale inverse problems are also present in biomedical imaging applications like diffuse optical tomography (DOT) [44, 45], ultrasound tomography (UST) [46], X-ray computed tomography (CT) [47], and magnetic resonance imaging (MRI) [48]; but also in seismic imaging [49]. Because of this wealth of imaging techniques that can be described in a similar framework, this thesis will not solely focus on PAT, but also on image reconstruction and the solution of inverse problems in general.

Within the field of image reconstruction, we are at the frontier of using artificial intelligence for solving inverse problems. In the past five years, classical state-of-the-art methods have been rapidly surpassed by deep learning methods [50, 51, 52] due to the emergence of artificial neural networks [53]. However, knowledge of their quality and accuracy depends mainly on empirical results and are often not yet supported by mathematical theory. An important and open question is how to combine well-understood classical methods with high-quality learned methods; how to create a hybrid technique that has the best of both worlds.

Promising attempts [54] range from learning parameters [55] or regularisers [56, 57, 58] in classical methods, to post-processing classical methods with neural networks [59, 60], to using the forward model in a neural network [61, 62, 63, 64], to fully learned networks [65] that might be a generalisation of classical methods. This shows that on the one hand researchers trust in their knowledge about physics, but on the other hand it is evident that this knowledge is incomplete and simulation of a physics process is only approximate. Moreover, it is well understood that measurements always contain uncertainty, often expressed in

the form of noise. To reduce the effect of uncertainty on the results, researchers try to separate important information from noise, sometimes by nonlinearly transforming the data into its principal elements. In this thesis, we investigate these topics and explore the whole range from non-learned to fully learned image reconstruction, with photoacoustic tomography as the main application.

1.1 Main challenges

The promising and interesting technique of PAT comes with several challenges. Here we mention the main challenges related to PAT reconstruction and indicate promising solution directions, of which some are explored in this thesis.

The main challenge of photoacoustic imaging in general is its imaging depth. Although PAT provides deeper imaging depth than some purely optical methods, it is still lacking compared to MRI or CT. This creates problems in signal-to-noise ratio (SNR) and contrast deep inside tissue. A solution is not evident: the penetration depth of light is intrinsically limited by the absorbing and scattering nature of light in tissue. Therefore, to some extent, researchers will simply have to cope with it. On the other hand, the depth can be slightly improved by changing the wavelength, at the cost of contrast. To exploit the advantages of multiple wavelengths, spectroscopic imaging is a promising technique.

Spectroscopic or quantitative reconstruction in PAT is a challenge on itself, especially in-vivo. After performing an acoustic reconstruction, the highly nonlinear optical inverse problem needs to be solved, which takes the acoustic reconstruction as input. This leads to a propagation of errors in the inversion procedure, especially when little information is available. To diminish these errors, information about the forward process and the desired reconstructions needs to be as accurate as possible.

This requires exact knowledge about the physiology of the tissue under investigation, but also about the propagation of light and sound in tissue, with parameters like scattering and sound speed. Even if such information is available, it is challenging to incorporate it in a numerical model with high accuracy and resolution. Artificial neural networks provide a potential solution to this problem. Improvements to the numerical forward model may be implemented explicitly, but also implicitly by incorporating the networks in the reconstruction procedure. It is an open research question which way of utilising artificial neural networks leads to accurate, stable and reliable solutions of inverse problems.

In this thesis, these challenges are directly or indirectly addressed. The goal of a larger imaging depth is pursued in Chapters 3 and 4. Multiple options to combine classical reconstruction methods for solving inverse problems with artificial neural networks are explored in Chapters 4, 5 and 6. Since this is a relatively new research direction, this thesis does not focus on applying

these methods to the quantitative optical inverse problem, which is generally considered to be more involved than the acoustic problem. However, the general quest for accurate and reliable data-driven reconstruction provides solutions that may allow for the extension to quantitative PAT reconstruction.

1.2 Thesis outline

This thesis combines the topics of photoacoustic tomography, inverse problems and artificial neural networks. The fundamental concepts of each of these topics and some connections between them are explained in Chapter 2. Figure 1.2.1 shows the connection between the chapters of this thesis and these general topics.

Figure 1.2.1: Visual overview of the connection between the general topics in this thesis and the chapters indicated with their chapter numbers. 3) A reconstruction framework for photoacoustic tomography. 4) Partially learned joint reconstruction and segmentation. 5) Data-consistent networks for nonlinear inverse problems. 6) Learned SVD for solving inverse problems.

- Chapter 3 treats the variational problem in PAT and provides a framework in which hand-crafted regularisers can easily be compared. For a PAT problem with structures resembling vasculature, directional methods as well as higher-order total variation methods give improved results over direct methods.
- Chapter 4 provides a method to jointly solve the PAT reconstruction and segmentation problem for absorbing structures resembling vasculature.

Artificial neural networks are embodied in the algorithmic structure of primal-dual methods, which are a popular way to solve variational problems. It is shown that a diverse training set is of utmost importance to solve multiple problems with one learned algorithm.

- Chapter 5 provides a convergence analysis for data-consistent networks, which combine classical regularisation methods with artificial neural networks. Numerical results are shown for an inverse problem that couples the Radon transform with a saturation problem for biomedical images.
- Chapter 6 explores the idea of fully learned reconstruction by connecting two nonlinear autoencoders. By enforcing a dimensionality reduction in the artificial neural network, a joint manifold for measurements and images is learned. The method, coined learned SVD, provides advantages over other fully learned methods in terms of interpretability and generalisation. Numerical results show high-quality reconstructions, even in the case where no information on the forward process is used.

After these chapters, a general conclusion of this thesis and outlook for future research directions is given in Chapter 7. The thesis closes with a summary and some final words.

References

- [1] L. V. Wang, "Multiscale photoacoustic microscopy and computed tomography," *Nature photonics*, vol. 3, no. 9, pp. 503–509, 2009.
- [2] J. Yao and L. V. Wang, "Photoacoustic microscopy," *Laser & photonics reviews*, vol. 7, no. 5, pp. 758–778, 2013.
- [3] D. Piras, C. Grijzen, P. Schutte, W. Steenbergen, and S. Manohar, "Photoacoustic needle: minimally invasive guidance to biopsy," *Journal of biomedical optics*, vol. 18, no. 7, p. 070502, 2013.
- [4] M. K. A. Singh, V. Parameshwarappa, E. Hendriksen, W. Steenbergen, and S. Manohar, "Photoacoustic-guided focused ultrasound for accurate visualization of brachytherapy seeds with the photoacoustic needle," *Journal of biomedical optics*, vol. 21, no. 12, p. 120501, 2016.
- [5] M. K. A. Singh, W. Steenbergen, and S. Manohar, "Handheld probe-based dual mode ultrasound/photoacoustics for biomedical imaging," in *Frontiers in Biophotonics for Translational Medicine*, pp. 209–247, Springer, 2016.
- [6] P. K. Upputuri and M. Pramanik, "Recent advances toward preclinical and clinical translation of photoacoustic tomography: a review," *Journal of biomedical optics*, vol. 22, no. 4, p. 041006, 2016.
- [7] L. V. Wang and S. Hu, "Photoacoustic tomography: in vivo imaging from organelles to organs," *Science*, vol. 335, no. 6075, pp. 1458–1462, 2012.
- [8] J. Xia and L. V. Wang, "Small-animal whole-body photoacoustic tomography: a review," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 5, pp. 1380–1389, 2013.
- [9] S. Manohar and M. Dantuma, "Current and future trends in photoacoustic breast imaging," *Photoacoustics*, 2019.
- [10] Y. Zhou, J. Yao, and L. V. Wang, "Tutorial on photoacoustic tomography," *Journal of biomedical optics*, vol. 21, no. 6, p. 061007, 2016.
- [11] X. Deán-Ben, S. Gottschalk, B. Mc Larney, S. Shoham, and D. Razansky, "Advanced photoacoustic methods for multiscale imaging of in vivo dynamics," *Chemical Society Reviews*, vol. 46, no. 8, pp. 2158–2198, 2017.
- [12] K. J. Francis, Y. E. Boink, M. Dantuma, M. K. A. Singh, S. Manohar, and W. Steenbergen, "Tomographic imaging with an ultrasound and led-based photoacoustic system," *Biomedical Optics Express*, vol. 11, no. 4, pp. 2152–2165, 2020.
- [13] D. Chamberland, Y. Jiang, and X. Wang, "Optical imaging: new tools for arthritis," *Integrative Biology*, vol. 2, no. 10, pp. 496–509, 2010.
- [14] P. van Es, S. K. Biswas, H. J. B. Moens, W. Steenbergen, and S. Manohar, "Initial results of finger imaging using photoacoustic computed tomography," *Journal of biomedical optics*, vol. 19, no. 6, p. 060501, 2014.
- [15] P. J. van den Berg, K. Daoudi, H. J. B. Moens, and W. Steenbergen, "Feasibility of photoacoustic/ultrasound imaging of synovitis in finger joints using a point-of-care system," *Photoacoustics*, vol. 8, pp. 8–14, 2017.

- [16] J. Jo, C. Tian, G. Xu, J. Sarazin, E. Schioppa, G. Gandikota, and X. Wang, "Photoacoustic tomography for human musculoskeletal imaging and inflammatory arthritis detection," *Photoacoustics*, vol. 12, pp. 82–89, 2018.
- [17] M. Heijblom, D. Piras, M. Brinkhuis, J. C. van Hespén, F. M. van den Engh, M. van der Schaaf, J. M. Klaase, T. G. van Leeuwen, W. Steenbergen, and S. Manohar, "Photoacoustic image patterns of breast carcinoma and comparisons with magnetic resonance imaging and vascular stained histopathology," *Scientific reports*, vol. 5, p. 11778, 2015.
- [18] M. Toi, Y. Asao, Y. Matsumoto, H. Sekiguchi, A. Yoshikawa, M. Takada, M. Kataoka, T. Endo, N. Kawaguchi-Sakita, M. Kawashima, *et al.*, "Visualization of tumor-related blood vessels in human breast by photoacoustic imaging system with a hemispherical detector array," *Scientific reports*, vol. 7, p. 41970, 2017.
- [19] S. M. Schoustra, D. Piras, R. Huijink, T. J. Op't Root, L. Alink, W. M. Kobold, W. Steenbergen, and S. Manohar, "Twente photoacoustic mammoscope 2: system overview and three-dimensional vascular network images in healthy breasts," *Journal of biomedical optics*, vol. 24, no. 12, p. 121909, 2019.
- [20] P. K. Upputuri and M. Pramanik, "Fast photoacoustic imaging systems using pulsed laser diodes: a review," *Biomedical engineering letters*, vol. 8, no. 2, pp. 167–181, 2018.
- [21] H. Zhong, T. Duan, H. Lan, M. Zhou, and F. Gao, "Review of low-cost photoacoustic sensing and imaging based on laser diode and light-emitting diode," *Sensors*, vol. 18, no. 7, p. 2264, 2018.
- [22] J. Leskinen, A. Pulkkinen, J. Tick, and T. Tarvainen, "Photoacoustic tomography setup using led illumination," in *European Conference on Biomedical Optics*, p. 11077_25, Optical Society of America, 2019.
- [23] B. E. Treeby and B. T. Cox, "k-Wave: Matlab toolbox for the simulation and reconstruction of photoacoustic wave fields," *Journal of biomedical optics*, vol. 15, no. 2, p. 021314, 2010.
- [24] G. Bal and K. Ren, "On multi-spectral quantitative photoacoustic tomography in diffusive regime," *Inverse Problems*, vol. 28, no. 2, p. 025010, 2012.
- [25] B. T. Cox, J. G. Laufer, P. C. Beard, and S. R. Arridge, "Quantitative spectroscopic photoacoustic imaging: a review," *Journal of biomedical optics*, vol. 17, no. 6, p. 061202, 2012.
- [26] A. Pulkkinen, B. T. Cox, S. R. Arridge, J. P. Kaipio, and T. Tarvainen, "A Bayesian approach to spectral quantitative photoacoustic tomography," *Inverse Problems*, vol. 30, no. 6, p. 065012, 2014.
- [27] C. Huang, K. Wang, L. Nie, L. V. Wang, and M. A. Anastasio, "Full-wave iterative image reconstruction in photoacoustic tomography with acoustically inhomogeneous media," *IEEE transactions on medical imaging*, vol. 32, no. 6, pp. 1097–1110, 2013.
- [28] S. R. Arridge, M. M. Betcke, B. T. Cox, F. Lucka, and B. E. Treeby, "On the adjoint operator in photoacoustic tomography," *Inverse Problems*, vol. 32, no. 11, p. 115012, 2016.
- [29] M. Haltmeier and L. V. Nguyen, "Analysis of iterative methods in photoacoustic tomography with variable sound speed," *SIAM Journal on Imaging Sciences*, vol. 10, no. 2, pp. 751–781, 2017.

- [30] K. Bredies, R. Nuster, and R. Watschinger, "TGV-regularized inversion of the Radon transform for photoacoustic tomography," *Biomedical Optics Express*, vol. 11, no. 2, pp. 994–1019, 2020.
- [31] F. Lucka, N. Huynh, M. Betcke, E. Zhang, P. Beard, B. Cox, and S. Arridge, "Enhancing compressed sensing 4D photoacoustic tomography by simultaneous motion estimation," *SIAM Journal on Imaging Sciences*, vol. 11, no. 4, pp. 2224–2253, 2018.
- [32] A. Hauptmann, F. Lucka, M. Betcke, N. Huynh, J. Adler, B. Cox, P. Beard, S. Ourselin, and S. Arridge, "Model-based learning for accelerated, limited-view 3-D photoacoustic tomography," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1382–1393, 2018.
- [33] A. Hauptmann, B. Cox, F. Lucka, N. Huynh, M. Betcke, P. Beard, and S. Arridge, "Approximate k-space models and deep learning for fast photoacoustic reconstruction," in *International Workshop on Machine Learning for Medical Image Reconstruction*, pp. 103–111, Springer, 2018.
- [34] S. Antholzer, M. Haltmeier, and J. Schwab, "Deep learning for photoacoustic tomography from sparse data," *Inverse problems in science and engineering*, vol. 27, no. 7, pp. 987–1005, 2019.
- [35] N. Davoudi, X. L. Deán-Ben, and D. Razansky, "Deep learning optoacoustic tomography with sparse data," *Nature Machine Intelligence*, vol. 1, no. 10, pp. 453–460, 2019.
- [36] S. Guan, A. Khan, S. Sikdar, and P. Chitnis, "Fully dense UNet for 2-D sparse photoacoustic tomography artifact removal," *IEEE journal of biomedical and health informatics*, 2019.
- [37] Z. Yuan and H. Jiang, "A calibration-free, one-step method for quantitative photoacoustic tomography," *Medical physics*, vol. 39, no. 11, pp. 6895–6899, 2012.
- [38] T. Tarvainen, A. Pulkkinen, B. T. Cox, J. P. Kaipio, and S. R. Arridge, "Bayesian image reconstruction in quantitative photoacoustic tomography," *IEEE transactions on medical imaging*, vol. 32, no. 12, pp. 2287–2298, 2013.
- [39] M. Haltmeier, L. Neumann, and S. Rabanser, "Single-stage reconstruction algorithm for quantitative photoacoustic tomography," *Inverse Problems*, vol. 31, no. 6, p. 065005, 2015.
- [40] H. Gao, J. Feng, and L. Song, "Limited-view multi-source quantitative photoacoustic tomography," *Inverse Problems*, vol. 31, no. 6, p. 065004, 2015.
- [41] T. Ding, K. Ren, and S. Vallélian, "A one-step reconstruction algorithm for quantitative photoacoustic imaging," *Inverse Problems*, vol. 31, no. 9, p. 095005, 2015.
- [42] P. Kuchment, "Mathematics of hybrid imaging: A brief review," in *The mathematical legacy of Leon Ehrenpreis*, pp. 183–208, Springer, 2012.
- [43] G. Bal, "Hybrid inverse problems and internal functionals," *Inverse problems and applications: inside out. II*, vol. 60, pp. 325–368, 2013.
- [44] A. Gibson, J. Hebden, and S. R. Arridge, "Recent advances in diffuse optical imaging," *Physics in Medicine & Biology*, vol. 50, no. 4, p. R1, 2005.
- [45] S. R. Arridge and J. C. Schotland, "Optical tomography: forward and inverse problems," *Inverse problems*, vol. 25, no. 12, p. 123010, 2009.

- [46] A. Goncharsky, S. Y. Romanov, and S. Y. Seryozhnikov, "Inverse problems of 3D ultrasonic tomography with complete and incomplete range data," *Wave motion*, vol. 51, no. 3, pp. 389–404, 2014.
- [47] F. Natterer, *The mathematics of computerized tomography*. SIAM, 2001.
- [48] Z.-P. Liang and P. C. Lauterbur, *Principles of magnetic resonance imaging: a signal processing perspective*. SPIE Optical Engineering Press, 2000.
- [49] N. Bleistein, J. K. Cohen, and J. W. Stockwell Jr, *Mathematics of Multidimensional Seismic Imaging, Migration, and Inversion*. Interdisciplinary Applied Mathematics 13, Springer-Verlag New York, 2001.
- [50] 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, 2017.
- [51] 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, 2018.
- [52] S. Ravishankar, J. C. Ye, and J. A. Fessler, "Image reconstruction: from sparsity to data-adaptive methods and machine learning," *Proceedings of the IEEE*, vol. 108, no. 1, pp. 86–109, 2019.
- [53] J. Schmidhuber, "Deep learning in neural networks: an overview," *Neural networks*, vol. 61, pp. 85–117, 2015.
- [54] S. Arridge, P. Maass, O. Öktem, and C.-B. Schönlieb, "Solving inverse problems using data-driven models," *Acta Numerica*, vol. 28, pp. 1–174, 2019.
- [55] S. Banert, A. Ringh, J. Adler, J. Karlsson, and O. Oktem, "Data-driven nonsmooth optimization," *SIAM Journal on Optimization*, vol. 30, no. 1, pp. 102–131, 2020.
- [56] J. Rick Chang, C.-L. Li, B. Póczos, B. Vijaya Kumar, and A. C. Sankaranarayanan, "One network to solve them all—solving linear inverse problems using deep projection models," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 5888–5897, 2017.
- [57] S. Lunz, O. Öktem, and C.-B. Schönlieb, "Adversarial regularizers in inverse problems," in *Advances in Neural Information Processing Systems*, pp. 8507–8516, 2018.
- [58] H. Li, J. Schwab, S. Antholzer, and M. Haltmeier, "NETT: solving inverse problems with deep neural networks," *Inverse Problems*, 2020.
- [59] 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, 2017.
- [60] J. Schwab, S. Antholzer, and M. Haltmeier, "Deep null space learning for inverse problems: convergence analysis and rates," *Inverse Problems*, vol. 35, no. 2, p. 025008, 2019.
- [61] E. Kobler, T. Klatzer, K. Hammernik, and T. Pock, "Variational networks: connecting variational methods and deep learning," in *German conference on pattern recognition*, pp. 281–293, Springer, 2017.

- [62] T. Meinhardt, M. Moller, C. Hazirbas, and D. Cremers, "Learning proximal operators: using denoising networks for regularizing inverse imaging problems," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1781–1790, 2017.
- [63] J. Adler and O. Öktem, "Learned primal-dual reconstruction," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1322–1332, 2018.
- [64] H. Chen, Y. Zhang, Y. Chen, J. Zhang, W. Zhang, H. Sun, Y. Lv, P. Liao, J. Zhou, and G. Wang, "LEARN: learned experts' assessment-based reconstruction network for sparse-data CT," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1333–1347, 2018.
- [65] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, "Image reconstruction by domain-transform manifold learning," *Nature*, vol. 555, no. 7697, pp. 487–492, 2018.

Chapter 2

Technical background

In this chapter, we give an introduction to the photoacoustic process, to inverse problems in imaging and to artificial neural networks. It is not the goal to provide a comprehensive overview of the three research fields, but to provide the fundamental concepts on which we build in this thesis. For this reason, we mainly focus on the technical aspects of the topics mentioned.

2.1 Photoacoustic process

The goal of photoacoustic tomography (PAT) is to visualise optically high absorbing regions within a surrounding medium that is optically scattering and has low absorption. Often, this medium is soft biological tissue, where a tomography setup is useful to create a two- or three-dimensional image of part of the (human) body. PAT is a hybrid modality in which light is emitted and ultrasound is detected. For certain tissue types this provides greater imaging depth and higher resolution than most purely optical modalities and higher contrast than purely ultrasound modalities. The photoacoustic process can be divided into three steps:

1. *optical process*: the propagation of light and its absorption in tissue;
2. *photoacoustic effect*: the thermalisation of absorbed energy and formation of an initial pressure;
3. *acoustic process*: the propagation of ultrasound waves and their detection.

In this section we describe these three steps and uncover their relations by presenting their mathematical equations. A very clear overview figure of the process was created by Cox *et al.* [1], which is reprinted in Figure 2.1.1. We explain the process by moving from the top and right of Figure 2.1.1 to the bottom.

Figure 2.1.1: The photoacoustic process consists of an optical process, a coupling and an acoustic process. Reprint from [1], published under Creative Commons (CC BY 4.0).

2.1.1 Optical process

The photoacoustic process starts with the emittance of electromagnetic waves in the optical spectral range (i.e. light). This is usually done by a short-pulsed nanosecond source such as a laser, emitting coherent monochromatic light. It is transported using free-beam optics or through optical fibres to illuminate the tissue, possibly from different directions and locations. Often it first propagates through an acoustic coupling medium like water. Within the

tissue it propagates further while undergoing scattering and absorption. Light scattering is generically described [2] by Maxwell's equations. It can be divided into two regimes in the tissue: Rayleigh scattering, due to particles much smaller than the wavelength of light, and Mie scattering, due to particles comparable or larger than the wavelength of light. Absorption of light arises from the presence of chromophores, which exist in tissue as endogenous components or can be added in the form of contrast agents. The ultimate goal of PAT reconstruction is to recover either the image of chromophore concentrations or the absorption coefficient at a chosen wavelength of light, by which the existing tissue components can be deduced (see Section 2.2.3). The wavelength of light is optimally chosen such that on the one hand absorption in water and scattering in tissue is low, causing a large penetration depth. On the other hand, the absorption coefficient of the tissue of interest should be larger than that of its surroundings, such that large contrast is obtained. With blood vessels as the tissue of interest, this combination of objectives often leads to a wavelength in the near-infrared range (690-1100 nm) [1, 3].

The transport of light is described by the radiative transfer equation (RTE), an integro-differential equation. Since light transport takes place on a much smaller time scale than the subsequent acoustic problem, the RTE is often only described in its time-independent steady-state form:

$$\theta \cdot (\nabla_{\mathbf{r}} + \mu_a(\mathbf{r}) + \mu_s(\mathbf{r})) \phi(\mathbf{r}, \theta) - \mu_s(\mathbf{r}) \int_{4\pi} k(\theta, \tilde{\theta}) \phi(\mathbf{r}, \tilde{\theta}) d\tilde{\theta} = q(\mathbf{r}, \theta). \quad (2.1.1)$$

Here ϕ is the light radiance at location $\mathbf{r}$ moving in direction θ , μ_a and μ_s are the absorption and scattering coefficient respectively, and $k(\theta, \tilde{\theta})$ is the scattering function describing the probability that a photon travelling in direction θ scatters to direction $\tilde{\theta}$. Note that we left out the dependency on the wavelength of light, which all parameters and variables possess. Also, suitable boundary conditions need to be applied in case of a bounded region. The left hand side of (2.1.1) describes the loss of radiance in direction θ due to a gradient $\nabla_{\mathbf{r}}\phi$, due to absorption μ_a or scattering μ_s in a different direction. It describes a gain of radiance in direction θ due to scattering from a different direction $\tilde{\theta}$ according to the scattering function $k(\theta, \tilde{\theta})$. The right hand side describes the light source q . The RTE can be implemented numerically by a finite element method (FEM) [4, 5, 6] or a Monte Carlo simulation (MC) [7, 8], where the former is usually used in the inverse problem setup, due to the availability of an adjoint equation (see Section 2.2.3). The photoacoustic forward process is further described via the fluence rate Φ , which is obtained by integrating the radiance over all angles θ :

$$\Phi(\mathbf{r}) := \int_{4\pi} \phi(\mathbf{r}, \theta) d\theta. \quad (2.1.2)$$

As an alternative model to the RTE, sometimes the diffusion approximation (DA) is used, because it is easier to implement and computationally less intensive. This is mainly due to the fact that the radiance, modelled in the

RTE, is essentially a five-dimensional variable (three for location and two for direction), while the fluence rate, modelled in the DA, is a three-dimensional variable. The DA is a parabolic partial differential equation (PDE) given by

$$\mu_a(\mathbf{r})\Phi(\mathbf{r}) - \nabla \cdot (\kappa(\mathbf{r})\nabla\Phi(\mathbf{r})) = q_0(\mathbf{r}), \quad (2.1.3)$$

where suitable boundary conditions need to be applied in case of a bounded region. In (2.1.3), $\kappa := 1/(3(\mu_a + \mu'_s))$ is the diffusion coefficient, $\mu'_s := (1 - g)\mu_s$ is the reduced scattering coefficient, and g is an anisotropy factor related to the scattering function k . The reduced scattering coefficient μ'_s can be seen as a description of the scattering as if it were isotropic, therefore correcting for the forward scattering behaviour of light in tissue. The DA is obtained by writing the radiance ϕ as a sum of spherical harmonics and truncating after the first term. This approximation is only valid if $\mu_a \ll \mu_s$ and in weakly anisotropic light (g not too close to 1), which makes the DA inaccurate close to the tissue boundary [9]. A better approximation close to the boundary is acquired when the δ -Eddington approximation is used within the DA [10]. As an alternative, the DA can be coupled with the RTE close to the boundary, which results in a less computationally intensive algorithm than the RTE, while preserving its accuracy [5]. This and other numerical implementations of the DA are usually implemented by a FEM [5, 11].

Once the fluence rate Φ is obtained, the absorbed energy image H can be computed by multiplying it with the absorption coefficient:

$$H(\mathbf{r}) := \mu_a(\mathbf{r})\Phi(\mathbf{r}). \quad (2.1.4)$$

The absorbed energy image H provides the starting point for the photoacoustic effect, the step that couples the optical and acoustic process.

2.1.2 Photoacoustic effect

After light is absorbed, tissue molecules are raised to an excited state. The energy in these excited molecules is mostly distributed in the vibrational modes of these and neighbouring molecules, which is called thermalisation. In other words, most absorbed light energy is converted into heat. This heat causes a local increase in temperature and a corresponding increase in pressure p_0 . The process from absorbed light to increased pressure is known as the photoacoustic effect. It is usually assumed [1] that this pressure increase, often called the initial pressure, is linear with the absorbed energy:

$$p_0(\mathbf{r}) = \hat{\Gamma}(\mathbf{r})H(\mathbf{r}). \quad (2.1.5)$$

Here $\hat{\Gamma}$ is the photoacoustic efficiency, which depends on the molecular environment in the tissue. It comprises multiple thermal coefficients of the local tissue [1, 12]; loosely speaking, it represents the balance between how much energy is used to do work, such as increasing the mass density, and how much

energy is converted into heat. The photoacoustic efficiency $\hat{\Gamma}$ is often referred to as the Grüneisen parameter Γ . However, in general $\hat{\Gamma} < \Gamma$, where the two are only close in a perfect scenario where the absorber and surrounding tissue have the same thermodynamic properties, where there is no radiative decay and where the source pulse of light is very short in time [1].

2.1.3 Acoustic process

The propagation of light, absorption of energy and the photoacoustic effect happen on such a small time-scale that they are often considered instantaneous for the acoustic process. Due to the elasticity of tissue, the initial pressure p_0 from (2.1.5) is propagated through the tissue as an ultrasound wave described [13] by

$$\begin{cases} \partial_t^2 p(\mathbf{r}, t) = c(\mathbf{r})^2 \Delta_{\mathbf{r}} p(\mathbf{r}, t) & \text{for } t \geq 0, \\ p(\mathbf{r}, t = 0) = p_0(\mathbf{r}), \\ \partial_t p(\mathbf{r}, t = 0) = 0. \end{cases} \quad (2.1.6)$$

where $c(\mathbf{r})$ is the heterogeneous (i.e. spatially varying) sound speed. Acoustic attenuation is not included in (2.1.6). However, depending on the wavelength of the ultrasound and the tissue in which it propagates, attenuation can be an important factor. There are many ways to model acoustic attenuation, for an overview we refer to [14].

A different representation from (2.1.6) is obtained in case of a homogeneous (i.e. constant) sound speed c_0 : the Poisson-Kirchhoff formulas [13, 15] provide a projection equation in the form of

$$p(\mathbf{r}, t) = \frac{1}{4\pi} \frac{\partial}{\partial t} \left(t \int_{|\mathbf{r}-\tilde{\mathbf{r}}|=c_0 t} p_0(\tilde{\mathbf{r}}) d\tilde{\mathbf{r}} \right). \quad (2.1.7)$$

This equation makes use of the spherical mean, since (2.1.7) computes the mean initial pressure on the surface of a sphere with radius $|\mathbf{r} - \tilde{\mathbf{r}}| = c_0 t$. This is also known as the spherical Radon transform.

In case of a heterogeneous sound speed, the projection representation of (2.1.7) is sometimes still used: the mean is then not computed over the surface of a sphere, but over a surface that has the same arrival time to the location $\mathbf{r}$ from all its points $\tilde{\mathbf{r}}$. It is possible to compute this surface with the fast marching method [16, 17, 18]. It should be noted that in this case the model (2.1.7) is limited in the sense that the physical phenomena of diffraction, refraction and reflection are not taken into account. However, if the heterogeneity of the sound speed is limited, it can still provide adequate results. A third representation is obtained in case the illumination profile is taken into account. This version of the projection model makes use of a calibration before each measurement. For details we refer to Section 3.4.2.

In PAT, different type of detectors are used to acquire the photoacoustic signals. The detection of absorbers in a wide range of sizes requires a

large bandwidth in the detectors. Moreover, their sensitivity is crucial for the quality and resolution of the desired photoacoustic reconstruction [19]. Classically, ultrasound transducers containing piezoelectric elements are used. A piezoelectric material has the ability to generate electrical signals from incoming (ultrasound) pressure waves and vice versa [19, 20]. Many piezoelectric materials with different properties are available, such as fine grain ceramics, polymers and composites. For a review we refer to [20]. These type of detectors are usually damped to resolve resonance of piezoelectric materials, which results in a trade-off between bandwidth and sensitivity. Moreover, piezoelectric detectors are not transparent and relatively large, which makes it a challenge to combine them with other optical imaging modalities and place them in the desired configuration. To overcome these difficulties, more recent research has focused towards optical detection of ultrasound, which can achieve a wide bandwidth and high sensitivity simultaneously, while being smaller and optically transparent [19]. For tomography in particular, Mach–Zehnder and Fabry–Perot interferometers are researched and implemented [21, 22].

Ultrasound waves are not measured exactly in the way they physically exist and propagate. For any detector, its bandwidth and sensitivity play an important role. Often these are modelled directly, but their effect can also be incorporated by convolving the waves with an impulse response [15]. Certain effects from the system electronics can also be incorporated in this manner. Finally, measurements are affected by noise from system electronics and thermal noise in case of piezoelectric detectors, which are usually considered of additive Gaussian type [23].

Different numerical implementations of the acoustic process exist. Roughly, one can identify finite difference (FD) methods [24, 25], pseudo-spectral time domain methods [26, 27] and projection-based methods that make use of a Green’s function [15, 17, 18, 28]. For specific detection configurations the fast Fourier transform (FFT) can be applied to numerically model the acoustic process [27, 29, 30, 31].

In this section, we provided a technical overview of the forward photoacoustic process in a tomography setup. The next section presents a general introduction of inverse problems and variational methods. The reconstruction problem of PAT can be seen as an inverse problem and is therefore discussed within that section (Section 2.2.3).

2.2 Inverse problems in imaging

As can be seen in Section 2.1, the forward photoacoustic process is well understood, which makes the aim for its reconstruction from measured signals an inverse problem. In this section, we provide a short introduction to that topic. We start with the definition of ill-posedness, which explains the difficulty of inverse problems. Next we consider popular solution methods that take

a variational approach. Finally, we describe photoacoustic reconstruction as an inverse problem and provide well-known solution methods for this specific problem.

For a more thorough explanation, many excellent books about the theory, but also the practical handling of inverse problems are available. Some books focus on the theory of inverse problems and the regularisation methods that are developed to solve them [32, 33, 34]. Other books treat the theory from a Bayesian perspective [35, 36]. Additionally, there are excellent books about inverse problems that specifically focus on imaging [33, 37] or partial differential equations (PDEs) [38], which can be useful for the topic of photoacoustic tomography. Computational methods for inverse problems are discussed in [39]. For solutions of mathematical imaging problems that are not stated as inverse problems, such as segmentation, we refer to [40].

2.2.1 Definition of ill-posed inverse problems

The name ‘inverse problems’ stems from the fact that these problems arise as an inverse to some forward problem or forward process. Often, these forward processes are governed by physics that are well understood. However, due to limited measurements and uncertainty, it is generally more difficult to solve the inverse problem than the forward problem.

In a mathematical setting, we consider model parameters (signals) $x \in \mathcal{X}$ and observations (measurements) $y \in \mathcal{Y}$, where $\mathcal{X}$ and $\mathcal{Y}$ are both Banach spaces. Signals and measurements are connected by a (non)linear mapping $\mathcal{A} : \mathcal{X} \rightarrow \mathcal{Y}$ via

$$y = \mathcal{A}(x). \quad (2.2.1)$$

While for the forward problem x and $\mathcal{A}$ are given and y is unknown, for the inverse problem y and $\mathcal{A}$ are given and x is unknown. According to Hadamard [41], (2.2.1) is called well-posed if the following properties hold [34]:

1. *Existence*: For every $y \in \mathcal{Y}$ there is an $x \in \mathcal{X}$ such that $\mathcal{A}(x) = y$.
2. *Uniqueness*: For every $y \in \mathcal{Y}$ there is at most one $x \in \mathcal{X}$ for which $\mathcal{A}(x) = y$.
3. *Stability*: The solution $x \in \mathcal{X}$ depends continuously on $y \in \mathcal{Y}$.

If one of these properties does not hold, the inverse problem is called ill-posed. Of all the reasons for ill-posedness, the following are arguably the most important ones:

- *Noisy measurements*: In almost all cases, there is no access to the exact y , but only to an inaccurate measurement $y^\delta \approx y$. Often this inaccuracy is modelled via a noise term. In case of additive noise, we model $y^\delta = \mathcal{A}(x) + \eta^\delta$, while in case of multiplicative noise, we model $y^\delta = \mathcal{A}(x) \cdot \eta^\delta$, where η^δ is taken from some statistical distribution and δ depicts the noise level. More involved models are given in [33]. Noise has an effect on existence, uniqueness, and stability of the inverse problem.

- *Inaccurate knowledge of forward process:* Inaccurate measurements can also be the result of a lack of knowledge or of simplifications of the forward process. Although many physics processes are understood relatively well, simplifications are always made, both in the assumptions and in the numerical implementation of the forward model. Since many solution methods make use of this numerical implementation, the existence of a solution of the inverse problem is not guaranteed with given forward mapping.
- *Limited measurements:* The number of measurements is limited, for instance due to time (it takes at least the time of the forward process to take a measurement), resources (each measurement takes energy and potentially other resources) and health regulations (some medical imaging techniques make use of harmful radiation). Also the locations at which the measurements are taken can be physically constrained, which is the case in seismic imaging (one can only measure on the surface or in cavities) and medical imaging (tomography generally measures around the body). Finally the limited bandwidth and finite size of detectors cause a limitation in what can be measured. All these factors have an effect on the uniqueness and stability of the inverse problem.

The above reasons only shed a bit of light on the very complex interplay between the desired signal x , the forward mapping $\mathcal{A}$ and the limited measurements y^δ . For more information and a solid mathematical theory, we refer to one of the many books mentioned at the start of this section.

2.2.2 Variational regularisation methods

Next, we explain the basics of popular solution methods for ill-posed inverse problems that make use of a variational approach. Although variational regularisation methods can be viewed from a Bayesian perspective (see Chapter 6 and [35, 36]), we only focus on its deterministic setting here. For brevity, we do not define all the function spaces and type of norms. For a more rigorous treatment of regularisation theory, we refer to one of the books [32, 33, 34].

We consider the problem where inaccurate measurements y^δ are given with an error $\delta > 0$, i.e. $\|y - y^\delta\| \leq \delta$ for some given norm. In case the error is a result of noise, we call δ the noise level. Since y^δ most likely does not lie in the range of $\mathcal{A}$, it is pointless to search for a solution x^δ of the equation $\mathcal{A}(x^\delta) = y^\delta$. Instead, we search for a solution of the regularised variational problem

$$x_\alpha := \operatorname{argmin}_{\tilde{x}} \left\{ D(\mathcal{A}(\tilde{x}), y^\delta) + \alpha R(\tilde{x}) \right\}. \quad (2.2.2)$$

Here D is the data-fidelity: some distance that measures how close the solution x_α is to the measurements y^δ when mapped by $\mathcal{A}$. The regularisation term R provides stability in the inversion procedure and encodes knowledge about the

prior of the desired signal, such as smoothness or anisotropy. A main idea in regularisation theory appears through the regularisation parameter α : for each regularisation method written as (2.2.2) there exists a parameter choice function $\alpha^*(\delta, y^\delta)$ such that

$$\begin{aligned}\limsup_{\delta \rightarrow 0} \alpha^*(\delta, y^\delta) &= 0, \\ \limsup_{\delta \rightarrow 0} \|x_{\alpha^*(\delta, y^\delta)} - x_0\| &= 0,\end{aligned}$$

where x_0 is the true solution x if it lies in the range of the minimisation problem (2.2.2). This means that the value of the regularisation parameter α can be chosen smaller for smaller noise levels, at least in the limit. In other words: as the measurements are less corrupted by noise, the solution method needs less additional stability by regularisation. Selecting the best regularisation parameter in practice is a difficult task. Some methods are described in [39, Chapter 7].

Obviously, (2.2.2) only provides a very general description. The choice of D mainly depends on the type of noise in the measurement: for additive Gaussian noise, an L^2 -distance is adequate, while a log-based distance should be chosen for multiplicative Gamma noise [42]. The choice of R mainly depends on prior knowledge of the signals one aims to reconstruct. A classical choice is that of Tikhonov-regularisation explained below. More popular choices involve non-smooth functionals, such as the popular total variation (TV) regularisation, which promotes reconstructions with sparse edges [43]. Second order total generalised variation (TGV) [44] and wavelet regularisation [45] are investigated in Chapter 3.

Pseudo-inverse and Tikhonov regularisation

The pseudo-inverse and Tikhonov regularised pseudo-inverse are two classical solution methods, which can be written as special cases of (2.2.2). If the mapping $\mathcal{A}$ is a linear operator A , the least-squares solution (which is the maximum likelihood estimator from a Bayesian perspective, see Section 6.2) coincides with applying the pseudo-inverse $A^\dagger$:

$$x_{\text{LS}} := \operatorname{argmin}_{\tilde{x}} \|A\tilde{x} - y^\delta\|_{L^2}^2 = (A^*A)^{-1}A^*y^\delta = A^\dagger y^\delta. \quad (2.2.3)$$

Tikhonov regularisation shifts the singular values in the pseudo-inverse closer to zero to create a solution that is more stable against noise:

$$x_{\text{LS-T}} := \operatorname{argmin}_{\tilde{x}} \|A\tilde{x} - y^\delta\|_{L^2}^2 + \alpha \|\tilde{x}\|_{L^2}^2 = (A^*A + \alpha \text{Id})^{-1}A^*y^\delta. \quad (2.2.4)$$

These and other classical regularisation methods can also be seen from the perspective of the singular value decomposition (SVD), which is shown in Section 6.2.

Solving the variational problem

Unlike the special cases (2.2.3) and (2.2.4), most variational methods do not yield a direct approach to solve them. Variational problems are usually solved by iterative approaches [39], which follow some discrete or continuous optimisation method [46]. These optimisation methods generally require many evaluations of the forward mapping and possibly derivatives or adjoint equations. Below we compare a few properties of the forward mapping $\mathcal{A}$, the data-fidelity D and the regulariser R and discuss their effect or restriction on the optimisation method of choice. This list is only intended to provide a general overview of the optimisation of variational problems. For more information, we refer to [39] and [46].

- *Smooth vs nonsmooth:* The smoothness of (2.2.2) depends on the smoothness of $\mathcal{A}$, D and R . For instance, choosing an L^2 -loss for D and R will result in a smooth problem, assumed $\mathcal{A}$ is linear or otherwise smooth. A sparsity-promoting L^1 -loss, such as in TV-regularisation, results in nonsmooth problems. Smooth problems are generally solved by derivative-based methods, such as (stochastic) gradient descent (GD), the conjugate gradient method (CG), Newton's method or the Broyden-Fletcher-Goldfarb-Shanno method (BFGS). Nonsmooth problems can be solved by smoothing the problem [47] or by using nonsmooth methods such as proximal gradient descent. Often the problem (2.2.2) is rewritten as a saddle point problem and solved with a primal-dual method [48], which makes use of proximal operators (see Section 3.3).
- *Linear vs nonlinear:* The linearity of the forward mapping $\mathcal{A}$ has effects on the optimisation method of choice. Linear mappings (operators A) ensure that a convex data-fidelity D stays convex. Moreover, for linear mappings it is straightforward to compute gradients or proximals of (2.2.2), which is not the case for all nonlinear mappings. For those cases, it might be necessary to linearise the problem at multiple instances of the iteration procedure, which can be computationally intensive. For this reason, nonlinear problems often make use of optimisation methods that require less iterations: they make use of line-search or are accelerated by taking information of previous iterates into account [46]. Smooth nonlinear problems are often solved by (quasi) higher-order derivative methods, such as Newton's method or BFGS, which incorporate part of the nonlinearity in their (quasi) higher-order derivative.
- *Convex vs nonconvex:* Since D and R are often chosen to be convex, the main source of nonconvexity is usually in $\mathcal{A}$. This nonconvexity often arises as a result of nonlinearity in $\mathcal{A}$ or a lack of measurements, which creates nonuniqueness in the inverse problem (see Section 2.2.1). Optimisation of nonconvex problems is difficult due to the many local optima in the 'loss landscape'. There is no general solution to overcome

the nonconvexity. Sometimes the problem is convexified, which means a similar convex problem is solved. Methods to solve the nonconvex problem directly are stochastic optimisation methods, where only part of the gradient is computed to avoid local optima, or higher-order derivative methods to detect whether a local optimum is a minimum or a saddle point. Additionally, acceleration or line-search can be used to overcome local minima. Finally, multiple random initialisations can be used to choose the best of multiple optima.

2.2.3 PAT as an inverse problem

In this section, we summarise inversion procedures for photoacoustic tomography, whose forward process is described in Section 2.1. The division of the photoacoustic process in three steps in an optical process, a coupling process and an acoustic process is also made here. Figure 2.2.1, taken from Cox *et al.* [1], shows possible steps in the inversion procedure; which steps are taken depends on the goal of the inversion.

Figure 2.2.1: Overview of the photoacoustic inverse problems. Solid lines indicate linear inverse problems, dotted lines indicate nonlinear inverse problems. Reprint from [1], published under Creative Commons (CC BY 4.0).

Most research is focused towards finding a good reconstruction for the initial pressure p_0 , given some measured pressure $p(\mathbf{r}_d, t)$ at detector locations $\mathbf{r}_d$ and time steps t . The main reason for this is that p_0 already provides qualitative information on geometrical structures in tissue, such as blood vessel networks, while the inverse problem is linear. This means that the inverse problem has a linear forward mapping, which makes reconstruction easier, although most reconstruction procedures are still nonlinear.

The research that focuses on obtaining optical parameters is known as quantitative photoacoustic tomography (QPAT). As indicated in Figure 2.2.1, there are essentially two ways to solve the full nonlinear QPAT inverse problem:

- *Two-step reconstruction:* This approach aims at first reconstructing the initial pressure p_0 , followed by a reconstruction of the absorbed energy H and finally reconstructing one or more of the optical parameters: chromophore concentrations C_k , absorption coefficient μ_a , scattering coefficient μ_s . When the diffusion approximation is used, the diffusion coefficient $\kappa := 1/(3(\mu_a + \mu'_s))$ is desired instead of μ_s . The photoacoustic efficiency $\hat{\Gamma}$, related to the Grüneisen parameter Γ (see Section 2.1), is often assumed to be known and spatially constant. Although this is not correct [1], it removes one step in the reconstruction process, such that after the acoustic inverse problem, only the optical inverse problem has to be solved.

A drawback of a two-step reconstruction procedure is that errors in the reconstruction of p_0 directly create errors in the data H for the optical inverse problem. Often these errors are ignored, but they can also be taken into account, for instance in a Bayesian setting [49].

- *One-step reconstruction:* This approach aims at combining the acoustic and optical process and reconstructing C_k , μ_a and μ_s or potentially κ or $\hat{\Gamma}$.

A benefit of a one-step reconstruction procedure is that there is no propagation of errors from the acoustic inverse problem to the optical one. A drawback is that the full inverse problem becomes nonlinear, which potentially leads to a longer computation time. However, when this nonlinearity is not too severe, it is possible to solve the full inverse problem as fast as the two decoupled inverse problems [50, 51].

The acoustic and optical processes are usually approached by different methods. The acoustic process is often solved by projection methods or pseudo-spectral methods, while the optical process is generally solved by finite element methods. This is due to the difference in type of PDE: the acoustic wave-equation is a hyperbolic PDE, while the RTE is an integro-differential equation and the DA an elliptic PDE. To solve both processes together (two-step or one-step), it is useful to make use of the same grid or even the same type of finite elements.

Besides the distinction between two-step methods and one-step methods, there are many modelling choices that need to be made to solve the complete QPAT inverse problem. Some important modelling choices are:

- whether the sound speed is assumed to be known or not, either homogeneous (spatially constant) or not;
- whether the photoacoustic efficiency $\hat{\Gamma}$ or Grüneisen parameter Γ is assumed to be known or not, either homogeneous or not;
- whether to use the RTE or the DA in the inversion of the optical process;
- whether to use a single-wavelength or a multi-wavelength light source.

All these modelling choices are intertwined and therefore difficult to discuss separately. Below we provide a short overview of different reconstruction approaches for the acoustic and optical inverse problems separately, in which these modelling choices play an important role.

Acoustic reconstruction methods

Probably the two most widely used methods for acoustic reconstruction in PAT are filtered backprojection (FBP) and time reversal (TR). FBP for PAT was first proposed by Kruger *et al.* [12] and has gained attention by multiple researchers in the field of mathematics over the years (see Section 3.1 for references). FBP is a standard technique in CT reconstruction [52] and it shows many similarities with FBP for PAT. The main idea is that the detected signals are first filtered, after which they are projected over straight rays (CT) or spherical rays (PAT). In case of a heterogeneous sound speed it is still possible to use the FBP, although the rays are not exactly spherical anymore (see Section 2.1.3). Time reversal was first proposed by Finch and Patch [53] and as the name suggests, it models the acoustic waves as if they are reversed in time from the detectors. Similar to the forward problem (see Section 2.1.3), TR can be implemented by finite difference (FD) methods [24, 25] or pseudo-spectral time domain methods [26, 27]. Both FBP and TR are proven to give the exact result in the theoretical case of an infinite amount of infinitesimally small detectors with infinite bandwidth enclosing the region of interest. Obviously, this requirement can never be met in practice, which has the consequence that both FBP and TR suffer from artefacts, especially when the noise level is high or the number of detectors is low [54].

A second approach to solve the acoustic inverse problem makes use of variational methods. With an L^2 -norm as data-fidelity, the detected photoacoustic pressure $p(\mathbf{r}_d, t)$ as measurements and the initial pressure p_0 as the desired solution, (2.2.2) takes the form

$$\min_{p_0} \left\{ \|\mathcal{A}(p_0) - p(\mathbf{r}_d, t)\|_{L^2}^2 + \alpha R(p_0) \right\}, \quad (2.2.5)$$

where $\mathcal{A}$ is the forward mapping in the acoustic problem. In case of a linear A , iterative methods often make use of the adjoint A^* to solve (2.2.5). If A is a projection operator, its adjoint is equal to the backprojection step in the FBP. As shown in [55], the analytical adjoint is very close to time-reversal, although not exactly the same. The selection of an adequate regulariser $R(p_0)$ is not

straightforward: this problem is considered in Chapter 3. There, the variational setting for PAT and related literature are treated in more detail.

In the variational setting, the forward mapping is often assumed to be known exactly. However in reality, due to modelling approximations, uncertainties in system design and unknown parameters in the tissue of interest, this is not the case.

Many researchers proposed ideas to overcome some of these problems. A well-studied approach is the estimation or reconstruction of a heterogeneous sound speed [56, 17, 57, 58, 59, 60]. Other works consider a statistical framework where errors due to variation in sound speed [61, 62] or detector locations [63] are overcome. In many recent papers the effects of these errors are negated by applying the forward mapping in a neural network architecture (see Section 2.3.4). A very recent paper provides an approach where a neural network improves the linear forward mapping, after which it is applied in a variational setting [64].

Optical reconstruction methods

Many papers on the optical inverse problem make use of the DA instead of the RTE. An early work [65] solves the problem of recovering μ_a by making an initial guess, solving the forward DA, dividing the absorbed energy by the obtained fluence rate and repeating this process. Other works [66, 67] aim for reconstructing both μ_a and κ by minimising the variational problem

$$\min_{(\mu_a, \kappa)} \|A(\mu_a, \kappa) - H_{\text{meas}}\|_{L^2}^2. \quad (2.2.6)$$

Here $\mathcal{A}$ is the forward mapping that maps the absorption and diffusion coefficients μ_a and κ to the absorbed energy H , given the light source q_0 (see (2.1.3)). Optimisation is done with the BFGS algorithm, a quasi-Newton method. Although possible, this procedure is relatively unstable, especially with respect to κ , which makes a good initial guess necessary. Bal and Ren [68] theoretically showed that it is possible to reconstruct all three parameters μ_a , κ and $\tilde{\Gamma}$. However, to reconstruct all three parameters at the same time multi-wavelength measurements are necessary, while two parameters can be reconstructed with single-wavelength measurements. Some stability estimates on the inversion process using the DA are given in [69]. There are many newer works that investigate variants of the DA problem. For instance, works that explore the DA problem in a Bayesian setting [49] or that assume piecewise constant tissue parameters [70].

There are significantly less papers that apply the RTE for solving the optical inverse problem. The first paper that considered the RTE used a similar approach to earlier mentioned [65], and used the RTE as a forward solver to reconstruct μ_a [71]. Later works reconstruct μ_a and μ_s by solving a regularised variational problem using the Gauss-Newton method [72] or a quasi-Newton

method [73, 74]. In [51], the regularised one-step QPAT problem is solved using a proximal gradient method. All above methods make use of FEM implementations of the RTE, which is computationally intensive for the 3D problem due to the additional angular dimensions. A recent paper provided an adjoint Monte Carlo method for solving the optical problem and showed results in 3D [75]. All previous papers consider a known scattering function $k(\theta, \tilde{\theta})$ (see Section 2.1.1), often via a constant anisotropy factor g . Bal *et al.* showed that a heterogeneous $g(x)$ can in principle be reconstructed [76]; their paper provides many theoretical results for the reconstruction of QPAT using the RTE.

2.3 Artificial neural networks

In this thesis, we combine model-based algorithms for inverse problems (Section 2.2) with learned parts in the form of artificial neural networks. In this section, we first introduce the fields of artificial intelligence and machine learning, of which the area of artificial neural networks (ANNs) forms a subfield. After that, the basic technical concepts of ANNs are explained. We conclude with a short explanation of how ANNs can be incorporated in methods for solving inverse problems.

2.3.1 Neural networks as part of artificial intelligence

The field of artificial intelligence (AI) is concerned with systems that perform tasks that are considered ‘intelligent’. In other words, these are tasks that can not be solved with ‘classical’ algorithms and often require many inputs to come to a decision or action. AI is a very broad term that has many subfields, some of which require learning, and others that do not, like robots that are programmed with decision rules.

Machine learning (ML) is a subfield of AI in which a system learns to make decisions or predictions based on input data. This learning can have many forms, but the important aspect is that the learned patterns change as soon as the training data set is changed. For complex tasks, even experts can not determine the features upon which they make the decisions or predictions themselves. ML is useful for these tasks, since they do not require the exact description of these features. In general, one can distinguish between supervised learning, unsupervised learning and reinforcement learning.

- In *supervised learning*, the training data consists of input data along with target data. An example is image classification, in which pairs of images and associated labels are available for training. Examples of algorithms are regression, decision tree, k-nearest neighbours and random forest.
- In *unsupervised learning*, the training data only consists of input data without target data. It may be the goal to cluster or compress the input

data. Examples of algorithms are k-means clustering or the (truncated) singular value decomposition.

- In *reinforcement learning*, the goal is to find the actions that provide the maximum reward in a specific situation. The optimal outputs are not given, but have to be found by a process of trial and error. Reinforcement learning typically needs a balance of exploration (of the environment) and exploitation (of the current situation). Examples of algorithms are Markov decision process, deep Q-network and deep deterministic policy gradient. We refer to [77] for an overview.

For a more detailed and mathematical description of ML, we refer to the book by Bishop [78]. A more probabilistic view is given in the book by Murphy [79].

Artificial neural networks are inspired by the biological neural network (NN) in a human brain [80, 81]. A biological brain consists of an estimated 100 billion neurons that communicate via electrical signals. Each neuron typically receives signals from thousands of other neurons. Loosely speaking, if the sum of signals exceeds some threshold, the neuron will generate an impulse signal by itself, which travels to other neurons. This explanation is an oversimplification of the complex biological processes that take place in the brain, but it forms the basis of ANNs. An ANN, like a biological NN, consists of neurons (nodes) that are connected to each other. Activation functions model the behaviour that resembles the threshold behaviour of biological NNs.

The first ANNs [82] (see also [80, 83, 81] for an overview) only contained several neurons with very simple activation functions. Throughout the last decades, most research did not aim for a strong resemblance of ANNs with their biological counterpart. Instead, research focused on ANNs as modification or extension of the first attempts, solving specific challenging tasks.

2.3.2 Building blocks

In this and the following section, the technical details of ANNs are briefly explained. For a complete overview on neural networks, we refer to the books [81, 84, 85] that form the basis of the explanation below.

Artificial neural networks are nonlinear functions that obtain their expressiveness by stacking and parallelising very simple building blocks. Generally, but with exceptions, neurons are stacked in several layers, where information can only flow from a previous layer to a next layer. The in- and outputs of ANNs are vectors or multidimensional matrices. The general mathematical structure of a single layer is given as

$$h^{l+1} = \sigma(K^l h^l + b^l), \quad (2.3.1)$$

where h^l and h^{l+1} are representations for two consecutive layers, K^l is a linear operator working on h^l , and b^l is a bias, creating the affine function $K^l h^l + b^l$. The function σ is called a nonlinearity or activation function and is necessary to create an expressive ANN. The input and output of the ANN are depicted by

h^0 and h^L respectively, where L is the number of layers. The layers between the input and output are called hidden layers.

In (2.3.1), the exact function or vector space of each component is not given explicitly, since they depend on the exact problem and the way the building blocks are put together (architecture). Below, we give two main architecture choices for ANNs and their building blocks.

- In a *fully connected network (FCN)*, each node is a scalar, which means that each layer h^l , consisting of c^l nodes, is a vector in $\mathbb{R}^{c^l}$. The layers are connected via a linear matrix $K^l \in \mathbb{R}^{c^{l+1} \times c^l}$ with entries (weights) $k_{i,j}^l$ and a bias $b^l \in \mathbb{R}^{c^{l+1}}$. FCNs are often used for relatively small-scale problems with the goal to extract nonlocal information, e.g. a classification or clustering problem. They are also used in combination with convolutional neural networks, where they form the final couple of layers, for instance in classification problems.

Figure 2.3.1: Visualisation of an FCN with red input neurons, blue hidden neurons and green output neurons. In the second set of connections, all weights (matrix entries) $k_{i,j}^l$ are specified. Biases are not visualised here.

- In a *convolutional neural network (CNN)*, each node is a vector or multidimensional matrix, e.g. a discretised 2D image. All nodes (images) in one layer are combined via several channels c , similar to a colour image that consists of three colour channels (RGB). By doing so, we obtain for our 2D-example a layer $h^l \in \mathbb{R}^{m \times n \times c^l}$, where m and n are the dimensions of each image, and c the number of channels, or nodes per layer. A CNN makes use of convolutions, whose kernels are described with the discrete kernels $k_{i,j}^l$, where $i \in \{0, \dots, c^l\}$ and $j \in \{0, \dots, c^{l+1}\}$. If the padding is chosen such that the output images are also of size $m \times n$, the linear operator is defined

$$(K^l h^l)_j := \sum_i k_{i,j}^l * h_i^l.$$

Here the subscript i in h^l indicates the i 'th channel of h^l . Each $k_{i,j}^l$ is a 2D-kernel of predefined size. The bias b^l has a scalar value for each output

channel j , i.e.

$$(K^l h^l + b^l)_j := \sum_i (k_{i,j}^l * h_i^l) + b_j^l. \quad (2.3.2)$$

Each convolutional kernel in a CNN is used for all pixels in one channel of the image. This is memory efficient compared to using a unique parameter for each pixel, as is done in a FCN. CNNs are often applied when the ANN should extract local information, for instance in computer vision problems. When a more global behaviour is required, one can choose to use dilated kernels or apply the kernel with a stride, i.e. skip pixels in performing the convolution. Moreover, the basic building blocks can be combined with pooling and up-sampling to reduce the dimension in some of the layers of the CNN.

Figure 2.3.2: Visualisation of a CNN with red input images, hidden layers in blue and green output images. In the second set of connections, all kernels $k_{i,j}^1$ are specified. Biases are not visualised here.

In both FCNs and CNNs, the activation function σ acts elementwise on its input, meaning that it does not change its shape. Popular choices for the activation function are shown in Figure 2.3.3. It is an open research question which activation functions work best for which architectures. However, most activation functions are monotone functions and are (close to) linear at some part of the function. However, as shown in [86], this is not necessary: the authors apply automatic techniques to search for new activation functions.

Figure 2.3.3: Visualisation of four widely used activation functions.

2.3.3 Training and architecture choices

The strength of ANNs lies in the stacking and parallelising of building blocks explained in Section 2.3.2. In the case that multiple layers (2.3.1) are simply stacked, the following expression for an ANN is obtained:

$$h^L = \mathcal{N}(h^0) = \sigma(K^{L-1}(\dots \sigma(K^0 h^0 + b^0) + \dots) + b^{L-1}). \quad (2.3.3)$$

Although most ANNs make use of such stacking of layers, (2.3.3) is not a general expression for all ANNs. For instance, one can devise ANNs where one input is split into different branches that create multiple outputs or come together again after several layers to provide one output. Moreover, the number of layers in the separate branches need not be the same, which is for instance the case in a residual neural network (ResNet):

$$h^L = \mathcal{N}(h^0) + h^0. \quad (2.3.4)$$

For readability, in the remainder of this section we consider the general formulation $h^L = \mathcal{N}(h^0)$ for any ANN.

Loss function

Training of an ANN is done by minimising some loss function

$$\min_{\Theta} \sum_{t \in \text{train}} \mathcal{L}(h_{(t)}^L, h_{(t)}^{\text{GT}}) \quad (2.3.5)$$

where $\mathcal{L}$ is some distance between the ANN output $h_{(t)}^L$ and some ground truth signal $h_{(t)}^{\text{GT}}$, where t indicates the sample from the training set. The parameters over which (2.3.5) is minimised are indicated with Θ . Often only the weights and biases are trainable (not fixed), i.e. $\Theta = (K^{\{0, \dots, L-1\}}, b^{\{0, \dots, L-1\}})$. However, some weights and biases may be fixed, or other parameters, such as the activation function, may also be trainable. The loss function $\mathcal{L}$ can take many forms. Often used loss functions are:

- *L²-loss*: $\mathcal{L}(h_{(t)}^L, h_{(t)}^{\text{GT}}) = \|h_{(t)}^L - h_{(t)}^{\text{GT}}\|_2^2$ is a convex and differentiable loss function, often used for regression or reconstruction problems.
- *Binary cross-entropy*: $\mathcal{L}(h_{(t)}^L, h_{(t)}^{\text{GT}}) = h_{(t)}^{\text{GT}} \log(h_{(t)}^L) + (1 - h_{(t)}^{\text{GT}}) \log(1 - h_{(t)}^L)$ is a convex and differentiable loss function, often used for classification or segmentation problems.
- *Learned loss*: instead of describing the loss function explicitly, it can also be learned by a separate ANN. For instance, in generative adversarial networks (GANs) [87], one network is trained to generate signals h^L , while a second network is trained to discriminate between real signals h^{GT} and fake signals h^L . The benefit of such a loss function is that it can have a different behaviour at different parts of the signal and it may fit better to

human perception than predescribed loss functions. On their turn, these learned loss functions are trained with other losses, such as the binary cross-entropy loss.

Training and algorithms

Before an ANN can be trained, the available data set first has to be split in a training, validation and test set. The training set is the largest set and will be used to do the actual optimisation. The validation set is used to check the loss function (2.3.5) while training, and to handcraft the architecture and other non-trainable parameters. After this is fixed, the trained network is tested on previously unseen data: the test set. This ensures that the communicated test loss is not biased towards the training or validation set, which are both used to train the ANN. So even though the training set is used for optimisation, the training goal is to minimise the validation loss, after which the test loss is communicated. This description shows the difference between ‘optimisation of the training loss’ and ‘training an ANN’.

For the training of ANNs, the backpropagation algorithm is applied. It makes use of the chain rule and computes the gradients of the loss with respect to certain trainable parameters (so not the input). Each gradient has a large number of components, because many paths can be followed from each parameter to the loss. However, this can be computed efficiently by using dynamic programming and a graphical processing unit (GPU) in the computer hardware. A detailed explanation can be found in [85, Section 1.3].

The minimisation of the loss function (2.3.5) is a rather difficult task for several reasons. First, all trainable parameters form a high dimensional space in which to optimise. Second, the depth (number of layers) of ANNs result in a very nonconvex problem (see Section 2.2.2). Third, the potentially large training set makes it difficult to optimise over all training samples at the same time.

The standard technique for training ANNs is mini-batch optimisation, which solves the third problem and at the same time mitigates the second problem. Instead of computing gradients with respect to all training samples, the gradients are computed with respect to the samples in one mini-batch at the time. After all samples have been handled in several mini-batches, the order is randomised and new mini-batches are created for a next epoch. Besides being less memory consuming, this way of optimisation has the effect that local minima and saddle points can be overcome. This can be explained as follows: it might be the case that an optimisation algorithm is stuck in a local minimum when the loss function (2.3.5) sums over all samples in the training set. However, when only the samples in a mini-batch are considered in (2.3.5), it is likely that the same location is not a minimum, but has a nonzero gradient. By propagating far enough in the direction of a negative gradient, the optimisation algorithm moves away from the local minimum and continues to search for a better value of the loss function.

The basic optimisation algorithm for optimising ANNs is gradient descent (GD). When GD is applied to a single sample from the training set, it is called stochastic gradient descent (SGD); when multiple samples are chosen in a randomised way, it is called mini-batch SGD. As described, SGD is a very simple and powerful technique to make nonconvex optimisation (training) of parameters possible. However, SGD can still get stuck in local minima and saddle points, especially when the learning rate (optimisation step size) is chosen small. One solution is to scale the learning rates of the individual gradients, which is done in the AdaGrad algorithm [88]. A second solution is to not follow the gradient of the latest mini-batch completely, but partly keep the direction of previous optimisation steps. This is also known as a Nesterov momentum, and is exploited in the RMSProp algorithm [89]. A very popular algorithm that combines both is the Adam algorithm [90]. Other techniques that might improve the training of ANNs, both in speed and in the final parameters learned, are second-order optimisation and parameter regularisation (weight regularisation). For this and more information on training ANNs, we refer to [81, Chapter 8] and [85, Chapter 3].

Architecture choices

Each ANN has a different aim, for instance classification, reconstruction or image generation. This aim asks for an effective combination of the architecture and loss function, which dictates the optimisation problem. In this part, we mention several architecture choices that are often used for various aims. For an overview of most architectures, we refer to ‘a mostly complete chart of Neural Networks’ by the Asimov Institute¹. In general, the distinction between supervised and unsupervised learning can be made:

- *Supervised learning*: in classification, segmentation and reconstruction problems, data is available in pairs: one signal h^0 , that acts as input to the ANN, and one signal or label h^{GT} , that acts as the desired ground truth. For instance, in the inverse problem setup of Section 2.2.1, a function $\mathcal{N} : \mathcal{Y} \rightarrow \mathcal{X}$ could be trained to map a measurement to the desired signal; a function $\mathcal{N} : \mathcal{X} \rightarrow \mathcal{X}$ could be trained to map an imperfect reconstruction to the desired signal.
- *Unsupervised learning*: in compression or image generation, a function $\mathcal{N}$ is trained with a single data set of signals, without the supervision of a second signal or label. For instance, an autoencoder (AE) is trained to give the same output as input, i.e. $h^L \approx h^0$. An AE can learn a nonlinear compression, if one of the middle layers (latent space) has a smaller dimension than the input. Such an architecture can also be used for unsupervised clustering. Another example of unsupervised learning is that of image generation. Two main architectures that achieve this are the

¹<https://www.asimovinstitute.org/neural-network-zoo/>

GAN, which makes use of a learned discriminator, and the variational autoencoder (VAE) [91], which is trained with a loss function where $h^{\text{GT}} = h^0$.

- *Semi-supervised learning*: Sometimes training an ANN has a supervised goal, such as classification or reconstruction, but not enough training pairs are available. In that case the loss function may consist of a supervised and an unsupervised part. An example is an autoencoder, where the part of the network with a supervised loss is attached to the latent space [92]. Another example is the learned SVD proposed in Chapter 6, where two AE are attached via their latent code. The advantage of such semi-supervised architectures over purely supervised ones is that part of the parameters can still be trained if only the unsupervised training signal is available and not a training pair.

As explained in Section 2.3.2, the building blocks of ANNs can be put together in many ways. A noteworthy architectural choice is that of residual blocks [93]. For a block with $\tilde{L}$ layers, this means that besides performing the standard operations in each layer (2.3.1), there is a ‘skip-connection’ between the first and last layer of such a block. This means that the input of a block is added to the output of this block: $h^{\tilde{L}} = \sigma(K^{\tilde{L}-1}(\dots\sigma(K^0h^0 + b^0) + \dots) + b^{\tilde{L}-1}) + h^0$. Such an architecture choice has shown to be very useful for ANNs that need to learn a function close to the identity: the parameters only have to be trained such that they capture the deviation from the identity, instead of learning the entire mapping. A well-known usage of this technique is in the post-processing of imperfect reconstructions in CT [94] or PAT (Chapter 4). A second advantage is a technical one: ANNs with residual blocks suffer less from vanishing gradients. This is the problem where gradients in the backpropagation algorithm get close to zero for very deep networks. This is because gradients are computed by running through many layers, which all contain an activation function that has small gradient far away from zero (tanh and sigmoid) or smaller than zero (ReLU), see Figure 2.3.3.

This section displayed that the effect of the loss function, architecture, activation functions and building blocks are all linked together. In general it is challenging to fully understand all these links, which is why most ANNs are designed by trial-and-error: one takes a combination of components that has shown to be effective for similar problems, after which certain parts are modified for the problem at hand. When the data set is very similar, the ANN is sometimes even pre-trained with the similar data set, after which it is fine-tuned for the actual data set of interest (transfer learning). Nowadays, there are many techniques that allow users to design their ANN more systematically. However, because these techniques arise from several different research fields, it is not obvious how to design the ultimate ANN for a specific problem from scratch. We refer to the books [81, 84, 85] for a general overview of these methods.

2.3.4 Neural networks for solving inverse problems

In this section, we provide an overview of how ANNs can be used to solve inverse problems in imaging, as introduced in Section 2.2. We categorise different methods based on the amount of learning that is required. The amount of learning depends on the way the ANN is embedded in or derived from classical (variational) reconstruction methods (see Section 2.2.2). Although this field of research is only a few years old, it is rapidly growing. For this reason, this section does not provide a comprehensive list of papers, but presents a selection of works that illustrate the different categories of methods. For a more elaborate reference list, we refer to Section 4.1 and Section 6.5.

- a. *Variational methods with learned parameters*: Methods of this type require a relatively small amount of learning. Many reconstruction methods require manual tuning of parameters, such as the regularisation parameter α (see (2.2.2)) or parameters related to the optimisation algorithm, like the step size or acceleration constant. They can be learned by placing the reconstruction method in an ANN architecture and making the parameters trainable. This has been done in [95], where the authors compared this approach with methods that involve more learned elements. This category is related to Chapter 3, where the regularisation parameters and combination of step sizes were chosen manually.
- b. *Variational method with a pre-trained regulariser*: These type of methods are motivated by the problem of manually designing good regularisers for different inverse problems in a variational setting. A network $\mathcal{N} : \mathcal{X} \rightarrow \mathbb{R}^+$ is created and acts as a regulariser, therefore replacing the manually designed counterpart. They are trained prior to the reconstruction procedure and should give a small output value for desired signals and a large value for undesired signals. The networks are designed in such a way that existence, stability and convergence is guaranteed when they are embedded in the variational function. Advantages of these methods are the relatively short training time and the fact that the learned regularisers can be used in different inverse problems without the need to train again. Two methods in this category are NETWORK Tikhonov (NETT) [96] and adversarial regularizers [97].
- c. *Data-consistent post-processing*: In this category, networks $\mathcal{N} : \mathcal{X} \rightarrow \mathcal{X}$ are trained. Post-processing an imperfect reconstruction was demonstrated by Jin *et al.* [94], who showed the effectiveness of improving a FBP reconstruction by an ANN in a CT setup. Although effective in practice, no theoretical certainty of finding a good reconstruction can be given. For linear operators A , this can be dealt with by restricting the post-processing ANN to the nullspace of A [98]. This means data-consistency is preserved, but knowledge of A is necessary for this method. The extension to data-consistent neural networks for nonlinear mappings $\mathcal{A}$ is investigated in Chapter 5, where a convergence analysis with respect to noise is also given.

- d. *Mimicking iterative optimisation in unrolled scheme:* This category considers networks $\mathcal{N} : \mathcal{Y} \rightarrow \mathcal{X}$ that include information on the forward mapping $\mathcal{A}$. Many methods start with a classic iterative reconstruction procedure, and exchange certain parts with an ANN. In some cases, the ANNs within different iterations have shared weights, but this is not necessary. Each iteration requires the application of $\mathcal{A}$, which is computationally intensive. For this reason, the number of ‘unrolled’ iterations is chosen to be small. The LEARN network [99] creates an unrolled version of gradient descent with an L^2 data-fidelity term and a learned regulariser. Variational networks [100] are based on proximal gradient methods, where kernel-function pairs of the regularisation term are learned. The learned primal-dual method [101] makes use of ANNs in both the primal domain $\mathcal{X}$ and the dual domain $\mathcal{Y}$. A Neumann network [102], which mimicks a Neumann series, has the special property that outputs of each iteration are summed, as in a residual neural network. The combination of iterative reconstruction methods, the mapping $\mathcal{A}$ and ANNs provides outstanding reconstruction quality, but theoretical results on convergence and stability are often not available. An exception is [95], where an approach is presented that learns parameters in a general set of convergent schemes involving iterative application of proximal operators. In Chapter 4, the learned primal-dual method is applied to photoacoustic reconstruction and segmentation, where both tasks are simultaneously solved with one iterative ANN.
- e. *Fully learned reconstruction:* A full inversion procedure can be learned with networks $\mathcal{N} : \mathcal{Y} \rightarrow \mathcal{X}$, without using any information of the forward mapping $\mathcal{A}$ that describes the physical process. These networks generally require a large training set, since no other information than training pairs (x, y) are provided. In [103] and [104], patch-based autoencoders are applied for solving the inverse problem of respectively superresolution and blurring. The more challenging problem of tomography reconstruction from undersampled data is considered in [105]. Fully learned reconstruction methods are especially useful if the forward mapping $\mathcal{A}$ is not exactly known or simply does not exist. By choosing specific network architectures, partial information of the forward mapping can be included, for instance via the singular value decomposition. This reduces the amount of learning and may provide theoretical guarantees. In Chapter 6, the learned singular value decomposition (L-SVD) is proposed: a fully learned reconstruction method with the possibility to exchange learned parts for non-learned parts known from the underlying physics.

Unfortunately, this categorisation does not show all subtleties within each category, neither does it show the connections between categories. For instance, if a regulariser is learned before the optimisation procedure, it belongs to

category b, while a regulariser of the same structure that is learned in an unrolled scheme belongs to category d. However, this list provides a rough separation between methods and training strategies that are closer to classical non-learned methods and those that have more learned elements and generally less theoretical guarantees.

Figure 2.3.4: Placement of the thesis chapters in the categorisation that is given in this section. 3) A reconstruction framework for photoacoustic tomography. 4) Partially learned joint reconstruction and segmentation. 5) Data-consistent networks for nonlinear inverse problems. 6) Learned SVD for solving inverse problems.

In Figure 2.3.4, each research chapter of this thesis is assigned to one of the categories. Since Chapter 3 does not contain learned elements, it falls left outside category a. However, the sliding bar indicates that it is easy to add learning of parameters as described above. The sliding bar with Chapter 6 indicates that certain parts of the fully learned network can be replaced with knowledge about the forward mapping $\mathcal{A}$, which is elaborated in Chapter 6.

References

- [1] B. T. Cox, J. G. Laufer, P. C. Beard, and S. R. Arridge, "Quantitative spectroscopic photoacoustic imaging: a review," *Journal of biomedical optics*, vol. 17, no. 6, p. 061202, 2012.
- [2] S. L. Jacques, "Optical properties of biological tissues: a review," *Physics in Medicine & Biology*, vol. 58, no. 11, p. R37, 2013.
- [3] M. Xu and L. V. Wang, "Photoacoustic imaging in biomedicine," *Review of scientific instruments*, vol. 77, no. 4, p. 041101, 2006.
- [4] V. Kisselev, L. Roberti, and G. Perona, "An application of the finite element method to the solution of the radiative transfer equation," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 51, no. 4, pp. 603–614, 1994.
- [5] T. Tarvainen, M. Vauhkonen, V. Kolehmainen, and J. Kaipio, "Finite element model for the coupled radiative transfer equation and diffusion approximation," *International journal for numerical methods in engineering*, vol. 65, no. 3, pp. 383–405, 2006.
- [6] J. Zhao and L. Liu, "Second-order radiative transfer equation and its properties of numerical solution using the finite-element method," *Numerical Heat Transfer, Part B: Fundamentals*, vol. 51, no. 4, pp. 391–409, 2007.
- [7] L. Wang, S. L. Jacques, and L. Zheng, "MCML—Monte Carlo modeling of light transport in multi-layered tissues," *Computer methods and programs in biomedicine*, vol. 47, no. 2, pp. 131–146, 1995.
- [8] Q. Fang and D. A. Boas, "Monte Carlo simulation of photon migration in 3D turbid media accelerated by graphics processing units," *Optics express*, vol. 17, no. 22, pp. 20178–20190, 2009.
- [9] M. Schweiger, S. Arridge, M. Hiraoka, and D. Delpy, "The finite element method for the propagation of light in scattering media: boundary and source conditions," *Medical physics*, vol. 22, no. 11, pp. 1779–1792, 1995.
- [10] J. H. Joseph, W. Wiscombe, and J. Weinman, "The delta-eddington approximation for radiative flux transfer," *Journal of the Atmospheric Sciences*, vol. 33, no. 12, pp. 2452–2459, 1976.
- [11] S. Arridge, M. Schweiger, M. Hiraoka, and D. Delpy, "A finite element approach for modeling photon transport in tissue," *Medical physics*, vol. 20, no. 2, pp. 299–309, 1993.
- [12] R. A. Kruger, P. Liu, Y. R. Fang, and C. R. Appledorn, "Photoacoustic ultrasound (PAUS)—reconstruction tomography," *Medical physics*, vol. 22, no. 10, pp. 1605–1609, 1995.
- [13] P. Kuchment and L. Kunyansky, "Mathematics of photoacoustic and thermoacoustic tomography," *arXiv:0912.2022*, 2009.
- [14] R. Kowar and O. Scherzer, "Attenuation models in photoacoustics," in *Mathematical Modeling in Biomedical Imaging II*, pp. 85–130, Springer, 2012.
- [15] Y. Wang, D. Xing, Y. Zeng, and Q. Chen, "Photoacoustic imaging with deconvolution algorithm," *Physics in Medicine & Biology*, vol. 49, no. 14, p. 3117, 2004.

- [16] J. A. Sethian, *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*, vol. 3. Cambridge university press, 1999.
- [17] R. Willeminck, *Image reconstruction in a passive element enriched photoacoustic tomography setup*. PhD thesis, University of Twente, 2010.
- [18] J. Jose, R. G. Willeminck, W. Steenbergen, C. H. Slump, T. G. van Leeuwen, and S. Manohar, "Speed-of-sound compensated photoacoustic tomography for accurate imaging," *Medical physics*, vol. 39, no. 12, pp. 7262–7271, 2012.
- [19] B. Dong, C. Sun, and H. F. Zhang, "Optical detection of ultrasound in photoacoustic imaging," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 1, pp. 4–15, 2016.
- [20] K. K. Shung, J. Cannata, and Q. Zhou, "Piezoelectric materials for high frequency medical imaging applications: A review," *Journal of Electroceramics*, vol. 19, no. 1, pp. 141–147, 2007.
- [21] G. Paltauf, R. Nuster, M. Haltmeier, and P. Burgholzer, "Photoacoustic tomography using a Mach-Zehnder interferometer as an acoustic line detector," *Applied optics*, vol. 46, no. 16, pp. 3352–3358, 2007.
- [22] E. Zhang, J. Laufer, and P. Beard, "Backward-mode multiwavelength photoacoustic scanner using a planar Fabry-Perot polymer film ultrasound sensor for high-resolution three-dimensional imaging of biological tissues," *Applied optics*, vol. 47, no. 4, pp. 561–577, 2008.
- [23] Y. Zhang, Y. Wang, and C. Zhang, "Total variation based gradient descent algorithm for sparse-view photoacoustic image reconstruction," *Ultrasonics*, vol. 52, no. 8, pp. 1046–1055, 2012.
- [24] D.-H. Huang, C.-K. Liao, C.-W. Wei, and P.-C. Li, "Simulations of optoacoustic wave propagation in light-absorbing media using a finite-difference time-domain method," *The Journal of the Acoustical Society of America*, vol. 117, no. 5, pp. 2795–2801, 2005.
- [25] Y.-L. Sheu and P.-C. Li, "Simulations of photoacoustic wave propagation using a finite-difference time-domain method with berenger's perfectly matched layers," *The Journal of the Acoustical Society of America*, vol. 124, no. 6, pp. 3471–3480, 2008.
- [26] T. D. Mast, L. P. Souriau, D.-L. Liu, M. Tabei, A. I. Nachman, and R. C. Waag, "A k-space method for large-scale models of wave propagation in tissue," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 48, no. 2, pp. 341–354, 2001.
- [27] B. E. Treeby and B. T. Cox, "k-Wave: Matlab toolbox for the simulation and reconstruction of photoacoustic wave fields," *Journal of biomedical optics*, vol. 15, no. 2, p. 021314, 2010.
- [28] M. Haltmeier, "Universal inversion formulas for recovering a function from spherical means," *SIAM Journal on Mathematical Analysis*, vol. 46, no. 1, pp. 214–232, 2014.
- [29] Y. Xu, D. Feng, and L. V. Wang, "Exact frequency-domain reconstruction for thermoacoustic tomography. i. planar geometry," *IEEE transactions on medical imaging*, vol. 21, no. 7, pp. 823–828, 2002.

- [30] Y. Xu, M. Xu, and L. V. Wang, "Exact frequency-domain reconstruction for thermoacoustic tomography. ii. cylindrical geometry," *IEEE transactions on medical imaging*, vol. 21, no. 7, pp. 829–833, 2002.
- [31] R. J. Schulze, O. Scherzer, G. Zangerl, M. Holotta, D. Meyer, F. Handle, R. Nuster, and G. Paltauf, "On the use of frequency-domain reconstruction algorithms for photoacoustic imaging," *Journal of biomedical optics*, vol. 16, no. 8, p. 086002, 2011.
- [32] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, vol. 375. Springer Science & Business Media, 1996.
- [33] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen, *Variational methods in imaging*. Springer, 2009.
- [34] A. Kirsch, *An introduction to the mathematical theory of inverse problems*, vol. 120. Springer Science & Business Media, 2011.
- [35] J. Kaipio and E. Somersalo, *Statistical and computational inverse problems*, vol. 160. Springer Science & Business Media, 2006.
- [36] A. M. Stuart, "Inverse problems: a Bayesian perspective," *Acta numerica*, vol. 19, pp. 451–559, 2010.
- [37] M. Bertero and P. Boccacci, *Introduction to inverse problems in imaging*. CRC press, 1998.
- [38] V. Isakov, *Inverse problems for partial differential equations*, vol. 127. Springer, 2006.
- [39] C. R. Vogel, *Computational methods for inverse problems*, vol. 23. SIAM, 2002.
- [40] G. Aubert and P. Kornprobst, *Mathematical problems in image processing: partial differential equations and the calculus of variations*, vol. 147. Springer Science & Business Media, 2006.
- [41] J. Hadamard, *Lectures on Cauchy's problem in linear partial differential equations*. Courier Corporation, 1923.
- [42] G. Aubert and J.-F. Aujol, "A variational approach to removing multiplicative noise," *SIAM journal on applied mathematics*, vol. 68, no. 4, pp. 925–946, 2008.
- [43] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: nonlinear phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.
- [44] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 492–526, 2010.
- [45] N. Kingsbury, "The dual-tree complex wavelet transform: a new efficient tool for image restoration and enhancement," in *9th European Signal Processing Conference (EUSIPCO 1998)*, pp. 1–4, IEEE, 1998.
- [46] J. Nocedal and S. Wright, *Numerical optimization*. Springer Science & Business Media, 2006.
- [47] X. Chen, "Smoothing methods for nonsmooth, nonconvex minimization," *Mathematical programming*, vol. 134, no. 1, pp. 71–99, 2012.
- [48] E. Esser, X. Zhang, and T. F. Chan, "A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science," *SIAM Journal on Imaging Sciences*, vol. 3, no. 4, pp. 1015–1046, 2010.

- [49] T. Tarvainen, A. Pulkkinen, B. T. Cox, J. P. Kaipio, and S. R. Arridge, "Bayesian image reconstruction in quantitative photoacoustic tomography," *IEEE transactions on medical imaging*, vol. 32, no. 12, pp. 2287–2298, 2013.
- [50] N. Song, C. Deumie, and A. Da Silva, "Considering sources and detectors distributions for quantitative photoacoustic tomography," *Biomedical optics express*, vol. 5, no. 11, pp. 3960–3974, 2014.
- [51] M. Haltmeier, L. Neumann, and S. Rabanser, "Single-stage reconstruction algorithm for quantitative photoacoustic tomography," *Inverse Problems*, vol. 31, no. 6, p. 065005, 2015.
- [52] F. Natterer, *The mathematics of computerized tomography*. SIAM, 2001.
- [53] D. Finch and S. K. Patch, "Determining a function from its mean values over a family of spheres," *SIAM journal on mathematical analysis*, vol. 35, no. 5, pp. 1213–1240, 2004.
- [54] A. Rosenthal, V. Ntziachristos, and D. Razansky, "Acoustic inversion in optoacoustic tomography: a review," *Current Medical Imaging*, vol. 9, no. 4, pp. 318–336, 2013.
- [55] S. R. Arridge, M. M. Betcke, B. T. Cox, F. Lucka, and B. E. Treeby, "On the adjoint operator in photoacoustic tomography," *Inverse Problems*, vol. 32, no. 11, p. 115012, 2016.
- [56] Z. Yuan, Q. Zhang, and H. Jiang, "Simultaneous reconstruction of acoustic and optical properties of heterogeneous media by quantitative photoacoustic tomography," *Optics express*, vol. 14, no. 15, pp. 6749–6754, 2006.
- [57] A. Kirsch and O. Scherzer, "Simultaneous reconstructions of absorption density and wave speed with photoacoustic measurements," *SIAM Journal on Applied Mathematics*, vol. 72, no. 5, pp. 1508–1523, 2012.
- [58] H. Liu and G. Uhlmann, "Determining both sound speed and internal source in thermo-and photo-acoustic tomography," *Inverse Problems*, vol. 31, no. 10, p. 105005, 2015.
- [59] C. Huang, K. Wang, R. W. Schoonover, L. V. Wang, and M. A. Anastasio, "Joint reconstruction of absorbed optical energy density and sound speed distributions in photoacoustic computed tomography: a numerical investigation," *IEEE transactions on computational imaging*, vol. 2, no. 2, pp. 136–149, 2016.
- [60] T. P. Matthews, J. Poudel, L. Li, L. V. Wang, and M. A. Anastasio, "Parameterized joint reconstruction of the initial pressure and sound speed distributions for photoacoustic computed tomography," *SIAM journal on imaging sciences*, vol. 11, no. 2, pp. 1560–1588, 2018.
- [61] X. L. Dean-Ben, R. Ma, D. Razansky, and V. Ntziachristos, "Statistical approach for optoacoustic image reconstruction in the presence of strong acoustic heterogeneities," *IEEE transactions on medical imaging*, vol. 30, no. 2, pp. 401–408, 2010.
- [62] J. Tick, A. Pulkkinen, and T. Tarvainen, "Modelling of errors due to speed of sound variations in photoacoustic tomography using a Bayesian framework," *Biomedical Physics & Engineering Express*, vol. 6, no. 1, p. 015003, 2019.
- [63] T. Sahlström, A. Pulkkinen, J. Tick, J. Leskinen, and T. Tarvainen, "Modeling of errors due to uncertainties in ultrasound sensor locations in photoacoustic tomography," *IEEE Transactions on Medical Imaging*, 2020.

- [64] S. Lunz, A. Hauptmann, T. Tarvainen, C.-B. Schönlieb, and S. Arridge, "On learned operator correction," *arXiv:2005.07069*, 2020.
- [65] B. T. Cox, S. R. Arridge, K. P. Köstli, and P. C. Beard, "Two-dimensional quantitative photoacoustic image reconstruction of absorption distributions in scattering media by use of a simple iterative method," *Applied Optics*, vol. 45, no. 8, pp. 1866–1875, 2006.
- [66] B. Cox, S. Arridge, and P. Beard, "Gradient-based quantitative photoacoustic image reconstruction for molecular imaging," in *Photons Plus Ultrasound: Imaging and Sensing 2007: The Eighth Conference on Biomedical Thermoacoustics, Optoacoustics, and Acousto-optics*, vol. 6437, p. 64371T, International Society for Optics and Photonics, 2007.
- [67] B. T. Cox, S. R. Arridge, and P. C. Beard, "Estimating chromophore distributions from multiwavelength photoacoustic images," *JOSA A*, vol. 26, no. 2, pp. 443–455, 2009.
- [68] G. Bal and K. Ren, "On multi-spectral quantitative photoacoustic tomography in diffusive regime," *Inverse Problems*, vol. 28, no. 2, p. 025010, 2012.
- [69] G. Bal and G. Uhlmann, "Inverse diffusion theory of photoacoustics," *Inverse Problems*, vol. 26, no. 8, p. 085010, 2010.
- [70] W. Naetar and O. Scherzer, "Quantitative photoacoustic tomography with piecewise constant material parameters," *SIAM Journal on Imaging Sciences*, vol. 7, no. 3, pp. 1755–1774, 2014.
- [71] L. Yao, Y. Sun, and H. Jiang, "Transport-based quantitative photoacoustic tomography: simulations and experiments," *Physics in Medicine & Biology*, vol. 55, no. 7, p. 1917, 2010.
- [72] T. Tarvainen, B. T. Cox, J. Kaipio, and S. R. Arridge, "Reconstructing absorption and scattering distributions in quantitative photoacoustic tomography," *Inverse Problems*, vol. 28, no. 8, p. 084009, 2012.
- [73] T. Saratoon, T. Tarvainen, B. Cox, and S. Arridge, "A gradient-based method for quantitative photoacoustic tomography using the radiative transfer equation," *Inverse Problems*, vol. 29, no. 7, p. 075006, 2013.
- [74] A. De Cezaro, F. T. De Cezaro, and J. S. Suarez, "Regularization approaches for quantitative photoacoustic tomography using the radiative transfer equation," *Journal of Mathematical Analysis and Applications*, vol. 429, no. 1, pp. 415–438, 2015.
- [75] J. Buchmann, B. A. Kaplan, S. Powell, S. Prohaska, and J. Laufer, "Three-dimensional quantitative photoacoustic tomography using an adjoint radiance Monte Carlo model and gradient descent," *Journal of biomedical optics*, vol. 24, no. 6, p. 066001, 2019.
- [76] G. Bal, A. Jollivet, and V. Jugnon, "Inverse transport theory of photoacoustics," *Inverse Problems*, vol. 26, no. 2, p. 025011, 2010.
- [77] Y. Li, "Deep reinforcement learning: an overview," *arXiv:1701.07274*, 2017.
- [78] C. M. Bishop, *Pattern recognition and machine learning*. springer, 2006.
- [79] K. P. Murphy, *Machine learning: a probabilistic perspective*. MIT press, 2012.
- [80] K. Gurney, *An introduction to neural networks*. CRC press, 1997.

- [81] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.
- [82] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics*, vol. 5, no. 4, pp. 115–133, 1943.
- [83] R. Lippmann, "An introduction to computing with neural nets," *IEEE Assp magazine*, vol. 4, no. 2, pp. 4–22, 1987.
- [84] M. A. Nielsen, *Neural networks and deep learning*, vol. 2018. Determination press San Francisco, CA, 2015.
- [85] C. C. Aggarwal *et al.*, *Neural networks and deep learning*. Springer, 2018.
- [86] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for activation functions," *arXiv:1710.05941*, 2017.
- [87] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, pp. 2672–2680, 2014.
- [88] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization.," *Journal of machine learning research*, vol. 12, no. 7, 2011.
- [89] T. Tieleman and G. Hinton, "Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude," *COURSERA: neural networks for machine learning*, vol. 4, no. 2, pp. 26–31, 2012.
- [90] D. P. Kingma and J. L. Ba, "Adam: a method for stochastic optimization," *ICLR 2015*, pp. 1–15, 2015.
- [91] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," *arXiv:1312.6114*, 2013.
- [92] L. L. Zeune, Y. E. Boink, G. van Dalum, A. Nanou, S. de Wit, K. C. Andree, J. F. Swennenhuis, S. A. van Gils, L. W. Terstappen, and C. Brune, "Deep learning of circulating tumour cells," *Nature Machine Intelligence*, vol. 2, no. 2, pp. 124–133, 2020.
- [93] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.
- [94] 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, 2017.
- [95] S. Banert, A. Ringh, J. Adler, J. Karlsson, and O. Oktem, "Data-driven nonsmooth optimization," *SIAM Journal on Optimization*, vol. 30, no. 1, pp. 102–131, 2020.
- [96] H. Li, J. Schwab, S. Antholzer, and M. Haltmeier, "NETT: solving inverse problems with deep neural networks," *Inverse Problems*, 2020.
- [97] S. Lunz, O. Öktem, and C.-B. Schönlieb, "Adversarial regularizers in inverse problems," in *Advances in Neural Information Processing Systems*, pp. 8507–8516, 2018.
- [98] J. Schwab, S. Antholzer, and M. Haltmeier, "Deep null space learning for inverse problems: convergence analysis and rates," *Inverse Problems*, vol. 35, no. 2, p. 025008, 2019.

- [99] H. Chen, Y. Zhang, Y. Chen, J. Zhang, W. Zhang, H. Sun, Y. Lv, P. Liao, J. Zhou, and G. Wang, "LEARN: learned experts' assessment-based reconstruction network for sparse-data CT," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1333–1347, 2018.
- [100] E. Kobler, T. Klatzer, K. Hammernik, and T. Pock, "Variational networks: connecting variational methods and deep learning," in *German conference on pattern recognition*, pp. 281–293, Springer, 2017.
- [101] J. Adler and O. Öktem, "Learned primal-dual reconstruction," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1322–1332, 2018.
- [102] D. Gilton, G. Ongie, and R. Willett, "Neumann networks for inverse problems in imaging," *arXiv:1901.03707*, 2019.
- [103] K. Zeng, J. Yu, R. Wang, C. Li, and D. Tao, "Coupled deep autoencoder for single image super-resolution," *IEEE transactions on cybernetics*, vol. 47, no. 1, pp. 27–37, 2017.
- [104] K. Gupta, B. Bhowmick, and A. Majumdar, "Motion blur removal via coupled autoencoder," in *2017 IEEE International Conference on Image Processing (ICIP)*, pp. 480–484, IEEE, 2017.
- [105] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, "Image reconstruction by domain-transform manifold learning," *Nature*, vol. 555, no. 7697, pp. 487–492, 2018.

Chapter 3

A framework for directional and higher-order reconstruction in photoacoustic tomography

Yoei E. Boink, Marinus J. Lagerwerf, Wiendelt Steenbergen,
Stephan A. van Gils, Srirang Manohar, Christoph Brune

Physics in Medicine & Biology 63.4 (2018): 045018

Abstract

Photoacoustic tomography is a hybrid imaging technique that combines high optical tissue contrast with high ultrasound resolution. Direct reconstruction methods such as filtered backprojection, time reversal and least squares suffer from curved line artefacts and blurring, especially in case of limited angles or strong noise. In the past years, there has been great interest in regularised iterative methods. These methods employ prior knowledge on the image to provide higher quality reconstructions. However, easy comparisons between regularisers and their properties are limited, since many tomography implementations heavily rely on the specific regulariser chosen. To overcome this bottleneck, we present a modular reconstruction framework for photoacoustic tomography. It enables easy comparisons between regularisers with different properties, e.g. nonlinear, higher-order or directional. We solve the underlying minimisation problem with an efficient first-order primal-dual algorithm. Convergence rates are optimised by choosing an operator dependent preconditioning strategy. A variety of reconstruction methods are tested on challenging 2D synthetic and experimental data sets. They outperform direct reconstruction approaches for strong noise levels and limited angle

measurements, offering immediate benefits in terms of acquisition time and quality. This chapter provides a basic platform for the investigation of future advanced regularisation methods in photoacoustic tomography.

3.1 Introduction

Photoacoustic tomography (PAT), also known as optoacoustic tomography, is an intrinsically hybrid imaging technique that combines the high spectroscopic contrast of tissue constituents to light, with the high resolution of ultrasound imaging techniques. Tissue is illuminated with nanosecond laser pulses, causing some of the optical energy to be absorbed and converted into heat. This leads to thermoelastic expansion and creates ultrasound waves that are detected at the boundary. These detected signals are employed to reconstruct the spatial distribution of absorbed optical energy inside the tissue. Common methods applied are filtered backprojection (FBP) [1, 2, 3, 4, 5, 6, 7, 8, 9] or time reversal [10, 11, 12], which was first proposed by Finch and Patch [2]. In this chapter, we develop a general variational framework that incorporates prior knowledge on the reconstruction, offering higher quality reconstructions than direct methods and providing robustness against noise and compressive sampling.

The optical absorption coefficient in tissue varies due to the relative presence of various components such as haemoglobin, oxyhaemoglobin, melanin, lipids and water. These chromophores have specific spectral signatures, and the use of multi-wavelength PAT can potentially reveal molecular specific information. This is directly linked to tissue physiopathology and function and has diagnostic potential. Photoacoustics is being researched for applications in various fields of biomedicine [13] such as brain imaging in small animals [14], breast cancer imaging in humans [15, 16], and imaging inflamed human synovial joints in rheumatoid arthritis [17, 18]. In most of these applications, detection of signals from haemoglobin and oxyhaemoglobin enables the visualisation of blood vessels. Several disorders are characterised by a dysregulation of blood vessel function, but also with uncontrolled creation of blood vessels or angiogenesis.

The standard methods of filtered backprojection (FBP) and time reversal suffer from curved line artefacts, especially when the noise level is high or the placement of detectors for measurements (sampling) is coarse [19]. In the past decade, there has been intense interest in solving the PAT reconstruction problem iteratively with a specific focus on regularised reconstruction. The regularisers used for these reconstructions have many different forms, depending on the prior assumptions on the image. Total variation (TV) regularisation is a powerful tool for noise removal [20], and generates the desired images with sharp edges. In [21], TV-regularised reconstructions on simulated data are presented, where an analytical adjoint operator for the k-space forward operator [12] is derived to model wave propagation. In [22] a similar reconstruction method is applied to compressively sampled simulated and experimental data. In [23], it

has been shown that the TV-basis and the measurement basis (spherical Radon transform) are not very incoherent, therefore it is not obvious to use the TV-basis as sparsifying transform. In practice however, the TV-basis appears to work well in the area of compressive sampling for the reconstruction of an absorbed energy density map with abrupt edges [24]. In two more recent works [25, 26], a different compressed sensing approach for PAT has been proposed: instead of taking less measurements, multiple pressure measurements have been added up to obtain the desired compression. Reconstruction proofs were given for a temporal transform that is sparse in space for any image that consists of piecewise constant regular domains. Other works that combine total variation with compressive sampling in PAT can be found in [27] and [28].

In case of a heterogeneous fluence rate, such as an exponentially decaying one, a piecewise constant prior on the reconstruction is not realistic. Higher-order total variation methods, such as total generalised variation (TGV) [29] are better suited to deal with such data. Tomographic algorithms with TGV regularisation have shown to be computationally feasible and give higher quality reconstructions for images that show approximate piecewise linear behaviour. Examples can be found in MRI [30], CT [31, 32], ultrasound waveform tomography [33] and diffusion tensor imaging [34]. Image denoising with TGV as a post-processing step has made its way into optical coherence tomography [35] and optical diffraction tomography [36]. However, solving these optical reconstruction problems with TGV regularisation within the reconstruction algorithm is still an open research topic. A recent conference proceeding [37] shows TGV-regularised reconstructions in PAT on vascular data. Here, measurements are taken by making snapshots of the acoustic waves using a CCD camera. Vascular structures show a strong anisotropy, that can be promoted by using directional wavelets [38] or curvelets [39] as ‘building blocks’ for the reconstruction. Provost and Lesage [23] showed that wavelets and curvelets are very sparse representations of the measurement data, and can help to recover anisotropic features in the reconstruction. This has been confirmed through new phantom and in vivo experiments in [24].

Several algorithms can be employed when a specific regularised reconstruction method is needed: a differentiable minimisation problem can for instance be solved using a smooth solver, such as gradient descent; problems containing nonsmooth regularisation terms such as TV or sparse wavelet regularisation can be solved using forward-backward splitting or proximal gradient descent [21]. Primal-dual algorithms have been employed for solving TGV-regularised problems [40], [37]. In [41], a series of example algorithms was given for CT reconstruction using a positivity constraint or total variation as regularisation term. The extension to infimal convolution type regularisers, such as TGV, was not given.

In this chapter, we give a variational reconstruction framework that is flexible to the specifications of the measurement system and the assumptions on the tissue to be imaged. Its applicability is demonstrated by using the framework

for three reconstruction models, which contain TV, TGV and directional wavelet regularisers. While others based their methods on specific measurement geometries, our framework makes no assumption on the geometry, requires no specific physics modelling and can be applied to the PAT reconstruction problem in a 2- or 3-dimensional setting. We develop not only a TV, but also a TGV regularised reconstruction model, which can deal with the more realistic assumption of a decaying heterogeneous fluence rate. While in [37], TGV has only been tested on vascular structures, we also quantitatively explore its usefulness in combination with other types of synthetic and experimental data. A reconstruction model using directional wavelet regularisation is developed for images containing anisotropic vascular structures. The numerical implementation with a primal-dual algorithm is a modular one: data-fidelities and regularisers can be chosen differently without having to change the structure of the algorithm. The algorithm has a convergence rate that is at least linear and that can be accelerated [42] for specific choices of the regulariser.

To summarise, we provide a framework for the photoacoustics community that extends the algorithms of Sidky *et al.*[41], such that infimal convolution type regularisers such as ICTV and TGV can be used within the same framework. We overcome typical algorithmic problems that arise in a tomography setup, such as large differences between the various operator norms, by using algorithm parameter selection. Besides explicitly describing the modular algorithm, we provide a number of regularisation options that naturally arise from the physics in PAT. The advantage of using regularised methods for PAT is quantified through many synthetic and experimental test cases: regularised reconstruction methods in this framework consistently give higher quality reconstructions under noise and compressive sampling, compared to direct methods.

The remainder of this chapter is organised as follows: in Section 3.2 the variational method for the regularised reconstruction of PAT is derived. After writing the reconstruction problem as a saddle point problem, the numerical implementation with the first-order primal-dual algorithm PDHGM is explained in Section 3.3. In Section 3.4, the experimental setup that is employed to generate our results is explained. Moreover, the specific forward model that is used for obtaining the synthetic data is derived. The digital and experimental phantoms that are created as test cases for our reconstruction framework are shown in Section 3.5. After an extensive analysis of the method on challenging synthetic data in Section 3.6 and experimental data in Section 3.7, we conclude and discuss the results in Section 3.8.

3.2 Variational methods using regularisation

We consider the image reconstruction problem of finding an estimate for the absorbed energy density map $x \in L^2(\Omega)$ from given (preprocessed) measurements $y^\delta \in L^2(\Sigma)$. Here $\Omega \subset \mathbb{R}^d$, where d is the dimension of the

space in which x must lie and $\Sigma = S \times (0, T)$ is the space-time domain of the detection surface S between times 0 and T . Mathematically, the inverse problem can be formulated as the solution to the equation

$$y^\delta = Ax + \eta^\delta, \quad (3.2.1)$$

where η^δ is an additive noise term with a noise level δ that is estimated and $A : L^2(\Omega) \mapsto L^2(\Sigma)$ is a bounded linear operator. This operator can be described in many ways, since there are multiple methods that model the forward process in photoacoustic imaging. For instance, it can be described as a k-space method that models linear acoustic propagation in time [12]. In our simulations and experiments, we use the spherical mean operator [1] to model the forward process.

Instead of using a direct reconstruction method such as FBP, the inverse problem is solved in a variational setting. We consider the following minimisation problem:

$$\min_{x \in L^2(\Omega)} \left\{ \frac{1}{2} \|Ax - y^\delta\|_{L^2(\Sigma)}^2 + \alpha R(x) \right\}. \quad (3.2.2)$$

The first term is the data-fidelity term, which ensures that the reconstructed image, after applying the forward operator A , is close to the noisy data y^δ . The L^2 -norm is chosen because the photoacoustic measurements are predominantly affected by noise from the system electronics and thermal noise from the transducers, which are known to be of additive Gaussian type [27]. The L^2 data-fidelity term can be exchanged for different types, for instance an L^1 -norm in case of strong data outliers (salt-and-pepper noise), or a Kullback-Leibler divergence in case of Poisson distributed noise. The second term is an α weighted regularisation term which becomes large when our prior assumptions on x are violated.

Among the regularisers that can be applied within our framework are linear convex operators that are bounded in the discrete setup. This is for instance the case for $R(x) = \|Kx\|_p^p$, with K a bounded linear operator and $p \in [1, 2]$. But also Huber-norms [43], which interpolate between L^1 - and L^2 -norms are possible. Furthermore, frame-based regularisation can be applied, where Kx represents the transform to a (tight) frame. This is demonstrated in Section 3.3 by making use of directional wavelet regularisation. Moreover, it is possible to promote sparsity with the use of learned dictionaries, if x in the data-fidelity is replaced with Pw and $\|w\|_1$ is used as a regulariser. In this case, P is a dictionary matrix that stores all learned elements, and w is a vector that tells which elements should be used on which location. Finally, regularisation with an infimal convolution type operator, such as ICTV [44] and TGV [29], can also be described within the framework, as is demonstrated with TGV in Section 3.3. Other regularisation choices generally depend on the requirements of the PDHGM algorithm, which is explained in Section 3.3.3.

Note that if there is no prior on the image and data and therefore $R(x) = 0$, the least squares (LS) reconstruction is obtained for (3.2.2). The LS reconstruction can be evaluated by using a gradient descent algorithm or, if A is a small enough matrix operator, by solving $x = (A^T A)^{-1} A^T y^\delta$ directly. However, since LS is very sensitive to noise in the data y^δ , Tikhonov regularised least squares (LS-T) is often used: the term $R(x)$ in (3.2.2) then takes the form $\frac{1}{2} \|x\|_{L^2(\Omega)}^2$. This can again be evaluated directly by solving $x = (A^T A + \alpha)^{-1} A^T y^\delta$. In Section 3.6, our models will be compared with LS-T.

3.2.1 Regularisation with total (generalised) variation

This chapter makes use of total generalised variation (TGV), as introduced by Bredies *et al.* [29]. TGV is a generalisation of TV, in which higher order derivatives are considered instead of only first order derivatives in TV. Bredies *et al.* proposed the following definition:

$$\text{TGV}_\beta^k(x) = \sup_q \left\{ \int_\Omega x \operatorname{div}^k(q) \, d\mathbf{r} \mid q \in \mathcal{C}_0^k(\Omega, \operatorname{Sym}^k(\mathbb{R}^d)), \right. \\ \left. \|\operatorname{div}^l(q)\|_\infty \leq \beta_l, \, l = 0, \dots, k-1 \right\}. \quad (3.2.3)$$

With the choice of k , the desired order of regularity can be obtained, while the parameter β gives a weight on every regularity level. By choosing $k = 1$ and $\beta = 1$, the definition of TGV coincides with the definition of TV

$$\text{TV}(x) = \sup_q \left\{ \int_\Omega x \operatorname{div}(q) \, d\mathbf{r} \mid q \in \mathcal{C}_0^1(\Omega, \mathbb{R}^d), \, \|q\|_\infty \leq 1 \right\}. \quad (3.2.4)$$

This chapter considers $\text{TV}(x)$ and $\text{TGV}_\beta^2(x)$ as regularisers. In our applications, the desired image x is a discretised one, on which a gradient or derivative is always numerically defined. Therefore, the assumption $x \in \mathcal{C}^1(\Omega)$ or $x \in \mathcal{C}^2(\Omega)$ is made, so that the expressions for $\text{TV}(x)$ and $\text{TGV}_\beta^2(x)$ respectively can be simplified. By substituting these expressions in (3.2.2) the following minimisation problems are obtained:

$$x_{\text{TV}} = \operatorname{argmin}_{x \in \text{BV}(\Omega)} \left\{ \frac{1}{2} \|Ax - y^\delta\|_{L^2(\Sigma)}^2 + \alpha \|\nabla x\|_{(L^1(\Omega))^d} \right\}, \quad (3.2.5)$$

$$x_{\text{TGV}} = \operatorname{argmin}_{x, z \in \text{BV}(\Omega)} \left\{ \frac{1}{2} \|Ax - y^\delta\|_{L^2(\Sigma)}^2 + \alpha \{\|\nabla x - z\|_{(L^1(\Omega))^d} \right. \\ \left. + \beta \|\mathcal{E}(z)\|_{(L^1(\Omega))^{d \times d}}\} \right\}, \quad (3.2.6)$$

where $\text{BV}(\Omega) = \{x \in L^2(\Omega) \mid \text{TV}(x) < \infty\}$ and $\mathcal{E}$ is the symmetrised gradient [29]. In the TV-regularised functional (3.2.5), the parameter α determines the smoothness of the solution. In the TGV-regularised functional (3.2.6) the

influence of first order and second order smoothness can be chosen by choosing combinations of α and β . The effect of different combinations of α and β on the reconstruction of one-dimensional TV and TGV eigenfunctions was analysed in [40]. In the remainder of this chapter, we write TGV instead of TGV_{β}^2 for conciseness.

3.2.2 Regularisation with directional wavelets

Blood vessel visualisation is an important aspect for the application of photoacoustic imaging in biomedicine. An image containing vascular structures can be sparsely represented by only a small number of elements in some basis or in a so called dictionary. There is a large amount of possibilities to represent this anisotropic data. Rubinstein *et al.*[45] made a distinction between analytic transforms, such as the Gabor and wavelet transform; analytic dictionaries, such as curvelets and contourlets; and trained dictionaries, such as sparse PCA and K-SVD.

The discrete wavelet transform is in essence a one-dimensional transform and the obvious extension to a multi-dimensional setting is not satisfactory, since it lacks directionality. In this chapter, we make use of the dual-tree wavelet transform [38], which has directionally selective filters and is thus better able to represent multi-dimensional directional data. However, since our framework is very general, it is easy to replace it with a transform or dictionary that suits the data best. We propose the following minimisation problem

$$x_{\text{wav}} = \underset{x}{\operatorname{argmin}} \left\{ \frac{1}{2} \|Ax - y^{\delta}\|_{L^2(\Sigma)}^2 + \alpha \|W(x)\|_{L^1(\mathcal{W})} \right\}, \quad (3.2.7)$$

where W is the complex dual-tree wavelet transform and $\mathcal{W}$ is the corresponding wavelet domain. Note the close similarity in structure between (3.2.5) and (3.2.7). Because of this similarity, we will not elaborate on the wavelet case in the numerical treatment of our models, but only mention the small changes that have to be made to the TV implementation to obtain the wavelet implementation.

3.3 Numerical framework and algorithm

In Section 3.2 the convex, possibly nonsmooth minimisation problem (3.2.2) was considered. TV and TGV are defined via a maximisation problem over their dual variable(s), which can be seen in (3.2.3) and (3.2.4). Substituting these definitions of TV and TGV in the minimisation problem changes it directly into a saddle point problem. However, also a general minimisation problem, such as (3.3.1) or (3.2.7), can be rewritten to a saddle point problem with the use of Fenchel duality: the general formulation

$$\min_{u \in \mathcal{U}} \{F(Bu) + G(u)\} \quad (3.3.1)$$

changes to the saddle point formulation

$$\min_{u \in \mathcal{U}} \max_{v \in \mathcal{V}} \{ \langle Bu, v \rangle + G(u) - F^*(v) \}, \quad (3.3.2)$$

where F^* is the convex conjugate of F .

Primal-dual algorithms are suited to solve such saddle point problems: by iteratively making a proximal descent step for the primal variable u and a proximal ascent step for the dual variable v , the solution is obtained. General forms of algorithms that can be used to solve saddle point problems are found in [46, 47], where clear connections between the algorithms are uncovered.

3.3.1 Description of the specific saddle point problem

The saddle point formulations for the TV, TGV and wavelet model are obtained by taking the general form of (3.3.2), setting $G = 0$ and write the functional F and operator B as follows:

$$\begin{aligned} \textbf{TV model: } v &= (q, r), \quad u = x, \quad Bu = (Ax, \nabla x), \\ F_1(\tilde{q}) &= \frac{1}{2} \|\tilde{q} - y^\delta\|_{L^2(\Sigma)}, \quad F_2(\tilde{r}) = \alpha \|\tilde{r}\|_{(L^1(\Omega))^d}, \\ \min_{x \in \mathcal{U}} \max_{(q, r) \in \mathcal{V}} \{ &\langle Ax, q \rangle + \langle \nabla x, r \rangle - F_1^*(q) - F_2^*(r) \}. \end{aligned} \quad (3.3.3)$$

Note: the wavelet model is obtained by changing ∇ to W , causing F_2 to change to $F_2(\tilde{r}) = \alpha \|\tilde{r}\|_{L^1(\mathcal{W})}$.

$$\begin{aligned} \textbf{TGV model: } v &= (q, r, s), \quad u = (x, z), \quad Bu = (Ax, \nabla x - z, \mathcal{E}z), \\ F_1(\tilde{q}) &= \frac{1}{2} \|\tilde{q} - y^\delta\|_{L^2(\Sigma)}, \quad F_2(\tilde{r}) = \alpha \|\tilde{r}\|_{(L^1(\Omega))^d}, \quad F_3(\tilde{s}) = \alpha \beta \|\tilde{s}\|_{(L^1(\Omega))^{d \times d}}, \\ \min_{(x, z) \in \mathcal{U}} \max_{(q, r, s) \in \mathcal{V}} \{ &\langle Ax, q \rangle + \langle \nabla x - z, r \rangle + \langle \mathcal{E}z, s \rangle - F_1^*(q) - F_2^*(r) - F_3^*(s) \}. \end{aligned} \quad (3.3.4)$$

Here the splitting $F(\cdot) = \sum_i F_i(\cdot)$ was made, similar to [41]. Note that tildes are used when the operators are in their convex conjugate form in the primal-dual problem. The convex conjugate, also known as Fenchel-Legendre transform, is defined as

$$F^*(v) = \max_{\tilde{v}} \{ \langle v, \tilde{v} \rangle - F(\tilde{v}) \}. \quad (3.3.5)$$

3.3.2 Proximal operators for TV, TGV and wavelet model

A key tool for the use of primal-dual algorithms is the proximal splitting method. The proximal operator for a general scaled convex functional $\gamma H(u)$ is defined as

$$\text{prox}_{\gamma H}(u) = \underset{\hat{u}}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\hat{u} - u\|_2^2 + \gamma F(\hat{u}) \right\}, \quad (3.3.6)$$

which gives a small value of $\gamma H(\hat{u})$, while keeping $\hat{u}$ close to u . The right hand side consists of a strongly convex quadratic part and $\gamma F(\hat{u})$, which is convex. Therefore, the right hand side is strongly convex thus (3.3.6) has a unique solution.

As can be seen in Section 3.3.3, the saddle point problems (3.3.3) and (3.3.4) now have the correct structure to be numerically implemented. Therefore, the proximal operators for all F_i^* 's are needed, which are derived below. The convex conjugate for F_1 is

$$F_1^*(q) = \max_{\tilde{q}} \left\{ \langle q, \tilde{q} \rangle - \frac{1}{2} \|\tilde{q} - y^\delta\|_{L^2(\Sigma)}^2 \right\} = \frac{1}{2} \|q\|_{L^2(\Omega)}^2 + \langle q, y^\delta \rangle. \quad (3.3.7)$$

The last equality was obtained by checking the first and second order optimality conditions. Using (3.3.7), we calculate the proximal operator

$$\text{prox}_{\gamma F_1^*}(q) = \underset{\hat{q}}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\hat{q} - q\|_{L^2(\Omega)}^2 + \gamma F_1^*(\hat{q}) \right\} = \frac{q - \gamma y^\delta}{1 + \gamma}. \quad (3.3.8)$$

The convex conjugate for F_2 is

$$F_2^*(r) = \max_{\tilde{r}} \left\{ \langle r, \tilde{r} \rangle - \alpha \|\tilde{r}\|_{(L^1(\Omega))^d} \right\} = \begin{cases} 0 & \text{if } |r|_2 \leq \alpha \text{ pointwise,} \\ \infty & \text{if } |r|_2 > \alpha \text{ pointwise.} \end{cases} \quad (3.3.9)$$

Here the pointwise 2-norm of r is chosen, which means $|r|_2 = |(r_1, r_2, \dots, r_d)|_2 = \sqrt{r_1^2 + r_2^2 + \dots + r_d^2}$ for $r \in (L^1(\Omega))^d$. This choice has the effect that the TV-eigenfunction will be a d -dimensional sphere instead of a d -dimensional cube (1-norm) or a d -dimensional diamond (∞ -norm) [48, Theorem 4.1]. For the wavelet model, $\tilde{r}$ is a scalar function instead of a vector function, which means that all pointwise p -norms are equal. The proximal operator for F_2^* is calculated

$$\text{prox}_{\gamma F_2^*}(r) = \underset{\hat{r}}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\hat{r} - r\|_{(L^2(\Omega))^d}^2 + \gamma F_2^*(\hat{r}) \right\} = \frac{\alpha r}{\max\{\alpha, |r|_2\}}. \quad (3.3.10)$$

Similarly, the proximal operator for F_3^* reads

$$\text{prox}_{\gamma F_3^*}(s) = \frac{\alpha \beta s}{\max\{\alpha \beta, |s|_2\}}. \quad (3.3.11)$$

3.3.3 First-order primal dual algorithm

We make use of the modified primal-dual hybrid gradient algorithm (PDHGM) as proposed by Chambolle and Pock [49], which can be seen as a generalisation

to the PDHG algorithm [50]. PDHGM has the advantage that it has a direct solution at every step. Furthermore it can be shown that PDHGM has a convergence rate of at least $\mathcal{O}(n)$ and can even be $\mathcal{O}(n^2)$, where n is the number of steps in the algorithm [49, 42]. The use of PDHGM requires that the all operators are bounded, which, for instance for the gradient operator, is not satisfied in the continuous setting. However, this requirement is fulfilled in the discretised problem, where all operators are bounded.

In Section 3.3.1 the functional F was split in multiple parts. Therefore, the proximal operator also needs to be splitted and evaluated separately. We propose two algorithms, which both possess the structure of [41, Algorithm 4], but have a different size in terms of steps within each iteration.

Algorithm 3.1 First-order primal-dual algorithm for TV reconstruction

Parameters: choose $\tau, \sigma_i > 0$ s.t. $\tau\sigma_i L_i^2 < 1$; choose $\theta \in [0, 1]$.

Initialisation: set $x^0 = 0, q^0 = 0, r^0 = 0, \bar{x}^0 = 0$.

for $n \leftarrow 0$ to $N-1$ **do**

$$q^{n+1} = \text{prox}_{\sigma_1 F_1^*}(q^n + \sigma_1 A \bar{x}^n),$$

$$r^{n+1} = \text{prox}_{\sigma_2 F_2^*}(r^n + \sigma_2 \nabla \bar{x}^n),$$

$$x^{n+1} = x^n - \tau(A^* q^{n+1} - \text{div } r^{n+1}),$$

$$\bar{x}^{n+1} = x^{n+1} + \theta(x^{n+1} - x^n).$$

end for

return x^N

The algorithm for wavelet reconstruction is obtained by changing ∇ to W and div to W^{-1} .

Algorithm 3.1 shows the implementation of the TV model. Two separate dual steps are performed, which are both used as input for the succeeding primal step. The implementation of the wavelet model is obtained by changing ∇ and div to W and W^{-1} respectively. The third line in the for-loop does not show a proximal operator, since $G = 0$ for this model in the original saddle point formulation (3.3.2). This is easily changed to a proximal operator in case one wishes to minimise a problem in which $G \neq 0$.

Algorithm 3.2 shows the implementation of the TGV model, which is built in the same spirit as the TV algorithm. Again, the proximal operators that would depend on G in the saddle point formulation (3.3.2) reduce to the identity. This extension to Algorithm 1 can also be used for other infimal convolution type operators, such as ICTV [44].

Discretisation

The measurements y^δ are sampled in both space and time: the detectors are modelled as point detectors at specific locations and measure the pressure at

Algorithm 3.2 First-order primal-dual algorithm for TGV reconstruction

Parameters: choose $\tau, \sigma_i > 0$ s.t. $\tau\sigma_i L_i^2 < 1$; choose $\theta \in [0, 1]$.

Initialisation: set $x^0 = 0, z^0 = 0, q^0 = 0, r^0 = 0, s^0 = 0, \bar{x}^0 = 0, \bar{z}^0 = 0$.

for $n \leftarrow 0$ to $N-1$ **do**

$$q^{n+1} = \text{prox}_{\sigma_1 F_1^*}(q^n + \sigma_1 A \bar{x}^n),$$

$$r^{n+1} = \text{prox}_{\sigma_2 F_2^*}(r^n + \sigma_2 (\nabla \bar{x}^n - \bar{z}^n)),$$

$$s^{n+1} = \text{prox}_{\sigma_2 F_3^*}(s^n + \sigma_2 \mathcal{E} \bar{z}^n),$$

$$x^{n+1} = x^n - \tau(A^* q^{n+1} - \text{div } r^{n+1}),$$

$$z^{n+1} = z^n - \tau(\mathcal{E}^* s^{n+1} - r^{n+1}),$$

$$\bar{x}^{n+1} = x^{n+1} + \theta(x^{n+1} - x^n),$$

$$\bar{z}^{n+1} = z^{n+1} + \theta(z^{n+1} - z^n),$$

end for

return x^N

sampled time instances with interval Δt . The desired reconstructed image is also a discretised one. In case these discretisations do not perfectly match (which is often the case), significant errors may occur. To avoid these errors, the interpolation procedure as explained in [6, Chapter 5] is applied.

Since an iterative procedure is performed, it is computationally expensive to evaluate the discrete integral operator A at every step. Therefore, the discretised matrix A is computed once and used iteratively. Instead of analytically determining A^* and then discretising it, the adjoint matrix A^T is applied. Accordingly, the operators ∇ and $\mathcal{E}$ are also applied in their matrix form. For the complex dual-tree wavelet transform and its adjoint (which in this case is equivalent to its inverse), the implementation by Cai and Li¹ is used.

The discrete operators $A, \nabla, \mathcal{E}$ and W are implemented such that the value $\|Ax - y^\delta\|^2$ does not change much when the number of detectors is altered. Similarly, the value $R(u)$ does not change significantly when the discretisation of the image domain (resolution) is modified. Because of this, the regularisation parameter α only has to be found once: after changing the resolution or number of detectors, α does not substantially change.

Algorithm parameter selection

In [49] it is shown that the PDHGM algorithm always converges if $\sigma\tau\|B\|^2 < 1$. Here B is the combined operator as defined in (3.3.3) and (3.3.4). A uniform bound for τ and σ might be undesired if the separate operators A and $\nabla, \mathcal{E}$ or W have very different norms: all the proximal steps have a small step size, while this might not be needed for the steps related to only one of these operators.

¹available at <http://eeweb.poly.edu/iselesni/WaveletSoftware/dt2D.html>

In [42] this problem has been solved for matrix operators. Instead of using one value for σ and τ in the PDGHM algorithm, one can use diagonal matrices Σ and T that satisfy $\|\Sigma^{1/2}BT^{1/2}\| < 1$. With this choice, the sequence generated by the algorithm weakly converges to an optimal solution of the saddle point problem.

The discretisation of operators is chosen such that α remains approximately constant for changing image resolution and number of detectors. This has the effect that $\|A\|$ scales linearly with the resolution and with the square root of the number of detectors. For high resolution and many detectors, $\|A\| \gg \|\nabla\| \approx \|\mathcal{E}\| \approx \|W\| \approx 1$, although they show a very similar structure within each operator. Therefore, two sets of two parameters are chosen, since they empirically have shown to give a smoother convergence plot than with the parameter choice in [42], while the convergence rate is similar. One set of two parameters is chosen for the operator A and one set for the combination of other operators. More precisely, σ_1, σ_2, τ are chosen such that $\sigma_1\tau L_1^2 < \frac{1}{4}$ and $\sigma_2\tau L_2^2 < \frac{1}{4}$. Here $L_1 = \|A\|$, $L_2 = \|\nabla\|$ or $L_2 = \|W\|$ in the TV or wavelet model and L_2 is the norm of the combined other operators in the TGV model.

It is easily shown that the bound $\|\Sigma^{1/2}BT^{1/2}\| < 1$ is satisfied by this choice if we write B as a concatenation of the matrix version of all operators. For the TGV model,

$$B = \begin{bmatrix} A & 0 \\ \nabla & -I \\ 0 & \mathcal{E} \end{bmatrix} \in \mathbb{R}^{(m_1+m_2) \times n},$$

and the diagonal matrices

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} = \begin{bmatrix} \sigma_1 I_{m_1} & 0 \\ 0 & \sigma_2 I_{m_2} \end{bmatrix}, \quad T = \tau I_n.$$

This gives us the estimate

$$\begin{aligned} \|\Sigma^{1/2}BT^{1/2}\| &\leq \|\Sigma_1^{1/2} \begin{bmatrix} A & 0 \end{bmatrix} T^{1/2}\| + \|\Sigma_2^{1/2} \begin{bmatrix} \nabla & -I \\ 0 & \mathcal{E} \end{bmatrix} T^{1/2}\| \\ &= \sqrt{\sigma_1\tau}L_1 + \sqrt{\sigma_2\tau}L_2 < \frac{1}{2} + \frac{1}{2} = 1. \end{aligned} \tag{3.3.12}$$

A similar estimate can be given for the TV and wavelet model.

Algorithm validation

The PDGHM algorithm with its specific discretisation and parameter choices is validated by looking at two criteria. Firstly, the duality gap is taken into consideration. The duality gap gives the (nonnegative) difference between the primal and the dual functional. As the duality gap approaches zero, the solution gets asymptotically close to the desired solution for the primal-dual problem [51]. Hence the algorithm is validated by checking if the duality gap approaches

zero. Moreover, a stopping criterion for the algorithm can be constructed based on the duality gap. In the formulation of (3.3.2), the duality gap reads

$$D^n = F(Bu^n) + G(u^n) + F^*(v^n) + G^*(-B^*v^n). \quad (3.3.13)$$

The second criterion is the residual of the variables, defined as

$$u_{\text{res}}^n = \frac{\|u^n - u^{n-1}\|}{u^{n-1}}, \quad (3.3.14)$$

where any variable can be filled in instead of u . The residual $u_{\text{res}}^n \rightarrow 0$ for $n \rightarrow \infty$ whenever $(u^n - u^{n-1}) \rightarrow 0$, which means that the sequence u^n converges.

3.4 Experimental setup and forward model

The proposed methods are defined for solving the general problem $y^\delta = Ax + \eta^\delta$. In this section, we switch to the noise-free problem $y = Ax$, where x is the desired reconstruction, y is noise-free (preprocessed) data and A is some operator that maps x to y . Because of this generic structure, the methods can be applied to reconstruct images for many photoacoustic tomographs with many different forward models.

3.4.1 Experimental setup

In our experimental setup, we make use of the finger joint imager as specified in [52]. A side-illumination laser delivers pulses of optical energy with a wavelength of 532 nm. For the recording of pressure waves a 1D piezoelectric detector array with 64 elements in a half-circle is used. The detector array has a central frequency of 7.5 MHz. Both the detectors and fibre bundles are simultaneously rotatable over 360 degrees. They can also be translated in order to image multiple slices. The detectors have a narrow focus (0.6 mm plane thickness) in one dimension, making it suitable for 2D slice based imaging. A schematic overview of the experimental setup is shown in Figure 3.4.1. For a more extensive explanation of the setup, we refer to [52].

Figure 3.4.1: Schematic overview of the experimental setup, where a phantom is measured.

3.4.2 Forward model

Based on the work of Kruger *et al.*[1] and Wang *et al.*[53], our forward model is written as the projection of the absorbed energy distribution $x(\mathbf{r})$ over spherical surfaces. Under the assumption of a homogeneous speed of sound c_0 , the following expression is obtained for the pressure measured by the detector at location $\mathbf{r}$:

$$p(\mathbf{r}, t) = \frac{\beta}{4\pi C_p} \left(\frac{1}{t} \iint_{|\mathbf{r}-\tilde{\mathbf{r}}|=c_0 t} x(\tilde{\mathbf{r}}) d\tilde{\mathbf{r}} \right) *_t \frac{\partial I(t)}{\partial t} *_t h_{IR}(t), \quad (3.4.1)$$

where β is the thermal expansion coefficient, C_p is the specific heat, $I(t)$ is the illumination profile of the laser, $h_{IR}(t)$ is the impulse response of the detectors and $x(\mathbf{r})$ is the absorbed energy. Instead of finding an explicit expression for $\frac{\partial I(t)}{\partial t}$ and $h_{IR}(t)$, we make use of the calibration measurement p_{cal} of an approximate delta pulse [53] located at $\mathbf{r}_p$, i.e. $x_{\text{cal}}(\mathbf{r}) \approx \delta(\mathbf{r} - \mathbf{r}_p)$. After finding an analytic expression for $p_{\text{cal}}(\mathbf{r}, t)$ in terms of $\frac{\partial I(t)}{\partial t}$ and $h_{IR}(t)$ [6], we obtain

$$p(\mathbf{r}, t) = |\mathbf{r} - \mathbf{r}_p| \left(\frac{1}{t} \iint_{|\mathbf{r}-\tilde{\mathbf{r}}|=c_0 t} x(\tilde{\mathbf{r}}) d\tilde{\mathbf{r}} \right) *_t p_{\text{cal}}(\mathbf{r}, t'), \quad (3.4.2)$$

where t' is the retarded time $t - \frac{|\mathbf{r}-\mathbf{r}_p|}{c}$.

3.4.3 Preprocessing for reconstruction

Expression (3.4.2) can be written concisely as

$$p = \left(\frac{|\mathbf{r} - \mathbf{r}_p|}{t} (Ax) \right) *_t p_{\text{cal}},$$

$$\text{with } Ax := \iint_{|\mathbf{r}-\tilde{\mathbf{r}}|=c_0 t} x(\tilde{\mathbf{r}}) d\tilde{\mathbf{r}}. \quad (3.4.3)$$

Here the dependencies on $\mathbf{r}$, t and t' were left out for an uncluttered notation; the convolution with $\tilde{p}_{\text{cal}}$ is performed in the time domain. Recall that the goal is to solve the problem $y = Ax$, where in this specific setup, A is defined as in (3.4.3). In order to get the preprocessed data y , a deconvolution problem is first solved:

$$p = y_t *_t p_{\text{cal}}, \quad (3.4.4)$$

$$\text{with } y_t := \frac{|\mathbf{r} - \mathbf{r}_p|}{t} y. \quad (3.4.5)$$

After applying the convolution theorem to (3.4.4), $\mathcal{F}\{p\} = \mathcal{F}\{y_t\} \mathcal{F}\{p_{\text{cal}}\}$ is obtained, which for $\mathcal{F}\{p_{\text{cal}}\} > 0$ is equivalent to

$$y_t = \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}\{p\}}{\mathcal{F}\{p_{\text{cal}}\}} \right\}. \quad (3.4.6)$$

Unfortunately, the right hand side of (3.4.6) is undefined when $\mathcal{F}\{p_{\text{cal}}\} = 0$. Moreover any noise on p_{cal} can have a big influence on the expression. For this reason, a regularised deconvolution filter is constructed, i.e.

$$h_{\text{dec}} := \mathcal{F}^{-1} \left\{ \frac{\overline{\mathcal{F}\{p_{\text{cal}}\}}}{|\mathcal{F}\{p_{\text{cal}}\}|^2 + \varepsilon} \right\},$$

such that

$$p * h_{\text{dec}} = \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}\{p\} \overline{\mathcal{F}\{p_{\text{cal}}\}}}{|\mathcal{F}\{p_{\text{cal}}\}|^2 + \varepsilon} \right\} \approx y_t. \quad (3.4.7)$$

Here ε is a small parameter such that it suppresses noise on p_{cal} , but leaves higher parts of it intact. By consecutively applying (3.4.7) and (3.4.5), the preprocessed measurement data for (3.2.1) is obtained.

3.5 Test objects

In order to test our reconstruction framework with its regularisers, both digital and experimental phantoms are created. For each of the three regularisers a specific digital phantom is created, which fulfils the prior assumption on the image to be reconstructed. An experimental phantom with a vascular structure is created to test the effectiveness of the different regularisers on real data.

3.5.1 Digital phantoms

The first digital phantom is shown in Figure 3.5.1a. For this phantom, it is assumed that the fluence rate in the illuminated object is homogeneous. This means that sharp changes in the absorbed energy are expected and it thus consists of piecewise constant values. This image consists of both large and small objects with a variety of shapes, among which squares, discs and elongated rectangles. The larger discs could resemble dermis in the finger, while the small discs and rectangles could resemble blood vessels respectively perpendicular and in the 2D-plane of focus. Because of the sharp discontinuities, this image is reconstructed with TV regularisation and compared with FBP and LS-T reconstructions.

The second digital phantom is shown in Figure 3.5.1b. For this phantom, it is assumed that the fluence rate is decaying from the edge of the disc to the middle of the disc. It resembles tissue that absorbs enough light such that the absorbed energy decays deeper inside the tissue. Since the fluence rate is assumed heterogeneous, TGV regularisation is used for the reconstruction.

The third digital phantom, shown in Figure 3.5.1c, is a vascular structure, which was obtained by preprocessing part of a retinal image from the DRIVE database [54]. Such a structure can be expected when blood vessels lie in the plane of focus. Moreover, such a phantom is similar to 3D vascular structures

Figure 3.5.1: (a) Ground truth digital phantom containing multiple piecewise constant structures of various shapes, sizes and intensities. (b) Digital phantom containing disc with smoothly decaying intensity from outside to inside. (c) Digital phantom containing a vascular structure.

that can be expected in photoacoustic breast imaging. It is not directly clear which of our proposed regularisers is most suitable for such a phantom: since the structure is almost piecewise constant, TV regularisation could be useful; because it does not really contain sharp discontinuities, but rather smooth structure edges, TGV regularisation might give a better result; on the other hand, the phantom has a strong anisotropy, which asks for comparison with directional wavelet regularisation. All three regularised reconstruction methods are compared with FBP and LS-T reconstruction.

3.5.2 Synthetic data acquisition

The acquisition of the preprocessed data y from the measured pressure $\tilde{p}$ has a strong dependency on the calibration measurement $\tilde{p}_{\text{cal}}$, as can be seen in Section 3.4.3. To simulate the acquisition of $\tilde{p}$, a calibration measurement $\tilde{p}_{\text{cal}}$ needs to be used, as can be seen in (3.4.2). In order to avoid an inverse crime, a different calibration measurement is used for the forward model as for the reconstruction. Moreover, the discretisation of the ground truth image x_{GT} is different from the discretisation of the reconstructed image $\hat{x}$.

A cylindrical calibration phantom (diameter 24 mm) was made of Agar gel with Intralipid to provide optical scattering. To create the gel, first 3% (v/v) Agar was dissolved in water by boiling it in the microwave until a clear solution had been formed. Next, the temperature of the mixture was decreased to 40°C under continuous stirring. A 3% (v/v) aqueous solution of 20% stock Intralipid was added drop-wise with stirring. This provides a reduced scattering coefficient (μ'_s) of 9.7 cm⁻¹ at 532 nm. Three separate black USP 6-0 sutures (thickness 70-100 μm) were tensed in a tube, along its longitudinal direction. These sutures function as photoacoustic point sources when measured in a 2D slice. The Agar solution was poured into the tube and was allowed to harden. After this, the calibration phantom was removed.

Calibration measurements were obtained with the setup as explained in

Section 3.4, by rotating stepwise around the calibration phantom and recording time of flight measurements from detector elements to the photoacoustic point sources. From these measurements the center of rotation and impulse response of the transducers are recovered. For a more detailed explanation of the calibration measurement and processing, we refer to [6, Chapter 2].

3.5.3 Experimental phantoms

In order to experimentally test the capability of the regularisers, we developed a phantom with absorbers that resemble a vascular structure. The vessel-shaped absorbers (filaments) were constructed of sodium alginate (SA) gel carrying iron oxide nanoparticles to impart absorption. For the latter we used commercially available superparamagnetic iron oxide (SPIO) nanoparticles (Endorem - Guerbet, Villepinte, France). A dilute Endorem dispersion was mixed with SA solution in distilled water to arrive at a final 2% (w/v) of SA solution. The filaments were fabricated by extruding the SA solution through a syringe with a 30g needle and allowing to fall into 0.7 M calcium chloride (CaCl_2) solution. SA undergoes gel formation in the presence of Ca^{2+} ions in water. The resulting gel sinks to the bottom and hardens for 15 minutes. Finally, the filaments are isolated and washed three times with distilled water to ensure removal of residual Ca^{2+} .

Figure 3.5.2: (a) Experimental phantom containing multiple anisotropic vascular-like structures before pouring the second layer of Agar. (b) Experimental phantom resembling structure after pouring the second layer of Agar. (c) Thresholded version of Figure 3.5.2b.

To prepare the phantom, the same Agar solution as for the calibration phantom was poured halfway into a tube (diameter 24 mm) as mould and allowed to harden. A number of filaments were then laid on the surface of the stiff Agar gel (Figure 3.5.2a). The tube was then topped up with more Agar solution by carefully pouring. When the Agar solution had set, the phantom was removed from the mould.

As the reconstructions in Figure 3.7.1 show, some of the filaments moved during the pouring of the Agar solution. To enable later comparisons, an image

deformation algorithm² was applied to the image in Figure 3.5.2a, such that the filaments in the photograph are aligned with the filaments in the reconstruction. The resulting image is shown in Figure 3.5.2b and resembles the phantom after the pouring of the second layer of Agar. Finally, this image was segmented to enable future analysis. The segmented image is shown in Figure 3.5.2c. Some thin lines in Figure 3.5.2b are artefacts from the image deformation algorithm and were therefore removed. The segmentation serves as a (quasi) ground truth reference for validating the morphology of our experimental reconstructions.

3.5.4 Quality measures

For the synthetic phantoms, a digital ground truth image is available to compare the reconstructions with. Besides a visual comparison, we make use of the peak signal-to-noise ratio (PSNR), defined as

$$\text{PSNR}(\hat{x}, x_{\text{GT}}) = 10 \log_{10} \left(\frac{\max(x_{\text{GT}})}{\|\hat{x} - x_{\text{GT}}\|_2^2 / N_p} \right),$$

where $\hat{x}$ is the reconstructed image, x_{GT} the ground truth image and N_p the number of pixels in the image.

A quality measure based on image intensities is not possible for the experimental reconstructions, since no digital ground truth image is available. However, when interested in the geometry and location of the vascular structure, only the segmentation of the reconstruction is of importance. This reconstruction segmentation can be compared with the ground truth segmentation (Figure 3.5.2c). We follow this idea and make use of the receiver operating characteristic (ROC) curve. The ROC curve is obtained by plotting the false positive rate (one minus specificity) against the true positive rate (sensitivity) of the thresholded reconstruction for various threshold values. The true positive rate gives the fraction of pixels inside the ground truth segmentation that are correctly classified as such by the reconstruction segmentation. The false positive rate gives the fraction of pixels outside the ground truth segmentation that are incorrectly classified as such by the reconstruction segmentation. As the threshold for the reconstruction varies, more pixels will be correctly segmented, at the cost of incorrectly segmented pixels. An ideal situation would be one where the true positive rate increases, without changing the false positive rate, i.e. an almost vertical line, followed by an almost horizontal line. The area under the curve (AUC) is often used to give the quality in a single number. An AUC-value of 1 corresponds to a perfect segmentation of the reconstruction.

²Affine and B-spline grid based registration. Available at <https://nl.mathworks.com/matlabcentral/fileexchange/20057-b-spline-grid-image-and-point-based-registration>

3.6 Synthetic results

Our reconstruction framework is tested on both synthetic and experimental data. It is compared with two standard direct reconstruction methods: filtered backprojection (FBP) and Tikhonov-regularised least squares (LS-T), cf. Section 3.2. The specific FBP algorithm that was used is explained in [6]. In the synthetic case, for the LS-T, TV, TGV and wavelet methods, the regularisation parameters are chosen such that the PSNR between the ground truth and the reconstruction are best. In case of data with additive Gaussian noise, the optimal regularisation parameters can be found with an L-curve method [55].

3.6.1 Robustness under compressive sampling

Synthetic preprocessed measurement data were obtained according to (3.4.2) with a uniform sampling of 64 detectors over a 172 degree array. Six rotations of 60 degrees were performed, providing an almost uniform sampling in a total of 384 detector locations. It is unclear if uniform compressive sampling gives the best reconstruction quality with respect to the different reconstruction methods (see e.g. [56] on experimental design). However, it intuitively makes sense to place the detectors uniformly. Moreover, since the detectors need some space, they are often placed in such a fashion. For our first comparison, different reconstruction methods under uniform compressive sampling are considered.

Figure 3.6.1: Reconstructions of the ‘piecewise constant’ digital phantom in Figure 3.5.1a with different uniform samplings of the detectors.

In Figure 3.6.1 the results of three reconstruction methods (FBP, LS-T, TV) are shown for different number of detectors. All methods provide a good reconstruction in case of 386 detectors, although every method has its own reconstruction bias: TV gives contrast loss in smaller structures, LS-T gives blurred structure edges, whereas FBP gives curved line artefact in the whole image. The TV method is able to keep important features for a longer time as we sample coarser: using 64 detectors, it is still possible to detect the minor contrast changes in the upper right square, whereas this is not possible for the other methods; using only 16 detectors, the TV method still gives sharp edges and structures with the right intensity, while other methods fail to provide this information.

Figure 3.6.2: Reconstructions of the ‘smooth disc’ digital phantom in Figure 3.5.1b with a uniform sampling of 16 detectors.

For the second digital phantom, we again compare with FBP and LS-T. The reconstructions using a uniform sampling of 16 detectors are shown in Figure 3.6.2. The TGV method provides the desired smooth intensity within the disc, while keeping the sharp discontinuities. Moreover, it does not show the curved line artefacts that are visible in the FBP reconstruction.

Figure 3.6.3: Reconstructions of the ‘vascular structure’ digital phantom in Figure 3.5.1c with a uniform sampling of 32 detectors.

In Figure 3.6.3, the reconstructions of the third phantom using a uniform sampling of 32 detectors are shown. FBP performs poorly for this vascular data set, since it is not completely clear which parts of the reconstruction are vascular structures and which are curved line artefacts. All other reconstruction methods show less pronounced curved line artefacts than the FBP reconstruction. Similar to earlier reconstructions (Figures 3.6.1 and 3.6.2), LS-T shows a lower intensity within the vessels and is not very smooth in the background. Regularisation with TV gives the desired sharp edges at the cost that discontinuities also appear within the vessels. It is effective in removing the curved line artefacts: the strongest remaining artefact (around $y = 13$ and $x > 0$) has a much lower intensity than in the FBP and LS-T result. Remarkably, the TGV reconstruction is almost identical to the TV reconstruction: the highest PSNR value was obtained by setting the parameter β in (3.2.6) very high, which means that no linear parts z are obtained and the TGV reduces to the same model as TV. The wavelet reconstruction gives a smoother reconstruction, both inside the vascular structure and in the background. Although the curved line artefacts are less pronounced, there are many small artefacts that have the same anisotropic structure as the vessels. The probable reason for this is that not only the vascular structure, but also the curved line artefacts can be sparsely represented in the directional wavelet basis. Note that this is only the case in a 2D setting, since in 3D, a backprojection artefact is a part of the surface of a sphere instead of a curved line. Directional wavelets might be much more effective in removing these 3D artefacts, as long as a suitable transform is used that can sparsely represent anisotropic vascular structures.

In Figure 3.6.4, a plot of the PSNR values for different reconstruction methods under compressive sampling is shown. For the first two phantoms, both the TV and the TGV method perform better than FBP and LS-T. For the third phantom, the lines for the TV and TGV method overlap, since they give an almost identical reconstruction (cf. Figure 3.6.3). All regularised reconstruction methods provide a smaller improvement compared to the first two phantoms. This is probably due to the problem that artefacts are very similar to the vascular structure and thus harder to remove. Moreover, the wavelet method gives a lower quality reconstruction than TV and TGV under compressive sampling. Since artefacts are sparse in the wavelet basis, it is very difficult to remove them using directional wavelet regularisation. No noise has been added to the data for this comparison.

3.6.2 Robustness against noise

As explained in Section 3.2, additive Gaussian noise is expected due to the system electronics and thermal noise from the transducers. After the simulated measured pressure $p(\mathbf{r}, t)$ was obtained, Gaussian noise with zero mean and standard deviation σ was added to the measured pressure. In Figure 3.6.5, a comparison between reconstruction methods for the noisy data is shown. The

Figure 3.6.4: PSNR values for different reconstruction methods applied to compressively sampled data. Simulated data from digital phantom in (a) Figure 3.5.1a; (b) Figure 3.5.1b; (c) Figure 3.5.1c.

TV and TGV method outperform FBP and LS-T for both low and high noise levels for the first two phantoms. For the third phantom, the TV method seems to outperform both the TGV method, which fails for higher noise levels, and the wavelet method, which has a slightly lower PSNR value at all noise levels. The results were obtained with a uniform sampling of 192 detectors.

3.7 Experimental results

For the acquisition of experimental data, the cylindrical phantom was scanned in six rotations, giving a total uniform sampling of 384 detectors. We compare the TV, TGV and wavelet reconstruction methods with FBP by using both the full data (384 detectors) and using 16.7% compressively sampled data (64 detectors) for the reconstructions. The LS-T reconstruction is omitted, since under low noise and middle to high sampling, FBP has shown to outperform LS-T for the synthetic case. The regularisation parameters for the regularised reconstruction were chosen similar to the parameters found in the synthetic tests, with a small deviation such that the reconstruction is visually best.

Figure 3.6.5: PSNR values for different reconstruction methods applied to data with additive Gaussian noise. Simulated data from digital phantom in (a) Figure 3.5.1a; (b) Figure 3.5.1b; (c) Figure 3.5.1c.

In Figure 3.7.1 it can be seen that FBP gives a sharp reconstruction with the expected curved line artefacts in the background. Due to these artefacts, the intensity of the anisotropic structure is not uniform along the structure and undesired low intensity gaps are created. All three regularisation methods provide a much smoother reconstruction, both in the background and in the anisotropic structure. Moreover, the small gaps of the FBP reconstruction are filled.

Figure 3.7.1: Reconstructions of the experimental phantom in Figure 3.5.2 with a uniform sampling of 384 detectors (full reconstruction).

All methods present the expected result: a piecewise constant reconstruction for TV, a piecewise linear reconstruction for TGV and an anisotropy enhanced reconstruction for wavelet. Since the dual-tree wavelet approach makes use of smooth elongated structure as it's 'building blocks', this reconstruction looks very similar to the TGV reconstruction, although the anisotropic structure shows a slightly sharper boundary in the wavelet reconstruction.

Figure 3.7.2: Reconstructions of the experimental phantom in Figure 3.5.2 with a uniform sampling of 64 detectors (compressive sampling reconstruction).

For the compressive sampling reconstruction (Figure 3.7.2) we can draw similar conclusions: the overall reconstructions are smoother and many gaps that are apparent in the FBP reconstruction are filled in the other reconstructions. If we focus on the region below $y = 5$ and between $x = 0$ and $x = 5$, it is clear that FBP gives strong curved line artefacts, while with the other methods these artefacts are clearly reduced. In these reconstructions, there is much less similarity between the TGV and wavelet results. Due to the use of directional elements, the wavelet reconstruction presents a much better connected structure than the TGV reconstruction.

Figure 3.7.3: ROC curves and area under curve (AUC) values of different reconstruction methods for full and compressive sampling reconstruction. Curves were obtained by varying the segmentation threshold of the reconstructions and comparing it with the ground truth segmentation of Figure 3.5.2c.

As can be seen in Figure 3.7.3a, the ROC curves of the regularised reconstructions show an almost ideal behaviour. The ROC curve of the FBP

reconstruction lies much lower, which shows that this method is a poor choice when it comes to thresholding capability. In Figure 3.7.3b, we see a similar relation between the reconstruction methods, although all curves lie lower than in Figure 3.7.3a. It is interesting to see that all regularised reconstruction methods give a similar ROC curve. TGV and wavelet almost completely overlap, while TV is a bit higher in the steep part, but continues slightly lower in the second part of the graph. This tells us that either method will help us to correctly segment the photoacoustic reconstruction for anisotropic vascular structures. This is also reflected in the values of the area under the curve (AUC).

3.8 Summary and outlook

In this chapter we derived a general variational framework for regularised PAT reconstruction where no specific measurement geometry or physics modelling is assumed. The framework can be applied to a PAT reconstruction problem in a 2- or 3-dimensional setting. The primal-dual implementation is a modular one and the optimisation parameters are chosen such that the algorithm is efficient. Specific choices for the regulariser made in this chapter are TV, second order TGV and the complex dual-tree wavelet transform. These regularisers can respectively deal with the prior assumptions of either sharp discontinuities in combination with a homogeneous and heterogeneous fluence rate, or anisotropic structures. The methods have been compared on both synthetic and experimental data with two widely used direct reconstruction methods.

We showed that the modularity of the variational framework enables easy exchange of data-fidelity and regulariser, which arise from different prior assumptions on the type of noise and the reconstructed image respectively. Using this framework, the terms within can be chosen as desired, without having to change the structure of the algorithm.

Furthermore it is shown that our TV and TGV methods perform better than direct reconstruction methods: they are more effective in handling both low and high noise levels and provide better reconstructions under uniform compressive sampling. In a clinical setup, it might not be preferred to change the detector locations during a full scan, because of limited view or movement of the tissue under consideration. For this reason, the use of TV and TGV methods are promising. In this chapter, reconstructions were made using uniform compressive sampling. Future efforts could lie in finding the best sampling strategy for various data sets and regularisers.

Finally, research could be focused towards the reconstruction of anisotropic structures such as blood vessels and the difference between this reconstruction in 2D and 3D. It is unclear if the similarity between backprojection artefacts and vascular structures is preventing better reconstructions in the 2D case and if this problem is absent in the 3D case. For this reason, it is useful to develop or learn a dictionary that allows sparse reconstruction of vascular structures in 3D.

References

- [1] R. A. Kruger, P. Liu, Y. R. Fang, and C. R. Appledorn, "Photoacoustic ultrasound (PAUS)—reconstruction tomography," *Medical physics*, vol. 22, no. 10, pp. 1605–1609, 1995.
- [2] D. Finch and S. K. Patch, "Determining a function from its mean values over a family of spheres," *SIAM journal on mathematical analysis*, vol. 35, no. 5, pp. 1213–1240, 2004.
- [3] M. Haltmeier, T. Schuster, and O. Scherzer, "Filtered backprojection for thermoacoustic computed tomography in spherical geometry," *Mathematical methods in the applied sciences*, vol. 28, no. 16, pp. 1919–1937, 2005.
- [4] M. Xu and L. V. Wang, "Universal back-projection algorithm for photoacoustic computed tomography," *Physical Review E*, vol. 71, no. 1, p. 016706, 2005.
- [5] D. Finch, M. Haltmeier, and Rakesh, "Inversion of spherical means and the wave equation in even dimensions," *SIAM Journal on Applied Mathematics*, vol. 68, no. 2, pp. 392–412, 2007.
- [6] R. Willeminck, *Image reconstruction in a passive element enriched photoacoustic tomography setup*. PhD thesis, University of Twente, 2010.
- [7] F. Natterer, "Photo-acoustic inversion in convex domains," *Inverse Problems & Imaging*, vol. 6, no. 2, p. 315, 2012.
- [8] M. Haltmeier, "Universal inversion formulas for recovering a function from spherical means," *SIAM Journal on Mathematical Analysis*, vol. 46, no. 1, pp. 214–232, 2014.
- [9] L. Kunyansky, "Inversion of the spherical means transform in corner-like domains by reduction to the classical Radon transform," *Inverse Problems*, vol. 31, no. 9, p. 095001, 2015.
- [10] P. Burgholzer, G. J. Matt, M. Haltmeier, and G. Paltauf, "Exact and approximative imaging methods for photoacoustic tomography using an arbitrary detection surface," *Physical Review E*, vol. 75, no. 4, p. 046706, 2007.
- [11] Y. Hristova, P. Kuchment, and L. Nguyen, "Reconstruction and time reversal in thermoacoustic tomography in acoustically homogeneous and inhomogeneous media," *Inverse Problems*, vol. 24, no. 5, p. 055006, 2008.
- [12] B. E. Treeby and B. T. Cox, "k-wave: Matlab toolbox for the simulation and reconstruction of photoacoustic wave fields," *Journal of biomedical optics*, vol. 15, no. 2, p. 021314, 2010.
- [13] Y. Zhou, J. Yao, and L. V. Wang, "Tutorial on photoacoustic tomography," *Journal of biomedical optics*, vol. 21, no. 6, p. 061007, 2016.
- [14] X. Deán-Ben, S. Gottschalk, B. Mc Larney, S. Shoham, and D. Razansky, "Advanced optoacoustic methods for multiscale imaging of in vivo dynamics," *Chemical Society Reviews*, vol. 46, no. 8, pp. 2158–2198, 2017.
- [15] M. Heijblom, D. Piras, M. Brinkhuis, J. C. van Hespén, F. M. van den Engh, M. van der Schaaf, J. M. Klaase, T. G. van Leeuwen, W. Steenbergen, and S. Manohar, "Photoacoustic image patterns of breast carcinoma and comparisons with magnetic resonance imaging and vascular stained histopathology," *Scientific reports*, vol. 5, p. 11778, 2015.

- [16] M. Toi, Y. Asao, Y. Matsumoto, H. Sekiguchi, A. Yoshikawa, M. Takada, M. Kataoka, T. Endo, N. Kawaguchi-Sakita, M. Kawashima, *et al.*, "Visualization of tumor-related blood vessels in human breast by photoacoustic imaging system with a hemispherical detector array," *Scientific reports*, vol. 7, p. 41970, 2017.
- [17] D. Chamberland, Y. Jiang, and X. Wang, "Optical imaging: new tools for arthritis," *Integrative Biology*, vol. 2, no. 10, pp. 496–509, 2010.
- [18] P. van Es, S. K. Biswas, H. J. B. Moens, W. Steenbergen, and S. Manohar, "Initial results of finger imaging using photoacoustic computed tomography," *Journal of biomedical optics*, vol. 19, no. 6, p. 060501, 2014.
- [19] A. Rosenthal, V. Ntziachristos, and D. Razansky, "Acoustic inversion in optoacoustic tomography: a review," *Current Medical Imaging*, vol. 9, no. 4, pp. 318–336, 2013.
- [20] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: nonlinear phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.
- [21] S. R. Arridge, M. M. Betcke, B. T. Cox, F. Lucka, and B. E. Treeby, "On the adjoint operator in photoacoustic tomography," *Inverse Problems*, vol. 32, no. 11, p. 115012, 2016.
- [22] S. Arridge, P. Beard, M. Betcke, B. Cox, N. Huynh, F. Lucka, O. Ogunlade, and E. Zhang, "Accelerated high-resolution photoacoustic tomography via compressed sensing," *Physics in Medicine & Biology*, vol. 61, no. 24, p. 8908, 2016.
- [23] J. Provost and F. Lesage, "The application of compressed sensing for photo-acoustic tomography," *IEEE transactions on medical imaging*, vol. 28, no. 4, pp. 585–594, 2008.
- [24] Z. Guo, C. Li, L. Song, and L. V. Wang, "Compressed sensing in photoacoustic tomography in vivo," *Journal of biomedical optics*, vol. 15, no. 2, p. 021311, 2010.
- [25] M. Sandbichler, F. Krahmer, T. Berer, P. Burgholzer, and M. Haltmeier, "A novel compressed sensing scheme for photoacoustic tomography," *SIAM Journal on Applied Mathematics*, vol. 75, no. 6, pp. 2475–2494, 2015.
- [26] M. Haltmeier, T. Berer, S. Moon, and P. Burgholzer, "Compressed sensing and sparsity in photoacoustic tomography," *Journal of Optics*, vol. 18, no. 11, p. 114004, 2016.
- [27] Y. Zhang, Y. Wang, and C. Zhang, "Total variation based gradient descent algorithm for sparse-view photoacoustic image reconstruction," *Ultrasonics*, vol. 52, no. 8, pp. 1046–1055, 2012.
- [28] J. Meng, L. V. Wang, L. Ying, D. Liang, and L. Song, "Compressed-sensing photoacoustic computed tomography in vivo with partially known support," *Optics Express*, vol. 20, no. 15, pp. 16510–16523, 2012.
- [29] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 492–526, 2010.
- [30] F. Knoll, K. Bredies, T. Pock, and R. Stollberger, "Second order total generalized variation (TGV) for MRI," *Magnetic resonance in medicine*, vol. 65, no. 2, pp. 480–491, 2011.
- [31] J. Yang, H. Yu, M. Jiang, and G. Wang, "High-order total variation minimization for interior tomography," *Inverse problems*, vol. 26, no. 3, p. 035013, 2010.

- [32] S. Niu, Y. Gao, Z. Bian, J. Huang, W. Chen, G. Yu, Z. Liang, and J. Ma, "Sparse-view x-ray CT reconstruction via total generalized variation regularization," *Physics in Medicine & Biology*, vol. 59, no. 12, p. 2997, 2014.
- [33] Y. Lin and L. Huang, "Ultrasound waveform tomography with the second-order total-generalized-variation regularization," in *Medical Imaging 2016: Physics of Medical Imaging*, vol. 9783, p. 978356, International Society for Optics and Photonics, 2016.
- [34] T. Valkonen, K. Bredies, and F. Knoll, "Total generalized variation in diffusion tensor imaging," *SIAM Journal on Imaging Sciences*, vol. 6, no. 1, pp. 487–525, 2013.
- [35] J. Duan, W. Lu, C. Tench, I. Gottlob, F. Proudlock, N. N. Samani, and L. Bai, "Denoising optical coherence tomography using second order total generalized variation decomposition," *Biomedical Signal Processing and Control*, vol. 24, pp. 120–127, 2016.
- [36] W. Krauze, P. Makowski, M. Kujawińska, and A. Kuś, "Generalized total variation iterative constraint strategy in limited angle optical diffraction tomography," *Optics Express*, vol. 24, no. 5, pp. 4924–4936, 2016.
- [37] K. Hammernik, T. Pock, and R. Nuster, "Variational photoacoustic image reconstruction with spatially resolved projection data," in *Photons Plus Ultrasound: Imaging and Sensing 2017*, vol. 10064, p. 100643I, International Society for Optics and Photonics, 2017.
- [38] N. Kingsbury, "The dual-tree complex wavelet transform: a new efficient tool for image restoration and enhancement," in *9th European Signal Processing Conference (EUSIPCO 1998)*, pp. 1–4, IEEE, 1998.
- [39] E. J. Candès and D. L. Donoho, "Curvelets: A surprisingly effective nonadaptive representation for objects with edges," tech. rep., Stanford Univ Ca Dept of Statistics, 2000.
- [40] M. Benning, C. Brune, M. Burger, and J. Müller, "Higher-order TV methods - enhancement via Bregman iteration," *Journal of Scientific Computing*, vol. 54, no. 2-3, pp. 269–310, 2013.
- [41] E. Y. Sidky, J. H. Jørgensen, and X. Pan, "Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle - Pock algorithm," *Physics in Medicine & Biology*, vol. 57, no. 10, p. 3065, 2012.
- [42] T. Pock and A. Chambolle, "Diagonal preconditioning for first order primal-dual algorithms in convex optimization," in *2011 International Conference on Computer Vision*, pp. 1762–1769, IEEE, 2011.
- [43] P. J. Huber *et al.*, "Robust regression: asymptotics, conjectures and Monte Carlo," *The annals of statistics*, vol. 1, no. 5, pp. 799–821, 1973.
- [44] A. Chambolle and P.-L. Lions, "Image recovery via total variation minimization and related problems," *Numerische Mathematik*, vol. 76, no. 2, pp. 167–188, 1997.
- [45] R. Rubinstein, A. M. Bruckstein, and M. Elad, "Dictionaries for sparse representation modeling," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1045–1057, 2010.
- [46] E. Esser, X. Zhang, and T. F. Chan, "A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science," *SIAM Journal on Imaging Sciences*, vol. 3, no. 4, pp. 1015–1046, 2010.

- [47] D. A. Lorenz and T. Pock, "An inertial forward-backward algorithm for monotone inclusions," *Journal of Mathematical Imaging and Vision*, vol. 51, no. 2, pp. 311–325, 2015.
- [48] S. Esedoglu and S. J. Osher, "Decomposition of images by the anisotropic Rudin-Osher-Fatemi model," *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, vol. 57, no. 12, pp. 1609–1626, 2004.
- [49] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of mathematical imaging and vision*, vol. 40, no. 1, pp. 120–145, 2011.
- [50] M. Zhu and T. Chan, "An efficient primal-dual hybrid gradient algorithm for total variation image restoration," *UCLA CAM Report*, vol. 34, 2008.
- [51] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.
- [52] P. van Es, R. C. Vlieg, E. Hondebrink, J. C. van Hespén, H. J. B. Moens, W. Steenbergen, and S. Manohar, "Coregistered photoacoustic and ultrasound tomography of healthy and inflamed human interphalangeal joints," in *European Conference on Biomedical Optics*, p. 95390C, Optical Society of America, 2015.
- [53] Y. Wang, D. Xing, Y. Zeng, and Q. Chen, "Photoacoustic imaging with deconvolution algorithm," *Physics in Medicine & Biology*, vol. 49, no. 14, p. 3117, 2004.
- [54] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. Van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE transactions on medical imaging*, vol. 23, no. 4, pp. 501–509, 2004.
- [55] P. C. Hansen, "The L-curve and its use in the numerical treatment of inverse problems," *Computational Inverse Problems in Electrocardiology*, vol. 5, pp. 119–142, 2001.
- [56] E. Haber, L. Horesh, and L. Tenorio, "Numerical methods for experimental design of large-scale linear ill-posed inverse problems," *Inverse Problems*, vol. 24, no. 5, p. 055012, 2008.

Chapter 4

A partially learned algorithm for joint photoacoustic reconstruction and segmentation

Yoei E. Boink, Srirang Manohar, Christoph Brune

IEEE transactions on medical imaging 39.1 (2019): 129-139

Abstract

In an inhomogeneously illuminated photoacoustic image, important information like vascular geometry is not readily available when only the initial pressure is reconstructed. To obtain the desired information, algorithms for image segmentation are often applied as a post-processing step. In this chapter, we propose to jointly acquire the photoacoustic reconstruction and segmentation, by modifying a recently developed partially learned algorithm based on a convolutional neural network. We investigate the stability of the algorithm against changes in initial pressures and photoacoustic system settings. These insights are used to develop an algorithm that is robust to input and system settings. Our approach can easily be applied to other imaging modalities and can be modified to perform other high-level tasks different from segmentation. The method is validated on challenging synthetic and experimental photoacoustic tomography data in limited angle and limited view scenarios. It is computationally less expensive than classical iterative methods and enables higher quality reconstructions and segmentations than state-of-the-art learned and non-learned methods.

4.1 Introduction

Photoacoustic tomography (PAT) is a hybrid imaging technique that combines high optical absorption contrast of tissues with high resolution from ultrasound detection. It is in the focus of research towards applications in various fields of biomedicine [1], in particular for breast cancer imaging in humans [2, 3, 4]. One of the main objectives in PAT is to acquire information on the vascular geometry in soft tissue: many diseases, such as breast cancer and rheumatoid arthritis can be characterised with increased blood vessel density and irregular vessel structure. A common practice is to first reconstruct the initial pressure image, after which a segmentation algorithm is applied to obtain the segmented vascular geometry. A big disadvantage of such ‘two-step’ approaches is that reconstruction errors due to inaccurate physics modelling, noise or a lack of data will propagate to the subsequent segmentation step. In this chapter, we employ a partially learned algorithm to solve the reconstruction and segmentation task jointly. By utilising the same physics model for both problems, we can achieve higher quality segmentation results than classical non-learned or even learned two-step approaches (Figure 4.1.1). The algorithm is a partially learned one in the sense that the physics model is given, but other parts of the algorithm are learned.

Figure 4.1.1: Example of a two-step approach, where a segmentation is acquired by first applying FBP, after which U-Net is used as a second step. In our approach, we incorporate the physics modelling in the joint learned algorithm, providing both reconstruction and segmentation.

Common reconstruction methods for PAT are filtered backprojection (FBP) [5, 6, 7, 8] and time reversal [9, 10]. These methods apply a variant of the adjoint acoustic operator on the measurements, which has proven to give the exact result in case of infinite measurements without noise. Empirically, these approaches are still effective if ‘enough’ measurements with limited noise are available. If this is not the case, these methods fail to provide adequate results and iterative approaches utilising regularisation [11, 12, 13, 14, 15, 16] are needed. This topic was investigated in Chapter 3 of this thesis. We showed that it is not always clear which regulariser should be selected, which limits the practical use of

those approaches. Moreover, iterative approaches often need many iterations to converge, which can make them computationally expensive.

In the past years there has been a big interest in using deep neural networks for various tomography modalities to enable faster and higher quality reconstructions. Very deep networks like U-Net [17] are applied to post-process FBP-results and achieve a better reconstruction [18, 19, 20, 21], but can also be used to obtain a segmentation of the tissue under consideration. These networks are not very effective when the quality of the input data is low, as is the case for under-sampled FBP reconstructions. For this reason, other algorithms incorporate the operator that describes the physics in the learning process. DALnet [22] is a non-iterative method that learns the weights of an operator that acts as a pseudo-inverse on the measurements. In [23], the authors first solve a non-learned reconstruction problem, followed by a learned smoothing step; these steps are iteratively executed in the resulting algorithm. Many other learned iterative methods are based on classical iterative methods, but execute far fewer steps in their reconstruction procedure: variational networks [24, 25, 26] can be seen as a learned variant of proximal gradient methods, where kernel-function pairs of the regularisation term are learned. Learning the regularisation term is also the idea behind [27] and [28]. In [29], a learned variant of gradient-descent is applied, while learned proximal operators are investigated in [30]. The learned primal-dual (L-PD) algorithm [31] can be seen as the learned variant of many popular and effective primal-dual methods [32]. The key idea in [29], [30] and [31] is that instead of taking a pre-defined gradient or proximal step with respect to the regularisation function, an algorithm learns the steps that have to be taken to obtain a high-quality reconstruction. In [33], L-PD was applied to 3D photoacoustic data. Since their 3D implementation of the acoustic operator is computationally expensive, the authors chose to not learn the whole ‘unrolled’ algorithm, but learn per iteration instead. In case the acoustic operator is expensive, one can also choose to use an approximation of this operator [34], which still provides results superior to non-learned methods. Although empirically strong in their reconstruction capacity, most learned iterative methods lack a mathematical analysis of their convergence and stability. This makes them precarious for implementation in many applications, especially in clinical practice. A convergence proof for a method in which the regulariser was learned is given in [28]. Recently, Banert *et al.* [35] presented a convergence proof for learned algorithms that make use of proximal operators.

Segmentation in biomedical imaging is a well-studied problem; see [36] for an overview. New learned methods for segmentation are developed abundantly, of which U-Net [17] is currently a frequently used one. There are not many works on segmentation specifically designed for PAT; a simple segmentation procedure is given in [37] and a complete ‘pipeline’ from FBP reconstruction to segmentation is presented in [38]. In the past years, solving the reconstruction and segmentation problem jointly has become increasingly popular: by treating

both problems simultaneously, reconstruction can benefit from information available in the segmentation and vice versa. One such method is given in [39], which also contains an overview on other non-learned joint methods. A recent work in which the segmentation is jointly solved with motion-estimation in an unsupervised learning approach is given in [40]. A mathematical framework for solving high-level tasks together with reconstruction in a learned manner is provided in [41]. In that paper, the network architectures for reconstruction is not necessarily the same as that for the high-level task.

In this chapter, we make use of the partially learned L-PD method and modify it for the goal of joint reconstruction and segmentation in two-dimensional PAT. With this method, accurate information on soft tissue vasculature is directly available, without the necessity for post-processing. By using the acoustic model within the learning of the algorithm, both reconstruction and segmentation are of higher quality than in other learned algorithms. To obtain a robust method, the sensitivity of L-PD to image changes and changes in PAT system settings is investigated and solutions are provided. The method is tested on challenging synthetic and experimental data sets.

To the best of our knowledge, this is the first time the joint problem of photoacoustic reconstruction and segmentation is addressed. Since a major interest in photoacoustic imaging is the visualisation of blood vessels, a segmentation of the vascular geometry is of high importance. Our method provides this segmentation, including a reliability estimate, which can be interpreted in an easy way. An analysis of the sensitivity of the learned iterative algorithm is performed, which has not been done before.

The remainder of this chapter is organised as follows: in Section 4.2 we give a background to photoacoustic tomography, explain our forward model, and introduce the learned primal-dual method. In Section 4.4, we first present the neural network architecture that is used, before explaining how its sensitivity is investigated. In the same section, the joint segmentation-reconstruction method is provided, as well as the explanation of the experimental data on which it is tested. Quantitative and visual results of these experiments are given in Section 4.5. In Section 4.6 we conclude with some remarks and outlook for the future.

4.2 Theory of photoacoustic tomography

Photoacoustic signals are generated by illuminating tissue with nanosecond laser pulses, which causes the tissue to heat up at locations of optical absorption. Thermoelastic expansion causes the generation of pressure, which propagates in the form of ultrasound waves. An array of ultrasound detectors is placed around the tissue to detect this signal. For chosen wavelengths, certain tissue constituents such as haemoglobin have a higher optical absorption coefficient than surrounding tissue, giving the desired high contrast. The generated initial

pressure p_0 depends linearly on the combination of optical absorption coefficient μ_a and fluence rate Φ :

$$p_0 \propto \mu_a \Phi. \quad (4.2.1)$$

Quantitative PAT [42] has the goal to reconstruct μ_a , but has the drawback of being computationally expensive or inaccurate, depending on the light propagation model. The objective of the acoustic PAT inverse problem is to reconstruct p_0 . This is computationally less expensive than the quantitative problem and the result already gives insight in the vascular geometry. However, a decaying fluence rate can cause problems for analysing this qualitative image, since the same blood vessel structure can have a lower intensity deeper inside the tissue. Therefore, there is a need for a computationally inexpensive method which does not only provide the reconstructed initial pressure, but also gives a segmentation of the vascular geometry, insensitive to fluence rate reduction in depth.

4.2.1 Forward model

In this section the two-dimensional acoustic PAT forward problem is described. Throughout this chapter, with the exception of Appendix B, the sound speed is assumed to be constant and known. For the derivation of the model with non-constant sound speed, we refer to [7, Section 5.3]. From this point onwards the initial pressure is written as x instead of p_0 to improve readability of formulas and algorithms. We make use of a projection model with calibration as described in [43]:

$$p(\mathbf{r}, t) = |\mathbf{r} - \mathbf{r}_p| \left(\frac{1}{t} \int_{|\mathbf{r} - \tilde{\mathbf{r}}| = c(\mathbf{r})t} x(\tilde{\mathbf{r}}) d\tilde{\mathbf{r}} \right) *_t p_{\text{cal}}(\mathbf{r}, t'). \quad (4.2.2)$$

Here $p(\mathbf{r}, t)$ is the measured pressure at location $\mathbf{r}$, which is the result of a spherical mean transform [5] applied to the initial pressure x convolved with a calibration measurement $p_{\text{cal}}(\mathbf{r}, t')$. The last is the measurement of a point source located at $\mathbf{r}_p$, which is experimentally approximated. After pre-processing, we obtain the simplified equation

$$y = Ax := \int_{|\mathbf{r} - \tilde{\mathbf{r}}| = c(\mathbf{r})t} x(\tilde{\mathbf{r}}) d\tilde{\mathbf{r}}, \quad (4.2.3)$$

where A is the acoustic forward operator mapping the initial pressure to the preprocessed measurements y^δ . For a more detailed explanation of the calibration measurement and pre-processing, we refer to Section 3.4.3 and [7, Chapter 2]. For computational reasons, we make use of a matrix implementation of A . However, it should be noted that our algorithm can also be applied when a matrix implementation of the acoustic operator is not available.

4.3 From non-learned to learned primal-dual methods

Many linear inverse problems, including the PAT reconstruction problem (4.2.3), can be written as

$$y^\delta = Ax + \eta^\delta, \quad (4.3.1)$$

where noisy measurements y^δ are given, an approximate forward model A is available and the noise level δ of the additive noise η^δ is estimated. In case a direct reconstruction method does not give an adequate result, variational methods are employed to solve the minimisation problem

$$\min_u F_{y^\delta}(Bx) + G(x). \quad (4.3.2)$$

Here F_{y^δ} is an operator that acts on the dual domain and handles the measurements, and G is an operator that acts on the primal domain. For instance, writing down a classical $L^2 - TV$ model [44] in discrete form, it reads

$$\min_x \frac{1}{2} \|Ax - y^\delta\|_2^2 + \alpha \|\nabla x\|_1. \quad (4.3.3)$$

This can be put in the form of (4.3.2) by choosing $B := [A, \nabla]$ and for F_{y^δ} choose a combination of the 2-norm and 1-norm acting on Bx , while $G(x) := 0$. The reason of writing it down in such a form is that any primal-dual algorithm can be applied to solve this minimisation problem [32]. The modified primal-dual hybrid gradient (PDGHM) [45] is a popular choice for tomography problems [46], see also Chapter 3 of this thesis. A key element in primal-dual algorithms is updating via the proximal operator, which for a general scaled functional $\gamma H(x)$ is defined as

$$\text{prox}_{\gamma H}(x) = \underset{\hat{x}}{\text{argmin}} \left\{ \frac{1}{2} \|\hat{x} - x\|_2^2 + \gamma F(\hat{x}) \right\}. \quad (4.3.4)$$

This proximal operator takes a descent step in the functional $H(\hat{x})$, while staying close to a previous iterate x . A more extensive explanation of primal-dual methods and proximal operators is given in Section 3.3.

Algorithm 4.1 PDHGM (non-learned).

Parameters: choose $\tau, \sigma > 0$ s.t. $\tau\sigma L^2 < 1$; choose $\theta \in [0, 1]$.

Initialisation: set $x^0 = 0, q^0 = 0$.

for $n \leftarrow 0$ to $\tilde{N}-1$ **do**

$$q^{n+1} = \text{prox}_{\sigma F_{y^\delta}^*} \left(q^n + \sigma A \left[(1 + \theta)x^n - \theta x^{n-1} \right] \right),$$

$$x^{n+1} = \text{prox}_{\tau G} \left(x^n - \tau A^* q^{n+1} \right).$$

end for

return $x^{\tilde{N}}$

In Algorithm 4.1 PDHGM is stated, where the θ -acceleration step is applied in the first update (c.f. Section 3.3.3). Note that this is the PDHGM algorithm for the case where B is chosen to be A in minimisation problem (4.3.2). Step sizes τ and σ can be freely chosen, as long as $\tau\sigma L^2 < 1$, where $L = \|A\|$. The first update of Algorithm 4.1 takes a proximal step with respect to the convex conjugate of F and acts in the dual domain; its inputs are the previous dual update q^n , measurements y^δ and the acoustic operator A applied to the primal updates x^{n-1} and x^n . The second update takes a proximal step with respect to G and acts on the previous primal update x^n and the adjoint operator A^* applied to the dual update q^{n+1} . The adjoint operator A^* is obtained by first discretising A in a matrix and then taking the numerical adjoint.

Algorithm 4.2 L-PD (learned). See Figure 4.3.1 for a visualisation.

Parameters: learned weights Θ_n and Ψ_n .

Initialisation: set $x^0 = 0, q^0 = 0$.

for $n \leftarrow 0$ to $N-1$ **do**

$$q_{\{1,\dots,k\}}^{n+1} = q_{\{1,\dots,k\}}^n + \mathcal{N}_{\Theta_n}^q(q_{\{1,\dots,k\}}^n, Ax_1^n, y^\delta),$$

$$x_{\{1,\dots,k\}}^{n+1} = x_{\{1,\dots,k\}}^n + \mathcal{N}_{\Psi_n}^x(x_{\{1,\dots,k\}}^n, A^*q_1^{n+1}).$$

end for

return x_1^N

In this chapter, we make use of the learned primal-dual (L-PD) algorithm [31], which can be derived from PDGGM. The idea for the L-PD approach (Algorithm 4.2) is to not choose the functionals F and G explicitly, but learn the best update steps for each iteration. This is achieved by a CNN, here represented by the nonlinear functions $\mathcal{N}_{\Theta_n}^q$ and $\mathcal{N}_{\Psi_n}^x$, where Θ_n and Ψ_n describe the learned weights for iteration n . The key idea is that the inputs of the learned steps of Algorithm 4.2 are still the same as in Algorithm 4.1, such that the primal-dual structure is preserved. Note that weights of the network can be different for every iteration $n \in \{1, \dots, N\}$. Moreover, instead of updating one channel of the primal and dual, we allow the network to use k channels. This could encode some kind of ‘history’, which is similar to the acceleration of PDGGM, where not only x^n , but also x^{n-1} is an input in the first proximal step. However, the different channels can represent more general structures in the primal and dual domain that are useful for reconstruction. Channel 1 of the final iteration N , i.e. x_1^N , gives our desired reconstruction. This leaves room for the other channels to represent other useful information; this is leveraged in Section 4.4.2, where channel 2 is exploited to provide a non-binary segmentation. In Section 4.5 it is shown that the L-PD algorithm provides higher quality reconstructions than PDGGM applied to a TV regularised model (4.3.3). An additional advantage is that L-PD can be enforced to only use a small amount of iterations, whereas in PDGGM, one has to wait till convergence, which generally takes many more

iterations. A visualisation of the algorithm as a neural network is given in Figure 4.3.1.

Figure 4.3.1: Visualisation of the learned primal-dual architecture as described in Algorithm 4.2. The number of channels is indicated below the layers on the left. On the right, every node corresponds to a 3×3 convolution followed by a ReLU-activation function.

4.4 Methodology

A major interest in photoacoustic tomography is the visualisation of blood vessels in tissue as a marker of diseases, for instance in breast imaging. For this reason, we made use of retinal blood vessel images from the openly available DRIVE data set [47]. A total of 768 training images and 192 test images were obtained by preprocessing patches with a size of 192×192 pixels from this data set. These training and test images were further processed to contribute to our two goals: obtaining robustness to variations in images and system settings, and obtaining an additional segmentation in a realistic photoacoustic scenario. This is explained in detail in Section 4.4.1 and Section 4.4.2 respectively. The discrete size of the synthetic measurements is $1450 \times n_s$ pixels, where n_s is the number of sensors used.

For the L-PD algorithm, both a small network and a larger network were trained. In the first one, the number of iterations was chosen to be $N = 10$ and the number of primal- and dual channels $k = 5$ (cf. Algorithm 4.2). In the second one, $N = 5$ and $k = 2$ were chosen. In both networks, there are 32 channels in 2 hidden layers. The filter size for the convolutions is 3×3 and ReLUs are chosen as activation functions. The Adam optimiser [48] with a mean squared error (MSE) loss on the difference between ground truth and reconstruction is used:

$$\mathcal{L} = \|x_{GT} - x_1^N\|_2^2. \quad (4.4.1)$$

For stability in the optimisation, the batch size increases from 2 to 16 in three steps [49] during 200 Epochs. We make use of TensorFlow, executed on a single

GTX 1080 TI GPU. Training took between 1 and 4 hours, depending on the amount of detectors used. It should be noted that this depends heavily on the specific operator A used: for training, one has to perform $2N$ iterations per backprojection, which makes the application of A the limiting factor for most tomography problems.

After training, the computational cost for applying the algorithm also depends mostly on the operator A . For each iteration, A and A^* have to be applied once. This means that the computational cost is approximately $2N$ times that of FBP. For a TV-reconstruction, the final iteration $\tilde{N}$ depends on the desired accuracy, but is usually in the order of 100 or 1000 (c.f. Section 4.4.3), while we choose $N = 5$ or $N = 10$ for L-PD.

4.4.1 Training for robustness

In this experiment we investigate robustness of the L-PD network to changes in the image and system settings, and give solutions to improve this robustness. Digital phantoms are created by setting the background of the vascular images to 0 and setting the maximum to 1. Synthetic measurements are generated by applying the forward model as explained in Section 4.2.1 to these phantoms. For the standard data set, a setting of 32 detectors is used, uniformly placed around the phantom (c.f. right of Figure 4.4.3). The L-PD algorithm is trained with the architecture as specified before.

Image uncertainty

In Table 4.1, 8 classes of images are defined, in which one or multiple image properties were changed. For a visual impression, one instance of every class is shown in Figure 4.4.1. To analyse the sensitivity of L-PD, the algorithm was trained on class 0 and applied to the test set of class i , with $i = 1, 2, \dots, 7$. Then the network was retrained on a training set of class i and again applied to the test set. We investigate how this affects the reconstruction quality.

System uncertainty

To investigate the sensitivity of L-PD to changes in system settings, the algorithm is first trained on data from 32 uniformly placed detectors. The trained algorithm is applied to test data with 16 or 64 uniformly placed detectors and 32 detectors in a limited view scenario (c.f. right of Figure 4.4.3). Next, the network is retrained on a training set where the detectors were randomly chosen from a possibility of 128 detectors. We investigate how retraining on this more diverse training set affects the reconstruction quality. It is possible to explicitly treat the limited view problem by weighing the pixels at different locations differently [50, 22]. Instead, we choose to let the L-PD algorithm learn to correct for the limited view problem.

class (c_i)	feature change	explanation
0	nothing	<i>For an example image see top left in Figure 4.4.1.</i>
1	diameter	Pixel-wise increase/decrease of vessel diameter randomly chosen from the set $\{-2, -1, 0, 1, 2\}$.
2	contrast	Intensity of the vessel randomly chosen from the set $\{0.5, 0.7, 1, 1.4, 2\}$.
3	background	With probability $p = \frac{1}{2}$ a uniformly increased background intensity, with $p = \frac{1}{2}$ a speckled background.
4	coverage	With probability $p = \frac{1}{3}$ a removal of all smallest vessels, with $p = \frac{1}{3}$ nothing, with $p = \frac{1}{3}$ a doubling of vessels.
5	structure	Inhomogeneous vessel intensity scaled between $[m, 1]$, with m randomly chosen from the set $\{0.2, 0.4, 0.6, 0.8, 1\}$.
6	noise level	Gaussian noise level randomly chosen from $[0, 3\delta]$, where δ is the standard noise level.
7	all the above	Application of all the above changes with given probabilities.

Table 4.1: Explanation of image classes where one feature (class 1-6) or multiple features (class 7) are changed. See Figure 4.4.1 for a visual impression.

Figure 4.4.1: Example image where one feature (class 1-6) or more features (class 7) are changed. See Table 4.1 for detailed information.

In Appendix B and Appendix C, robustness to heterogeneous sound speed changes and uncertainty in the calibration measurements are analysed, respectively.

4.4.2 Joint reconstruction and segmentation

The experiments in Section 4.4.1 are meant to analyse the robustness of the L-PD algorithm to changes in image or in system settings. In these experiments, a homogeneous internal fluence rate is assumed, so that the initial pressure within the blood vessels is uniform.

This is not a realistic scenario, especially for breast imaging, since the fluence rate drops rapidly with depth from the illuminated surface due to scattering and absorption in tissue. To get more realistic digital phantoms, we model light propagation with the diffusion approximation (DA) of the radiative transfer equation (RTE) [51]. This model is applied to the digital phantoms of class 0 in 4.4.1, which serve as the optical absorption coefficient μ_a . The DA is implemented in the FEniCS framework [52] in Python. In the test set, 48 images have homogeneous internal fluence rate and 144 images have a heterogeneous internal fluence rate modelled with the DA with illumination from either 1, 2 or 4 sides.

For the test set, values for absorption μ_a and reduced scattering μ'_s were selected from literature [53] to represent the values of whole blood for the vessels and glandular tissue or lipid for the background. For the training set, vessel absorption and background scattering were chosen in a wider range to improve reconstruction robustness, as can be seen in Section 4.5.1. The exact values are stated in Table 4.2.

data set	part of phantom	value
Test set	vessel	$\mu_a = 0.40 \text{ mm}^{-1}$ $\mu'_s = 0.45 \text{ mm}^{-1}$
	background	$\mu_a = 0.004 \text{ mm}^{-1}$ $\mu'_s = 0.97 \text{ mm}^{-1}$
Training set	vessel	$\mu_a \in [0.20 \text{ mm}^{-1}, 0.60 \text{ mm}^{-1}]$ $\mu'_s = 0.45 \text{ mm}^{-1}$
	background	$\mu_a = 0.004 \text{ mm}^{-1}$ $\mu'_s \in [0.50 \text{ mm}^{-1}, 2.00 \text{ mm}^{-1}]$

Table 4.2: Optical properties in the test and training set.

For these more realistic phantoms with spatially varying fluence rate it is difficult to detect the vascular geometry, since blood vessels deep in the tissue will give a lower initial pressure than shallow ones. Moreover, low intensity parts of the reconstruction can be invisible due to artefacts originating from higher intensity parts. To accurately reveal the vascular geometry at all tissue depths, we propose a method to compute the segmentation jointly with the photoacoustic reconstruction: instead of solely training for the best reconstruction, the neural network architecture is also trained for a segmentation

output. The same network as in Section 4.4.1 is used, which was explained in Section 4.3. Now however, not only a reconstruction output x_1^N in the first channel is required, but also a segmentation output x_2^N in the second channel of iteration N . This is achieved by extending the reconstruction-based loss function (4.4.1) with a binary cross-entropy loss on a segmentation output compared to the ground truth segmentation x_S .

$$\mathcal{L} = \left\{ \|x_{GT} - x_1^N\|_2^2 - \beta \left(x_S \log(x_2^N) + (1 - x_S) \log(1 - x_2^N) \right) \right\}. \quad (4.4.2)$$

Here β is a parameter that determines the weighting between reconstruction quality and segmentation quality. For all our experiments, $\beta = 0.5$ was chosen, as empirically it is seen that solutions are not very sensitive to this parameter. Note that all the layers before the output layer are the same, meaning that the same information is used for the reconstruction and for the segmentation, which makes it a joint algorithm.

Because the neural network needs to be differentiable, the segmentation output x_2^N is not binary, but in the range $[0, 1]$. This can be interpreted as some kind of reliability estimate: when a pixel value is very close to 1, it is almost surely part of a vessel, while it is almost surely not when its value is very close to 0. To obtain a truly binary segmentation, the non-binary output will be thresholded at the value that gives the highest segmentation accuracy and highest Dice-coefficient [54]. The interpretation of the output as a reliability estimate implies that this threshold should be similar for all outputs, which is verified by experiments. For this reason, we choose one threshold which provides the best binary segmentation, averaged over the whole training set.

4.4.3 Comparison with other methods

The reconstruction output of the joint L-PD algorithm is compared with the following other methods:

- Direct FBP method as described in [7].
- Iterative TV-regularised method (4.3.3) using PDGHM as described in Algorithm 3.1. To ensure that the algorithm is converged to high precision, we use 1000 iterations. Visually, the results do not change after the first 100 iterations. The regularisation parameter α was chosen such that the PSNR between ground truth and reconstruction is highest. Moreover, a scaling was introduced to compensate for contrast loss that is well known for TV. That is, the reconstruction is multiplied with a scalar such that the outcome gives the highest PSNR.
- Learned U-Net approach [18], where a neural network is used to post-process the FBP-reconstruction for a higher quality reconstruction. Parameters are chosen as in [18].

For segmentation, the following comparisons are made:

- Thresholding the FBP-reconstruction and TV-reconstruction.
- A spectral globally convex segmentation (S-GCS) [55], which can be seen as a convexified version of the level set method [56] applied with different scales. We use the numerical implementation as described in [55], where we choose 15 scales of $\alpha \in \{2, 3, \dots, 16\}$. The output scales are normalised to make comparison to the L-PD output possible. The algorithm is applied to the FBP-reconstruction and uses an L^1 data-fidelity term: this can potentially account for the decaying intensity in the reconstruction.
- Learned U-Net approach [17], where we ask for a segmentation with the FBP-reconstruction as input.

4.4.4 Application to experimental data

We apply the L-PD algorithm that was trained on the training set as explained in Section 4.4.2 to experimental data. With this we check the potential of L-PD to be applied to real measurements and the uncertainty that comes with it.

Experimental setup

We make use of a tomographic photoacoustic imager as specified in [57]. A laser delivers 5 ns pulses of optical energy with a wavelength of 532 nm. For the recording of pressure waves a 1D piezoelectric detector array with 64 elements in a half-circle is used. The detector array has a central frequency of 7.5 MHz with a fractional bandwidth of 85%. The acquired measurements are discretised with a time sampling of 25 ns, for a total time of approximately 45 μ s. Both the detectors and fibre bundles are simultaneously rotatable over 360 degrees. They can also be translated in order to image multiple slices. The detector array has a narrow focus (0.6 mm slice thickness) in one dimension, making it suitable for 2D slice based imaging. A schematic overview of the experimental setup is shown in Figure 4.4.2. For a more extensive explanation of the setup, we refer to [57].

Figure 4.4.2: Schematic overview of the experimental setup, where a phantom is measured.

Vascular phantom

For experimental data, a phantom with vessel-shaped absorbers in an optically scattering medium was created. Absorbing filaments were gently placed on stiff agar gel, after which a second layer of agar solution was poured on the whole and allowed to harden. A photograph of the filaments before pouring the second layer of agar is shown in Figure 4.4.3. The experimental agar phantom has the property that the sound speed is almost identical to that of water of the same temperature. For details on the phantom creation, we refer to Section 3.5.3.

Figure 4.4.3: **Left:** Experimental phantom consisting of absorbing filaments in a scattering medium. **Right:** Schematic of detector settings used in experiments. Black: 32 detectors uniform. Red: 64 detectors uniform. Blue: 64 detectors limited-view.

Measurements are obtained in a limited-view and a uniform detector setting: in the former, fibre bundles and detector array are placed on the top of the phantom; in the latter, fibre bundles and detector array are rotated over an angle of 180 degrees, which means that the detector array covers the whole circle. We point out that by this rotation not only acquisition is uniform, but also light is more homogeneously spread over the surface of the phantom. A schematic drawing of the detector setting is given in Figure 4.4.3, with two uniform settings (32 and 64 detectors) and one limited-view setting (64 detectors).

4.5 Results

In Figure 4.5.1 reconstructions of a digital phantom from class 0 are shown, where measurements were obtained by simulating for a uniform placement of 32 detectors. In line with the results in [31] it can be seen that the non-learned methods FBP and TV show artefacts in the background due to limited measurements and noise, while L-PD gives an almost perfect reconstruction with clear contrast and zero background.

Figure 4.5.1: Reconstructions of a vascular phantom of class 0 using a uniform placement of 32 detectors.

4.5.1 Robustness against uncertainty

Image uncertainty

Both the large and the small L-PD network were trained on all the image classes as defined in Table 4.1 and shown in Figure 4.4.1. We compared reconstructions of L-PD trained on class 0 to L-PD trained on the same class as the test class.

One visual comparison is provided in Figure 4.5.2. Here it is seen that the L-PD network trained on class 0 incorrectly removes most of the background when applied to test data from class 3: only two spots of high background intensity remain. One might argue that this is a positive property, since out-of-focus elements or low-absorbing structures without diagnostic value can be removed. However, spurious high-intensity background elements can be misclassified as a part of the vascular geometry, which is undesired. After training on class 3, the complete background is correctly identified, which makes it less likely that background elements are misclassified as vessels.

Figure 4.5.2: Example reconstructions of L-PD trained on different classes, both applied to the same test data (class 3).

We compare peak signal-to-noise ratios (PSNR) of the small and large network for the whole test set for all classes. In Table 4.3 it can be seen that for almost all image classes, it helps to train on the specific image class instead of the class 0. This is not very surprising, since the specific class is tailored to the test set on which the trained network is applied. It is however interesting to see that the degree of improvement was very different for the different classes: geometrical changes such as diameter (class 1) and coverage (class 4) did not

show any significant improvement, while intensity changes such as contrast (class 2) and background (class 3) showed that it was necessary to retrain the network for this specific class. Furthermore it was surprising that retraining for different amounts of noise (class 6) was not really necessary, indicating that the denoising capacity of L-PD seems to be quite robust. It is not clear to us why for a test set of class 1 the PSNR value is higher when trained on class 0 instead of on class 1; we hypothesise that this is due to the stochasticity of training the neural network in terms of the optimisation.

network type	class 0: nothing	class 1: diameter	class 2: contrast	class 3: background
large (10 iterations) trained on class i	41.00	39.81	39.53	38.27
large (10 iterations) trained on class 0	41.00	40.78	28.77	31.86
small (5 iterations) trained on class i	39.32	39.40	37.33	36.22
small (5 iterations) trained on class 0	39.32	39.09	25.50	31.04

network type	class 4: coverage	class 5: structure	class 6: noise level	class 7: everything
large (10 iterations) trained on class i	39.18	39.09	37.72	31.86
large (10 iterations) trained on class 0	38.00	35.97	36.64	21.64
small (5 iterations) trained on class i	37.20	37.46	36.38	30.21
small (5 iterations) trained on class 0	36.61	34.72	35.48	18.88

Table 4.3: Performance comparison (PSNR) between training on standard class and training on class with more variety in image properties.

It can be seen as well that taking a larger network gives a minor improvement over the small network, but no extreme changes are seen. Finally, it is also apparent that the class that contains all the variety of all the other classes benefits from retraining on the same class, but due to this variety, the PSNR value is lower than for other classes. It is unclear to which extent this could be solved by simply taking a bigger training set.

System uncertainty

Next we compare reconstructions of L-PD trained on a fixed setting of 32 detectors to L-PD trained on a random number of detectors, as described in Section 4.4.1. Here only the results for a test set with 64 detectors are shown: the full results are provided in Appendix A. In Figure 4.5.3, it can be seen that without training for invariance in detector placement, there are a lot of undesired local intensity changes in the reconstruction. When training is done with data from a random detector placement, we obtain the desired high-quality

reconstruction. This means that L-PD can be applied to data from different number of detectors, without having to perform a time-consuming retraining of the algorithm.

Figure 4.5.3: Reconstructions of L-PD trained on 32 or a random number of detectors, applied to test data from 64 detectors.

In Appendix B and Appendix C it is shown that a trained L-PD algorithm is robust against heterogeneous sound speed changes and uncertainty in the calibration measurements, respectively.

4.5.2 Joint reconstruction and segmentation

In this section, the joint reconstruction-segmentation algorithm as explained in Section 4.4.2 is compared to four different well-known methods. In Figure 4.5.4 a comparison between the different reconstruction methods is shown for synthetic data using a uniform placement of 32 detectors.

Figure 4.5.4: Reconstructions of a phantom using a uniform placement of 32 detectors: L-PD gives an improved reconstruction.

In this specific phantom, light was modelled to come from the left hand side of the image, resulting in a decaying fluence rate from left to right. It can be seen that TV and U-Net remove part of the background and partly smooth the vascular structure, but only L-PD is able to correctly identify the entire background while retaining most of the vascular structure.

The same synthetic data was used for Figures 4.5.5 and 4.5.6, where S-GCS and U-Net are compared to L-PD. Both the non-binary output (Figure 4.5.5) and the binary output 4.5.6 indicate the superior performance of the L-PD method. The L-PD method provides the correct vascular geometry, while both S-GCS and U-Net fail in the region where the initial pressure is lower. Furthermore, S-GCS, which is based on a level-set method, connects different structures which are not necessarily connected. It can be seen that the L-PD segmentation gives somewhat ‘fuzzy’ results in the right part of the image, while it is very sharp in the left part. This indicates that the L-PD segmentation is pixel-wise less reliable when the fluence rate drops. However, it is still clear that four blood vessels are migrating to the right; this possibly important diagnostic information is completely absent in the other segmentation methods.

Figure 4.5.5: Segmentations of the same phantom as in Figure 4.5.4. L-PD gives better geometrical information than other methods.

Figure 4.5.6: Thresholded segmentations from the result shown in Figure 4.5.5. L-PD shows connected vascular structures.

Compressive sampling is a useful way to limit measurement time, thereby minimising influences due to patient movement. We quantitatively show the superior performance of L-PD in both reconstruction and segmentation under uniform compressive samplings. In Figure 4.5.7 it can be seen that both learned reconstruction methods (U-Net and L-PD) give higher quality reconstructions

than non-learned methods (FBP and TV). Moreover, it can be seen that L-PD is significantly better than U-Net, especially in case of substantial compressive sampling, meaning a smaller number of detectors. This can be explained by the fact that the FBP-reconstruction with less than 32 detectors sometimes misses important features in the vascular geometry, which cannot be regained with the U-Net approach. With 64 detectors, the FBP-reconstruction does not really miss any important features, so U-Net can be adequately applied as an artefact removal tool.

Figure 4.5.7: Reconstruction quality for four reconstruction methods under uniform compressive sampling: L-PD outperforms other methods in all cases.

In Figure 4.5.8 the segmentation quality is assessed with the area under ROC-curve. The ROC-curve is obtained by comparing the false positive rate to the true positive rate of the thresholded reconstruction for various threshold values. The area under the curve (AUC) is then used as a single value reflecting the segmentation accuracy. The input image for computing the AUC is a non-binary image. In Figure 4.5.8 it can be seen that the segmentation quality is very much in line with the reconstruction quality that was shown before: learned methods give more accurate segmentations than non-learned methods. Again, the L-PD method is particularly interesting in case of highly limited number of detectors: it helps to add the acoustic operator in the training, since otherwise previously created artefacts can be difficult to remove.

Figure 4.5.8: Segmentation quality for five segmentation methods under uniform compressive sampling: L-PD outperforms other methods in all cases.

4.5.3 Experimental results

All methods are tested on experimental data, which was obtained by measuring the experimental phantom as explained in Section 4.4.4. Reconstructions and segmentations of the experimental phantom for a uniform sampling of 64 detectors are shown in Figures 4.5.9 and 4.5.10, respectively. Although the sampling of 64 detectors allows for a reasonable reconstruction with either TV or U-Net, the reconstruction with L-PD shows a clearer vascular structure with less artefacts in the background. This is also seen in the segmentation, where a larger part of the vascular geometry is identified.

Figure 4.5.9: Reconstructions of an experimental phantom with uniform placement of 64 detectors.

Figure 4.5.10: Segmentations of an experimental phantom with uniform placement of 64 detectors.

Figure 4.5.11: Segmentations of experimental phantom with limited view placement of 64 detectors at the top of the image.

Figure 4.5.12: Reconstructions of experimental phantom with limited view placement of 64 detectors at the top of the image.

In Figures 4.5.12 and 4.5.11, reconstructions and segmentations for the limited-view sampling of 64 detectors are shown. Recall that in this experimental

setup, light deposition and ultrasound detection is at the top half of the phantom. Also here it can be seen that the L-PD approach gives superior reconstructions to the other methods. The L-PD segmentation does not create the spurious parts that are visible in the U-Net segmentation.

4.6 Conclusion and outlook

In this chapter we developed a partially learned joint reconstruction and segmentation method for PAT. The sensitivity of the algorithm for reconstruction was investigated. When suitable training data is considered, the method has shown to be robust to changes in image and system settings and variety in sound speed and calibration measurements. For this, it is assumed that the expected variety can be estimated before training. Compared to state-of-the-art, our method gives higher quality reconstruction and segmentation results with less computational costs, both in simulation and experiment. More specifically, applying our method in a limited-angle scenario results in less streaking artefacts compared to the U-Net approach and can handle smooth intensity gradients. Thus, by improving reconstruction and segmentation quality, photoacoustic tomography can achieve deeper penetration in tissue. The approach can be modified for imaging modalities with different underlying physics or can be extended to other high-level tasks beyond segmentation.

In our experiments only 2D reconstruction and segmentation were considered, while both 2D [28] and 3D systems [3] are in use. Our method can be applied to 3D without changing the structure of the algorithm. If the 3D photoacoustic operator is computationally too expensive, one might have to change the learning of the algorithm to a step-by-step approach, as was done in [33], but this does not change the structure of the convolutional neural network.

One approach to receive the desired robust algorithm is to train on a big variety of input images and system settings. However, when this variety is large, training will take a considerable amount of time. Moreover, it is never possible to train for all possible settings. Therefore, future research could focus on learning a representation of the manifold in which the reconstructions and segmentations should lie. This could be done by combining L-PD with a generative adversarial network [58] or variational autoencoder [59] which learns a low-dimensional representation of the ground truth images. These networks could then learn from a finite number of settings and interpolate on the learned manifold to cover all reasonable inputs and system settings.

References

- [1] Y. Zhou, J. Yao, and L. V. Wang, "Tutorial on photoacoustic tomography," *Journal of biomedical optics*, vol. 21, no. 6, p. 061007, 2016.
- [2] M. Heijblom, D. Piras, M. Brinkhuis, J. C. van Hespén, F. M. van den Engh, M. van der Schaaf, J. M. Klaase, T. G. van Leeuwen, W. Steenbergen, and S. Manohar, "Photoacoustic image patterns of breast carcinoma and comparisons with magnetic resonance imaging and vascular stained histopathology," *Scientific reports*, vol. 5, p. 11778, 2015.
- [3] M. Toi, Y. Asao, Y. Matsumoto, H. Sekiguchi, A. Yoshikawa, M. Takada, M. Kataoka, T. Endo, N. Kawaguchi-Sakita, M. Kawashima, *et al.*, "Visualization of tumor-related blood vessels in human breast by photoacoustic imaging system with a hemispherical detector array," *Scientific reports*, vol. 7, p. 41970, 2017.
- [4] L. Lin, P. Hu, J. Shi, C. M. Appleton, K. Maslov, L. Li, R. Zhang, and L. V. Wang, "Single-breath-hold photoacoustic computed tomography of the breast," *Nature communications*, vol. 9, no. 1, pp. 1–9, 2018.
- [5] R. A. Kruger, P. Liu, Y. R. Fang, and C. R. Appledorn, "Photoacoustic ultrasound (PAUS)-reconstruction tomography," *Medical physics*, vol. 22, no. 10, pp. 1605–1609, 1995.
- [6] D. Finch and S. K. Patch, "Determining a function from its mean values over a family of spheres," *SIAM journal on mathematical analysis*, vol. 35, no. 5, pp. 1213–1240, 2004.
- [7] R. Willeminck, *Image reconstruction in a passive element enriched photoacoustic tomography setup*. PhD thesis, University of Twente, 2010.
- [8] M. Haltmeier, "Universal inversion formulas for recovering a function from spherical means," *SIAM Journal on Mathematical Analysis*, vol. 46, no. 1, pp. 214–232, 2014.
- [9] P. Burgholzer, G. J. Matt, M. Haltmeier, and G. Paltauf, "Exact and approximative imaging methods for photoacoustic tomography using an arbitrary detection surface," *Physical Review E*, vol. 75, no. 4, p. 046706, 2007.
- [10] B. E. Treeby and B. T. Cox, "k-Wave: Matlab toolbox for the simulation and reconstruction of photoacoustic wave fields," *Journal of biomedical optics*, vol. 15, no. 2, p. 021314, 2010.
- [11] X. L. Dean-Ben, A. Buehler, V. Ntziachristos, and D. Razansky, "Accurate model-based reconstruction algorithm for three-dimensional optoacoustic tomography," *IEEE transactions on medical imaging*, vol. 31, no. 10, pp. 1922–1928, 2012.
- [12] C. Huang, K. Wang, L. Nie, L. V. Wang, and M. A. Anastasio, "Full-wave iterative image reconstruction in photoacoustic tomography with acoustically inhomogeneous media," *IEEE transactions on medical imaging*, vol. 32, no. 6, pp. 1097–1110, 2013.
- [13] S. Arridge, P. Beard, M. Betcke, B. Cox, N. Huynh, F. Lucka, O. Ogunlade, and E. Zhang, "Accelerated high-resolution photoacoustic tomography via compressed sensing," *Physics in Medicine & Biology*, vol. 61, no. 24, p. 8908, 2016.
- [14] S. R. Arridge, M. M. Betcke, B. T. Cox, F. Lucka, and B. E. Treeby, "On the adjoint operator in photoacoustic tomography," *Inverse Problems*, vol. 32, no. 11, p. 115012, 2016.

- [15] M. Haltmeier and L. V. Nguyen, "Reconstruction algorithms for photoacoustic tomography in heterogeneous damping media," *Journal of Mathematical Imaging and Vision*, vol. 61, no. 7, pp. 1007–1021, 2019.
- [16] J. Frikel and M. Haltmeier, "Efficient regularization with wavelet sparsity constraints in photoacoustic tomography," *Inverse Problems*, vol. 34, no. 2, p. 024006, 2018.
- [17] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241, Springer, 2015.
- [18] 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, 2017.
- [19] S. Antholzer, M. Haltmeier, and J. Schwab, "Deep learning for photoacoustic tomography from sparse data," *Inverse problems in science and engineering*, vol. 27, no. 7, pp. 987–1005, 2019.
- [20] A. Kofler, M. Haltmeier, C. Kolbitsch, M. Kachelrieß, and M. Dewey, "A U-Nets cascade for sparse view computed tomography," in *International Workshop on Machine Learning for Medical Image Reconstruction*, pp. 91–99, Springer, 2018.
- [21] D. Waibel, J. Gröhl, F. Isensee, T. Kirchner, K. Maier-Hein, and L. Maier-Hein, "Reconstruction of initial pressure from limited view photoacoustic images using deep learning," in *Photons Plus Ultrasound: Imaging and Sensing 2018*, vol. 10494, p. 104942S, International Society for Optics and Photonics, 2018.
- [22] J. Schwab, S. Antholzer, R. Nuster, and M. Haltmeier, "Real-time photoacoustic projection imaging using deep learning," *arXiv:1801.06693*, 2018.
- [23] B. Kelly, T. P. Matthews, and M. A. Anastasio, "Deep learning-guided image reconstruction from incomplete data," *arXiv:1709.00584*, 2017.
- [24] E. Kobler, T. Klatzer, K. Hammernik, and T. Pock, "Variational networks: connecting variational methods and deep learning," in *German conference on pattern recognition*, pp. 281–293, Springer, 2017.
- [25] K. Hammernik, T. Würfl, T. Pock, and A. Maier, "A deep learning architecture for limited-angle computed tomography reconstruction," in *Bildverarbeitung für die Medizin 2017*, pp. 92–97, Springer, 2017.
- [26] K. Hammernik, T. Klatzer, E. Kobler, M. P. Recht, D. K. Sodickson, T. Pock, and F. Knoll, "Learning a variational network for reconstruction of accelerated MRI data," *Magnetic resonance in medicine*, vol. 79, no. 6, pp. 3055–3071, 2018.
- [27] H. Chen, Y. Zhang, Y. Chen, J. Zhang, W. Zhang, H. Sun, Y. Lv, P. Liao, J. Zhou, and G. Wang, "LEARN: learned experts' assessment-based reconstruction network for sparse-data CT," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1333–1347, 2018.
- [28] H. Li, J. Schwab, S. Antholzer, and M. Haltmeier, "NETT: solving inverse problems with deep neural networks," *Inverse Problems*, 2020.
- [29] J. Adler and O. Öktem, "Solving ill-posed inverse problems using iterative deep neural networks," *Inverse Problems*, vol. 33, no. 12, p. 124007, 2017.

- [30] T. Meinhardt, M. Moller, C. Hazirbas, and D. Cremers, "Learning proximal operators: using denoising networks for regularizing inverse imaging problems," in *Proc. of the IEEE International Conference on Computer Vision*, pp. 1781–1790, 2017.
- [31] J. Adler and O. Öktem, "Learned primal-dual reconstruction," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1322–1332, 2018.
- [32] E. Esser, X. Zhang, and T. F. Chan, "A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science," *SIAM Journal on Imaging Sciences*, vol. 3, no. 4, pp. 1015–1046, 2010.
- [33] A. Hauptmann, F. Lucka, M. Betcke, N. Huynh, J. Adler, B. Cox, P. Beard, S. Ourselin, and S. Arridge, "Model-based learning for accelerated, limited-view 3-D photoacoustic tomography," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1382–1393, 2018.
- [34] A. Hauptmann, B. Cox, F. Lucka, N. Huynh, M. Betcke, P. Beard, and S. Arridge, "Approximate k-space models and deep learning for fast photoacoustic reconstruction," in *International Workshop on Machine Learning for Medical Image Reconstruction*, pp. 103–111, Springer, 2018.
- [35] S. Banert, A. Ringh, J. Adler, J. Karlsson, and O. Öktem, "Data-driven nonsmooth optimization," *SIAM Journal on Optimization*, vol. 30, no. 1, pp. 102–131, 2020.
- [36] S. T. Acton and N. Ray, "Biomedical image analysis: segmentation," *Synthesis Lectures on Image, Video, and Multimedia Processing*, vol. 4, no. 1, pp. 1–108, 2009.
- [37] B. Soetikno, S. Hu, E. Gonzales, Q. Zhong, K. Maslov, J.-M. Lee, and L. V. Wang, "Vessel segmentation analysis of ischemic stroke images acquired with photoacoustic microscopy," in *Photons Plus Ultrasound: Imaging and Sensing 2012*, vol. 8223, p. 822345, International Society for Optics and Photonics, 2012.
- [38] P. Raunonen and T. Tarvainen, "Segmentation of vessel structures from photoacoustic images with reliability assessment," *Biomedical optics express*, vol. 9, no. 7, pp. 2887–2904, 2018.
- [39] V. Corona, M. Benning, M. J. Ehrhardt, L. F. Gladden, R. Mair, A. Reci, A. J. Sederman, S. Reichelt, and C.-B. Schönlieb, "Enhancing joint reconstruction and segmentation with non-convex bregman iteration," *Inverse Problems*, vol. 35, no. 5, p. 055001, 2019.
- [40] C. Qin, W. Bai, J. Schlemper, S. E. Petersen, S. K. Piechnik, S. Neubauer, and D. Rueckert, "Joint motion estimation and segmentation from undersampled cardiac MR image," in *International Workshop on Machine Learning for Medical Image Reconstruction*, pp. 55–63, Springer, 2018.
- [41] J. Adler, S. Lunz, O. Verdier, C.-B. Schönlieb, and O. Öktem, "Task adapted reconstruction for inverse problems," *arXiv:1809.00948*, 2018.
- [42] B. T. Cox, J. G. Laufer, P. C. Beard, and S. R. Arridge, "Quantitative spectroscopic photoacoustic imaging: a review," *Journal of biomedical optics*, vol. 17, no. 6, p. 061202, 2012.
- [43] Y. Wang, D. Xing, Y. Zeng, and Q. Chen, "Photoacoustic imaging with deconvolution algorithm," *Physics in Medicine & Biology*, vol. 49, no. 14, p. 3117, 2004.
- [44] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: nonlinear phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.

- [45] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of mathematical imaging and vision*, vol. 40, no. 1, pp. 120–145, 2011.
- [46] E. Y. Sidky, J. H. Jørgensen, and X. Pan, "Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm," *Physics in Medicine & Biology*, vol. 57, no. 10, p. 3065, 2012.
- [47] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. Van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE transactions on medical imaging*, vol. 23, no. 4, pp. 501–509, 2004.
- [48] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," *arXiv:1412.6980*, 2014.
- [49] S. L. Smith, P.-J. Kindermans, C. Ying, and Q. V. Le, "Don't decay the learning rate, increase the batch size," *arXiv:1711.00489*, 2017.
- [50] G. Paltauf, R. Nuster, M. Haltmeier, and P. Burgholzer, "Experimental evaluation of reconstruction algorithms for limited view photoacoustic tomography with line detectors," *Inverse Problems*, vol. 23, no. 6, p. S81, 2007.
- [51] S. Arridge, M. Schweiger, M. Hiraoka, and D. Delpy, "A finite element approach for modeling photon transport in tissue," *Medical physics*, vol. 20, no. 2, pp. 299–309, 1993.
- [52] M. S. Alnæs, A. Logg, K. B. Ølgaard, M. E. Rognes, and G. N. Wells, "Unified form language: a domain-specific language for weak formulations of partial differential equations," *ACM Transactions on Mathematical Software (TOMS)*, vol. 40, no. 2, pp. 1–37, 2014.
- [53] S. L. Jacques, "Optical properties of biological tissues: a review," *Physics in Medicine & Biology*, vol. 58, no. 11, p. R37, 2013.
- [54] A. P. Zijdenbos, B. M. Dawant, R. A. Margolin, and A. C. Palmer, "Morphometric analysis of white matter lesions in MR images: method and validation," *IEEE transactions on medical imaging*, vol. 13, no. 4, pp. 716–724, 1994.
- [55] L. Zeune, S. A. van Gils, L. W. Terstappen, and C. Brune, "Combining contrast invariant l1 data fidelities with nonlinear spectral image decomposition," in *International Conference on Scale Space and Variational Methods in Computer Vision*, pp. 80–93, Springer, 2017.
- [56] T. Goldstein, X. Bresson, and S. Osher, "Geometric applications of the split Bregman method: segmentation and surface reconstruction," *Journal of Scientific Computing*, vol. 45, no. 1-3, pp. 272–293, 2010.
- [57] P. van Es, R. C. Vlieg, E. Hondebrink, J. C. van Hespén, H. J. B. Moens, W. Steenbergen, and S. Manohar, "Coregistered photoacoustic and ultrasound tomography of healthy and inflamed human interphalangeal joints," in *European Conference on Biomedical Optics*, p. 95390C, Optical Society of America, 2015.
- [58] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, pp. 2672–2680, 2014.
- [59] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," *arXiv:1312.6114*, 2013.

Appendices

A Robustness to detector settings: full results

In this section the full results of the experiments described in Section 4.4.1 are shown: we compare reconstructions of L-PD trained on a fixed setting of 32 detectors to L-PD trained on a random number of detectors. In Figure A.1, it can be seen that the L-PD algorithm trained on a fixed detector setting can not directly be applied to a different detector setting. All reconstructions show significant artefacts, however it is not clear to us why different detector settings give very different artefacts.

Figure A.1: L-PD trained on a fixed setting of 32 detectors, tested on different detector settings.

In Figure A.2, it can be seen that it helps to train on many detector settings, randomly chosen from a possibility of 128 detectors. The reconstruction quality is good for all images.

Figure A.2: L-PD trained on a variable detector setting, tested on different detector settings.

B Robustness to varying heterogeneous sound speed

In this section we analyse the robustness of a trained L-PD algorithm to changes in the sound speed. The algorithm was trained for a uniform detector setting with 32 detectors, as explained in Section 4.4.1. Next, the test set is changed to a scenario where the synthetic phantom has a higher sound speed than the surrounding medium (Figure B.1). This means that the sound speed is still assumed to be known, but not constant throughout the medium. Derivation of the model with nonconstant speed of sound is done by calculating the time-of-flight values, taking refraction of rays into account. The approach is based on solving the Eikonal equation. Details can be found in [7, Section 5.3].

Figure B.1: Sound speed distribution for a uniform detector setting with 32 detectors: the orange region (synthetic phantom) has a higher sound speed than the blue region (water).

In Figure B.2 it can be seen that the reconstruction quality does not become worse when a heterogeneous sound speed is assumed: even if the sound speed is considerably higher within the synthetic phantom (+80 m/s), the reconstruction has the same quality as in the homogeneous sound speed case.

Figure B.2: L-PD trained for a homogeneous sound speed, tested on data obtained with a heterogeneous sound speed distribution.

C Robustness to different calibrations

Here the robustness of a trained L-PD algorithm to different calibrations is analysed. As explained in Section 4.2.1, a deconvolution is applied to the measured data, which gives us the pre-processed data y . This pre-processed data is then used as input for the L-PD algorithm. A different calibration measurement p_{cal} gives us a different deconvolution, which results in different input data for our L-PD algorithm.

We have measured four different calibration phantoms at 4 different times. The photoacoustic imaging system was shut down between the separate measurements to cover all system uncertainty that can be expected between any two measurements. The L-PD algorithm was trained for a uniform detector setting with 32 detectors, as explained in Section 4.4.1; this was done with calibration 1. Four test sets were created: one used the same calibration 1, the others had a different calibration in their pre-processing.

Figure C.1: L-PD trained on data from calibration 1 and tested on data from four different calibrations.

In Figure C.1, it can be seen that a different calibration has a very minor effect on the reconstruction quality: calibration 2 gives a visually identical result to calibration 1; calibration 3 shows two higher intensity spots in the background; calibration 4 shows a slightly increased overall intensity. The latter can be explained by the fact that the absorbed energy might have been different in the fourth calibration phantom, giving a different scaling when performing a deconvolution.

Chapter 5

Data-consistent neural networks for solving nonlinear inverse problems

Yoeri E. Boink, Markus Haltmeier,
Sean Holman, Johannes Schwab

submitted to SIAM Journal on Imaging Sciences (2020)
preprint available: arXiv 2003.11253

Abstract

Data-assisted reconstruction algorithms, incorporating trained neural networks, are a novel paradigm for solving inverse problems. One approach is to first apply a classical reconstruction method and then apply a neural network to improve its solution. Empirical evidence shows that plain two-step methods provide high-quality reconstructions, but they lack a convergence analysis. In this chapter we formalise the use of such two-step approaches with classical regularisation theory. We propose data-consistent neural networks that we combine with classical regularisation methods. This yields a data-driven regularisation method for which we provide a convergence analysis with respect to noise. Numerical simulations show that compared to standard two-step deep learning methods, our approach provides better stability with respect to structural changes in the test set, while performing similarly on test data similar to the training set. Our method provides a stable solution of inverse problems that exploits both the known nonlinear forward model as well as the desired solution manifold from data.

5.1 Introduction

In recent years there has been massive interest in new machine learning based approaches to reconstruction in inverse problems. In this chapter we consider a class of two-step methods for ill-posed nonlinear inverse problems which combine a known inversion technique with a second step given by a trained deep neural network. Many other works looked at similar methods, but our particular view is that the neural network should be designed so that individual reconstructions remain consistent with observed data. This means specifically that the neural network should preserve level sets of the forward operator, at least approximately if the data are noisy, and in this case we call the neural network *data-consistent*. Intuitively, data-consistent networks help us to pick among the various possible reconstructions that are all consistent with the observed data based on a known forward model. Traditionally, regularisation methods or a-priori assumptions perform the same function. However, we believe it is natural to let this choice be informed by a set of training data containing correct reconstructions. Because of this our approach, using data-consistent networks may be considered a type of learned data-driven regularisation for ill-posed problems.

This chapter contains both theoretical and numerical results on data-consistent networks for ill-posed nonlinear inverse problems. On the theoretical side we prove convergence, under certain hypotheses, of reconstructions obtained by data-consistent networks to the correct solution as the noise level goes to zero. This proves that data-consistent networks can provide a data-driven regularisation method for ill-posed problems in the standard technical sense. On the numerical side we apply data-consistent networks to the problem of saturation as well as reconstruction from severely undersampled and saturated Radon transform data. We compare the results of our data-consistent network to reconstructions using U-Nets [1, 2] without data-consistency.

Background

Here, we introduce the technical framework for our results. Let $(\mathcal{X}, \|\cdot\|)$ and $(\mathcal{Y}, \|\cdot\|)$ be Banach spaces¹ and let $\mathcal{A}: \mathcal{D} \subseteq \mathcal{X} \rightarrow \mathcal{Y}$ be a continuous, possibly nonlinear, mapping. We study the stable solution of the inverse problem problem:

$$\text{recover } x \in \mathcal{D} \text{ from } \mathcal{A}(x) = y, \quad (5.1.1)$$

where $y \in \mathcal{A}(\mathcal{D})$ are exact measurements. Furthermore, we are especially interested in the noisy case where measurements $y^\delta \in \mathcal{Y}$ are given with $\|y - y^\delta\| \leq \delta$. An inversion method for exact measurements is a right inverse $G_0: \mathcal{A}(\mathcal{D}) \rightarrow \mathcal{D}$ for $\mathcal{A}$,

$$\forall y \in \mathcal{A}(\mathcal{D}): \quad \mathcal{A}G_0(y) = y. \quad (5.1.2)$$

¹You can take $\mathcal{X}$ and $\mathcal{Y}$ as finite dimensional spaces $\mathbb{R}^n$ and $\mathbb{R}^q$ with the Euclidian norm if you are not familiar with infinite dimensional Banach spaces.

The inversion method G_0 therefore recovers elements in $G_0\mathcal{A}(\mathcal{D}) = \text{Fix}(G_0\mathcal{A})$, the set of fixed points of $G_0\mathcal{A}: \mathcal{D} \rightarrow \mathcal{D}$. We are mainly interested in the case where (5.1.1) is ill-posed, where no continuous right inverse G_0 exists. If noisy measurements $y^\delta \in \mathcal{Y}$ are given with $\|\mathcal{A}(x) - y^\delta\| \leq \delta$ for $x \in G_0\mathcal{A}(\mathcal{D})$, then $G_0(y^\delta)$ is either not well defined or arbitrary far away from x . In this case one has to apply regularisation methods to the measurements, which are stable approximations to G_0 and continuous on all of $\mathcal{Y}$.

Well established regularisation methods are quadratic Tikhonov regularisation [3, 4, 5] and iterative regularisation [6, 7]. Both methods are designed to approximate minimum norm solutions $G_0(y) \in \text{argmin} \{\|x - x_0\| \mid \mathcal{A}(x) = y\}$ with fixed $x_0 \in \mathcal{X}$. However, for most applications minimum norm solutions are not the desired ones. One way to overcome this issue is by convex variational regularisation [8], where one takes

$$G_\alpha(y^\delta) \in \text{argmin} \left\{ \frac{1}{2} \|\mathcal{A}(x) - y^\delta\|^2 + \alpha \mathcal{R}(x) \mid x \in \mathcal{D} \right\}, \quad (5.1.3)$$

which approximate $\mathcal{R}$ -minimising solutions $G_0(y) \in \text{argmin} \{\mathcal{R}(x) \mid \mathcal{A}(x) = y\}$. Here the regularisation functional $\mathcal{R}: \mathcal{X} \rightarrow [0, \infty]$ incorporates a-priori information and acts as criterion for selecting particular solutions. There are still several challenges related to variational regularisation techniques. For example, computing $\mathcal{R}$ -minimising solutions requires time consuming iterative minimisation schemes. Moreover, finding a regularisation functional that well models solutions of interest is a difficult issue. Typical choices such as total variation or the ℓ^q -norm with respect to some frame enforce strong handcrafted prior assumptions that are often not met in practical applications.

Recently, several deep learning methods to solve inverse problems were proposed. Some works apply iterative methods, where operators are replaced by neural networks [9, 10, 11, 12, 13]. This approach was also taken in Chapter 4 of this thesis. Others aim for a fully learned reconstruction scheme, like in [14] or Chapter 6 of this thesis. Other approaches are inspired by solving the regularised linear or nonlinear inverse problem by means of truncated Neumann series with learned components [15, 16] or by minimising a variational problem incorporating a learned regularisation functional [17, 18, 19]. Another popular approach for imaging problems is to use a neural network as a second step after some initial reconstruction. Several such post processing methods have been considered in the literature [20, 2, 12, 21]. In these two-step approaches, the reconstruction network takes the form $R = \Phi G$ where $G: \mathcal{Y} \rightarrow \mathcal{X}$ maps the data to the reconstruction space (reconstruction layer or backprojection) and $\Phi: \mathcal{X} \rightarrow \mathcal{X}$ is a neural network (NN) whose parameters are adjusted to the training data. In particular, so called residual networks $\Phi = \text{Id}_{\mathcal{X}} + \mathcal{N}$, where only the residual part $\mathcal{N}$ is trained [2, 22, 23, 24] show very accurate results for solving inverse problems.

In order to address the ill-posedness of linear inverse problems, regularising networks of the form $R_\alpha = \Phi_\alpha G_\alpha$ were introduced in [25, 26]. Here $G_\alpha: \mathcal{Y} \rightarrow$

$\mathcal{X}$ defines any regularisation and $\Phi_\alpha: \mathcal{X} \rightarrow \mathcal{X}$ are trained neural networks approximating data-consistent networks. In the linear case these networks are named nullspace networks because they only add parts in the kernel of $\mathcal{A}$ as proposed in [27, 26]. A somewhat similar approach was proposed by Bubba et al. [28], combining model-based sparse regularisation and deep learning, so that only the part that is not in the information of the data is learned and the remaining part is handled by a regularisation method.

In this chapter we extend the concept of nullspace networks and derive convergence and convergence rates for data-consistent network families $(R_\alpha)_{\alpha>0}$ for nonlinear problems and provide some numerical examples.

Regularising networks for nonlinear problems

Let G_0 be any right inverse of $\mathcal{A}$ and $(G_\alpha)_{\alpha>0}$ a regularisation of G_0 , for example classical Tikhonov regularisation. Further, let $(\Phi_\alpha)_{\alpha>0}$ be a family of Lipschitz continuous mappings $\Phi_\alpha: \mathcal{X} \rightarrow \mathcal{X}$ (networks). In this chapter we show that under suitable assumptions the reconstruction networks

$$R_\alpha := \Phi_\alpha G_\alpha: \mathcal{Y} \rightarrow \mathcal{X}, \quad (5.1.4)$$

define a convergent regularisation method. Additionally, we derive convergence rates for the reconstruction error.

A main condition for these results is that $\Phi_\alpha G_\alpha \mathcal{A}$ converges pointwise to a network of the form $\Phi_0 G_0 \mathcal{A}$ where $\Phi_0: \mathcal{X} \rightarrow \mathcal{X}$ is data-consistent in the sense that $\mathcal{A}\Phi_0 z = \mathcal{A}z$ for $z \in G_0 \mathcal{A}(\mathcal{D})$. The latter property implies that Φ_0 preserves data-consistency of G_0 , meaning that if $G_0(y)$ is a solution of (5.1.1), then $\Phi_0 G_0(y)$ is a solution of (5.1.1) too. Hence the goal of the learned mapping Φ_0 is to improve solutions of the inverse problem, while keeping data-consistency of the initial solution. We prove that the family of reconstruction networks $(\Phi_\alpha G_\alpha)_{\alpha>0}$ is a convergent regularisation method, and we provide convergence rates.

The benefits of $(\Phi_\alpha G_\alpha)_{\alpha>0}$ over $(G_\alpha)_{\alpha>0}$ are twofold. First, in the limit $\alpha \rightarrow 0$, the network $\Phi_0 G_0 \mathcal{A}$ selects solutions in $\Phi_0 G_0 \mathcal{A}(\mathcal{D})$ that can be trained to better reflect the desired signal class than $G_0 \mathcal{A}(\mathcal{D})$. Second, for $\alpha > 0$, the networks Φ_α can be trained to undo the smoothing effect of $G_\alpha \mathcal{A}$ and thereby allow for obtaining convergence rates for less regular elements than the original regularisation method $(G_\alpha)_{\alpha>0}$. The operator $\Phi_0 G_0$ can be seen as a right inverse that is learned from a suitable class of training data.

Outline

The rest of this chapter is organised as follows. In Section 5.2 we describe the regularisation of nonlinear inverse problems and we define the proposed two-step data-driven regularisation method. We introduce data-consistent networks, which allow the definition of *regularising networks*, a regularisation method

which approximates a data-driven right inverse. We investigate under which assumptions these networks generate a convergent regularisation method and we give examples of how such regularising networks can be constructed. In Section 5.3 we present a convergence analysis and derive convergence rates for the proposed method. Section 5.4 presents the mathematical description of the inverse problems considered in the numerical simulations, after which the implementation is explained in detail. Results for the simulations are shown in section 5.5. Additional numerical results can be found in the appendices. The chapter concludes with a short summary of the established theory and the numerical simulations.

Glossary

To make this chapter easier to read, we provide a table that gives an overview of the nomenclature of terms and their meaning.

Nomenclature	Description	Definition
$\mathcal{X}$	reconstruction space	Banach space (norm $\ \cdot\ $)
$\mathcal{Y}$	data space	Banach space (norm $\ \cdot\ $)
$\mathcal{D}$	domain of $\mathcal{A}$	$\mathcal{D} \subseteq \mathcal{X}$
$\mathcal{A}$	forward operator	$\mathcal{A}: \mathcal{D} \subseteq \mathcal{X} \rightarrow \mathcal{Y}$
δ	noise level	$\delta \geq 0$
G_0	right inverse of $\mathcal{A}$	$G_0: \mathcal{A}(\mathcal{D}) \rightarrow \mathcal{D}$
$\mathcal{X}_0$	reconstruction space with source condition	$\mathcal{X}_0 \subseteq G_0\mathcal{A}(\mathcal{D})$
$\mathcal{R}$	regularisation functional	$\mathcal{R}: \mathcal{X} \rightarrow [0, \infty]$
α	regularisation parameter	$\alpha \in (0, \infty)$
α^*	parameter choice function	$\alpha^*: (0, \infty) \times \mathcal{Y} \rightarrow (0, \infty)$
Φ_0	data-consistent neural network	$\Phi_0: \mathcal{X} \rightarrow \mathcal{X}$ s.t. $\mathcal{A}\Phi_0 = \mathcal{A}$
$(G_\alpha)_{\alpha>0}$	regularisation of G_0	$G_\alpha: \mathcal{Y} \rightarrow \mathcal{X}$
$(\Phi_\alpha)_{\alpha>0}$	family of Lipschitz mappings (networks)	$\Phi_\alpha: \mathcal{X} \rightarrow \mathcal{X}$
$(R_\alpha)_{\alpha>0}$	regularising networks	$R_\alpha := \Phi_\alpha G_\alpha$
G, G^δ, R^δ	reconstruction algorithm	$G, G^\delta, R^\delta: \mathcal{Y} \rightarrow \mathcal{X}$

Furthermore it is our convention to write x as the variable in the domain of the forward mapping, i.e. $x \in \mathcal{D}$, and to write z as the variable after some right inverse has been applied, i.e. $z \in G_0\mathcal{A}(\mathcal{D})$.

To increase readability, we write a combination of mappings, e.g. $\mathcal{A}$ and G_0 , as $G_0\mathcal{A}(x)$ instead of $G_0 \circ \mathcal{A}(x)$ or $G_0(\mathcal{A}(x))$ throughout this chapter.

5.2 Convergence of regularised data-consistent networks

Throughout the rest of this chapter, let $\mathcal{A}: \mathcal{D} \subseteq \mathcal{X} \rightarrow \mathcal{Y}$ be a continuous mapping between Banach spaces $\mathcal{X}$ and $\mathcal{Y}$. We study the stable solution of the inverse problem (5.1.1). In this section we introduce the regularising networks and present the convergence analysis.

5.2.1 Regularisation of inverse problems

Let $G_0: \mathcal{A}(\mathcal{D}) \rightarrow \mathcal{D}$ be a right inverse for $\mathcal{A}$. If (5.1.1) is ill-posed and y^δ are noisy data with $\|\mathcal{A}(x) - y^\delta\| \leq \delta$, then the reconstruction method $G_0(y^\delta)$ is unstable, meaning arbitrarily far away from $G_0\mathcal{A}(x)$ or even not defined. To obtain meaningful approximations of $G_0\mathcal{A}(x)$, one has to apply regularisation methods defined as follows.

Definition 5.2.1 (Regularisation method). Let $(G_\alpha)_{\alpha>0}$ be a family of continuous mappings $G_\alpha: \mathcal{Y} \rightarrow \mathcal{X}$. If for all $x \in G_0\mathcal{A}(\mathcal{D})$ there exists a parameter choice function $\alpha^*: (0, \infty) \times \mathcal{Y} \rightarrow (0, \infty)$ such that

$$\begin{aligned} 0 &= \limsup_{\delta \rightarrow 0} \{\alpha^*(\delta, y^\delta) \mid y^\delta \in B_\delta(\mathcal{A}(x))\} \\ 0 &= \limsup_{\delta \rightarrow 0} \{\|x - G_{\alpha^*(\delta, y^\delta)}(y^\delta)\| \mid y^\delta \in B_\delta(\mathcal{A}(x))\}, \end{aligned}$$

where $B_\delta(\mathcal{A}(x))$ is the ball with radius δ around $\mathcal{A}(x)$, we call $((G_\alpha)_{\alpha>0}, \alpha^*)$ a regularisation method for G_α . If $((G_\alpha)_{\alpha>0}, \alpha^*)$ is a regularisation method for G_0 , we call $(G_\alpha)_{\alpha>0}$ a regularisation of G_0 and α^* an admissible parameter choice.

Probably the best known regularisation is quadratic Tikhonov regularisation in Hilbert spaces [3]. Under the assumption that $\mathcal{A}$ is weakly sequentially closed, one shows that there exist solutions of (5.1.1) with minimal distance to a given point $x_0 \in \mathcal{X}$ and that

$$\mathcal{T}_{\alpha, y^\delta}(x) := \frac{1}{2} \|\mathcal{A}(x) - y^\delta\|^2 + \frac{\alpha}{2} \|x - x_0\|^2 \quad \text{for } x \in \mathcal{D} \quad (5.2.1)$$

has at least one minimiser. We can define $G_\alpha(y^\delta)$ as any minimiser of $\mathcal{T}_{\alpha, y^\delta}$. If the solution of (5.1.1) with minimal distance to x_0 is unique and denoted by $G_0(y)$, then $((G_\alpha)_{\alpha>0}, \alpha^*)$ with a parameter choice satisfying $\delta^2/\alpha^* \rightarrow 0$ and $\alpha^* \rightarrow 0$ as $\delta \rightarrow 0$ is a regularisation method for G_0 [3, 4]. Research indicates that solutions with minimal distance to a fixed initial guess $x_0 \in \mathcal{X}$ are too simple in many applications. The use of nonquadratic penalties has demonstrated to often give better results. Recently, deep learning methods showed outstanding performance. Here solutions are defined by a neural network that maps the given measurements to desired signals.

5.2.2 Data-consistent networks

The first ingredient for constructing regularising two-step networks are data-consistent networks.

Definition 5.2.2 (Data-consistent network). We call $\Phi_0: \mathcal{X} \rightarrow \mathcal{X}$ a data-consistent network if Φ_0 is Lipschitz continuous and $\forall z \in G_0\mathcal{A}(\mathcal{D}): \mathcal{A}\Phi_0(z) = \mathcal{A}(z)$.

In data-consistent networks, if $z \in G_0\mathcal{A}(\mathcal{D})$ is a solution of (5.1.1), then $\Phi_0(z)$ is solution of (5.1.1) too. In particular, $\Phi_0 G_0$ is a right inverse for $\mathcal{A}$ with solution set $\Phi_0 G_0 \mathcal{A}(\mathcal{D}) = \text{Fix}(\Phi_0 G_0 \mathcal{A})$. Data-consistent networks can be constructed by

$$\Phi_0(z) = P_{z,0}\mathcal{N}(z), \quad (5.2.2)$$

where $\mathcal{N}: \mathcal{X} \rightarrow \mathcal{X}$ is a Lipschitz continuous trained neural network, and $P_{z,0}: \mathcal{X} \rightarrow \mathcal{X}$ a Lipschitz continuous mapping with $P_{z,0}(x) \in \mathcal{A}^{-1}(\{\mathcal{A}(z)\}) = \{x \in \mathcal{D} \mid \mathcal{A}(x) = \mathcal{A}(z)\}$. The mapping $P_{z,0}$ can be seen as a generalised projection on the preimage $\mathcal{A}^{-1}(\{\mathcal{A}(z)\})$. In the special case where $\mathcal{A}$ is a linear mapping, $\Phi_0(z)$ can be chosen as $\Phi_0(z) = z + P_{\ker(\mathcal{A})}\mathcal{N}(z)$, where $P_{\ker}$ is the projection on the kernel of $\mathcal{A}$ [26].

Definition 5.2.3 (Regularising networks). We call $(R_\alpha: \mathcal{Y} \rightarrow \mathcal{X})_{\alpha>0}$ defined by $R_\alpha := \Phi_\alpha G_\alpha$ a family of regularising networks if the following hold:

- (R1) $(G_\alpha: \mathcal{Y} \rightarrow \mathcal{X})_{\alpha>0}$ is a regularisation of G_0 .
- (R2) $(\Phi_\alpha: \mathcal{X} \rightarrow \mathcal{X})_{\alpha>0}$ are uniformly L -Lipschitz continuous mappings.
- (R3) For some data-consistent network $\Phi_0: \mathcal{X} \rightarrow \mathcal{X}$ we have

$$\forall x \in \mathcal{D}: \lim_{\alpha \rightarrow 0} \Phi_\alpha G_\alpha \mathcal{A}(x) = \Phi_0 G_0 \mathcal{A}(x). \quad (5.2.3)$$

In practice an important issue is to design networks that converge to a data-consistent limiting network as the noise level goes to zero. Next we give examples for a possible strategy to train such networks.

Example 5.2.4. Let $x_1, \dots, x_N \in \mathcal{X}$ be training signals, and $y_i = \mathcal{A}(x_i)$ and y_i^δ be the corresponding exact and noisy data, respectively. Further define the vectors of reconstructions $z := (z_1, \dots, z_N)$ and $z^\alpha := (z_1^\alpha, \dots, z_N^\alpha)$ where $z_i = G_0(y_i)$ and $z_i^\alpha = G_\alpha(y_i^\delta)$. The weights of the neural network are denoted by $\theta \in \Theta$. We write $(\Phi^\theta)_{\theta \in \Theta}$ for a neural network with given architecture, whose weights have not yet been fixed.

- One approach is to take the networks $\Phi_\alpha := \Phi_0$ for all $\alpha > 0$, where Φ_0 is the network obtained by minimising the functional

$$\min_{\theta} \sum_{i=1}^N \|\Phi^\theta(z_i) - x_i\|^2 + \mathcal{R}(\theta). \quad (5.2.4)$$

Here $\mathcal{R}$ denotes some regularisation functional for the weights θ that may be used to ensure a small Lipschitz constant. Clearly, since Φ_0 is Lipschitz continuous and $(G_\alpha)_{\alpha>0}$ is a regularisation method, we have the desired limit in (5.2.3) for all $x \in \mathcal{D}$. Now if the data-consistency is incorporated in the network architecture, the condition (R3) is satisfied.

- A more sophisticated approach is to choose the sequence of networks depending on the regularisation parameter α . Here the networks Φ_α are obtained by minimising

$$\min_{\theta} \sum_i^N \|\Phi^\theta(z_i^\alpha) - x_i\|^2 + \mathcal{R}(\theta).$$

To enforce the data-consistency of the limiting network Φ_0 one could either choose the network architecture to be data-consistent, meaning $\forall \theta \in \Theta \forall x \in \mathcal{D}: \mathcal{A}(\Phi^\theta(x)) = \mathcal{A}(x)$, or taking networks increasingly data-consistent of the form

$$\Phi_\alpha(z) = P_{z,\alpha} \mathcal{N}_\alpha(z). \quad (5.2.5)$$

Here $\mathcal{N}_\alpha: \mathcal{X} \rightarrow \mathcal{X}$ is a trained network and $P_{z,\alpha}$ is a Lipschitz continuous mapping with $\text{Im}(P_{z,\alpha}) \subseteq E_{\alpha,z} := \{x \mid \|\mathcal{A}(x) - \mathcal{A}(z)\| \leq r(\alpha)\}$ with $\lim_{\alpha \rightarrow 0} r(\alpha) = 0$. Data-consistency is obtained in the limit. One example for $P_{z,\alpha}$ is the metric projection on $E_{\alpha,z}$, which is Lipschitz continuous if $E_{\alpha,z}$ is convex. Note that in (5.2.5) there are no restrictions on the particular choice of the architecture of the networks $\mathcal{N}_\alpha$.

- Another network architecture guaranteeing data-consistency is given by

$$\Phi_\alpha(z) = \Phi^{dec}(S_0 + \alpha S_1) \Phi^{enc}(z). \quad (5.2.6)$$

Here Φ^{enc} and Φ^{dec} denote an encoder and decoder network respectively, S_0 denotes an α -independent network such that $\Phi^{dec} S_0 \Phi^{enc}$ is data-consistent and S_1 denotes a network that is allowed to depend on α .

5.2.3 Convergence analysis

We have the following convergence result for regularising networks according to Definition 5.2.3.

Theorem 5.2.5 (Regularising networks). *Any family of regularising networks $(R_\alpha = \Phi_\alpha G_\alpha)_{\alpha>0}$ (see Definition 5.2.3) is a regularisation for $\Phi_0 G_0$ in the sense of Definition 5.2.1.*

Proof. Let $x \in \Phi_0 G_0 \mathcal{A}(\mathcal{D})$, $y^\delta \in \mathcal{Y}$ with $\|\mathcal{A}(x) - y^\delta\| \leq \delta$ and set $x_\alpha^\delta := \Phi_\alpha G_\alpha(y^\delta)$. Then

$$\begin{aligned} \|x - x_\alpha^\delta\| &= \|\Phi_0 G_0 \mathcal{A}(x) - \Phi_\alpha G_\alpha y^\delta\| \\ &\leq \|\Phi_0 G_0 \mathcal{A}(x) - \Phi_\alpha G_\alpha \mathcal{A}(x)\| + \|\Phi_\alpha G_\alpha \mathcal{A}(x) - \Phi_\alpha G_\alpha y^\delta\| \end{aligned}$$

$$\leq \|\Phi_\alpha G_\alpha \mathcal{A}(x) - \Phi_0 G_0 \mathcal{A}(x)\| + L \|G_\alpha \mathcal{A}(x) - G_\alpha y^\delta\|.$$

Now if $\alpha^*(\delta, y^\delta)$ is an admissible parameter choice for $(G_\alpha)_{\alpha>0}$, then

$$\begin{aligned} \sup_{y^\delta} \{ \|x - R_{\alpha^*(\delta, y^\delta)}(y^\delta)\| \mid y^\delta \in B_\delta(\mathcal{A}(x)) \} \\ \leq \|\Phi_{\alpha^*(\delta, y^\delta)} G_{\alpha^*(\delta, y^\delta)} \mathcal{A}(x) - \Phi_0 G_0 \mathcal{A}(x)\| \\ + L \sup_{y^\delta} \{ \|G_{\alpha^*(\delta, y^\delta)} \mathcal{A}(x) - G_{\alpha^*(\delta, y^\delta)} y^\delta\| \mid y^\delta \in B_\delta(\mathcal{A}(x)) \}. \end{aligned} \quad (5.2.7)$$

According to (R3) in Definition 5.2.3 the first term converges to zero. Because of $\|G_{\alpha^*(\delta, y^\delta)} \mathcal{A}(x) - G_{\alpha^*(\delta, y^\delta)} y^\delta\| \leq \|G_{\alpha^*(\delta, y^\delta)} \mathcal{A}(x) - G_0 \mathcal{A}(x)\| + \|G_0 \mathcal{A}(x) - G_{\alpha^*(\delta, y^\delta)} y^\delta\|$ and the fact that $(G_\alpha)_{\alpha>0}$ is a regularisation method for G_0 with parameter choice rule $\alpha = \alpha^*(\delta, y^\delta)$, the second term converges to zero as $\delta \rightarrow 0$. ■

5.3 Convergence rates

Another important issue for regularisation methods is the rate of approximation. This means specifically that there exists a decreasing function $f: (0, \infty) \rightarrow (0, \infty)$ such that $\lim_{\delta \rightarrow 0} f(\delta) = 0$ and $\|R_{\alpha^*(\delta, y^\delta)}(y^\delta) - x\| \leq f(\delta)$ uniformly for all $y^\delta \in \mathcal{Y}$ with $\|\mathcal{A}(x) - y^\delta\| \leq \delta$.

5.3.1 Reconstruction algorithms and convergence rates

In the following we call $G: \mathcal{Y} \rightarrow \mathcal{X}$ a reconstruction algorithm.

Definition 5.3.1 (Reconstruction error of an algorithm). Let $\mathcal{X}_0 \subseteq \mathcal{X}$, $\delta > 0$ and $G: \mathcal{Y} \rightarrow \mathcal{X}$ be a reconstruction algorithm. We call

$$\mathcal{E}(G, \delta, \mathcal{X}_0) = \sup_{x, y^\delta} \{ \|x - G(y^\delta)\| \mid x \in \mathcal{X}_0 \wedge y^\delta \in \overline{B_\delta(\mathcal{A}(x))} \} \quad (5.3.1)$$

the reconstruction error of G over $\mathcal{X}_0$.

Definition 5.3.2 (Convergence rate of an algorithm). Let $\mathcal{X}_0 \subseteq \mathcal{X}$, $r \in (0, 1]$ and for any $\delta > 0$, let G^δ be a reconstruction algorithm. We say that $(G_{\alpha^*(\delta, y^\delta)})_{\delta>0}$ converges at rate δ^r over $\mathcal{X}_0$ if $\mathcal{E}(G_{\alpha^*(\delta, y^\delta)}, \delta, \mathcal{X}_0) = \mathcal{O}(\delta^r)$ as $\delta \rightarrow 0$.

The concept of convergence rates in particular applies for reconstruction algorithms defined by regularisation methods. In general, no convergence rate over $G_0 \mathcal{A}(\mathcal{D})$ is possible; they require restricting to proper subsets $\mathcal{X}_0 \subsetneq G_0 \mathcal{A}(\mathcal{D})$ [3, Proposition 3.11].

Source conditions define suitable sets $\mathcal{X}_0$ for classical Tikhonov regularisation and related methods based on minimal norm solutions. We investigate the source conditions (transformed source sets) and convergence rates for regularising networks where $(G_\alpha)_{\alpha>0}$ is given by Tikhonov regularisation in Example 5.3.5.

5.3.2 Rates for the regularising networks

Our aim is to prove a convergence rate for $R_{\alpha^*(\delta, y^\delta)}$ assuming a convergence rate for $G_{\alpha^*(\delta, y^\delta)}$. Let $(\Phi_\alpha G_\alpha)_{\alpha>0}$ be a regularising network and α^* a parameter choice function. For any $\delta > 0$ we define the reconstruction algorithms $G^\delta, R^\delta: \mathcal{Y} \rightarrow \mathcal{X}$ and $\Phi^\delta: \mathcal{X} \rightarrow \mathcal{X}$ by

$$\begin{aligned} G^\delta(y) &:= G_{\alpha^*(\delta, y)}(y), \\ \Phi^{\delta, y^\delta}(x) &:= \Phi_{\alpha^*(\delta, y^\delta)}(x), \\ R^\delta(y) &:= \Phi_{\alpha^*(\delta, y)} G_{\alpha^*(\delta, y)}(y), \end{aligned}$$

for $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. We will derive convergence rates for regularising networks under the following assumptions.

Assumption 5.3.3 (Convergence rate conditions). *Let $\mathcal{X}_0 \subseteq G_0\mathcal{A}(\mathcal{D})$ satisfy the following for some $r \in (0, 1]$:*

- (R1) $\mathcal{E}(G_{\alpha^*(\delta, y^\delta)}, \delta, \mathcal{X}_0) = \mathcal{O}(\delta^r)$ as $\delta \rightarrow 0$.
- (R2) $\sup_{x, y^\delta} \{ \|G_{\alpha^*(\delta, y^\delta)}(y^\delta) - G_{\alpha^*(\delta, y^\delta)}\mathcal{A}(z)\| \mid z \in \mathcal{X}_0 \wedge y^\delta \in \overline{B_\delta(\mathcal{A}(z))} \} = \mathcal{O}(\delta^r)$.
- (R3) $\sup_{x, y^\delta} \{ \|\Phi_{\alpha^*(\delta, y^\delta)}(z) - \Phi_0(z)\| \mid z \in \mathcal{X}_0 \wedge y^\delta \in \overline{B_\delta(\mathcal{A}(z))} \} = \mathcal{O}(\delta^r)$.

The first condition (R1) means that $(G_{\alpha^*(\delta, y^\delta)})_{\delta>0}$ converges at rate δ^r . Condition (R2) is a stability estimate for $G_{\alpha^*(\delta, y^\delta)}$. Condition (R3) gives a relation between $\Phi_{\alpha^*(\delta, y^\delta)}$, applied in noisy cases, and Φ_0 , applied in the noiseless case.

Theorem 5.3.4 (Convergence rate for regularising networks). *Let $\mathcal{M}_0 = \Phi_0(\mathcal{X}_0)$. Under Assumption 5.3.3 we have $\mathcal{E}(R_{\alpha^*(\delta, y^\delta)}, \delta, \mathcal{M}_0) = \mathcal{O}(\delta^r)$.*

Proof. Let $x \in \mathcal{M}_0$, $\|\mathcal{A}(x) - y^\delta\| \leq \delta$, and $z \in \mathcal{X}_0$ s.t. $\Phi_0(z) = x$. Then

$$\begin{aligned} \|R_{\alpha^*(\delta, y^\delta)}(y^\delta) - x\| &\leq \|R_{\alpha^*(\delta, y^\delta)}(y^\delta) - R_{\alpha^*(\delta, y^\delta)}\mathcal{A}(x)\| + \|R_{\alpha^*(\delta, y^\delta)}\mathcal{A}(x) - x\| \\ &\leq L\|G_{\alpha^*(\delta, y^\delta)}(y^\delta) - G_{\alpha^*(\delta, y^\delta)}\mathcal{A}(x)\| + \|\Phi_0 G_{\alpha^*(\delta, y^\delta)}\mathcal{A}(z) - \Phi_0(z)\| \\ &\quad + \|R_{\alpha^*(\delta, y^\delta)}\mathcal{A}(x) - \Phi_0 G_{\alpha^*(\delta, y^\delta)}\mathcal{A}(z)\| \\ &\leq L\|G_{\alpha^*(\delta, y^\delta)}(y^\delta) - G_{\alpha^*(\delta, y^\delta)}\mathcal{A}(x)\| + L\|G_{\alpha^*(\delta, y^\delta)}\mathcal{A}(z) - z\| \\ &\quad + \|\Phi_{\alpha^*(\delta, y^\delta)} G_{\alpha^*(\delta, y^\delta)}\mathcal{A}(x) - \Phi_0 G_{\alpha^*(\delta, y^\delta)}\mathcal{A}(x)\| \\ &\leq L\|G_{\alpha^*(\delta, y^\delta)}(y^\delta) - G_{\alpha^*(\delta, y^\delta)}\mathcal{A}(x)\| + L\|G_{\alpha^*(\delta, y^\delta)}\mathcal{A}(z) - z\| \\ &\quad + \|\Phi_{\alpha^*(\delta, y^\delta)} G_{\alpha^*(\delta, y^\delta)}\mathcal{A}(x) - \Phi_{\alpha^*(\delta, y^\delta)} G_0\mathcal{A}(x)\| \\ &\quad + \|\Phi_{\alpha^*(\delta, y^\delta)} G_0\mathcal{A}(x) - \Phi_0 G_0\mathcal{A}(x)\| \\ &\quad + \|\Phi_0 G_0\mathcal{A}(x) - \Phi_0 G_{\alpha^*(\delta, y^\delta)}\mathcal{A}(x)\| \\ &\leq L\|G_{\alpha^*(\delta, y^\delta)}(y^\delta) - G_{\alpha^*(\delta, y^\delta)}\mathcal{A}(x)\| + L\|G_{\alpha^*(\delta, y^\delta)}\mathcal{A}(z) - z\| \end{aligned}$$

$$\begin{aligned}
& + \|\Phi_{\alpha^*(\delta, y^\delta)} G_0 \mathcal{A}(x) - \Phi_0 G_0 \mathcal{A}(x)\| \\
& + 2L \|G_{\alpha^*(\delta, y^\delta)} \mathcal{A}(x) - G_0 \mathcal{A}(x)\|.
\end{aligned}$$

Each of the above terms are $\mathcal{O}(\delta^r)$: the first term due to the stability estimate (R2), the second term due to (R1), and the third term due to (R3). For the fourth term we use that $G_0 \mathcal{A}(x) = G_0 \mathcal{A}(z) = G_0 \mathcal{A} G_0 \mathcal{A}(w) = G_0 \mathcal{A}(w) = z \in \mathcal{X}_0$ for some $w \in \mathcal{D}$. This implies that the fourth term is $\mathcal{O}(\delta^r)$, due to (R1) again. ■

In the following we give an explicit example of a classical regularisation method combined with a sequence of regularising networks, where Assumption 5.3.3 is satisfied and therefore Theorem 5.3.4 can be applied.

Example 5.3.5 (Regularising networks combined with Tikhonov regularisation). *Given a Gâteaux differentiable forward operator $\mathcal{A}$ we consider $(G_\alpha)_{\alpha>0}$ defined by classical Tikhonov regularisation (5.2.1), a data-consistent network $\Phi_0: \mathcal{X} \rightarrow \mathcal{X}$ and a sequence of regularising networks $(\Phi_\alpha: \mathcal{X} \rightarrow \mathcal{X})_{\alpha>0}$ satisfying (R3) for $r = 1/2$.*

Corollary 5.3.6. *If we consider the set $\mathcal{M}_0 := \Phi_0(\mathcal{X}_0)$ where $\mathcal{X}_0$ is the source set of classical Tikhonov regularisation and we assume that the networks $\Phi_{\alpha^*(\delta, y^\delta)}$ converge pointwise to Φ_0 on $\mathcal{X}$ at rate $\mathcal{O}(\delta^{1/2})$ as $\delta \rightarrow 0$, then $\mathcal{E}(R_{\alpha^*(\delta, y^\delta)}, \delta, \mathcal{M}_0) = \mathcal{O}(\delta^{1/2})$.*

Proof. The convergence rate condition (R1) holds according to [3]. Since for $x = \Phi_0(z) \in \mathcal{M}_0$ we have $\mathcal{A}(x) = \mathcal{A}(z)$ (Definition 5.2.2) and because of the stability estimate for Tikhonov regularisation [3] we have

$$\begin{aligned}
& \sup_{x, y^\delta} \{ \|G_{\alpha^*(\delta, y^\delta)}(y^\delta) - G_{\alpha^*(\delta, y^\delta)} \mathcal{A}(x)\| \mid x \in \mathcal{M}_0 \wedge y^\delta \in \overline{B_\delta(\mathcal{A}x)} \} \\
& = \sup_{x, y^\delta} \{ \|G_{\alpha^*(\delta, y^\delta)}(y^\delta) - G_{\alpha^*(\delta, y^\delta)} \mathcal{A}(z)\| \mid z \in \mathcal{X}_0 \wedge y^\delta \in \overline{B_\delta(\mathcal{A}x)} \} = \mathcal{O}(\delta^{1/2}),
\end{aligned} \tag{5.3.2}$$

which shows (R2). Finally (R3) holds by assumption and therefore the conditions of Theorem 5.3.4 are satisfied for $r = 1/2$. ■

This shows one of the main benefits of the concept of regularising networks, namely transforming the set $\mathcal{X}_0$ on which the basic regularisation converges at a certain rate, to a different data-dependent set $\Phi_0(\mathcal{X}_0)$ with possibly less regularity, while preserving the convergence rate.

5.4 Examples and numerical setup

In this section we provide a mathematical description of two general classes of nonlinear inverse problems to which we can apply data-consistent networks, as well as specific examples of these problems that will be used for our simulation experiments. In each case, we define a nonlinear forward mapping $\mathcal{A}$ and derive a right inverse G_0 and data-consistent network Φ_0 . We also describe the network architecture for our neural networks in section 5.4.5.

5.4.1 Projection onto a convex set

For the first class of nonlinear inverse problems, we consider the case where $\mathcal{A} := P_C: \mathcal{D} \rightarrow C$ is a metric projection on a closed convex set $C \subseteq \mathcal{D}$, i.e.

$$y = P_C(x) := \operatorname{argmin}_{\tilde{x} \in C} \{\|\tilde{x} - x\|\} . \quad (5.4.1)$$

The affine normal cone to C at x is defined as

$$N_C(x) := \{\tilde{x} \in \mathcal{D} \mid P_C(\tilde{x}) = x\} .$$

It is easily shown that any mapping that maps $x \in C$ to an element in the normal cone $N_C(x)$ is a right inverse of P_C . In particular, the projection mapping $P_{N_C(x)}: C \rightarrow \mathcal{D}$,

$$P_{N_C(x)}(\hat{x}) := \operatorname{argmin}_{\tilde{x} \in N_C(x)} \{\|\tilde{x} - \hat{x}\|\}$$

defines a right inverse G_0 of P_C , since $\forall x \in C: P_C P_{N_C(x)}(x) = x$. According to Definition 5.2.2, a data-consistent network Φ_0 satisfies that $\forall z \in G_0 \mathcal{A}(\mathcal{D}): \mathcal{A}\Phi_0(z) = \mathcal{A}(z)$. We define

$$\Phi_0(z) := P_{N_C(P_C(z))}(\mathcal{N}(z)), \quad (5.4.2)$$

so this requirement is satisfied. Here $\mathcal{N}: \mathcal{X} \rightarrow \mathcal{X}$ is any Lipschitz continuous trained neural network (c.f. definition 5.2.2). See Figure 5.4.1 for a visual illustration.

Figure 5.4.1: Two-dimensional visualisation of the data-consistent network for the ‘projection on convex set’ problem explained in Section 5.4.1. The blue region indicates the convex set C , the green region (that extends infinitely to the right) indicates the affine normal cone to C at $P_C(z)$. The output of the data-consistent network is indicated by $\Phi_0(z)$; it can be seen that it is obtained by taking the input z , applying a Lipschitz continuous neural network $\mathcal{N}(z)$ and projecting it to the normal cone N_C at $P_C(z)$. This ensures that $P_C(\Phi_0(z)) = P_C(z)$, as required by the definition of a data-consistent network.

5.4.2 Spatially dependent saturation of multivariate Gaussians

As an example of the ‘projection on convex set’ problem explained in Section 5.4.1, we consider the inverse problem of recovering images of multivariate Gaussians which are nonhomogeneously saturated. Formally we define the domain $\mathcal{D} = \mathcal{X} := \ell^2(\Omega)$ and we define the saturation mapping as a projection on a convex set, as described in Section 5.4.1. This means $\mathcal{A}(x) := P_C(x)$, where

$$C := \left\{ x \in \ell^2(\Omega) \mid \forall r \in \Omega: x(r) \leq M(r) \right\},$$

where $M(r) \geq 0$ is the saturation value at location r . The corresponding right inverse for $y \in C$ is defined as $G_0(y) = P_{N_C(y)}(y) = \text{Id}(y)$. The projection defined by (5.4.1) is explicitly given by

$$[P_C(x)](r) := \min \{x(r), M(r)\}.$$

Since $N_C P_C(z) = \{x \mid P_C(x) = P_C(z)\}$, (5.4.2) can be written pointwise as

$$[\Phi_0(z)](r) = \begin{cases} z(r) & \text{for } z(r) < M(r), \\ \max \{[\mathcal{N}(z)](r), M(r)\} & \text{for } z(r) \geq M(r). \end{cases} \quad (5.4.3)$$

We consider the square domain $\Omega := [-1, 1] \times [-1, 1]$. The spatially dependent saturation function is defined

$$M(r) := \begin{cases} 0.6 & \text{if } \|r\| \leq \frac{1}{2}, \\ 0 & \text{if } \|r\| > \frac{1}{2}. \end{cases} \quad (5.4.4)$$

Each image in the training or test set contains one centered multivariate Gaussian with diagonal covariance matrix, having standard deviations (σ_1, σ_2) independently randomly chosen in the interval $[0.24, 0.32]$. All images in the training and test set are scaled to obtain maximum values randomly chosen in the interval $[0.75, 1]$. Opposed to standard networks, one of the benefits of using a data-consistent network is that it is more robust to changes in the data. For this reason, a modified test set is created, where the Gaussians have standard deviations in the interval $[0.12, 0.20]$ with maximum intensities in the interval $[0.6, 0.8]$. For the numerical implementation we consider the discretised domain $\bar{\Omega} := \mathbb{R}^{128 \times 128}$ as discretisation of Ω .

The data-consistent network $\Phi_0(z)$, as described in (5.4.3), is compared with the neural network $\mathcal{N}(z)$ without data-consistency. We compare reconstruction quality for both the regular test set and modified test set. A description of the neural network architecture and training details are provided in Section 5.4.5.

5.4.3 Composition of mappings

As a second inverse problem, we consider the mapping $\mathcal{A}: \mathcal{D} \rightarrow \mathcal{Y}$ that is defined as a composition of two (possibly nonlinear) mappings:

$$\mathcal{A}(x) := \mathcal{A}_2(\mathcal{A}_1(x)), \quad \text{where } \mathcal{A}_1: \mathcal{D} \rightarrow E \text{ and } \mathcal{A}_2: E \rightarrow \mathcal{Y},$$

where E is a Banach space. Furthermore we impose the restriction that the mapping $\mathcal{A}_2$ provides a data-consistent network that can be written as a projection

$$\Phi_0^{(2)}(z) := P_S(\mathcal{N}_2(z)),$$

where $S \subseteq E$. In particular, this is true for the mapping described in Section 5.4.1, where the data-consistent networks makes use of a projection on a normal cone, as shown in (5.4.2). Finally we assume that $S \cap \text{Im}(\mathcal{A}_1) \neq \emptyset$. We define the data-consistent network Φ_0 for the full mapping as

$$\Phi_0(z) = \Phi_0^{(1)} G_0^{(1)} P_{S \cap \text{Im}(\mathcal{A}_1)} \Phi_0^{(2)} G_0^{(2)} \mathcal{A}_2 \mathcal{A}_1(z) \quad (5.4.5)$$

where $\Phi_0^{(i)}$ and $G_0^{(i)}$ are defined as the data-consistent network for $\mathcal{A}_i$ and the right inverse for $\mathcal{A}_i$ respectively. We check the data-consistent property (Definition 5.2.2) by

$$\begin{aligned} \mathcal{A}\Phi_0(z) &= \mathcal{A}_2 \mathcal{A}_1 \Phi_0^{(1)} G_0^{(1)} P_{S \cap \text{Im}(\mathcal{A}_1)} \Phi_0^{(2)} G_0^{(2)} \mathcal{A}_2 \mathcal{A}_1(z) \\ &= \mathcal{A}_2 \mathcal{A}_1 G_0^{(1)} P_{S \cap \text{Im}(\mathcal{A}_1)} \Phi_0^{(2)} G_0^{(2)} \mathcal{A}_2 \mathcal{A}_1(z) \\ &= \mathcal{A}_2 \Phi_0^{(2)} G_0^{(2)} \mathcal{A}_2 \mathcal{A}_1(z) \\ &= \mathcal{A}_2 G_0^{(2)} \mathcal{A}_2 \mathcal{A}_1(z) \\ &= \mathcal{A}_2 \mathcal{A}_1(z) \\ &= \mathcal{A}(z), \end{aligned}$$

where we used in order: the data-consistent property of $\Phi_0^{(1)}$; the definition of a right-inverse $G_0^{(1)}$ in combination with the projection on the range of $\mathcal{A}_1$; the data-consistent property of $\Phi_0^{(2)}$; and the definition of the right-inverse $G_0^{(2)}$.

We note that this is not the only data-consistent network possible for such an inverse problem: one could also design a network that only makes use of either $\Phi_0^{(1)}$ or $\Phi_0^{(2)}$. However, (5.4.5) provides a network that is intuitively clear: an initial solution z is obtained by a classical regularisation method, after which first a better ‘guess’ is made by applying a neural network on E , followed by a neural network that makes a better guess on the reconstruction space $\mathcal{D}$, while keeping the solutions data-consistent throughout.

5.4.4 Saturation of Radon transformed human chest images

As an example of the ‘composition of mapping’ problem explained in Section 5.4.3, we consider the inverse problem of reconstructing images of the human chest, from saturated and highly limited angle Radon measurements. We

consider the composition of two mappings as described in Section 5.4.3, where $\mathcal{A}_1$ is a linear mapping that acts as the discrete Radon transform and $\mathcal{A}_2$ is a nonlinear saturation mapping that saturates the Radon signals at a constant value M . For conciseness, we define our domains and mappings in the discretised setup: $\mathcal{D} := \mathbb{R}^{n_x \times n_x}$, $\mathcal{Y} = \mathbb{R}^{n_a \times \frac{3}{2}n_x}$, where $n_x = 192$ is the number of pixels in each direction of the image and $n_a = 8$ is the number of angles in the Radon transform, uniformly sampled in the interval $[0, \pi]$.

We now define all elements that are needed to obtain the data-consistent network (5.4.5). A matrix representation $\mathcal{A}_1 \in \mathbb{R}^{n_a \cdot \frac{3}{2}n_x \times n_x^2}$ of the Radon transform is obtained as described in [29]. Its right inverse is taken as the pseudo-inverse of $\mathcal{A}_1$, i.e. $G_0^{(1)} := \mathcal{A}_1^\dagger$. Since the mapping is linear, the corresponding data-consistent network is a null-space network [26], i.e. $\Phi_0^{(1)}(z) = z + P_{\ker(\mathcal{A}_1)}\mathcal{N}(z)$, where $P_{\ker(\mathcal{A}_1)} = \text{Id} - \mathcal{A}_1^\dagger \mathcal{A}_1$. The saturation mapping $\mathcal{A}_2 = P_C$, its right inverse $G_0^{(2)}$ and the data-consistent network $\Phi_0^{(2)}$ are chosen as described in Section 5.4.2, this time with constant saturation level $M = 8$. Finally, for this particular choice of $\mathcal{A}_2$, the projection on the intersection of convex sets reads $P_{N_C(P_C(z)) \cap \text{Im}(\mathcal{A}_1)}$, which can be achieved by the ‘projection onto convex sets’ (POCS) algorithm [30]: by alternatingly performing $P_{N_C(P_C(z))}$ and $P_{\text{Im}(\mathcal{A}_1)} = \mathcal{A}_1 \mathcal{A}_1^\dagger$, the resulting iteration converges linearly to a point on the intersection.

Training and test images were obtained from the LoDoPaB-CT data set [31], which on its turn makes use of the LIDC/IDRI data set [32]. In our work, we only make use of the high-quality CT reconstructions in the LoDoPaB-CT data set that we use as ‘ground truth’ for our setup. The images are scaled to 192×192 pixels, after which the mappings $\mathcal{A}_1$ and $\mathcal{A}_2$ are applied to obtain simulated sinograms. After that, pseudo-inverses $G_0^{(2)}$ and $G_0^{(1)}$ are applied to obtain the input for our data-consistent network. For this simulation experiment we also created a modified test set to investigate how the trained networks generalise towards slightly modified data. For conciseness, the procedure to get from the regular test data to the modified test data is not explained in full detail. In short, the test set consists of images in the range of $\mathcal{A}_1^\dagger$ that produce sinograms that have a maximum below or around the saturation level. This means that the saturation mapping $\mathcal{A}_2$ will not have a big effect on the unsaturated sinograms. Images in the modified test set look very similar to the ones in the regular test set, but they often show a small gradient at locations where the regular images show a piecewise constant structure. Some samples from the modified test set are shown in Figure 5.5.2 and Appendix B.

The data-consistent network $\Phi_0(z)$, as described in (5.4.5), is compared with two other networks: the first one applies a single neural network to the pseudo-inverse reconstruction; the second one first applies a neural network in the sinogram domain, then applies the pseudo-inverse of the Radon-transform, followed by a neural network in the image domain. For completeness, we

summarise the three networks below:

- One neural network: $\mathbf{N}_1(z) = \mathcal{N}_1(z)$.
- Two neural networks: $\mathbf{N}_2(z) = \mathcal{N}_1^{(1)} \mathcal{N}_2 \mathcal{A}_1(z)$.
- Data-consistent network: $\Phi_0(z) = \Phi_0^{(1)} G_0^{(1)} P_{N_C(P_C(z)) \cap \text{Im}(\mathcal{A}_1)} \Phi_0^{(2)} G_0^{(2)} \mathcal{A}_2 \mathcal{A}_1(z)$.

We emphasise that the data-consistent networks, in terms of architecture, make use of the same neural networks $\mathcal{N}_1$ and $\mathcal{N}_2$ as the first two networks, i.e. $\Phi_0^{(1)}(z) = z + P_{\ker(\mathcal{A}_1)} \mathcal{N}_1(z)$ and $\Phi_0^{(2)}$ makes use of $\mathcal{N}_2$ as defined in (5.4.3). A more detailed description of the neural network architecture and training details are provided in Section 5.4.5.

Ideally, the data-consistent network is trained ‘end-to-end’, meaning that both $\mathcal{N}_1$ which is used in $\Phi_0^{(1)}$, and $\mathcal{N}_2$ which is used in $\Phi_0^{(2)}$, are trained at the same time. However, the application of the POCS algorithm is computationally intensive, since it requires iterative application of the operators $\mathcal{A}_1$ and $\mathcal{A}_1^\dagger$. For this reason, we choose to first train $\Phi_0^{(2)}$ to output the unsaturated sinogram, then perform the POCS algorithm and finally train $\Phi_0^{(1)}$ to output the reconstructed image.

5.4.5 Neural network architecture and training details

In this chapter, the popular U-Net [1] is implemented as a neural network. By combining max-pooling and upsampling with residual connections, U-Net has shown to be effective in restoring or improving images that possess both small- and large-scale artefacts [2]. By using standard nonlinearities such as rectified linear units (ReLU) and convolutions, the network is Lipschitz continuous, which is a requirement as described in Definition 5.2.2. The Lipschitz constant can be controlled by weight regularisation, such as adding an L^2 -loss on the weights in the loss function. We stress that instead of U-Net, any other Lipschitz continuous neural network can be chosen.

The U-Net was implemented as described in [2], although for each experiment some parameters were chosen slightly different to obtain optimal results. For all experiments, the network has a ‘depth’ of four, meaning four times max-pooling and upsampling. The U-Nets in the image domain perform the regular max-pooling and upsampling in two directions, while the U-Net in the sinogram domain performs these only in one direction, leaving the number of angles constant at 8. This is done because the neighbouring angles in the sinogram show very little resemblance to each other and there are only 8. The ‘width’, or the amount of convolutions at every depth is chosen to be two. As in [2] the number of convolution channels doubles after each max-pooling; the number of channels at the start is stated in Table 5.1, since this was chosen differently for every simulation experiment. In all experiments a residual structure that is also apparent in [2] is used. The U-Net uses 3×3 convolutions

	Exp. 1 ($\mathcal{N}/\Phi_0$): image domain	Exp. 2 ($\mathcal{N}_1/\Phi_0^{(1)}$): image domain	Exp. 2 ($\mathcal{N}_2/\Phi_0^{(2)}$): sinogram domain
#training samples	1024	35584	35584
#validation samples	256	3522	3522
#test samples	1024	3553	3553
depth	4	4	4
width	2	2	2
#channels in top layer	8	16	16
convolution size	3×3	3×3	3×3
nonlinearity	ReLU	ReLU	ReLU
start learning rate	10^{-3}	10^{-3}	10^{-3}
final learning rate	10^{-4}	$2 \cdot 10^{-4}$	$2 \cdot 10^{-4}$
batch size	64	32	32
#epochs	1000	25	25

Table 5.1: U-Net parameter details for all simulation experiments.

with biases and applied a ReLU-activation after each convolution, except the last one. In all experiments, an L^2 -loss function on the difference between output and ground truth is minimised, i.e. $\min_{\theta} \sum_i^N \|\mathcal{N}^{\theta}(z_i) - x_i\|_{\ell^2}^2$ for the U-Net $\mathcal{N}^{\theta}$, or $\min_{\theta} \sum_i^N \|\Phi_0^{\theta}(z_i) - x_i\|_{\ell^2}^2$ for the data-consistent network Φ_0^{θ} , both parametrised by the weights $\theta \in \Theta$. For optimisation, the ADAM optimiser with exponentially decaying learning rate is chosen. The learning rates and batch sizes are stated in Table 5.1. Training was performed in TensorFlow, executed on a single GTX 1080 TI GPU.

5.5 Numerical results

In this Section, the reconstruction quality of data-consistent networks is compared with the U-Nets that are not data-consistent. Besides visual comparison, the quality will be compared by means of peak signal-to-noise ratio (PSNR) and structural similarity (SSIM). For both experiments this will be done for the regular test set as well as the modified test set in order to investigate the generalisation capacity of the networks.

5.5.1 Spatially dependent saturation of multivariate Gaussians

Multivariate Gaussians are saturated with a spatially dependent saturation function (5.4.4), as described in Section 5.4.2. Results for one sample from the regular test set are shown in Figure 5.5.1. Here it can be seen that both U-Net and the data-consistent network provide a very accurate reconstruction. This is

also reflected in the PSNR and SSIM values shown in Table 5.1. The pseudo-inverse reconstruction, which in this case is just the measurement, is not a good one, since a lot of information is lost by applying the saturation mapping.

Figure 5.5.1: Reconstructions of a sample from the regular test set. In the bottom the horizontal central slice is shown. Both U-Net and data-consistent network provide an almost perfect reconstruction.

In Figure 5.5.2 the results for one sample are shown for the modified test set, which contains smaller Gaussians with a slightly lower intensity. Both U-Net and the data-consistent network are not perfectly able to fill in the missing information in the small Gaussians. This can be expected, since Gaussians of this size were not included in the training set. However, the data-consistent network does not deform the Gaussian at the location where it is not saturated, while U-Net does this slightly; for instance around -0.5 in the slice plot.

Figure 5.5.2: Reconstructions of a sample from the modified test set. In the bottom the horizontal central slice is shown. Data-consistency makes sure that intensity is only changed above the saturation level.

This behaviour is also reflected in the PSNR and SSIM values in Table 5.1. Interestingly, the pseudo-inverse behaves very well if we just look at the values in the table, because the saturation mapping did not destroy a lot of the information in the Gaussian. Visual results of three more samples in the modified test set are shown in Appendix A. It can be seen that U-Net tends to widen the Gaussians, since it was trained on Gaussians in the training set that were wider. Although the modified test set shows a very specific modification, it illustrates that a data-consistent network is beneficial over using an arbitrary neural network: by making use of the information that we have from the mapping $\mathcal{A}$, we obtain generalisation capacity.

	PSNR		
	Pseudo-inverse	U-Net	Data-consistent
Regular set	24.2 ± 2.2	60.6 ± 2.1	66.7 ± 1.6
Modified set	48.0 ± 7.8	36.9 ± 2.9	48.0 ± 4.4
	SSIM		
	Pseudo-inverse	U-Net	Data-consistent
Regular set	0.56 ± 0.08	1.00 ± 0.00	1.00 ± 0.00
Modified set	0.99 ± 0.01	0.92 ± 0.03	0.97 ± 0.01

Table 5.1: Comparison of PSNR and SSIM for all reconstruction methods on the saturated Gaussian data set.

5.5.2 Saturation of Radon transformed human chest images

For one selected sample in the regular test set and one in the modified test set, all reconstructions are shown in Figures 5.5.1 and 5.5.2. These specific samples were selected because their PSNR values for the U-Nets and the data-consistent network show a similar relation to each other as the average PSNR values of the whole test set (c.f. Table 5.1).

In the top of Figure 5.5.1, the inputs of the right inverse $G_0^{(1)}$ are shown. This corresponds to the saturated sinograms in case no or only one neural network is trained, and this corresponds to the output of the neural network in the sinogram domain in case two neural networks are trained. The sinogram signals are plotted in a different colour for each angle. The data-consistent network does not change the values of the sinogram that are below the saturation level ($M = 8$). Interestingly, the U-Net in the sinogram domain has learned not to do this as well to a large extent: some values just below the saturation level are changed (for instance the purple line around -0.5), but values much smaller

Figure 5.5.1: Reconstructions of a typical sample from the regular test set. Top: reconstructed sinograms with all 8 angles in different colours. Bottom: reconstructed images (grayscale from 0 to 1).

are not changed at all. In the bottom of Figure 5.5.1, all reconstructions for this sample are shown. It can be seen that although PSNR values are similar, there is a clear visual difference: the data-consistent network shows more streak artefacts that are typical for limited-angle Radon reconstructions, while the standard U-Nets provide over-smoothed reconstructions. In other words, data-consistency prevents the neural network to smooth out all image details, but the image details are ‘smeared out’ through the streak artefacts, due to the limited information of 8 angles. However, since there are no image details that are visible in one of the reconstructions and not in the others, it can be concluded that all networks have a similar reconstruction quality. This is also reflected in Table 5.1. Moreover, since only 8 angles were used in the Radon transform, only larger structures can be reconstructed. Some extreme samples, for which the PSNR value of the data-consistent solution is high, similar or low compared to the U-Net solutions, are shown in Appendix B.

In the top of Figure 5.5.2, again all sinogram reconstructions are plotted, now for the modified test set. As the modified sinograms contain many values just below or around the saturation level ($M = 8$), this is challenging for the regular U-Net. Indeed it can be seen that the data-consistent network does not change the saturated sinogram much, while the U-Net on the sinogram increases the purple line well beyond the saturation level. In the bottom of Figure 5.5.2, it can be seen that the ground truth possesses some ‘smooth’ regions, especially in the background on the left and right of the two dark inclusions. While both U-Nets create piecewise constant reconstructions that completely remove this gradient, the data-consistent network keeps these smooth regions: it generalises better to test images that are not found in the training set by making use of

Figure 5.5.2: Reconstructions of a typical sample from the modified test set. Top: reconstructed sinograms with all 8 angles in different colours. Bottom: reconstructed images (grayscale from 0 to 1).

the information in the operator. More extreme samples, where the PSNR value of the data-consistent solution is large, similar or small compared to the U-Net solutions, are shown in Appendix B. The same effect on smooth regions can be seen in these visualisations. In Table 5.1, it can be seen that for all networks the quality drops when the regular test set is replaced by the modified test set; however, this drop is only very small for the data-consistent network and is much bigger for the regular U-Nets. Note that the pseudo-inverse gives an increased PSNR for the modified test set, because the modified images were constructed to lie in the range of $\mathcal{A}$.

Finally in Table 5.1 we check the data-fidelity of the solutions from all networks by computing $\|\mathcal{A}(\tilde{x}) - \mathcal{A}(x)\|_{\ell^2}$, where $\tilde{x}$ is the solution of the respective reconstruction method and x is the ground truth. Ideally, the data-fidelity should be zero for the pseudo-inverse and the data-consistent network. It can be seen that indeed the data-fidelity is much smaller for these solutions than for the U-Nets, although not completely zero. This is most likely due to numerical issues (especially for the pseudo-inverse of the regular set) and due to the fact that we saved the images in a 16-bit format between consecutive steps of the data-consistent network to not overload working memory.

	PSNR			
	Pseudo-inverse	One U-Net	Two U-Nets	Data-consistent
Regular set	23.1 ± 2.3	30.5 ± 1.5	31.0 ± 1.5	30.1 ± 1.9
Modified set	29.1 ± 1.6	27.5 ± 1.7	28.3 ± 1.4	29.9 ± 1.2
	SSIM			
	Pseudo-inverse	One U-Net	Two U-Nets	Data-consistent
Regular set	0.50 ± 0.07	0.82 ± 0.04	0.83 ± 0.04	0.74 ± 0.07
Modified set	0.71 ± 0.07	0.74 ± 0.05	0.73 ± 0.05	0.75 ± 0.05
	Data-fidelity			
	Pseudo-inverse	One U-Net	Two U-Nets	Data-consistent
Regular set	6.1 ± 3.3	4.8 ± 1.5	3.9 ± 1.1	0.9 ± 0.4
Modified set	0.4 ± 0.2	11.9 ± 5.4	8.5 ± 2.8	0.6 ± 0.2

Table 5.1: Comparison of PSNR, SSIM and data-fidelity $\|\mathcal{A}(\tilde{x}) - \mathcal{A}(x)\|_{\ell^2}$ for the human chest data set. Here $\tilde{x}$ is the reconstruction and x is the ground truth.

5.6 Conclusion

In this chapter we introduced data-consistent networks for nonlinear inverse problems. We presented a convergent regularisation method by combining neural networks that converge to a data-consistent network with classical regularisation methods. With the proposed data-driven regularisation methods we are able to preserve convergence rates of classical methods over a transformed source set, which is adapted to some data set. This yields improved reconstructions for elements close to the training set, but at the same time data-consistent networks make use of the information from the forward mapping $\mathcal{A}$, which provides increased generalisation capacity. This is particularly useful when the physics process is understood, but exact knowledge on real measurements is not available or when it is not possible to create a training set that is similar to the real measurements. We showed that on a test set similar to the training set, our approach shows reconstruction results comparable to a classical post-processing network, whereas for instances not represented in the training set, the loss of performance is much less present. This demonstrates the generalisation ability of our approach.

References

- [1] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, 2015.
- [2] 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, 2017.
- [3] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, vol. 375. Springer Science & Business Media, 1996.
- [4] V. A. Morozov, *Methods for solving incorrectly posed problems*. New York, Berlin, Heidelberg: Springer Verlag, 1984.
- [5] A. N. Tikhonov and V. Y. Arsenin, *Solutions of ill-Posed problems*. Washington, D.C.: John Wiley & Sons, 1977.
- [6] A. B. Bakushinsky and M. Y. Kokurin, *Iterative methods for approximate solution of inverse problems*, vol. 577 of *Mathematics and Its Applications*. Dordrecht: Springer, 2004.
- [7] B. Kaltenbacher, A. Neubauer, and O. Scherzer, *Iterative regularization methods for nonlinear ill-posed problems*, vol. 6 of *Radon Series on Computational and Applied Mathematics*. Walter de Gruyter GmbH & Co. KG, Berlin, 2008.
- [8] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen, *Variational methods in imaging*. Applied Mathematical Sciences, Springer, 2009.
- [9] J. Sun, H. Li, Z. Xu, *et al.*, "Deep ADMM-Net for compressive sensing MRI," in *Advances in neural information processing systems*, pp. 10–18, 2016.
- [10] J. Adler and O. Öktem, "Solving ill-posed inverse problems using iterative deep neural networks," *Inverse Problems*, vol. 33, no. 12, p. 124007, 2017.
- [11] J. Rick Chang, C.-L. Li, B. Póczos, B. Vijaya, K., and A. C. Sankaranarayanan, "One network to solve them all - solving linear inverse problems using deep projection models," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 5888–5897, 2017.
- [12] E. Kobler, T. Klatzer, K. Hammernik, and T. Pock, "Variational networks: connecting variational methods and deep learning," in *German Conference on Pattern Recognition*, pp. 281–293, Springer, 2017.
- [13] A. Kofler, M. Haltmeier, C. Kolbitsch, M. Kachelrieß, and M. Dewey, "A U-Nets cascade for sparse view computed tomography," in *International Workshop on Machine Learning for Medical Image Reconstruction*, pp. 91–99, Springer, 2018.
- [14] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, "Image reconstruction by domain-transform manifold learning," *Nature*, vol. 555, no. 7697, pp. 487–492, 2018.
- [15] D. Gilton, G. Ongie, and R. Willett, "Neumann networks for inverse problems in imaging," *arXiv:1901.03707*, 2019.
- [16] M. V. de Hoop, M. Lassas, and C. A. Wong, "Deep learning architectures for nonlinear operator functions and nonlinear inverse problems," *arXiv:1912.11090*, 2019.

- [17] S. Lunz, O. Öktem, and C.-B. Schönlieb, "Adversarial regularizers in inverse problems," in *Advances in Neural Information Processing Systems*, pp. 8507–8516, 2018.
- [18] H. Li, J. Schwab, S. Antholzer, and M. Haltmeier, "NETT: solving inverse problems with deep neural networks," *Inverse Problems*, 2020.
- [19] S. Mukherjee, S. Dittmer, Z. Shumaylov, S. Lunz, O. Öktem, and C.-B. Schönlieb, "Learned convex regularizers for inverse problems," *arXiv:2008.02839*, 2020.
- [20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.
- [21] Y. Rivenson, H. Göröcs, Z. and Günaydin, Y. Zhang, H. Wang, and A. Ozcan, "Deep learning microscopy," *Optica*, vol. 4, no. 11, pp. 1437–1443, 2017.
- [22] D. Lee, J. Yoo, and J. C. Ye, "Deep residual learning for compressed sensing MRI," in *IEEE 14th Int. Symposium on Biomedical Imaging (ISBI 2017)*, pp. 15–18, IEEE, 2017.
- [23] S. J. Hamilton and A. Hauptmann, "Deep D-bar: real time electrical impedance tomography imaging with deep neural networks," *IEEE transactions on medical imaging*, pp. 2367–2377, 2018.
- [24] S. Antholzer, M. Haltmeier, and J. Schwab, "Deep learning for photoacoustic tomography from sparse data," *Inverse problems in science and engineering*, vol. 27, no. 7, pp. 987–1005, 2019.
- [25] J. Schwab, S. Antholzer, and M. Haltmeier, "Big in japan: regularizing networks for solving inverse problems," *Journal of Mathematical Imaging and Vision*, pp. 1–11, 2018.
- [26] J. Schwab, S. Antholzer, and M. Haltmeier, "Deep null space learning for inverse problems: convergence analysis and rates," *Inverse Problems*, vol. 35, no. 2, p. 025008, 2019.
- [27] M. Mardani, E. Gong, J. Y. Cheng, S. Vasanawala, G. Zaharchuk, M. Alley, N. Thakur, W. Han, S. and Dally, J. M. Pauly, and L. Xing, "Deep generative adversarial networks for compressed sensing automates MRI," *arXiv:1706.00051*, 2017.
- [28] T. A. Bubba, G. Kutyniok, M. Lassas, M. März, W. Samek, S. Siltanen, and V. Srinivasan, "Learning the invisible: A hybrid deep learning-shearlet framework for limited angle computed tomography," *Inverse Problems*, vol. 35, no. 6, p. 064002, 2019.
- [29] R. M. Lewitt, "Multidimensional digital image representations using generalized Kaiser-Bessel window functions," *Journal of the Optical Society of America*, vol. 7, no. 10, pp. 1834–1846, 1990.
- [30] H. H. Bauschke and J. M. Borwein, "On projection algorithms for solving convex feasibility problems," *SIAM review*, vol. 38, no. 3, pp. 367–426, 1996.
- [31] J. Leuschner, M. Schmidt, D. O. Baguer, and P. Maaß, "The LoDoPaB-CT dataset: a benchmark dataset for low-dose CT reconstruction methods," *arXiv:1910.01113*, 2019.
- [32] S. G. Armato III, G. McLennan, L. Bidaut, M. F. McNitt-Gray, C. R. Meyer, A. P. Reeves, B. Zhao, D. R. Aberle, C. I. Henschke, E. A. Hoffman, *et al.*, "The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans," *Medical physics*, vol. 38, no. 2, pp. 915–931, 2011.

Appendices

A Additional results for saturated Gaussians

In this appendix, three more samples from the modified test set of the saturated Gaussians are shown. It can be seen in Figures A.1 and A.2 that U-Net tends to slightly widen the Gaussian in some cases, since this was necessary for the wider Gaussians in the training set. Moreover it can be seen in Figures A.2 and A.3 that both U-Net and the data-consistent network sometimes fail to restore the top of the Gaussian adequately: the training on wider Gaussians is not directly generalised for smaller Gaussians.

Figure A.1

Figure A.2

Figure A.3

B Additional results for saturated Radon transform

In this appendix, three additional samples from the regular and the modified test set of human chest images are shown. The samples have been selected based on their PSNR values: we show the samples in the test set for which the data-consistent network yields the largest relative PSNR value (Figures B.1 and B.4), a similar PSNR value (Figures B.2 and B.5), and the smallest relative PSNR value (Figures B.3 and B.6) when compared to the U-Nets.

Regular test set

Below the additional results for the regular test set are shown.

Figure B.1: Sample for which the data-consistent PSNR value is relatively large compared to the U-Net PSNR values (grayscale from 0 to 1).

Figure B.2: Sample for which the data-consistent PSNR value is approximately the same as the U-Net PSNR values (grayscale from 0 to 1).

Figure B.3: Sample for which the data-consistent PSNR value is relatively small compared to the U-Net PSNR values (grayscale from 0 to 1).

Modified test set

The modified test set contains more nonconstant regions than the regular test set. The data-consistent network is better able to deal with these modifications in the images than the other networks. This can be seen particularly in Figure B.4: The U-Nets create a piecewise constant dark structure in the middle, while the data-consistent network keeps it smoother.

Figure B.4: Sample for which the data-consistent PSNR value is relatively large compared to the U-Net PSNR values (grayscale from 0 to 1).

Figure B.5: Sample for which the data-consistent PSNR value is approximately the same as the U-Net PSNR values (grayscale from 0 to 1).

Figure B.6: Sample for which the data-consistent PSNR value is relatively small compared to the U-Net PSNR values (grayscale from 0 to 1).

Chapter 6

Learned SVD: solving inverse problems via hybrid autoencoding

Yoeri E. Boink, Christoph Brune

submitted to Inverse Problems and Imaging (2020)
preprint available: arXiv 1912.10840

Abstract

Our world is full of physics-driven data where effective mappings between data manifolds are desired. There is an increasing demand for understanding combined model-based and data-driven methods. We propose a nonlinear, learned singular value decomposition (L-SVD), which combines autoencoders that simultaneously learn and connect latent codes for desired signals and given measurements. We provide a convergence analysis for a specifically structured L-SVD that acts as a regularisation method. In a more general setting, we investigate the topic of model reduction via data dimensionality reduction to obtain a regularised inversion. We present a promising direction for solving inverse problems in cases where the underlying physics are not fully understood or have very complex behaviour. We show that the building blocks of learned inversion maps can be obtained automatically, with improved performance upon classical methods and better interpretability than black-box methods.

6.1 Introduction

We are living in a world full of physics-driven data with an increasing demand for combining *model-based* and *data-driven* approaches in areas of science, industry and society. In many cases it is essential to reliably recover hidden multi-dimensional model parameters (signals) $x \in \mathcal{X}$ from noisy indirect observations (measurements) $y^\delta \in \mathcal{Y}$, e.g. in imaging or sensing technology in medicine, engineering, astronomy or geophysics. These *inverse problems*, $y^\delta = \mathcal{A}(x) + \eta^\delta$, are often ill-posed, suffering from nonuniqueness and instability in direct inversion. Classical model-based research on inverse problems has focused on variational regularisation methods to guarantee existence and stable approximation of solutions under uncertainty like noise η^δ in the measurements [1, 2]. For linear inverse problems the singular value decomposition (SVD) [3] is a classical tool to directly construct a regularised inverse, e.g. in the sense of Tikhonov regularisation [1]. Recent research in inverse problems has focused on combining deep learning with model-based approaches based on knowledge of the underlying physics [4]. Precise knowledge is often not available; for now we rely mainly on empirical evidence that such approaches can still be applied when one makes use of inexact operators that approximate the exact physical process [5]. The main limitation of such methods are that they require an iterative application of expensive, possibly nonlinear mappings. Moreover, they are hard to interpret due to the lack of connection between data structure and structure of the mapping.

Figure 6.1.1: L-SVD learns a nonlinear inversion mapping by connecting two data manifolds.

We propose the ‘learned singular value decomposition’ (L-SVD): a direct method that provides the inversion procedure with an explainable connection between measurements and signals. The method does not rely on an iterative application of expensive mappings; it does not need any information on the mapping at all. L-SVD makes use of two connected autoencoders: the first one encodes measurement y^δ to latent code z_y , while the second one encodes signal x to latent code z_x in a nonlinear way; both latent codes are connected with a linear ‘scaling’ layer. The training of all parameters is done simultaneously, which

enforces the latent codes to preserve as much information on the measurements and signals as possible, while making sure that the codes have very similar structure. After training, a reconstruction is obtained by consecutive application of encoding, scaling and decoding (see Figure 6.1.1).

In case a forward mapping is available, it can be used to extract an effective encoding and decoding, e.g. via the SVD. In such cases, the L-SVD can be shaped to a data-driven Tikhonov regularisation method, for which we prove convergence with respect to noise. In case a forward mapping is not available, L-SVD allows to learn the nonlinear inversion dynamics via two autoencoders and a linear scaling layer. This has the advantage that advances in interpretable autoencoding have a direct effect on the L-SVD method, making it easier to analyse than other fully learned inversion methods. Assuming that the autoencoders can be trained with high accuracy, finding the connection between both codes is a much lower dimensional problem than finding a nonlinear map between the measurements and signals directly. Moreover, in a semi-supervised scenario where not all training data is available in supervised pairs, the autoencoders still allow to learn from all samples, which is not possible in most supervised neural networks.

Current research in inverse problems is focused at developing theory for combining data-driven deep learning with model-based approaches [4]. Recently developed methods with theoretical guarantees such as convergence and stability include those in which a regularisation term is explicitly learned [6, 7] and those where an initial imperfect reconstruction is post-processed by a neural network [8], also shown in Chapter 5. These methods are two-step procedures and require knowledge of the forward mapping, while L-SVD provides a one-step procedure without this knowledge. A series of works about the optimal regularised inverse matrix (ORIM) investigated the problem of finding a linear inverse matrix for a noisy inverse problem, both for the case where the linear forward matrix is known [9, 10] and for the case where it is not known [9], i.e. a fully learned scenario. Similar to the L-SVD, the idea of data-driven approximation of nonlinear mappings via model reduction was proposed in [11]. In that paper, the interest is in approximating a forward mapping, while we are interested in solving an inverse problem. The idea of connecting two autoencoders was exploited in [12] and [13], where the authors solved a superresolution and a deconvolution problem with a patch-based method. In this chapter, we consider a more general method that does not assume identical domains for measurement and signal. An extensive literature review of similar methods and the embedding of L-SVD in its research context is given in Section 6.5.

6.1.1 Contribution

This chapter proposes the learned singular value decomposition (L-SVD), a data-driven method that nonlinearly encodes (compresses) data in two vector spaces

and connects them in an easy-to-understand way. The contributions of our method can be seen as the extension of existing methods in the following way:

1. **Data-driven solution of inverse problems:** L-SVD is a nonlinear generalisation of Tikhonov regularisation in Bayesian inverse problems [14, 15, 16] and piecewise linear estimates [17], which are linear data-driven variants of classical SVD approaches. We show that L-SVD can be shaped to a data-driven Tikhonov regularisation for which we provide a convergence analysis. In general, L-SVD requires no a-priori information on the forward mapping to achieve good reconstruction quality.
2. **Improved generalisation via nonlinear hybrid encoding:** Autoencoders show that nonlinear encoding provides better encodings than linear encoding [18]. L-SVD shows that this is also the case when encodings are used to solve inverse problems. An autoencoder can act as a regulariser when attached to a supervised neural network that is trained for a supervised task, e.g. classification [19, 20]. L-SVD makes use of two autoencoders for the task of solving an inverse problem, which gives improved generalisation performance. Moreover, it enables high-quality reconstruction in a semi-supervised setup.

6.1.2 Overview of this chapter

Throughout this chapter, we make use of the above two perspectives to show the advantage of using the L-SVD method for addressing the previously mentioned limitations. In Section 6.2, a brief overview of the classical SVD and inversion methods is given, which serves as motivation for the L-SVD method. Next, a precise definition of L-SVD is provided in Section 6.3, accompanied with various architecture choices that exploit its potential. One of these choices results in a data-driven Tikhonov regularisation, while a second one results in a fully learned L-SVD method. In Section 6.4, we analyse the L-SVD method by showing its connection with Bayesian inverse problems, by showing a convergence analysis of the data-driven Tikhonov regularisation and providing a stability and error estimate for the fully learned L-SVD method. After that, in Section 6.5, the connection of this chapter with several fields of research are discussed. In Section 6.6, we explain simulation experiments that show the transition from non-learned to learned, linear to nonlinear, and single to hybrid encoding. Results are provided in Section 6.7, where we visualise these transitions by looking at the latent space, the decoded representations of the latent space and the dependency of L-SVD on noise. The section is completed with a comparison between L-SVD and state-of-the art reconstruction methods applied on a biomedical CT data set. In Section 6.8, we conclude with some remarks and outlook for future work.

6.2 Motivation: SVD and inversion methods

The motivation of our L-SVD method can be found in the application of classical SVD and its variants in inversion methods. In this chapter we consider the finite dimensional version of the equation introduced in Section 6.1. That is, we make use of the ‘first discretise, then optimise’ approach. We define the inverse problem as

$$y^\delta = \mathcal{A}(x) + \eta^\delta, \quad (6.2.1)$$

where we wish to reconstruct the signals $x \in \mathcal{X} \subseteq \mathbb{R}^m$ from measurements $y^\delta \in \mathcal{Y} \subseteq \mathbb{R}^n$ corrupted by additive noise $\eta^\delta \sim N(0, \delta \text{Id})$. Here $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces. The mapping $\mathcal{A} : \mathcal{X} \mapsto \mathcal{Y}$ is in general a nonlinear one. For this section however, we assume a linear operator that we call A . Any A can be written in its singular value decomposition: $A = USV^*$, where $U \in \mathbb{R}^{n \times n}$ and $V \in \mathbb{R}^{m \times m}$ are unitary matrices and $S \in \mathbb{R}^{n \times m}$ is a diagonal matrix with nonnegative real numbers s_i (singular values) on the diagonal. We now summarise well-known inversion methods that can be written as the application of an SVD [3].

6.2.1 Maximum likelihood estimator (MLE)

The MLE is defined via the maximisation over x given measurements y^δ [16]. Its solution x_{MLE} is obtained by applying the Moore-Penrose inverse $(A^*A)^{-1}A^*$ to the measurements y^δ .

$$\begin{aligned} x_{\text{MLE}} &:= \underset{x}{\operatorname{argmax}} p(x|y^\delta) = \underset{x}{\operatorname{argmin}} \|Ax - y^\delta\|_{\ell^2}^2 \\ &= (A^*A)^{-1}A^*y^\delta = VS^{-1}U^*y^\delta. \end{aligned} \quad (6.2.2)$$

6.2.2 Tikhonov regularisation

Tikhonov regularisation is a method which puts a uniform variance prior on the desired solution x . It solves an α -weighted minimisation problem that can be solved directly via its regularised Moore-Penrose inverse:

$$\begin{aligned} x_\alpha &:= \underset{x}{\operatorname{argmin}} \|Ax - y^\delta\|_{\ell^2}^2 + \alpha \|x\|_{\ell^2}^2, \\ &= (A^*A + \alpha \text{Id})^{-1}A^*y^\delta = V \underbrace{(S^2 + \alpha \text{Id})^{-1}S}_{S_\alpha^{-1}} U^*y^\delta. \end{aligned} \quad (6.2.3)$$

The diagonal elements of S_α^{-1} are defined as $s_i / (s_i^2 + \alpha)$, which means that for smaller scales s_i , the new inverse scale goes to zero as α gets larger. The optimal α depends on the type of noise; usually α increases with noise level δ .

For a specific model choice (see Section 6.3.1), L-SVD becomes a data-driven Tikhonov regularisation method, for which we provide a convergence analysis in Section 6.4.2.

6.2.3 Truncated SVD

The best Frobenius approximation of A with rank r is given [21] by the truncated SVD (T-SVD):

$$A_r := U_r S_r V_r^* = \underset{\tilde{A}}{\operatorname{argmin}} \|A - \tilde{A}\|_{\text{Fro}} \text{ s.t. } \operatorname{rank}(\tilde{A}) = r. \quad (6.2.4)$$

Here we made use of the ‘thin’ representation, where $U_r \in \mathbb{R}^{n \times r}$ and $V_r \in \mathbb{R}^{m \times r}$ consist of the top r rows of U and V respectively. $S_r \in \mathbb{R}^{r \times r}$ is a diagonal square matrix that consists of the largest r singular values of A . With the thin representation, we lose one desirable property, namely that of unitary matrices: while $U_r^* U_r = \operatorname{Id}_r = V_r^* V_r$ still holds, generally $U_r U_r^* \neq \operatorname{Id} \neq V_r V_r^*$.

T-SVD can be applied in an inversion method instead of the standard SVD for noisy measurements y^δ : when s_i becomes small for i large, noise is amplified by $1/s_i$ in (6.2.2). This problem is mitigated by solving $x_{\text{trunc}} := V_r S_r^{-1} U_r^* y^\delta$ instead. It has the additional benefit that the thin decomposition is smaller than the full SVD, requiring less memory and computation time.

6.3 The learned singular value decomposition

In this section, we provide the general L-SVD method for solving inverse problems. It aims to solve the inverse problem as defined in (6.2.1), where the forward mapping $\mathcal{A}$ may be nonlinear. L-SVD can be seen as a nonlinear learned variant of the inversion methods in Section 6.2, where U^* is replaced by a nonlinear encoder and V by a nonlinear decoder.

6.3.1 Model statement

The L-SVD model (Figure 6.3.1) is a trained neural network that consists of a measurement autoencoder (green), a signal autoencoder (blue) and a reconstruction component (red). Reconstruction $\hat{x}$ from measurement y^δ is obtained via the latent representations $z_x \in \mathcal{Z}^x \subseteq \mathbb{R}^k$ and $z_y \in \mathcal{Z}^y \subseteq \mathbb{R}^k$, which are part of the autoencoders. The latent space $\mathbb{R}^k$ is a low-dimensional space, i.e. $k \leq \min\{m, n\}$.

Definition 6.3.1. We define the nonlinear functions

$$\varphi_{\text{enc}}^y : \mathcal{Y} \mapsto \mathcal{Z}^y, \quad \varphi_{\text{dec}}^y : \mathcal{Z}^y \mapsto \mathcal{Y}, \quad \varphi_{\text{enc}}^x : \mathcal{X} \mapsto \mathcal{Z}^x, \quad \varphi_{\text{dec}}^x : \mathcal{Z}^x \mapsto \mathcal{X}, \quad \Sigma : \mathcal{Z}^y \mapsto \mathcal{Z}^x.$$

and the variables

$$\begin{aligned} z_y &:= \varphi_{\text{enc}}^y(y^\delta), & z_x^{\text{AE}} &:= \varphi_{\text{enc}}^x(x), & z_x^\Sigma &:= \Sigma(z_y), \\ \hat{y}^{\text{AE}} &:= \varphi_{\text{dec}}^y(z_y), & \hat{x}^{\text{AE}} &:= \varphi_{\text{dec}}^x(z_x^{\text{AE}}), & \hat{x}^\Sigma &:= \varphi_{\text{dec}}^x(z_x^\Sigma). \end{aligned}$$

Figure 6.3.1: Schematic of the L-SVD method. **Green**: autoencoder for measurement y^δ . **Blue**: autoencoder for signal x . **Red**: reconstruction procedure. The standard network does not use the gray connections for training. Note that ϕ_{dec}^x is used multiple times, but has shared weights.

The L-SVD model is obtained by minimising a neural network loss function that consists of three parts:

$$\min_{\theta} \left\{ \sum_{i=1}^{\# \text{train}} \underbrace{\mathcal{L}_1(\hat{x}_{(i)}^\Sigma, x_{(i)})}_{\text{reconstruction}} + \alpha_y \underbrace{\mathcal{L}_2(\hat{y}_{(i)}^{\text{AE}}, y_{(i)})}_{\text{autoencoder}} + \alpha_x \underbrace{\mathcal{L}_3(\hat{x}_{(i)}^{\text{AE}}, x_{(i)})}_{\text{autoencoder}} \right\}, \quad (6.3.1)$$

where we minimise over all trainable parameters θ and over all samples (i) in the training set. The distance functions $\mathcal{L}_j(\cdot, \cdot)$ can be any metric; often used in neural networks are ℓ^2 , ℓ^1 and W^2 (Wasserstein) metrics. The L-SVD model encodes measurements y^δ into a representation that contains sufficient information to approximately reconstruct the clean data y , while being able to map to an encoded representation that can approximately be decoded to the desired signal x . Since the output of the data autoencoder is the clean data y , instead of the corrupted measurements y^δ , it can be seen as a denoising autoencoder (DAE) [22]. This means that noise will not necessarily be represented in the latent variable z_y .

Opportunities by various model choices

Below some specific model choices and variations on the standard model are discussed, which establish certain capabilities of the L-SVD model:

- *Data-driven Tikhonov regularisation*: With the SVD, a linear encoder and decoder can be derived from the operator A (see Section 6.2). If this is combined with the nonlinear scaling function $\Sigma = (S^2 + \alpha \mathcal{N}(y^\delta))^{-1} S$, with $0 < \mathcal{N}(y^\delta) < \infty$ a nonlinear function, a data-driven Tikhonov functional is obtained. A convergence analysis is given in Section 6.4.2.
- *Linearly connected nonlinear representations*: Results from autoencoding [18] motivate the search for a nonlinear encoding and decoding such that a

nonlinear representation of signals and measurements is obtained. The expressiveness of the nonlinear encoder and decoder allows us to restrict the scaling layer to be a square matrix $\Sigma \in \mathbb{R}^{k \times k}$, either full or diagonal. Generic stability and reconstruction estimates are given in Section 6.4.3. They depend on the reconstruction quality of the autoencoders, which can be freely chosen depending on the application at hand. This includes regularised autoencoders, such as sparse [23] or contractive autoencoders [24] of a fully connected or convolutional type.

- *Noise-aware Σ* : The autoencoder on the measurement side can be chosen to be a regular autoencoder instead of a denoising autoencoder, with the effect that noise is represented in the encoded version z_y . This means that the latent dimension should be large enough, since unstructured noise can not be compressed. Moreover, this means that Σ should be able to remove (part of) the noise, since the latent variable z_x is noise-free.
- *Structured latent space*: No specific structures of the latent spaces $\mathcal{Z}^x$ and $\mathcal{Z}^y$ are imposed. If control on these spaces is desired, one could sample from a desired set in the latent space $\tilde{z}_x \in E \subset \mathcal{Z}^x$ and add one of the following losses to (6.3.1):

$$\alpha_{z_x^{\text{AE}}} \mathcal{L}_4(\tilde{z}_x, z_x^{\text{AE}}) \quad \text{or} \quad \alpha_{z_x^{\text{AE}}} \mathcal{L}_5(\tilde{z}_x, z_x^{\text{AE}}).$$

This means the sampled latent code is decoded to a signal x (gray in Figure 6.3.1), after which it takes either the blue autoencoder path or the red reconstruction path, without the final decoder step φ_{dec}^x . Although not guaranteed, it is likely that due to this additional loss, the encoder φ_{enc}^x will map all samples $x_{(i)}$ in the training set to this subset E . If this is the case, it means that we have control over the latent space $\mathcal{Z}^x$. Moreover it turns out that having a bound on $\mathcal{L}_5(\tilde{z}_x, z_x^{\text{AE}})$ enables us to compute a uniform error bound for the reconstruction procedure (see Section 6.4.3).

6.4 Analysis

In this section, we first consider the case where the encoder and decoder are derived from the SVD of A . In Section 6.4.1, we show that training this L-SVD model with a linear scaling coincides with learning the covariance matrix of a prior in Bayesian inverse problems. In Section 6.4.2, we provide a convergence analysis with respect to noise in case a nonlinear scaling is used. Finally, in Section 6.4.3, we provide a stability and error estimate for the L-SVD model with nonlinear encoding and decoding.

6.4.1 Connection with Bayesian inverse problems

Here we provide an explicit connection between a linear L-SVD model and the solution of a Bayesian inverse problem with Gaussian noise, Gaussian prior

and known forward operator A . For an introduction to statistical and Bayesian inverse problems, we refer to [14, 15, 16].

Learning the prior covariance matrix

Proposition 6.4.1. *Let $x \in \mathcal{X} \subset \mathbb{R}^m$, $\tilde{y} \in \mathcal{Y} \subset \mathbb{R}^n$ and $\eta \sim N(0, B)$, where $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces. Consider the inverse problem*

$$\tilde{y} = Ax + \eta,$$

where $A \in \mathbb{R}^{n \times m}$ has full row-rank, i.e. $\text{rank}(A) = n \leq m$, with thin SVD decomposition $A = US_n V_n^*$. Moreover, let $\mu_0 \sim N(0, C_0)$ be a Gaussian prior measure on x . We define $\tilde{B} := U^* B U$ and restrict the covariance matrix C_0 to be of rank n that can be written as $C_0 = V_n C_{V_n} V_n^*$, where C_{V_n} is positive definite.

Then the maximum a posteriori (MAP) estimate $x_{\text{MAP}} := \arg\max_x p(y|x)p(x)$ can be written as an SVD inversion method in the following way:

$$\begin{aligned} x_{\text{MAP}} &= V_n \Sigma U^* \tilde{y} \\ \text{with } \Sigma &= \left[\tilde{B}(C_{V_n} S_n)^{-1} + S_n \right]^{-1}. \end{aligned} \quad (6.4.1)$$

For the proof we refer to Appendix A. The connection between (6.4.1) and L-SVD is clear if we define the linear measurement encoder to be $\varphi_{\text{enc}}^y := U^*$, the linear signal decoder to be $\varphi_{\text{dec}}^x := V_n$ and we assume the noise covariance matrix B to be known. Then it can be seen in (6.4.1) that learning a linear Σ is equivalent to optimising over the prior covariance matrix C_0 , defined via C_{V_n} .

Scale dependency on Gaussian noise level

For many inverse problems one assumes an additive noise term that originates from the Gaussian distribution $\eta^\delta \sim N(0, \delta \text{Id})$, where the noise level δ is either known or estimated. If the data covariance matrix B is replaced with δId , (6.4.1) is simplified to

$$\Sigma = \left[\delta(C_{V_n} S_n)^{-1} + S_n \right]^{-1}. \quad (6.4.2)$$

This implies a stronger regularisation $(C_{V_n} S_n)^{-1}$ by an increased noise level δ . Thus, by learning the scales Σ , we learn the prior distribution on x which regularises our inverse problem. For a prior distribution $C_{V_n} = \gamma \text{Id}$, it is easily shown that we get the formulation for classical Tikhonov-regularisation (6.2.3) back:

$$\Sigma = \left[\delta(\gamma \text{Id} S_n)^{-1} + S_n \right]^{-1} = \left[(\delta/\gamma + S_n^2) S_n^{-1} \right]^{-1} = S_n^{-1} \text{ for } \alpha := \delta/\gamma.$$

6.4.2 Data-driven Tikhonov regularisation with L-SVD

In case the forward operator A is known, it is possible to use the linear SVD encoder and decoder and train a nonlinear scaling Σ . By doing this in a structured manner, a data-driven Tikhonov regularisation can be learned, where the scales Σ nonlinearly depend on the measurements y^δ , potentially via $z_y = U^*y^\delta$.

Definition 6.4.2 (Data-driven Tikhonov regularisation). Let $x \in \mathcal{X} \subset \mathbb{R}^m$ and $y^\delta \in \mathcal{Y} \subset \mathbb{R}^n$, where $\mathcal{X}$ and $\mathcal{Y}$ are Hilbert spaces. Let $\mathcal{N} : \mathbb{R}^k \rightarrow \mathbb{R}^k$ be a nonlinear function s.t. for all $y \in \mathcal{Y}$: $0 < c_{\min} \leq \mathcal{N}(y) \leq C_{\max} < \infty$. Let $A = USV^*$ be the singular value decomposition as defined in Section 6.2. We define

$$\begin{aligned} x_\alpha^\delta &:= (A^*A + \alpha\mathcal{N}(y^\delta))^{-1}A^*y^\delta \\ &= V(S^2 + \alpha\mathcal{N}(y^\delta))^{-1}SU^*y^\delta. \end{aligned} \quad (6.4.3)$$

L-SVD provides a direct reconstruction through the encoder, scaling and decoder. The function $\mathcal{N}$ is a trained neural network that can be bounded by construction. In case it has the form of (6.4.3), the acquired solution is the unique minimiser of a convex functional:

Theorem 6.4.3 (Minimising functional). Let x_α^δ be defined as in (6.4.3). Then x_α^δ is the unique minimiser of the functional

$$J_\alpha(x) := \|Ax - y^\delta\|^2 + \alpha\mathcal{N}(y^\delta)\|x\|^2. \quad (6.4.4)$$

Proof. Since $0 < c_{\min} \leq \mathcal{N}(y) \leq C_{\max} < \infty$, for $\alpha > 0$, J_α is strictly convex. Moreover, $\lim_{\|x\| \rightarrow \infty} J_\alpha(x) = \infty$. Hence, J_α has a unique minimiser, which can be found by checking the first-order optimality condition. This yields the expression of (6.4.3). $\blacksquare$

L-SVD in the form of (6.4.3) provides a regularisation method: in the noisy case $\mathcal{N}(y^\delta)$ determines which scales should be regularised more and which less. In the limit of noise decreasing to zero, the solution converges to the unregularised solution, as shown in Theorem 6.4.4.

Theorem 6.4.4 (Convergence). Let $y \in \text{Im}(A)$, $\|y - y^\delta\| \leq \delta$ and x_α^δ defined as in (6.4.3). If $\alpha = \alpha(\delta)$ such that

$$\lim_{\delta \rightarrow 0} \alpha(\delta) = 0 \quad \text{and} \quad \lim_{\delta \rightarrow 0} \frac{\delta^2}{\alpha(\delta)} = 0, \quad (6.4.5)$$

then

$$\lim_{\delta \rightarrow 0} x_\alpha^\delta = A^\dagger y. \quad (6.4.6)$$

Proof. Define the sequence $\{\delta_n\}$ such that $\delta_n \rightarrow 0$, $\alpha_n := \alpha(\delta_n)$, $y_n := y^{\delta_n}$, $x_n := x_{\alpha_n}^{\delta_n}$. With J_n we define the functional (6.4.4) with variables as defined

above. By Theorem 6.4.3, x_n is the unique minimiser of J_n . Hence with $x^\dagger := A^\dagger y$,

$$\begin{aligned} \alpha_n \mathcal{N}(y_n) \|x_n\|^2 &\leq J_n(x_n) \leq J_n(x^\dagger) \\ &= \|Ax^\dagger - y_n\|^2 + \alpha_n \mathcal{N}(y_n) \|x^\dagger\|^2 \\ &= \|AA^\dagger y - y_n\|^2 + \alpha_n \mathcal{N}(y_n) \|x^\dagger\|^2 \\ &\leq \delta_n^2 + \alpha_n \mathcal{N}(y_n) \|x^\dagger\|^2. \end{aligned}$$

From this, we obtain

$$\|x_n\|^2 \leq \frac{\delta_n^2}{\alpha_n \mathcal{N}(y_n)} + \|x^\dagger\|^2 \leq \frac{\delta_n^2}{c_{\min} \alpha_n} + \|x^\dagger\|^2. \quad (6.4.7)$$

Because x_n is bounded, it has a convergent subsequence

$$x_{n_k} \rightarrow v \in \mathcal{X}.$$

Since the bounded linear operator A is sequentially continuous,

$$Ax_{n_k} \rightarrow Av \in \mathcal{Y}. \quad (6.4.8)$$

Again by Theorem 6.4.3, we obtain that

$$\|Ax_{n_k} - y_{n_k}\|^2 \leq J_{n_k}(x_{n_k}) \leq \delta_{n_k}^2 + \alpha_{n_k} C_{\max} \|x^\dagger\|^2 \rightarrow 0 \text{ as } k \rightarrow \infty. \quad (6.4.9)$$

Together with (6.4.8) this implies

$$Av = y. \quad (6.4.10)$$

Since any minimiser of J_n is in $\ker(A)^\perp$, $x_n \in \ker(A)^\perp$ for all x_n and therefore $v \in \ker(A)^\perp$. Together with (6.4.10), by [1, Theorem 2.5], this implies that $v = x^\dagger$, so that $x_{n_k} \rightarrow x^\dagger$. By applying the same argument to all subsequences, we obtain

$$x_n \rightarrow x^\dagger. \quad (6.4.11)$$

Since the sequence $\{\delta_n\}$ is arbitrarily chosen s.t. $\delta_n \rightarrow 0$, the desired expression (6.4.6) follows. $\blacksquare$

Finally, the convergence rate of (6.4.3) is derived. The proof of Theorem 6.4.5 makes use of several theorems in [1].

Theorem 6.4.5 (Convergence rate). *Let $w \in \mathcal{X}$ s.t. $\|w\|^2 \leq \rho$ and define $y := Aw$, $x^\dagger := A^\dagger y = A^\dagger Aw$. Then, the optimal parameter choice is $\alpha \sim \left(\frac{\delta}{\rho}\right)^{\frac{2}{3}}$, which provides the convergence rate*

$$\|x_\alpha^\delta - x^\dagger\|^2 = \mathcal{O}(\delta^{\frac{2}{3}}). \quad (6.4.12)$$

Proof. Let $g_{\alpha, y^\delta} : [0, \|A\|^2] \rightarrow \mathbb{R}$ be defined

$$g_{\alpha, y^\delta}(\lambda) := \frac{1}{\lambda + \alpha \mathcal{N}_\alpha(y^\delta)}. \quad (6.4.13)$$

This function meets the assumptions of [1, Theorem 4.1]:

$$|\lambda g_{\alpha, y^\delta}(\lambda)| = \frac{\lambda}{\lambda + \alpha \mathcal{N}_\alpha(y^\delta)} \leq 1$$

and

$$\lim_{\alpha \rightarrow 0} g_{\alpha, y^\delta}(\lambda) = \frac{1}{\lambda}$$

for all $\lambda \in (0, \|A\|^2]$. Next, for $\alpha > 0$ we define

$$G_{\alpha, y^\delta} := \sup_{\lambda \in [0, \|A\|^2]} |g_{\alpha, y^\delta}(\lambda)| = \sup_{\lambda \in [0, \|A\|^2]} \frac{1}{\lambda + \alpha \mathcal{N}_\alpha(y^\delta)} = \frac{1}{\alpha \mathcal{N}_\alpha(y^\delta)}.$$

Furthermore, we define

$$r_{\alpha, y^\delta}(\lambda) := 1 - \lambda g_{\alpha, y^\delta}(\lambda) = \frac{\alpha \mathcal{N}_\alpha(y^\delta)}{\lambda + \alpha \mathcal{N}_\alpha(y^\delta)}. \quad (6.4.14)$$

Finally, we define for $0 < \mu \leq 1$

$$\omega_\mu(\alpha) := \tilde{C}_{\max} \alpha^\mu, \quad \text{with } \tilde{C}_{\max} := \max\{1, C_{\max}\}. \quad (6.4.15)$$

For this ω_μ , $0 < \mu \leq 1$, the requirement in [1, Theorem 4.3] holds:

$$\lambda^\mu |r_{\alpha, y^\delta}| = \frac{\lambda^\mu \alpha \mathcal{N}_\alpha(y^\delta)}{\lambda + \alpha \mathcal{N}_\alpha(y^\delta)} \leq (C_{\max} \alpha)^\mu \leq \omega_\mu(\alpha). \quad (6.4.16)$$

An expanded derivation of (6.4.16) is provided in Appendix B. For $\mu \leq 1$, by [1, Corollary 4.4], the parameter choice $\alpha \sim \left(\frac{\delta}{\rho}\right)^{\frac{2}{2\mu+1}}$ is of optimal order in $\{x \in \mathcal{X} \mid x = (A^\dagger A)^\mu w, \ \|w\|^2 \leq \rho\}$. The best possible convergence rate is obtained with $\mu = 1$ for $x^\dagger = A^\dagger A w$, $\|w\|^2 \leq \rho$. This provides the convergence rate (6.4.12). ■

6.4.3 Fully learned L-SVD

In Section 6.4.2, it was shown that a data-driven Tikhonov regularisation is obtained by giving L-SVD the structure as shown in (6.4.3). The linear encoder and decoder are obtained from the SVD of A , which allows for the convergence analysis that was provided. In the current section we analyse the more general case where the L-SVD is fully learned, which means that the SVD of A can not be used. Furthermore, we consider the L-SVD with nonlinear encoder and decoder and diagonal matrix Σ , i.e. the second model choice of Section 6.3.1.

Definition 6.4.6 (Nonlinear L-SVD with diagonal scaling). Define the nonlinear functions and variables as in Definition 6.3.1. Let $\Sigma : \mathcal{Z}^y \rightarrow \mathcal{Z}^x$ be a diagonal matrix $\Sigma \in \mathbb{R}^{k \times k}$ with $\{\sigma_1, \dots, \sigma_k\}$ on the diagonal.

Theorem 6.4.7 (Stability estimate). *Let the variables and functions in L-SVD be defined as in Definition 6.4.6. Define two different measurements $y_{(1)}^\delta$ and $y_{(2)}^\delta$, s.t. $\|y_{(1)} - y_{(2)}\|_{\ell^2} \leq \varepsilon_y$ for some $\varepsilon_y \geq 0$. Then there is an $M > 0$ that depends on the weights and nonlinearities in the L-SVD network such that*

$$\|\hat{x}_{(1)}^\Sigma - \hat{x}_{(2)}^\Sigma\|_{\ell^2} \leq M\varepsilon. \quad (6.4.17)$$

Proof. From Definition 6.4.6 and on its turn Definition 6.3.1, we compute the bound

$$\begin{aligned} \|\hat{x}_{(1)}^\Sigma - \hat{x}_{(2)}^\Sigma\|_{\ell^2} &\leq \|\phi_{\text{dec}}^x\|_{\text{op}} \|\hat{z}_{x,(1)} - \hat{z}_{x,(2)}\|_{\ell^2} \\ &\leq |\sigma_{\max}| \|\phi_{\text{dec}}^x\|_{\text{op}} \|\hat{z}_{y,(1)} - \hat{z}_{y,(2)}\|_{\ell^2} \\ &\leq \|\phi_{\text{enc}}^y\|_{\text{op}} |\sigma_{\max}| \|\phi_{\text{dec}}^x\|_{\text{op}} \|y_{(1)} - y_{(2)}\|_{\ell^2} \\ &\leq \|\phi_{\text{enc}}^y\|_{\text{op}} |\sigma_{\max}| \|\phi_{\text{dec}}^x\|_{\text{op}} \varepsilon_y =: M\varepsilon_y, \end{aligned} \quad (6.4.18)$$

where $\|\cdot\|_{\text{op}}$ represents the operator norm. After training, the weights of the L-SVD model are fixed, which provides the desired bound (6.4.17). ■

Corollary 6.4.8. *For a nonlinear function*

$$\phi(x) := \tau(W_L \tau(W_{L-1} \dots \tau(W_1 x) \dots))$$

consisting of L layers with weight matrices W_l and pointwise nonlinearities $\tau(x)$ s.t. $\|\tau\|_{\text{op}} = C$, the following bound is achieved: $\|\phi\|_{\text{op}} \leq C^L \prod_{l=1}^L (\|W_l\|_{\ell^2})$. In case the nonlinearity is a ReLU, leaky ReLU or hyperbolic tangent, $C = 1$ and thus the norm only depends on the norms of the weight matrices. In this case the error estimate (6.4.18) can be written as

$$\|\hat{x}_{(1)}^\Sigma - \hat{x}_{(2)}^\Sigma\|_{\ell^2} \leq |\sigma_{\max}| \prod_{l=1}^L \left(\|W_{l,\text{enc}}^y\|_{\ell^2} \right) \prod_{l=1}^L \left(\|W_{l,\text{dec}}^x\|_{\ell^2} \right) \|y_{(1)} - y_{(2)}\|_{\ell^2}.$$

The error estimate can thus be controlled by assuring small ℓ^2 -norms of the weight matrices, which can be achieved by adding weight regularisation in the neural network objective function.

Next, we prove that an error estimate on the difference between any reconstructed signal and the true solution exists, provided that its associated measurement maps to a ball in the latent space. Before we provide this error estimate, we first prove a supporting Lemma.

Lemma 6.4.9. *Let $z \in \mathbb{R}^k$, let $F : \mathbb{R}^k \rightarrow \mathbb{R}^k$ be a continuous function. Assume $\forall z \in B_1, \|z - F(z)\|_{\ell^2} < \varepsilon$ for some given $0 \leq \varepsilon < 1$. Then $B_{1-\varepsilon} \subset F(B_1)$, where $B_r := \{z \in \mathbb{R}^k \mid \|z\|_{\ell^2} \leq r\}$ is the closed ball centered at 0 with radius r .*

Proof. Let us first define a scaled function $\tilde{F} : \mathbb{R}^k \rightarrow \mathbb{R}^k$, with $\tilde{F}(z) := \frac{1}{1+\varepsilon}F(z)$. For this scaled function, $\tilde{F}(B_1) \subset B_1$.

The closed unit ball B_1 is a contractible space. By definition of a contractible space [25]: $\exists G : B_1 \times [0, 1] \rightarrow B_1$ continuous such that for all z , $G(z, 0) = z_0$ and $G(z, 1) = z$, where z_0 is the contraction point. Since $\tilde{F}(B_1) \subset B_1$ for all $z \in \tilde{F}(B_1)$, it holds that $G(\tilde{F}(z), 0) = z_0$ and $G(\tilde{F}(z), 1) = \tilde{F}(z)$. Since both G and $\tilde{F}$ are continuous, its composition is continuous. From this it follows that $\tilde{F}(B_1)$ is a contractible space, which implies that $F(B_1)$ is a contractible space. In other words, $F(B_1)$ does not have any ‘holes’.

Left to show is that the boundary of $F(B_1)$ lies outside $B_{1-\varepsilon}$, which implies $B_{1-\varepsilon} \subset F(B_1)$, because $F(B_1)$ is contractible. For this we make use of [26, Theorem 4.22]: since F is a continuous mapping of a metric space $(\mathbb{R}^k, \|\cdot\|_{\ell^2})$ into a metric space $(\mathbb{R}^k, \|\cdot\|_{\ell^2})$, and the boundary of the unit ball (i.e. ∂B_1) is a connected subset of $\mathbb{R}^k$, this implies $F(\partial B_1)$ is connected. Moreover, $\forall z \in \partial B_1$, $F(z) \in B_{1+\varepsilon} \setminus B_{1-\varepsilon}$. This implies that the boundary of the unit ball lies completely outside $B_{1-\varepsilon}$, which completes the proof. ■

Theorem 6.4.10 (Reconstruction error estimate). *Let the variables and functions in L-SVD be defined as in Definition 6.4.6. Assume that for some $0 < \varepsilon_z < 1$, for all $\tilde{z}_x \in B_1$, $\|\tilde{z}_x - z_x^\Sigma\|_{\ell^2} < \varepsilon_z$. Then there is an $M > 0$ that depends on the weights and nonlinearities in the L-SVD network such that for all $\tilde{z}_x \in B_1$, $\|\phi_{\text{dec}}^x(\tilde{z}_x) - \phi_{\text{dec}}^x(\Sigma\phi_{\text{enc}}^y(A\phi_{\text{dec}}^x(\tilde{z}_x)))\|_{\ell^2} < M\varepsilon_z$. Moreover, for all $x \in \mathbb{R}^n$ for which $\phi_{\text{enc}}^x(x) \in B_{1-\varepsilon_z}$, we have the error estimate $\|x - \phi_{\text{dec}}^x(\Sigma\phi_{\text{enc}}^y(Ax))\|_{\ell^2} < M\varepsilon_z$.*

Proof. The first part of the proof can be obtained by combining the operator norm of the decoding function in x and the given bound:

$$\begin{aligned} & \|\phi_{\text{dec}}^x(\tilde{z}_x) - \phi_{\text{dec}}^x(\Sigma\phi_{\text{enc}}^y(A\phi_{\text{dec}}^x(\tilde{z}_x)))\|_{\ell^2} \\ & \leq \|\phi_{\text{dec}}^x\|_{\text{op}} \|\tilde{z}_x - \Sigma\phi_{\text{enc}}^y(A\phi_{\text{dec}}^x(\tilde{z}_x))\|_{\ell^2} \\ & < \|\phi_{\text{dec}}^x\|_{\text{op}} \varepsilon_z =: M\varepsilon_z \end{aligned} \tag{6.4.19}$$

For the second part of the proof, we make use of Lemma 6.4.9. We define $F(\tilde{z}_x) := \Sigma\phi_{\text{enc}}^y(A\phi_{\text{dec}}^x(\tilde{z}_x))$, which is a continuous function from $\mathbb{R}^k$ to $\mathbb{R}^k$. Since for all $\tilde{z}_x \in B_1$, $\|\tilde{z}_x - F(\tilde{z}_x)\| < \varepsilon_z$, we know that all elements in the ball $B_{1-\varepsilon_z}$ are in the range of $F(B_1)$. Therefore, for all $z_x^\Sigma \in B_{1-\varepsilon_z}$, there exists a $\tilde{z}_x \in B_1$, such that the same bound (6.4.19) holds. ■

Note that the reconstruction error estimate depends on $\|\phi_{\text{dec}}^x\|_{\text{op}}$ and ε_z . The former can be kept small by regularising the weights of the decoder in the training phase. The latter can be kept small by including $\|\tilde{z}_x - \Sigma\phi_{\text{enc}}^y(A\phi_{\text{dec}}^x(\tilde{z}_x))\|_{\ell^2}$ in the cost function, as described in the last point of Section 6.3.1. After training, points in the unit ball can be sampled uniformly and passed through the network to examine an actual value for ε_z .

6.5 Research context

This chapter presents the L-SVD method for solving inverse problems via hybrid autoencoding. Within the method, a low-dimensional (i.e. sparse) representation or manifold is explicitly learned. It is clear that this model has connections with many research fields, which are pointed out in this section.

6.5.1 Combining data and models for solving inverse problems

Recent research in inverse problems seeks to develop a mathematically coherent foundation for combining data-driven deep learning with model-based approaches based on physical-analytical domain knowledge [4]. A first class of methods are partially learned variational and iterative methods [27, 28, 5], see also Chapter 4. These methods can be seen as a learned variant of gradient, proximal or primal-dual methods. They require less iterations than their non-learned counterparts, but the demand on training time is substantial, while the mathematical analysis of these methods is limited. A second approach is to learn an explicit regularisation term [29, 30, 6, 7]. Signals affected by artefacts are penalised, while the desired signals are not. Reconstructions are of higher quality compared to classical regularisation choices, but their computation time is of the same order. A third approach is to perform learned post-processing of initial reconstructions obtained by classical methods, which may be affected by artefacts [31, 8]. Data-consistent reconstructions can be obtained without an iterative procedure [8]. However, the quality of the reconstructions heavily depends on the initial reconstruction, which is often obtained by applying a pseudo-inverse to the data.

The above methods depend on precise knowledge of the physical process, modelled by a forward mapping, which is not always available. However, empirically it was shown that learned iterative methods can still be used in case of inexact forward operators [5]. A recent work [32] aims to improve an inexact forward operator (linear mapping) explicitly with a neural network, after which it is applied in a variational framework. It was proven that small training losses ensure that the optimisation procedure finds a solution close to the one that would have been found with the true forward operator.

The optimal regularised inverse matrix (ORIM) method [9, 10] is a data-driven method that aims to find a linear inverse matrix for a noisy inverse problem. A global minimiser of the Bayes' risk is found when there is knowledge about the forward operator and about the probability distribution of the signals. A fully learned variant where the forward operator is not known is also available [9], albeit at a higher computational cost.

6.5.2 Fully learned image reconstruction

This chapter is closely related to the work of Zeng *et al.* [12]. In that paper, the task of superresolution is solved by autoencoding patches of both a low- and

high-resolution image and finding a nonlinear mapping between them. Gupta *et al.* [13] used this method for the task of removing motion blur, which is a specific case of a deconvolution problem. In both cases, a one-layered autoencoder was applied to patches of the distorted image (measurement) and desired image (signal). This is only possible if measurement and signal lie in the same domain and if the forward mapping has very little effect outside the local patch. We consider a more general method that does not work patch-based and therefore does not assume identical domains for measurement and signal. As a result, the forward mapping may have a global behaviour.

A different fully learned reconstruction method without this restriction is proposed by Zhu *et al.* [33], where the problem of finding a reconstruction from undersampled MRI data is considered. Their neural network consists of three fully connected layers, followed by three convolutional layers, which maps Fourier measurements directly to the desired image. A joint low-dimensional manifold is learned implicitly, since there is no explicit low-dimensional representation within the network architecture. In this chapter, an explicit representation of a joint manifold is learned in the form of two linearly connected latent codes. Results in [33] show a clear improvement over non-learned state-of-the-art methods for in-vivo data.

These works display the potential for fully learned methods in image reconstruction and inverse problems in general: high-quality reconstructions are obtained, while no exact knowledge of underlying physics or specifics of the measurement system is required.

6.5.3 Manifold learning

Many relevant inverse problems in medicine, engineering, astronomy or geophysics are large-scale in both signal and measurement space. However, seen from a statistical point of view, probability mass concentrates near manifolds of a much lower dimension than the original data spaces [34]. To detect linear manifolds, principal component analysis is a suitable and simple method. However, since manifolds for real data are expected to be strongly nonlinear [34], one needs to make use of nonlinear techniques. One of the best known methods that achieves this is kernel PCA [35]. In this method, data is mapped to a reproducing kernel Hilbert space by applying a nonlinear kernel, after which the linear PCA is applied. Other methods are principal geodesic analysis [36], which can be applied for Riemannian manifolds and geodesic PCA [37], which acts in a Wasserstein space, which is nonlinear.

Another approach to learn nonlinear manifolds is to use autoencoders, which can also be seen as a generalisation of linear PCA [18]. Autoencoders have shown to learn explicit representations of nonlinear manifolds [24] and provide better low-dimensional latent code in terms of clustering and reconstruction performance [18] than their linear counterpart. For the inverse problems (6.2.1) considered in this chapter, there is an explicit relationship between signals and

measurement via the forward mapping that models the physics: one shared data manifold is learned by connecting two autoencoders.

The observation of a shared manifold, or “an unknown underlying relationship between two domains” [38], is also the idea driving the cycle GAN. Unlike in this chapter, the goal of the cycle GAN is not to find a unique and supervised one-to-one mapping from one domain to the other, but to identify the shared parameters and add such elements so that the output is realistic in its respective domain. This cycle GAN was applied for inverse problems in different forms [39, 40]. In these works, the manifold is implicitly learned, unlike the explicitly learned representation that we study in this chapter.

6.5.4 Transfer learning with autoencoders

Transfer learning is used to exploit similarities between different tasks to share information necessary for both tasks. Representation learning, such as manifold learning, has a strong influence in transfer learning scenarios [34] since the learned representation can guide the supervised reconstruction task. Recent research has shown that autoencoders can be used as a regulariser for a supervised training task, such as classification [19, 20]. Such networks, coined supervised autoencoders (SAE), help to generalise the supervised problem. They are specifically useful in a semi-supervised scenario, where a lot of unsupervised training data is available, but supervised training pairs are scarce.

L-SVD profits from the same regularisation and generalisation effect by attaching two autoencoders to the supervised reconstruction problem. For inverse problems a semi-supervised scenario is also often encountered: imagine a training set of undersampled medical MRI or CT data sets and a set of high-quality reconstructions; not all pairs are available because not all patients have had a fully sampled scan that is needed for a high-quality reconstruction.

6.5.5 Model reduction and learning

Model reduction is a mathematical and computational field of study that derives low-dimensional models of complex systems [41, 42]. Via projections and decompositions it is possible to represent approximations of large-scale high-fidelity computational models resulting from discretisation of partial differential equations. Recent developments focus on data-driven learning of governing equations [43, 44, 45] and learned model reduction [46].

This chapter focuses on learning the inverse map for problems that are often physics-driven. The nonlinear equations or parameters of this map are implicitly learned through the latent representations via autoencoders.

6.5.6 Bayesian inversion and sparsity

The goal of the Bayesian approach to solving inverse problems is to find the posterior measure, given sampled data and a prior measure [16]. The posterior

measure will contain information about the relative probability of different outputs, given the data. Often the posterior is too complex to recover and the goal is shifted to finding a maximum a posteriori (MAP) estimate. In Section 6.4.1 we showed that a linear variant of L-SVD coincides with learning the covariance matrix of the prior measure.

Yu *et al.* [17] developed piecewise linear estimates (PLE), a method based on Gaussian mixture models (GMM), which are estimated via the MAP expectation-maximisation algorithm (MAP-EM). The method makes use of GMM as prior measures on local patches, which results in a linear reconstruction model for each patch. One could think of this procedure as finding a patchwork of locally linear tangent spaces which approximate a nonlinear manifold. If measurement and signal domain are the same, it can be shown that PLE is equivalent to learning a linear L-SVD method for a group of similar patches.

6.6 Experiments and implementation

In this section, we explain three simulation experiments: two that demonstrate the contributions as stated in Section 6.1.1 in a relatively low-resolution scenario and one that shows the application of L-SVD to a biomedical data-set in a higher resolution. The forward operator in all experiments is chosen to be the Radon transform [47], which is a nonlocal linear operator that produces an ill-posed inverse problem. We will refer to the measurements as ‘sinograms’ and to the signals as ‘images’ in their respective spaces $\mathcal{Y}$ and $\mathcal{X}$. In Appendix C, the neural network architectures and their parameter choices are provided, together with details of the training set.

6.6.1 Experiment 1: from model-based to data-driven

The goal of this simulation experiment is to demonstrate the first perspective of Section 6.1.1: data-driven solution of inverse problems. This is done by comparing classical non-learned methods to learned methods. For fair comparison and clarity, only linear inversion methods are considered.

This experiment makes use of the MNIST data set [48], after rescaling it to 64×64 pixels. We apply the Radon transform with 64 uniformly samples angles, i.e. a ‘full-angle’ Radon operator A , with a discretisation of 64 for each angular view. Moreover, there is no bottleneck latent space. This results in equally large spaces $\mathcal{Y} = \mathcal{Z}^y = \mathcal{Z}^x = \mathcal{X} = \mathbb{R}^{4096}$. For training, the ‘clean’ simulated full-angle measurements are normalised, after which Gaussian noise with a noise level $\delta = 0.05$ is added. The following linear reconstruction methods are compared:

- (a) Tikhonov-regularised reconstruction with optimally chosen α (Section 6.2.2).
- (b) T-SVD reconstruction with optimal truncation number r (Section 6.2.3).
- (c) Optimal regularised inverse matrix (ORIM) [9, Theorem 2.1].

- (d) Data-driven Tikhonov regularisation, i.e. U and V are from the SVD, while Σ is a learned diagonal matrix with the structure of (6.4.3) (Section 6.4.2).
- (e) Reconstruction from a learned covariance matrix of the prior, i.e. U and V are from the SVD, while Σ is a full matrix that is learned (Section 6.4.1).
- (f) Fully learned L-SVD: nonlinear L-SVD as in Definition 6.4.6. A regular autoencoder is used on the sinogram side, which means that noise should be reconstructed after the sinogram is encoded and decoded.

Besides comparing the reconstruction quality of above methods, we examine the ‘dictionary’ of elements that is obtained by decoding canonical basis vectors of the latent space: since all reconstruction methods are linear, the attainable reconstructions lie in the span of this dictionary. By this examination, the change from model-based to data-driven methods becomes evident.

Furthermore, by training on sinograms with noise levels between 0 and 0.2 instead of the aforementioned 0.05, the effect of noise on the matrix Σ is investigated. Instead of training for several noise levels individually, only one training set is created, where each sinogram has a randomly chosen noise level $\delta \in [0, 0.2]$. To process this data set adequately, the static scaling matrix Σ is exchanged for a noise-aware component that depends on the noisy input data (c.f. Section 6.3.1). This component is a nonlinear fully connected network which takes z_y as input and provides the diagonal scales $(\sigma_1, \sigma_2, \dots, \sigma_k)$ as output, which are multiplied with z_y .

6.6.2 Experiment 2: from linear autoencoding to hybrid nonlinear autoencoding

The goal of this simulation experiment is to demonstrate the second perspective of Section 6.1.1: nonlinear encoding is more effective than linear encoding; moreover, combining two nonlinear autoencoders has a regularising effect on the reconstruction and gives a more insightful latent representation than one autoencoder. For this experiment, we make use of the fully learned L-SVD.

This experiment again makes use of the MNIST data set [48], after rescaling it to 64×64 pixels. A limited-angle Radon transform of 8 uniformly sampled angles is applied, with a discretisation of 64 for each angular view. The latent space is chosen to have 64 dimensions, which means that it acts as a bottleneck. This results in the spaces $\mathcal{Y} = \mathbb{R}^{256}$, $\mathcal{Z}^y = \mathcal{Z}^x = \mathbb{R}^{64}$, $\mathcal{X} = \mathbb{R}^{4096}$, providing a dimensionality reduction of 12.5% and 1.56% compared to sinogram and image space respectively. Gaussian noise with a noise level $\delta = 0.05$ is added to the limited-angle measurements. We analyse the following methods:

- (a) linear autoencoder;
- (b) nonlinear autoencoder;
- (c) linear L-SVD;
- (d) nonlinear L-SVD ($\alpha = 0$);

(e) nonlinear L-SVD.

Here the first two methods are only applied on the image side and not on the sinogram side, meaning that the autoencoder only connects $\mathcal{X}$ and $\mathcal{Z}^x$. As can be seen in Figure 6.3.1, L-SVD connects the sinogram side with the image side, where a denoising autoencoder (DAE) is used on the sinogram side. This means that noise should not be reconstructed after decoding to $\mathcal{Y}$, which allows for a dimensionality reduction. The fourth method has the same network structure of nonlinear L-SVD, but without the autoencoders on either side (i.e. $\alpha_y = \alpha_x = 0$ in (6.3.1)), which allows us to investigate the regularising effect of the autoencoders. It is obvious that this experiment investigates the second perspective from Section 6.1.1, since the transition from linear to nonlinear is made, as well as the transition from a single to hybrid autoencoder.

Finally, we also compare nonlinear L-SVD for the following three cases:

- (a) All supervised training pairs (x, y^δ) are available.
- (b) Only 10% of the training samples is available, all in pairs (x, y^δ) .
- (c) All training samples x and y^δ are available, but only 10% is paired.

In the third case, if the training data is unpaired, then the encoders and decoders are only trained by the autoencoders. That is, the reconstruction loss $\mathcal{L}_1(\hat{x}_{(i)}^\Sigma, x_{(i)})$ in (6.3.1) is set to zero. We investigate how L-SVD can exploit the semi-supervised case in which additional unpaired training data is available.

6.6.3 Experiment 3: L-SVD on human chest CT images

The third simulation experiment demonstrates the capacity of L-SVD to reconstruct biomedical images without any knowledge on the forward operator. It exploits the Low-Dose Parallel Beam CT (LoDoPaB) data set [49], which is based on the LIDC/IDRI data set [50]. In this chapter, we only make use of the high-quality CT reconstructions in the LoDoPaB-CT data set that we use as ‘ground truth’ for our setup. The images are scaled to 128×128 pixels, after which a Radon transform of 36 uniformly sampled angles is applied with a discretisation of 192 for each angular view. The latent space of L-SVD is chosen to have 2048 dimensions. This results in the spaces $\mathcal{Y} = \mathbb{R}^{6912}$, $\mathcal{Z}^y = \mathcal{Z}^x = \mathbb{R}^{2048}$, $\mathcal{X} = \mathbb{R}^{16384}$, providing a dimensionality reduction of around 29.6% and 12.5% compared to sinogram and image space respectively. Gaussian noise with an signal-to-noise ratio (SNR) of 40dB is added to the measurements. The following reconstruction methods are compared:

- (a) T-SVD with optimal truncation number $k \leq 2048$;
- (b) optimal regularised inverse matrix (ORIM) [9];
- (c) total variation regularisation with optimal regularisation parameter α ;
- (d) fully learned nonlinear L-SVD.

Truncated SVD can be seen as the linear counterpart of L-SVD and makes use of the forward operator A . To obtain at least the same dimensionality reduction as L-SVD, the truncation number r of T-SVD is chosen smaller than 2048, but such that it gives the best reconstruction quality. ORIM is a linear data-driven method, for which we use the same training set as for L-SVD. For computational reasons, we make use of [9, equation (12)]: the ORIM method that requires knowledge about the mean and covariance of signals x and noise η^δ (see (6.2.1)). Moreover, the forward operator A is needed for this method. Total variation (TV) [51] is an adequate nonlinear regularised reconstruction method for this problem, since the ground truth CT-images consist of almost entirely piecewise constant structures. This method also needs the forward operator A : we minimise the functional $\min_x \|Ax - y^\delta\| + \alpha \text{TV}(x)$, implemented as described in Chapter 3.

6.7 Numerical results

In this section, the results of the simulation experiments explained in Section 6.6 are shown and discussed.

6.7.1 Experiment 1: from model-based to data-driven

The results of the first simulation experiments for a randomly chosen sample in the test set are shown in Figure 6.7.1. For this sample, Gaussian noise with a noise level $\delta = 0.05$ was added. Visually, the reconstruction improves gradually as we move from model-based methods with only one tunable parameter (b,c) to combinations of model-based and data-driven methods with tunable scaling matrix (d,e) to fully data-driven methods (f,g,h). This is most noticeable in the background of the reconstruction, which should be constant. The visual improvement is confirmed by the peak signal-to-noise ratio (PSNR), displayed above each reconstruction. These values represent the mean and standard deviation of the PSNR values for the first 1000 images in the test set.

To understand the transition from model-based to data-driven, canonical basis vectors in the latent space are decoded to the image space for both the regular SVD (Figure 6.7.1b-6.7.1e) and fully learned L-SVD with random initialisation (Figure 6.7.1h). Four selected elements from this ‘dictionary’ are shown in Figure 6.7.2.

SVD decomposes the Radon operator in different elements with a different geometrical scale. Moreover, it combines higher order harmonics in the image space and the sinogram space. For example, the second sinogram from the left in Figure 6.7.2c shows an approximate 2D sinusoidal structure, while Figure 6.7.2a provides its counterpart in the image space. For the third image from the left, it is the other way around. L-SVD shows similar behaviour for larger geometrical scales, but differences are also apparent: the sinusoids are only ‘active’ at the location of potential MNIST digits, and the sinogram space also encodes noise-

Figure 6.7.1: Reconstructions $\hat{x}^\Sigma$ (see Figure 6.3.1) for different models ranging from fully model-based (top left) to more data-driven (bottom right) with increasing amount of learning.

Figure 6.7.2: Selected elements in the latent space Z^x , decoded to the image space $\mathcal{X}$ and the sinogram space $\mathcal{Y}$. SVD only makes use of the operator, while L-SVD combines operator information with image and sinogram information. This results in more localised information in the decoded elements of the data-driven L-SVD approach.

like structures (first and second image in Figures 6.7.2b and 6.7.2d). Smaller scale elements (third image) show very localised geometrical structures in image space, while others (fourth image) only seem to capture noise.

Finally, we analyse the effect of noise on the diagonal scaling Σ with the network component as explained at the end of Section 6.6.1. We compare the networks where encoder and decoder are fixed as U^* and V with the fully learned networks. All networks are compared with Tikhonov regularisation, in which only one tunable parameter α is chosen such that the smallest MSE is

obtained. Gaussian noise is added with 6 different noise levels to all sinograms in the test set. The average scales per noise level are shown in Figure 6.7.3, where (a-c) have the same ordering as the classical SVD and (d) is ordered from large to small scales. For visualisation purposes, the graphs were smoothed by a Gaussian filter with scale 10.

Figure 6.7.3: Comparison of the scales σ_i depending on noise level for methods with increasing amount of learning. Although similar, methods (b-d), where the scales are learned individually, show a greater noise dependency than (a) Tikhonov regularisation. Note that methods (a-c) use the SVD ordering, while (d) is ordered from large to small scales at the highest noise level.

All methods show a decay of scales as the noise level grows. Tikhonov regularisation shows a very similar decay over all latent dimensions, while the decay is much more dimension dependent in (b-d). Here, the scales that coincide with large s_i in the SVD case, i.e. the first dimensions, are relatively noise independent. The scales that coincide with small s_i in the SVD case, i.e. dimensions 500-3000, show a relatively large decay, meaning that they are more affected by noise. Finally, the last part of all graphs show more than a thousand dimensions that have a small scale for all noise levels. To sum up, the behaviour of the scales to noise is similar for all methods in which the scales are learned individually, regardless of the encoding and decoding used. Moreover, most structural information is encoded in a limited number of latent dimensions, while the other dimensions encode of a substantial amount of noise: a compression is learned.

6.7.2 Experiment 2: from linear autoencoding to hybrid non-linear autoencoding

The results of the second simulation experiments for a randomly chosen sample in the test set are shown in Figure 6.7.1. For the conciseness of this section, only $\hat{x}^\Sigma$ is shown for all variants of L-SVD: not much difference between $\hat{x}^\Sigma$ and $\hat{x}^\Sigma$ was observed. For the autoencoders $\hat{x}^{\text{AE}}$ is shown, since $\hat{x}^\Sigma$ is not available there. Note that for $\hat{x}^{\text{AE}}$, no noise was added to their inputs x , while for the reconstruction outputs $\hat{x}^\Sigma$, noise with a noise-level of 0.05 was added to their inputs y^δ .

Figure 6.7.1: Comparison of the outputs $\hat{x}^{\text{AE}}$ and $\hat{x}^\Sigma$ (see Figure 6.3.1) for linear and nonlinear variants of the AE and fully learned L-SVD network.

It can be seen that the linear methods produce side-lobes to the main intensities, which are often observed in frequency-based compressions. The nonlinear methods provide a more homogeneous background in the reconstruction, especially L-SVD.

Type	Network	Output	Train loss	Test loss
linear	AE	$\hat{x}^{\text{AE}}$	10.2	10.1
	L-SVD	$\hat{x}^\Sigma$	13.7	13.4
nonlinear	AE	$\hat{x}^{\text{AE}}$	0.40	4.61
	L-SVD $_{\alpha=0}$	$\hat{x}^\Sigma$	1.22	5.34
	L-SVD	$\hat{x}^\Sigma$	2.00	3.94

Table 6.1: Generalisation performance analysis by comparing train and test losses (MSE). Nonlinear L-SVD with hybrid autoencoder shows a smaller difference between train and test loss than other nonlinear methods, indicating better generalisation performance.

In Table 6.1, the train and test losses of all compared methods are shown. The first thing that can be seen is that for the linear networks, errors are larger than for the nonlinear networks. The second thing is that there is no significant difference between train and test loss for the linear networks, which indicates that they generalise well. This difference is present in the nonlinear networks, but for L-SVD not to the same extent as for AE and L-SVD $_{\alpha=0}$. From the autoencoder point of view, it seems that L-SVD benefits from the sinogram

branch of the network in terms of generalisation. From a reconstruction point of view, L-SVD benefits from the incorporation of the two autoencoders, which contribute to its generalisation capacity in the reconstruction output $\hat{x}^\Sigma$. Finally, since the test loss for nonlinear L-SVD is smaller than for $\text{L-SVD}_{\alpha=0}$, we conclude that the two autoencoders act as regularisers for the reconstruction.

Next, canonical basis vectors in the latent space are decoded to the image space, to compose a ‘dictionary’ of elements. While in the linear case the output space is linearly spanned by dictionary elements, this is not true for the nonlinear case. This means that this dictionary only gives a partial view on the decoder.

Figure 6.7.2: Selected elements in the latent space $\mathcal{Z}^x$, decoded to the image space $\mathcal{X}$. Nonlinear $\text{L-SVD}_{\alpha=0}$ and nonlinear L-SVD learn a more interpretable representation than other methods by combining features from sinogram and image space in their ‘dictionary’. Due to its similarity to linear L-SVD, linear AE is not shown here. All 64 elements are shown in Appendix D.

Figure 6.7.2 shows four selected elements that are exemplary for the complete dictionaries, which are provided in Appendix D. Because the elements of linear AE are very similar to linear L-SVD, they are not shown here. In the top left, linear L-SVD shows an element that is very similar to the Euclidean mean of all training samples, while the top right and bottom images show low and high frequency components that are only active in the middle. Nonlinear AE shows much smaller structures in its elements, which do not seem to have a visual coherent structure. The nonlinear networks $\text{L-SVD}_{\alpha=0}$ and L-SVD do provide

this visually coherent structure, where L-SVD seems to provide somewhat ‘smoother’ and better connected structures than $\text{L-SVD}_{\alpha=0}$. Their dictionaries consist of various elements, of which one is similar to the Euclidean mean (top left), some are similar to digits (top right), some that show a combination of line segments (bottom left) and some with high-frequency components (bottom right) which were also visible in the linear dictionary. With this diversity, fully learned L-SVD combines information from images, sinograms and operator.

	#training pairs	#training samples	Test loss
supervised	60 000	60 000	3.94
supervised	6 000	6 000	13.3
semi-supervised	6 000	60 000	9.02

Table 6.2: Comparison of test losses (MSE) for supervised L-SVD and semi-supervised L-SVD, where 90% of the data is only trained by the autoencoders incorporated in L-SVD.

Finally, the capacity of L-SVD in a semi-supervised setup is shown in Table 6.2. Only 10% of the training samples are available in pairs (i.e. supervised), and the other 90% are available, but not in pairs (i.e. unsupervised). Table 6.2 shows that the semi-supervised setup provides a highly increased performance over the supervised case where only the pairs (i.e. 10%) are used. This demonstrates the advantage of adding the autoencoders in L-SVD: they help to efficiently encode and decode, even if pairs between data and signal are not available.

6.7.3 Experiment 3: L-SVD on human chest CT images

Here we demonstrate the capacity of L-SVD to reconstruct biomedical images without any knowledge on the forward operator. We compare it to T-SVD, ORIM and TV, which are all methods that make use of the forward operator A . For a randomly chosen sample in the test set, the outputs of all methods are shown in Figure 6.7.1. It can be seen that none of the methods can reconstruct all details that are apparent in the ground truth. This is probably due to the limited number of angles that were used and the noise added to the sinogram. The T-SVD and ORIM reconstructions give typical ringing artefacts, while the TV reconstruction has the typical staircase behaviour with clusters of piecewise constant structures. The L-SVD reconstruction is smoother than the other methods, which can be seen in the large piecewise constant structure in the bottom left of the image, as well as close to the boundary of the imaged body, where a low contrast between tissues should be reconstructed. To get a better idea of the reconstruction biases of all methods, three additional samples from the test set are provided in Appendix E.

Figure 6.7.1: Comparison between fully learned L-SVD and reconstruction methods that make use of the forward operator A .

In Table 6.1, the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) of all compared methods are given.

reconstruction method	PSNR	SSIM
truncated SVD (T-SVD)	33.86 ± 1.11	0.845 ± 0.023
optimal regularised inverse matrix (ORIM)	34.19 ± 1.12	0.865 ± 0.023
total variation (TV)	35.90 ± 1.40	0.917 ± 0.020
learned SVD (L-SVD)	36.34 ± 1.62	0.920 ± 0.023

Table 6.1: Comparison in PSNR and SSIM between L-SVD and non-learned reconstruction methods that make use of the forward operator A .

Firstly, it can be seen that TV and L-SVD, both nonlinear methods, provide a higher quality than T-SVD and ORIM, linear reconstruction methods. Secondly, from the methods that contain a bottleneck in their architecture, L-SVD is superior to T-SVD. This can be due to its nonlinearity or the fact that it is a learned method. Thirdly, from the data-driven methods, L-SVD shows better results than ORIM, which is probably due to the nonlinearity of L-SVD. Finally, TV and L-SVD are comparable in terms of PSNR and SSIM. However, to use TV as a reconstruction method, it is necessary to know the forward operator A exactly, which is not always possible. Moreover, this quality of TV comes at the expense of computation power: L-SVD is a direct reconstruction method, while

TV needs close to a 1000 iterations to converge to its solution. On the other hand, L-SVD requires a large training set and training time, while there is only one parameter to tune for TV.

6.8 Conclusion and outlook

We proposed the learned SVD for inverse problems: a reconstruction method that connects low-dimensional representations of corrupted measurements and desired signals. When the forward mapping is known, the L-SVD can be shaped into a data-driven Tikhonov regularisation method, for which we provided a convergence analysis. When the forward mapping is unknown, nonlinear representations of measurements and signals can be fully learned from data, with a connecting layer that is fully connected or sparse, linear or nonlinear, noise dependent or independent. One specific choice is to incorporate the necessary nonlinearity of the learned inverse function in the autoencoders, while the sparse diagonal scaling layer is chosen to be linear, making the connection between measurement and signal manifold easy to understand. In simulation experiments, it was shown that this nonlinear reconstruction gives superior performance to other methods, while providing interpretable autoencoding. Moreover, since the reconstruction error estimate depends on the autoencoding quality, L-SVD can benefit from general advances in nonlinear autoencoding.

Results show that L-SVD makes use of information from both measurements and signals; by doing so, it learns elements of the physics operator, although not explicitly provided. Therefore the method is especially promising in applications where the forward physics are not completely understood or computationally expensive to simulate. Learning a joint manifold by two connected autoencoders also enables the possibility of a semi-supervised setup: the autoencoders provide regularisation for reconstruction of the signal.

Due to its generic formulation, L-SVD is very flexible for other architecture choices in autoencoding. Therefore, future efforts will lie in investigating other architectures, such as convolutional autoencoders and ladder variational autoencoders [52], for their inclusion in L-SVD for large scale inverse problems. Furthermore, other loss functions such as the Wasserstein loss or a learned discriminator could be investigated. Finally, we will focus on other ways of incorporating (partial) information of the forward mapping in the architecture of the L-SVD method, to create a regularisation method with a convergence analysis that benefits from the advantages of nonlinear autoencoding.

References

- [1] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, vol. 375. Springer Science & Business Media, 1996.
- [2] M. Benning and M. Burger, "Modern regularization methods for inverse problems," *Acta Numerica*, vol. 27, pp. 1–111, 2018.
- [3] G. Golub and W. Kahan, "Calculating the singular values and pseudo-inverse of a matrix," *Journal of the Society for Industrial and Applied Mathematics, Series B: Numerical Analysis*, vol. 2, no. 2, pp. 205–224, 1965.
- [4] S. Arridge, P. Maass, O. Öktem, and C.-B. Schönlieb, "Solving inverse problems using data-driven models," *Acta Numerica*, vol. 28, pp. 1–174, May 2019.
- [5] A. Hauptmann, B. Cox, F. Lucka, N. Huynh, M. Betcke, P. Beard, and S. Arridge, "Approximate k-space models and deep learning for fast photoacoustic reconstruction," in *International Workshop on Machine Learning for Medical Image Reconstruction*, pp. 103–111, Springer, 2018.
- [6] S. Lunz, O. Öktem, and C.-B. Schönlieb, "Adversarial regularizers in inverse problems," in *Advances in Neural Information Processing Systems*, pp. 8507–8516, 2018.
- [7] H. Li, J. Schwab, S. Antholzer, and M. Haltmeier, "NETT: solving inverse problems with deep neural networks," *Inverse Problems*, 2020.
- [8] J. Schwab, S. Antholzer, and M. Haltmeier, "Deep null space learning for inverse problems: convergence analysis and rates," *Inverse Problems*, 2018.
- [9] J. Chung and M. Chung, "An efficient approach for computing optimal low-rank regularized inverse matrices," *Inverse Problems*, vol. 30, no. 11, p. 114009, 2014.
- [10] J. Chung, M. Chung, and D. P. O'Leary, "Optimal regularized low rank inverse approximation," *Linear Algebra and its Applications*, vol. 468, pp. 260–269, 2015.
- [11] K. Bhattacharya, B. Hosseini, N. B. Kovachki, and A. M. Stuart, "Model reduction and neural networks for parametric pdes," *arXiv preprint arXiv:2005.03180*, 2020.
- [12] K. Zeng, J. Yu, R. Wang, C. Li, and D. Tao, "Coupled deep autoencoder for single image super-resolution," *IEEE transactions on cybernetics*, vol. 47, no. 1, pp. 27–37, 2017.
- [13] K. Gupta, B. Bhowmick, and A. Majumdar, "Motion blur removal via coupled autoencoder," in *2017 IEEE International Conference on Image Processing (ICIP)*, pp. 480–484, IEEE, 2017.
- [14] S. N. Evans and P. B. Stark, "Inverse problems as statistics," *Inverse Problems*, vol. 18, p. 201, Aug 2002.
- [15] J. P. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, vol. 160 of *Applied Mathematical Sciences*. New York: Springer-Verlag, 2005.
- [16] A. M. Stuart, "Inverse problems: a Bayesian perspective," *Acta Numerica*, vol. 19, pp. 451–559, May 2010.
- [17] G. Yu, G. Sapiro, and S. Mallat, "Solving inverse problems with piecewise linear estimators: from Gaussian mixture models to structured sparsity," *IEEE transactions on image processing*, vol. 21, pp. 2481–2499, May 2012.

- [18] G. E. Hinton, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, pp. 504–507, Jul 2006.
- [19] Y. Zhang, K. Lee, and H. Lee, "Augmenting supervised neural networks with unsupervised objectives for large-scale image classification," in *International Conference on Machine Learning (ICML)*, vol. 48, (New York), pp. 612–621, JMLR, Jun 2016.
- [20] L. Le, A. Patterson, and M. White, "Supervised autoencoders: improving generalization performance with unsupervised regularizers," in *Neural Information Processing Systems (NeurIPS)*, pp. 107–117, Curran Associates, Inc., 2018.
- [21] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936.
- [22] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion," *Journal of Machine Learning Research*, vol. 11, pp. 3371–3408, 2010.
- [23] M. Ranzato, C. Poultney, S. Chopra, and Y. L. Cun, "Efficient learning of sparse representations with an energy-based model," in *Advances in neural information processing systems*, pp. 1137–1144, 2007.
- [24] S. Rifai, P. Vincent, X. Muller, X. Glorot, and Y. Bengio, "Contractive auto-encoders: explicit invariance during feature extraction," in *Proceedings of the 28th International Conference on International Conference on Machine Learning*, pp. 833–840, Omnipress, 2011.
- [25] T. W. Gamelin and R. E. Greene, *Introduction to topology*. Dover Publications, 1999.
- [26] W. Rudin, *Principles of Mathematical Analysis*. New York: McGraw-Hill, 3. Aufl. ed., 1976.
- [27] J. Adler and O. Öktem, "Solving ill-posed inverse problems using iterative deep neural networks," *Inverse Problems*, vol. 33, p. 124007, Dec 2017.
- [28] E. Kobler, T. Klatzer, K. Hammernik, and T. Pock, "Variational networks: connecting variational methods and deep learning," in *German conference on pattern recognition*, pp. 281–293, Springer, 2017.
- [29] H. Chen, Y. Zhang, Y. Chen, J. Zhang, W. Zhang, H. Sun, Y. Lv, P. Liao, J. Zhou, and G. Wang, "LEARN: learned experts' assessment-based reconstruction network for sparse-data CT," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1333–1347, 2018.
- [30] H. K. Aggarwal, M. P. Mani, and M. Jacob, "Modl: model-based deep learning architecture for inverse problems," *IEEE transactions on medical imaging*, vol. 38, no. 2, pp. 394–405, 2018.
- [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, 2017.
- [32] S. Lunz, A. Hauptmann, T. Tarvainen, C.-B. Schönlieb, and S. Arridge, "On learned operator correction," *arXiv:2005.07069*, 2020.
- [33] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, "Image reconstruction by domain-transform manifold learning," *Nature*, vol. 555, no. 7697, pp. 487–492, 2018.

- [34] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: a review and new perspectives," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.
- [35] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural computation*, vol. 10, no. 5, pp. 1299–1319, 1998.
- [36] P. T. Fletcher, C. Lu, S. M. Pizer, and S. Joshi, "Principal geodesic analysis for the study of nonlinear statistics of shape," *IEEE transactions on medical imaging*, vol. 23, no. 8, pp. 995–1005, 2004.
- [37] J. Bigot, R. Gouet, T. Klein, A. López, *et al.*, "Geodesic PCA in the Wasserstein space by convex PCA," in *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, vol. 53, pp. 1–26, Institut Henri Poincaré, 2017.
- [38] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *The IEEE International Conference on Computer Vision (ICCV)*, pp. 2223–2232, 2017.
- [39] O. Senouf, S. Vedula, T. Weiss, A. Bronstein, O. Michailovich, and M. Zibulevsky, "Self-supervised learning of inverse problem solvers in medical imaging," in *Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data*, pp. 111–119, Springer, 2019.
- [40] B. Sim, G. Oh, S. Lim, and J. C. Ye, "Optimal transport, cycleGAN, and penalized LS for unsupervised learning in inverse problems," *arXiv:1909.12116*, 2019.
- [41] T. Bui-Thanh, K. Willcox, and O. Ghattas, "Model reduction for large-scale systems with high-dimensional parametric input space," *SIAM Journal on Scientific Computing*, vol. 30, no. 6, pp. 3270–3288, 2008.
- [42] P. Benner, S. Gugercin, and K. Willcox, "A survey of projection-based model reduction methods for parametric dynamical systems," *SIAM Review*, vol. 57, pp. 483–531, Jan 2015.
- [43] S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proceedings of the National Academy of Sciences*, vol. 113, pp. 3932–3937, Apr 2016.
- [44] S. H. Rudy, S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Data-driven discovery of partial differential equations," *Science Advances*, vol. 3, p. e1602614, Apr 2017.
- [45] M. Raissi, P. Perdikaris, and G. Karniadakis, "Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, pp. 686–707, Feb 2019.
- [46] E. Qian, B. Kramer, A. N. Marques, and K. E. Willcox, "Transform & learn: a data-driven approach to nonlinear model reduction," in *AIAA Aviation 2019 Forum*, pp. 1–11, American Institute of Aeronautics and Astronautics, Jun 2019.
- [47] A. C. Kak, M. Slaney, and G. Wang, "Principles of computerized tomographic imaging," *Medical Physics*, vol. 29, no. 1, pp. 107–107, 2002.
- [48] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, pp. 2278–2324, Nov 1998.

- [49] J. Leuschner, M. Schmidt, D. O. Baguer, and P. Maaß, “The LoDoPaB-CT dataset: a benchmark dataset for low-dose CT reconstruction methods,” *arXiv:1910.01113*, 2019.
- [50] S. G. Armato III, G. McLennan, L. Bidaut, M. F. McNitt-Gray, C. R. Meyer, A. P. Reeves, B. Zhao, D. R. Aberle, C. I. Henschke, E. A. Hoffman, *et al.*, “The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans,” *Medical physics*, vol. 38, no. 2, pp. 915–931, 2011.
- [51] L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Physica D: nonlinear phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.
- [52] C. K. Sønderby, T. Raiko, L. Maaløe, S. K. Sønderby, and O. Winther, “Ladder variational autoencoders,” in *Advances in neural information processing systems*, pp. 3738–3746, 2016.
- [53] R. M. Lewitt, “Multidimensional digital image representations using generalized Kaiser-Bessel window functions,” *Journal of the Optical Society of America*, vol. 7, no. 10, pp. 1834–1846, 1990.

Appendices

A Proof of Proposition 6.4.1

Proof. We will first take a look at the more general case where $\mu_0 \sim N(m_0, C_0)$ before we set $m_0 = 0$. We determine the posterior measure μ_{post} for x given $\tilde{y}$ (see equation (3.4) in [16]) as

$$\begin{aligned} \mu_{\text{post}} &= N(m_{\text{post}}, C_{\text{post}}) \\ \text{where } m_{\text{post}} &= m_0 + C_0 A^* (B + A C_0 A^*)^{-1} (\tilde{y} - A m_0) \\ \text{and } C_{\text{post}} &= C_0 - C_0 A^* (B + A C_0 A^*)^{-1} A C_0. \end{aligned} \quad (6.A.1)$$

The maximum a posteriori (MAP) estimate $x_{\text{MAP}} := \operatorname{argmax}_x p(y|x)p(x)$ coincides with the mean of μ_{post} , i.e.

$$x_{\text{MAP}} = \operatorname{argmin}_x \left\{ \|Ax - \tilde{y}\|_B^2 + \|x - m_0\|_{C_0}^2 \right\} = m_{\text{post}}.$$

We substitute $A = U S_n V_n^*$ in (6.A.1), from which we obtain

$$\begin{aligned} m_{\text{post}} &= m_0 + C_0 V_n S_n^* U^* (B + U S_n V_n^* C_0 V_n S_n^* U^*)^{-1} (\tilde{y} - U S_n V_n^* m_0) \\ &= m_0 + C_0 V_n S_n^* U^* (U \tilde{B} U^* + U S_n V_n^* C_0 V_n S_n^* U^*)^{-1} (\tilde{y} - U S_n V_n^* m_0) \\ &= m_0 + C_0 V_n S_n^* (\tilde{B} + S_n V_n^* C_0 V_n S_n^*)^{-1} (U^* \tilde{y} - S_n V_n^* m_0) \\ &= m_0 + C_0 V_n (S_n^{-1} \tilde{B} (S_n^*)^{-1} + V_n^* C_0 V_n)^{-1} (S_n^{-1} U^* \tilde{y} - V_n^* m_0) \\ &= m_0 + V_n C_{V_n} (S_n^{-1} \tilde{B} (S_n^*)^{-1} + C_{V_n})^{-1} (S_n^{-1} U^* \tilde{y} - V_n^* m_0), \end{aligned}$$

where we used $\tilde{B} := U^* B U$ in the second equality and $C_0 = V_n C_{V_n} V_n^*$ in the last equality. In case of a prior distribution with zero-mean, i.e. $m_0 = 0$, we get the expression

$$m_{\text{post}} = V_n \underbrace{C_{V_n} (S_n^{-1} \tilde{B} (S_n^*)^{-1} + C_{V_n})^{-1} S_n^{-1}}_{\Sigma} U^* \tilde{y}. \quad (6.A.2)$$

The symmetric covariance matrix C_{V_n} is positive definite, hence it is also invertible. We simplify

$$\begin{aligned} \Sigma^{-1} &= S_n (S_n^{-1} \tilde{B} (S_n^*)^{-1} + C_{V_n}) C_{V_n}^{-1} \\ &= \tilde{B} (C_{V_n} S_n)^{-1} + S_n. \end{aligned}$$

By substituting this expression for Σ in (6.A.2), we obtain

$$\begin{aligned} m_{\text{post}} &= V_n \Sigma U^* \tilde{y} \\ \text{with } \Sigma &= \left[\tilde{B} (C_{V_n} S_n)^{-1} + S_n \right]^{-1}. \end{aligned}$$

■

B Inequality (6.4.16) written out

This section shows that the inequality (6.4.16) holds, given (6.4.14) and (6.4.15). We consider the function

$$\lambda^\mu |r_{\alpha, y^\delta}| = \frac{\lambda^\mu \alpha \mathcal{N}_\alpha(y^\delta)}{\lambda + \alpha \mathcal{N}_\alpha(y^\delta)}.$$

For $\mu < 1$ this function attains its maximum at $\lambda = (\mu \alpha \mathcal{N}_\alpha(y^\delta)) / (1 - \mu)$. Filling this in yields

$$\begin{aligned} \lambda^\mu |r_{\alpha, y^\delta}| &= \frac{\lambda^\mu \alpha \mathcal{N}_\alpha(y^\delta)}{\lambda + \alpha \mathcal{N}_\alpha(y^\delta)} \\ &\leq \frac{\left(\frac{(\mu \alpha \mathcal{N}_\alpha(y^\delta))^\mu}{(1-\mu)^\mu} \right) \alpha \mathcal{N}_\alpha(y^\delta)}{\left(\frac{(\mu \alpha \mathcal{N}_\alpha(y^\delta))}{(1-\mu)} \right) + \alpha \mathcal{N}_\alpha(y^\delta)} \\ &= \frac{(\mu^\mu (\alpha \mathcal{N}_\alpha(y^\delta))^{(\mu+1)}) / (1-\mu)^\mu}{(\alpha \mathcal{N}_\alpha(y^\delta)) / (1-\mu)} \\ &= (1-\mu)^{1-\mu} \mu^\mu (\alpha \mathcal{N}_\alpha(y^\delta))^\mu \\ &\leq (\alpha \mathcal{N}_\alpha(y^\delta))^\mu \leq (C_{\max} \alpha)^\mu \leq \tilde{C}_{\max} \alpha^\mu = \begin{cases} \alpha^\mu & \text{if } C_{\max} < 1, \\ C_{\max} \alpha^\mu & \text{if } C_{\max} \geq 1. \end{cases} \end{aligned}$$

For $\mu = 1$ the maximum is attained at $\lambda = \|A\|^2$ (the right end of the interval for λ). Filling this in yields

$$\begin{aligned} \lambda^\mu |r_{\alpha, y^\delta}| &= \frac{\lambda^\mu \alpha \mathcal{N}_\alpha(y^\delta)}{\lambda + \alpha \mathcal{N}_\alpha(y^\delta)} \\ &\leq \frac{\|A\|^{2\mu} \alpha \mathcal{N}_\alpha(y^\delta)}{\|A\|^2 + \alpha \mathcal{N}_\alpha(y^\delta)} \\ &\leq \|A\|^{2\mu-2} \alpha \mathcal{N}_\alpha(y^\delta) \\ &= \alpha \mathcal{N}_\alpha(y^\delta) \leq C_{\max} \alpha, \end{aligned}$$

where we filled in $\mu = 1$ in the last equality. Combined, this provides the desired inequality (6.4.16).

C Implementation details

This appendix provides the implementation details of the neural networks used in this chapter. Training is performed in Tensorflow. The first two experiments make use of the complete MNIST training set (60 000 training samples), where each image is rescaled to 64×64 using bilinear interpolation. Testing is done on the first 1000 samples of the MNIST test set, using the same interpolation. For these experiments, the Radon transform is applied using the ‘scikit-image’ toolbox in Python. The third experiment makes use of the complete LoDoPaB training set. A data-augmented training set is obtained by mirroring and rotating each image with multiples of 90° . This yields a training set with 284 672 samples: eight times the original number of samples. For this experiment, the Radon transform as described in [53] is applied.

The networks in the first and second experiment only make use of fully connected (FC) layers. The third experiment has additional convolutional layers on the image side: before the fully connected layers in the encoder and after the fully connected layers in the decoder. Details about the encoder and decoder are given in Table C.1. After each layer, except the last layer, a leaky ReLU (lReLU) with parameter γ as specified in Table C.1 is applied. All networks in the first two experiments are chosen without biases, the third experiment makes use of biases. Initial weights of all experiments are normally distributed with a standard deviation of 0.01.

Parameter	Experiment 1f	Experiment 2	Experiment 3
# FC layers $\varphi_{\text{enc}}^y / \varphi_{\text{dec}}^y$	1	1	1
# FC layers $\varphi_{\text{enc}}^x / \varphi_{\text{dec}}^x$	1	4	3
biases	no	no	yes
nonlinearity	-	lReLU $_{\gamma=0.1}$	lReLU $_{\gamma=0.2}$
# conv layers $\varphi_{\text{enc}}^y / \varphi_{\text{dec}}^y$	0	0	0
# conv layers $\varphi_{\text{enc}}^x / \varphi_{\text{dec}}^x$	0	0	3
# kernels per conv layer	-	-	32
kernel size	-	-	5×5

Table C.1: Architecture choices of the encoder and decoder for each of the experiments.

The fully learned L-SVD networks (experiments 1f, 2 and 3) make use of a linear scaling matrix Σ . Experiments 1d and 1e make use of a nonlinear scaling function $\Sigma(z_y)$ in the form of a neural network, which consists of five fully connected layers with biases. After each layer a leaky ReLU with parameter $\gamma = 0.1$ is applied, except for the last layer:

- Experiment 1d applies $((C_{\text{max}} - c_{\text{min}})\text{sig}(\cdot) + c_{\text{min}})$ as the final nonlinearity, where $\text{sig}(\cdot)$ denotes the sigmoid function. This nonlinearity ensures the bounds $c_{\text{min}} \leq \mathcal{N}(\cdot) \leq C_{\text{max}}$ (see Section 6.4.2). In this experiment,

$c_{\min} = 10^{-2}$ and $C_{\max} = 10$. The total scaling

$$\Sigma(z_y) = ((S^2 + \alpha \mathcal{N}(z_y))^{-1} S$$

is obtained by choosing $\alpha(\delta) = \delta^{\frac{2}{3}}$ and taking S from the SVD of A .

- Experiment 1e applies the softplus function as the final nonlinearity. This immediately yields the total scaling $\Sigma(z_y)$.

Initial weights are normally distributed with a standard deviation of 0.01.

For all experiments, each loss function $\mathcal{L}_j$ in (6.3.1) is of ℓ^2 -type (mean squared error) with α_y, α_x as specified in Table C.2. We apply the ADAM optimiser using a learning rate with exponential weight decay. The number of epochs, batch sizes and the start and final learning rates are stated in Table C.2. All other optimisation parameters are the default choices of ADAM in Tensorflow. Gradient norm clipping with a value of 10 is applied for training stability. No regularisation, dropout or batch normalisation are used.

Parameter	Experiment 1	Experiment 2	Experiment 3
loss parameters	$\alpha_y = 2, \alpha_x = 1$	$\alpha_y = 2, \alpha_x = 1$	$\alpha_y = 0.2, \alpha_x = 1$
# epochs	250	250	500
batch size	100	100	100
start learning rate	10^{-3}	10^{-3}	$2 \cdot 10^{-4}$
final learning rate	$2 \cdot 10^{-4}$	$2 \cdot 10^{-4}$	$2 \cdot 10^{-5}$

Table C.2: Optimisation choices for each of the described experiments.

D Visualisation of all latent space elements

Figure D.1 provides the complete dictionaries for all methods, from which a selection was shown in Figure 6.7.2. For a discussion on the results we refer to Section 6.7.2.

Figure D.1: All 64 elements of the dictionary of all methods, from which a selection was shown in Figure 6.7.2.

E Additional results for human chest CT images

Figure E.1

Figure E.2

Figure E.3

Chapter 7

Conclusion and outlook

The first part of this thesis focused on image reconstruction in the field of photoacoustic tomography (PAT). In the course of this work, the focus shifted from photoacoustic reconstruction in particular to image reconstruction and solving inverse problems in general. Simultaneously, reconstruction methods changed from purely model-based (non-learned) to combinations of model-based and data-driven (learned). Below we summarise our main findings per chapter. After that, we put these findings in a general perspective by discussing potential future steps in the research fields of photoacoustic imaging and image reconstruction.

In Chapter 3 we solved the acoustic PAT problem by a model-based primal-dual algorithm. Three hand-crafted regularisers were chosen based on the type of images to be reconstructed. All regularisers showed improved results over direct methods for PAT in cases where the noise level was high or the number of measurements was low. However, all results showed a distinctive bias induced by the regulariser. These findings indicate a potential gain by learning a regulariser that is tailored to the specific problem or by learning part of the reconstruction procedure.

In Chapter 4 we solved the acoustic PAT problem by the learned primal-dual algorithm. Here we found that an artificial neural network (ANN) trained on an inadequate training set can provide visually pleasing but wrong reconstructions. We showed that this danger can be mitigated by creating more variety in the training set, which creates generalisation performance of the algorithm and improved experimental results. Furthermore we investigated the possibility of performing the additional task of image segmentation, without changing the structure of the learned algorithm. Using an adequate training set, this algorithm provides reconstructions and segmentations of structures deeper in the tissue than other model-based and data-driven methods. The above findings motivate the search for reliable methods that are not only data-specific, but also possess a theoretical bound on the reconstruction error.

In Chapter 5 we developed data-consistent networks for nonlinear inverse problems, including a full convergence analysis. These networks combine an initial model-based reconstruction with an improvement by an ANN that enforces data-consistency in the limit of decreasing noise. Numerical results showed that these networks possess better generalisability than ANNs that do not ensure data-consistency: they performed better on data sets that are different from the training set.

In Chapter 6 we proposed the learned singular value decomposition (L-SVD), a fully learned reconstruction method for solving inverse problems. The method can be seen as a nonlinear data-driven extension of linear methods that consist of an encoding, scaling and decoding step. We showed that the use of an auto-encoder within is beneficial for the generalisability of the method and enables semi-supervised training. L-SVD does not require knowledge about the forward process, but its design allows to add parts of the forward mapping in case it is known. In that case a data-driven regularisation method is obtained, for which we proved convergence. We showed that fully learned L-SVD yields reconstructions with the same quality as a popular model-based approach. We gave a reconstruction error bound that depends on the autoencoder quality, which could provide insight in which signals can be reconstructed and which can not.

In this PhD research, it was found that the goal of improved photoacoustic tomography reconstructions is the product of many subtopics, all of which should be addressed. Below we elaborate on the main subtopics that could be addressed in future research. This is done via a general discussion on the topics of photoacoustic tomography and data-driven reconstruction.

General discussion and outlook

Many challenges in inverse problems are related to the presence of uncertainty. This uncertainty goes beyond the standard notions like additive or multiplicative noise: it also includes uncertainty in the forward model and in the type of signals that one tries to reconstruct.

Better characterisation and improvement of hardware already removes many sources of uncertainty. In PAT, electronic noise is reduced by using better ultrasound transducers or switching to a different way of detection [1]. Movement of tissue is avoided by fixating the tissue [2]. Light emitters and ultrasound detectors are characterised prior to a measurement and their locations are calibrated [3].

A source of uncertainty that currently can not be removed is in the prior knowledge of PAT reconstructions. For applications like breast imaging, the overall vascular anatomy of the breast is known, but the in-vivo appearance of vasculature associated with tumours is not completely clear [4, 5]. This gap of knowledge becomes critical as soon as regularisers for a variational problem are designed or training data is needed for data-driven reconstruction methods. In

Chapter 4 it was shown that good training data is of utmost importance. Without good knowledge of the pathophysiology and the manifestation of disordered vascularisation there is a serious danger of incorrect, but visually pleasing reconstructions.

Future efforts should lie in understanding the tissue under consideration. This can be done by investigating and segmenting MRI images [6] or PAT images themselves. A challenge in the latter is the fact that previous studies made use of various system geometries, types of detectors, wavelengths and reconstruction methods [7]. In-vivo images using MRI and PAT could be combined with histopathology, with the goal of creating digital and experimental phantoms for PAT. Recently the International Photoacoustic Standardisation Consortium (IPASC) was founded [8]. One of the goals is to reach consensus on photoacoustic imaging standardisation, which is beneficial for a general analysis of 3D PAT images. Another goal is to define the optical and acoustic properties for experimental phantoms that can be used to validate new and existing PAT systems. Being able to measure experimental phantoms in a PAT system is necessary for testing its hardware and software. However, the availability of good digital phantoms is equally important: the effectiveness of a new reconstruction method is first tested on simulated data from digital phantoms, before experimental or in-vivo measurements are considered. Moreover, a fully digital environment enables to investigate which system specifications, both in hardware and software, are necessary to solve the challenging problem of deep spectroscopic and quantitative imaging with PAT. Finally, for the development of new learned reconstruction methods, it is necessary to have access to a large and diverse training set. For these reasons, efforts should lie in creating a large open-source data bank for digital photoacoustic phantoms or creating open-source software for the automatic creation of such phantoms. In [6], three anatomically realistic breast phantoms were created from MRI images. Open-source code for the automatic creation of breast phantoms is available via [9]. This code could be extended to include breast tumours and their vascularisation.

Even with well-characterised improved hardware and with accurate knowledge about the tissue under consideration, there is often still some uncertainty left. One example in PAT is the sound speed map, which can be estimated from ultrasound transmission measurements [10]. Although it provides a clear improvement over a homogeneous sound speed, there is often room for improvement in resolution and accuracy. An explicit improvement of the sound speed map can be obtained by training an ANN [11]. The usage of inexact or inadequate forward mappings could give similar effects to using the wrong sound speed. For instance in quantitative PAT reconstruction, sometimes the diffusion approximation (DA) is considered instead of the radiative transfer equation (RTE), even though it is less accurate. This can be for computational reasons or because the DA is easier to implement than the RTE. It might be possible to explicitly improve the nonlinear forward mapping that models the DA with an ANN, similar to recent work for linear forward operators [12].

Another option is to treat the remaining inadequacy or uncertainty implicitly by classical methods using regularisation (Chapter 3) or reconstruction methods enriched by ANNs (Chapters 4, 5, 6), for which an overview was given in Section 2.3.4. All learned methods have their merits, but there is not a single predominant one in the field of deep learning for inverse problems. However, many recent efforts can be viewed as a way to perform manifold learning. In the remainder of this discussion and outlook, we take this viewpoint to consider the field of learned reconstruction methods.

An approach close to classical variational methods is that of training a regulariser before applying it in an iterative solution method. It is expected that such pre-trained regularisers encode prior knowledge better than hand-crafted ones, while the methods still benefit from classical regularisation theory. From the viewpoint of manifold learning: a regulariser provides small values for signals that are close to a manifold and large values for signals far away. Classical regularisers, as applied in Chapter 3, provide a generic manifold; other regularisers make use of explicit local data-dependent manifolds [13] of the desired signals. Regularisers that are learned in an ANN setup [14, 15] could implicitly define global data-dependent manifolds. In the first part of this thesis, the signal manifolds were found implicitly, through regularisation (Chapter 3) or a learned iterative scheme (Chapter 4). In the second part of this thesis explicit signal manifolds were used, by enforcing the ANN to operate in the zero-set of the forward mapping (Chapter 5) or by learning a reconstruction through an explicit low-dimensional manifold (Chapter 6).

Future efforts could lie in either explicitly forming the manifold or applying ANNs to learn implicit manifolds that can be analysed. Recent efforts towards this goal were made from several research directions. The inverse problem can be written in a Bayesian setting and solved by applying generative adversarial networks [16]. Graph neural networks [17] can be applied to learn a non-Euclidean nonlinear manifold, even if the respective inverse problem is generally solved in a Euclidean setting. Another promising research direction is that of nonlinear spectral analysis: ANNs can be viewed as a complex nonlinear operator and its nonlinear eigenvectors can be found [18].

The trend of finding solutions on highly data-dependent manifolds has clear advantages in terms of reconstruction quality. However, the ANNs that are used often possess too much freedom in their architecture and parameters to yield a manifold that is easy to understand and analyse. This creates the problem that the effect of regularisation on the reconstruction is not easily identified, something that is possible in classical methods. For instance, if one sees the staircasing behaviour of total variation, it is clear that this is an artefact created by the regulariser. With learned regularisation, the effects of the regulariser are much harder to recognise. This problem was observed in Chapter 4, where an imperfect training data set provides wrong reconstructions, even though the images themselves are realistic. A partial solution is to increase the training data set, which creates a smoother signal manifold. However, firstly this requires

good knowledge of what this training data set should look like. Secondly, we can never expect to create a training set that is large enough to embody all the anomalies that are expected in real medical imaging.

In the end there is a general question to be answered: do we wish to reconstruct all measurements reasonably well or do we wish to reconstruct some of the measurements very well, at the expense of being very bad in some cases? In other words: if a signal manifold is learned or chosen, how data-dependent should it be?

A future research direction that has not been explored often is that of examining in which cases which methods should be used. For instance, one might design a learned method that recognises when the input data is overly different from the training set. One might create a whole family of reconstruction methods, from which the best method is selected automatically.

Furthermore, future research could focus on the analysis of how an ANN influences the final reconstruction. This is relatively easy to see for a post-processing network, where a user can simply look at the input and output of the network. This is less evident in learned networks where the forward mapping and ANNs are intertwined, like in Chapter 4. It is these types of networks that are particularly effective for high-quality reconstruction. The field of nonlinear spectral analysis has made first steps in this direction [18].

Ultimately, it may be desirable to let the design of the learned reconstruction methods depend on the specific application. It might be possible to estimate the desired additional training data that is necessary to improve the results on a test set. Moreover, by combining training data with real-time reconstructions of incomplete measurements it might be possible to determine which additional measurements should be performed to reduce the uncertainty in the learned reconstruction. This requires knowledge of the application: if a learned method creates a PAT reconstruction that shows a tumour, more measurements could be focused on that specific spot, by changing the distribution of light and wavelength of the light to that specific area.

To summarise, this thesis explored various ways of combining ANNs with model-based reconstruction methods, each with its own drawbacks and advantages. It is not yet clear which combination provides a method that is both reliable and accurate for the field of PAT. Future research should focus on methods that are designed such that the effect of the ANN on the reconstruction can be analysed. This could be achieved by shaping or analysing the manifold on which desired solutions lie. Furthermore, efforts should lie in removing uncertainty in the PAT forward mapping, for which ANNs could play a role. Finally, better insight in the photoacoustic system and tissue under consideration is necessary. Only then good simulations can be set up to acquire training data, which is necessary for learned reconstruction methods to achieve deeper and higher quality photoacoustic imaging.

References

- [1] B. Dong, C. Sun, and H. F. Zhang, "Optical detection of ultrasound in photoacoustic imaging," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 1, pp. 4–15, 2016.
- [2] S. M. Schoustra, T. J. op 't Root, R. P. Pompe van Meerdervoort, L. Alink, W. Steenberg, and S. Manohar, "Pendant breast immobilization and positioning in photoacoustic tomographic imaging," *Photoacoustics*, in review.
- [3] Y. Wang, D. Xing, Y. Zeng, and Q. Chen, "Photoacoustic imaging with deconvolution algorithm," *Physics in Medicine & Biology*, vol. 49, no. 14, p. 3117, 2004.
- [4] G. L. Menezes, R. M. Pijnappel, C. Meeuwis, R. Bisschops, J. Veltman, P. T. Lavin, M. J. Van De Vijver, and R. M. Mann, "Downgrading of breast masses suspicious for cancer by using optoacoustic breast imaging," *Radiology*, vol. 288, no. 2, pp. 355–365, 2018.
- [5] G. L. Menezes, R. M. Mann, C. Meeuwis, B. Bisschops, J. Veltman, P. T. Lavin, M. J. van de Vijver, and R. M. Pijnappel, "Optoacoustic imaging of the breast: correlation with histopathology and histopathologic biomarkers," *European radiology*, vol. 29, no. 12, pp. 6728–6740, 2019.
- [6] Y. Lou, W. Zhou, T. P. Matthews, C. M. Appleton, and M. A. Anastasio, "Generation of anatomically realistic numerical phantoms for photoacoustic and ultrasonic breast imaging," *Journal of Biomedical Optics*, vol. 22, no. 4, p. 041015, 2017.
- [7] S. Manohar and M. Dantuma, "Current and future trends in photoacoustic breast imaging," *Photoacoustics*, 2019.
- [8] S. Bohndiek, "Addressing photoacoustics standards," *Nature Photonics*, vol. 13, no. 5, pp. 298–298, 2019.
- [9] C. G. Graff, "A new, open-source, multi-modality digital breast phantom," in *Medical Imaging 2016: Physics of Medical Imaging*, vol. 9783, p. 978309, International Society for Optics and Photonics, 2016.
- [10] C. Li, N. Duric, P. Littrup, and L. Huang, "In vivo breast sound-speed imaging with ultrasound tomography," *Ultrasound in medicine & biology*, vol. 35, no. 10, pp. 1615–1628, 2009.
- [11] E. M. A. Anas, A. Cheng, R. Seifabadi, Y. Wu, F. Aalamifar, B. Wood, A. Rahmim, and E. M. Boctor, "Cnn and back-projection: limited angle ultrasound tomography for speed of sound estimation," in *Medical Imaging 2019: Ultrasonic Imaging and Tomography*, vol. 10955, p. 109550M, International Society for Optics and Photonics, 2019.
- [12] S. Lunz, A. Hauptmann, T. Tarvainen, C.-B. Schönlieb, and S. Arridge, "On learned operator correction," *arXiv:2005.07069*, 2020.
- [13] G. Peyré, "Manifold models for signals and images," *Computer vision and image understanding*, vol. 113, no. 2, pp. 249–260, 2009.
- [14] H. Li, J. Schwab, S. Antholzer, and M. Haltmeier, "NETT: solving inverse problems with deep neural networks," *Inverse Problems*, 2020.
- [15] S. Lunz, O. Öktem, and C.-B. Schönlieb, "Adversarial regularizers in inverse problems," in *Advances in Neural Information Processing Systems*, pp. 8507–8516, 2018.

- [16] J. Adler and O. Öktem, "Deep bayesian inversion," *arXiv:1811.05910*, 2018.
- [17] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and S. Y. Philip, "A comprehensive survey on graph neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, 2020.
- [18] L. Bungert, E. Hait-Fraenkel, N. Papadakis, and G. Gilboa, "Nonlinear power method for computing eigenvectors of proximal operators and neural networks," *arXiv:2003.04595*, 2020.

Summary

Photoacoustic tomography (PAT) is an imaging technique with potential applications in various fields of biomedicine. By visualising vascular structures, PAT could help in the detection and diagnosis of diseases related to their dysregulation. An especially interesting application is breast imaging, since the uncontrolled creation of blood vessels is associated with breast cancer tumours.

In PAT, tissue is illuminated by monochromatic light. While travelling within the tissue, the light undergoes scattering and absorption. The absorbed light causes a local increase in temperature, which in turn causes an initial pressure to build up. This pressure then travels from the tissue to the surrounding detectors in the form of ultrasound. In short: by shining light on tissue, sound is created, which can be detected.

This thesis is concerned with the inverse problem of the described physical processes: what was the initial pressure in the tissue that gave rise to the detected pressure outside? The answer to this question is difficult to obtain when light penetration in tissue is not sufficient, the measurements are corrupted, or only a small number of detectors can be used in a limited geometry. For decades, the field of variational methods has come up with new approaches to tackle these kind of problems: the combination of new theory and clever algorithms has led to improved numerical results in many image reconstruction problems. In the past five years, previously state-of-the-art results were greatly surpassed by combining variational methods with artificial neural networks, a form of artificial intelligence. In this thesis we investigated several ways of combining data-driven artificial neural networks with model-based variational methods.

Variational methods often make use of regularisation, which suppresses the effect of noise and deals with limited measurements. An effective regulariser is generally chosen based on prior information on the images that one expects to reconstruct. For photoacoustic images showing vascular structures, it is not directly clear which regulariser is most effective. We developed a modular reconstruction framework, which provides a platform for comparing different regularisers. We showed that for PAT, both edge-preserving and directional

regularisers show improved results over methods not based on variational methods.

Minimising a regularised variational function results in an iterative procedure with predefined steps. By taking the structure of this iterative procedure and exchanging parts of it with a convolutional neural network, a partially learned reconstruction method is created. The training data set, which contains pairs of measurements and desired reconstructions, is important for the quality of the resulting algorithm. We investigated the properties that such a training set should have to obtain high-quality and robust reconstructions. Furthermore we showed that this partially learned method can be trained to not only provide a high-quality reconstruction, but also a segmentation of the vascular structure, without additional computation costs.

A different data-driven approach is to first create an imperfect reconstruction by a computationally cheap method, and then post-process it with an artificial neural network in order to improve it. We formalised this approach from a regularisation theory point of view to ensure data-consistency and convergence with respect to noise. Numerical results showed increased generalisation capacity of such data-consistent networks to data sets that are different from the training data set.

All previously mentioned methods need explicit knowledge about the physics of the forward process, including a numerical implementation. Sometimes this knowledge is not available, limited or the numerical evaluation is too computationally intensive. For such cases an alternative approach is to train a fully learned reconstruction method without using the mathematical forward model. We developed the learned singular value decomposition (L-SVD): an artificial neural network that connects low-dimensional nonlinear representations of measurements and reconstructions. We provided a reconstruction error estimate that depends on the quality of two connected autoencoders. The L-SVD has a flexible architecture, so that useful knowledge of the forward mapping could be incorporated. If this is done in a structured manner, convergence with respect to noise is obtained. For scenarios without this knowledge, we showed that L-SVD automatically learns elements of the forward mapping in a data-driven manner. For a tomography example with limited noisy measurements, L-SVD provided reconstructions that are of the same quality as a popular model-based reconstruction method.

In this thesis, several ways of combining model-based methods with data-driven artificial neural networks were investigated. The resulting hybrid methods showed improved tomography reconstructions. By allowing data to improve a structured method, deeper vascular structures could be imaged with photoacoustic tomography.

Samenvatting

Fotoakoestische tomografie (photoacoustic tomography, PAT) is een beeldvormingstechniek voor verschillende biomedische onderzoeksgebieden. Door vaatstructuren te visualiseren, zou PAT kunnen helpen bij het opsporen en diagnosticeren van ziekten die verband houden met hun ontregeling. Een bijzonder interessante toepassing is borstbeeldvorming, aangezien de ongecontroleerde aanmaak van bloedvaten geassocieerd wordt met borstkankertumoren.

In PAT wordt weefsel verlicht door monochromatisch licht. Terwijl het zich voortplant in het weefsel, ondergaat het licht verstrooiing en absorptie. Het geabsorbeerde licht veroorzaakt een lokale temperatuurstijging, wat op zijn beurt zorgt voor de opbouw van een initiële druk. Deze druk verplaatst zich dan als ultrasoon geluid van het weefsel naar de omliggende detectoren. Kortom: door licht op het weefsel te schijnen, ontstaat er geluid dat kan worden gedetecteerd.

Dit proefschrift behandelt het inverseprobleem van het zojuist beschreven fysische proces: wat was de initiële druk in het weefsel dat zorgde voor de gedetecteerde druk daarbuiten? Het antwoord op deze vraag is moeilijk te verkrijgen wanneer de lichtdoordringing in weefsel niet voldoende is, wanneer de metingen vervuild zijn of wanneer slechts een klein aantal detectoren gebruikt kan worden in een beperkte opstelling. Al decennia lang bedenkt het onderzoeksgebied van variationele methoden nieuwe manieren om dit soort problemen aan te pakken: de combinatie van nieuwe theorie en slimme algoritmes heeft geleid tot verbeterde numerieke resultaten bij vele beeldreconstructieproblemen. In de afgelopen vijf jaar werden voormalige state-of-the-art resultaten ruim overtroffen door het combineren van variationele methoden met kunstmatige neurale netwerken (KNN's), een vorm van kunstmatige intelligentie. In dit proefschrift hebben we verschillende manieren onderzocht om datagedreven KNN's te combineren met modelgestuurde variationele methoden.

Variationele methoden maken vaak gebruik van regularisatie, wat het effect van ruis onderdrukt en ervoor zorgt dat er met beperkte metingen omgegaan kan worden. Een effectieve regularisator wordt meestal gekozen op basis van het

type afbeelding dat men wil reconstrueren. Voor fotoakoestische afbeeldingen met vaatstructuren is het niet direct duidelijk welke regularisator het meest effectief is. Wij hebben een modulair reconstructiekader ontworpen, dat een platform biedt voor het vergelijken van verschillende regularisators. We hebben aangetoond dat voor PAT zowel randbehoudende als directionele regularisators verbeterde resultaten geven ten opzichte van niet-variationele methoden.

Het minimaliseren van een geregulariseerde variationele functie resulteert in een iteratief schema met vastliggende stappen. Door de structuur van dit iteratieve schema te nemen en bepaalde delen ervan te vervangen door een convolutioneel neurale netwerk, ontstaat een gedeeltelijk geleerde reconstructiemethode. De trainingsdataset, die paren van metingen en gewenste reconstructies bevat, is belangrijk voor de kwaliteit van het geleerde algoritme. Wij hebben onderzocht welke eigenschappen een dergelijke trainingsset moet hebben om robuuste reconstructies te krijgen van een hoge kwaliteit. Bovendien hebben we aangetoond dat deze gedeeltelijk geleerde methode getraind kan worden om niet alleen een hoogwaardige reconstructie te geven, maar ook een segmentatie van de vaatstructuur, zonder extra rekentijd.

Een andere datagedreven aanpak is om eerst een onvolmaakte reconstructie te maken met een snelle methode en deze vervolgens te verbeteren met een KNN. Wij hebben deze aanpak vanuit de regularisatietheorie geformaliseerd om dataconsistentie en convergentie met betrekking tot ruis te garanderen. Numerieke resultaten toonden aan dat de prestaties van zulke dataconsistente netwerken beter generaliseren tot datasets die er anders uitzien dan de trainingsset.

Alle hiervoor genoemde methoden hebben expliciete kennis nodig over het fysische proces, inclusief een numerieke implementatie. Soms is deze kennis niet beschikbaar, beperkt, of kost de numerieke evaluatie te veel rekentijd. Voor zulke gevallen bestaat er de alternatieve aanpak om een volledig geleerde reconstructiemethode te trainen, zonder gebruik te maken van het wiskundige voorwaartse model. Wij ontwikkelden de geleerde singulierewaardenontbinding (learned singular value decomposition, L-SVD): een KNN dat laagdimensionale representaties van metingen en reconstructies met elkaar verbindt. Voor de reconstructie hebben we een foutafschatting gegeven die afhangt van de twee verbonden autoencoders. De L-SVD heeft een flexibele architectuur die toestaat dat kennis van het voorwaartse proces ook meegenomen kan worden. We hebben echter aangetoond dat L-SVD ook zonder deze kennis elementen van dit proces op een datagedreven manier leert. In een voorbeeld met een beperkt aantal ruizige tomografiemetingen is aangetoond dat L-SVD dezelfde reconstructiekwaliteit levert als een populaire modelgestuurde methode.

In dit proefschrift zijn verschillende manieren onderzocht om modelgestuurde methoden te combineren met datagedreven KNN's. De hybride methoden die hieruit voortkwamen lieten verbeterde tomografiereconstructies zien. Door toe te staan dat data een gestructureerde methode verbetert, kunnen diepere vaatstructuren zichtbaar worden gemaakt met fotoakoestische tomografie.

List of publications

Journal articles

1. **A framework for directional and higher-order reconstruction in photoacoustic tomography**
Yoeri E. Boink, Marinus J. Lagerwerf, Wiendelt Steenbergen, Stephan A. van Gils, Srirang Manohar, Christoph Brune
Physics in Medicine & Biology 63.4 (2018): 045018
2. **A partially learned algorithm for joint photoacoustic reconstruction and segmentation**
Yoeri E. Boink, Srirang Manohar, Christoph Brune
IEEE transactions on medical imaging 39.1 (2019): 129-139
3. **Deep learning of circulating tumour cells**
Leonie L. Zeune, Yoeri E. Boink, Guus van Dalum, Afroditi Nanou, Sanne de Wit, Kiki C. Andree, Joost F. Swennenhuis, Stephan A. van Gils, Leon W.M.M. Terstappen, Christoph Brune
Nature Machine Intelligence 2.2 (2020): 124-133
4. **Tomographic imaging with an ultrasound and LED-based photoacoustic system**
Kalloor Joseph Francis, Yoeri E. Boink, Maura Dantuma, Mithun Kuniyil Ajith Singh, Srirang Manohar, Wiendelt Steenbergen
Biomedical Optics Express 11.4 (2020): 2152-2165
5. **Mechanisms underlying vascular endothelial growth factor receptor inhibition-induced hypertension – the HYPАЗ trial**
Kaisa Mäki-Petäjä, Adam McGeoch, Lucy Yang, Annette Hubsch, Carmel McEniery, Paul Meyer, Fraz Mir, Parag Gajendragadkar, Nicola Ramenatte, Gayathri Anandappa, Sara Santos-Franco, Simon Bond, Carola-Bibiane Schönlieb, Yoeri E. Boink, Christoph Brune, Ian Wilkinson, Duncan Jodrell, Joseph Cheriyan
Hypertension, in review (2020)
6. **Data-consistent neural networks for solving nonlinear inverse problems**
Yoeri E. Boink, Markus Haltmeier, Sean Holman, Johannes Schwab
SIAM Journal on Imaging Sciences, submitted (2020), preprint available: *arXiv* 2003.11253

7. **Learned SVD: solving inverse problems via hybrid autoencoding**
Yoeri E. Boink, Christoph Brune
Inverse Problems and Imaging, submitted (2020), preprint available: arXiv 1912.10840
8. **Clinical handheld laser speckle contrast perfusion imaging**
 Ata Chizari, Mirjam Schaap, Tom Knop, Yoeri E. Boink, Marieke Seyger, Wiendelt Steenbergen
Biomedical Optics Express, to be submitted (2021)

Book chapters

1. **Light Emitting Diodes Based Photoacoustic and Ultrasound Tomography: Imaging Aspects and Applications**
 Kalloor Joseph Francis, Yoeri E. Boink, Maura Dantuma, Mithun Kuniyil Ajith Singh, Srirang Manohar, Wiendelt Steenbergen
LED-Based Photoacoustic Imaging, Springer (2020): 245-266

Conference proceedings

1. **Sensitivity of a partially learned model-based reconstruction algorithm**
Yoeri E. Boink, Stephan A. van Gils, Srirang Manohar, Christoph Brune
Proceedings in Applied Mathematics and Mechanics 18.1 (2018): e201800222
2. **Robustness of a partially learned photoacoustic reconstruction algorithm**
Yoeri E. Boink, Christoph Brune, Srirang Manohar
Proceedings SPIE, Photons Plus Ultrasound: Imaging and Sensing 10878 (2019): 108781D

Conference abstracts

1. **Mechanisms of vascular endothelial growth factor inhibition induced hypertension**
 Kaisa Maki-Petaja, Adam McGeoch, Lucy Yang, Annette Hubsch, Carmel McEniery, Fraz Mir, Parag Gajendragadkar, Nicola Ramenatte, Gayathri Anandappa, Christoph Brune, Yoeri E. Boink, Carola Bibiane-Schönlieb, Paul Meyer, Simon Bond, Ian Wilkinson, Duncan Jodrell, Joseph Cheriyan
Abstract from Conference Artery Research (2018): 87-88

Conference presentations

1. **Combined modelling of optimal transport and segmentation revealing vascular properties**
Yoeri E. Boink, Christoph Brune
Workshop on Optimal Transport and Optimization in Imaging, October 2016, Paris, France
2. **Regularisation techniques in photoacoustic tomography enhancing vascular structures**
Yoeri E. Boink, Srirang Manohar, Wiendelt Steenbergen, Stephan A. van Gils, Christoph Brune
SIAM Conference on Imaging Science, June 2018, Bologna, Italy

3. **Regularisation and parameter learning for photoacoustic tomography**
Yoeri E. Boink, Srirang Manohar, Wiendelt Steenbergen, Stephan A. van Gils, Christoph Brune
IFIP Conference on System Modelling and Optimization, July 2018, Essen, Germany
4. **Joint photoacoustic reconstruction and segmentation with a partially learned algorithm**
Yoeri E. Boink, Srirang Manohar, Christoph Brune
Workshop on Deep Learning and Inverse Problems, January 2019, Stockholm, Sweden
5. **A partially learned algorithm for joint photoacoustic reconstruction and segmentation**
Yoeri E. Boink, Christoph Brune, Srirang Manohar
SPIE Photonics West, Photons Plus Ultrasound: Imaging and Sensing, February 2019, San Francisco, United States of America (Nominated for best paper award)
6. **Solving inverse problems with data-driven model-based methods**
Yoeri E. Boink, Srirang Manohar, Christoph Brune
NDNS+ PhD Days, May 2019, Lunteren, The Netherlands
7. **Learned SVD: nonlinear image reconstruction via hybrid autoencoding**
Yoeri E. Boink, Christoph Brune
SIAM Conference on Mathematics of Data Science (online), May 2020, Cincinnati, United States of America
8. **Learned SVD: nonlinear image reconstruction via hybrid autoencoding**
Yoeri E. Boink, Christoph Brune
SIAM Conference on Imaging Science (online), July 2020, Toronto, Canada

Conference posters

1. **Higher-order total variation for photoacoustic image reconstruction**
Yoeri E. Boink, Marinus J. Lagerwerf, Srirang Manohar, Christoph Brune
International Graduate Summer School on Biophotonics, June 2017, Ven, Sweden
2. **Directional and higher-order regularisation for photoacoustic image reconstruction**
Yoeri E. Boink, Marinus J. Lagerwerf, Srirang Manohar, Christoph Brune
Isaac Newton Institute Workshop: Generative Models, parameter learning and sparsity, November 2017, Cambridge, United Kingdom
3. **Sensitivity of a partially learned model-based reconstruction algorithm**
Yoeri E. Boink, Stephan A. van Gils, Srirang Manohar, Christoph Brune
Lorentz workshop: Image Reconstruction from Millimetres to the Globe: Bridging the Gap Between Seismic and Medical Imaging, June 2018, Leiden, The Netherlands
4. **Sensitivity of a partially learned model-based reconstruction algorithm**
Yoeri E. Boink, Srirang Manohar, Christoph Brune
CIRM conference: Mathematical and Numerical Approaches for Multi-Wave Inverse Problems, April 2019, Marseille, France
5. **Learned SVD: solving inverse problems via hybrid autoencoding**
Yoeri E. Boink, Christoph Brune
Autumn school: Deep Learning and Inverse Problems, November 2019, Bremen, Germany

Acknowledgements

The past four years at the University of Twente have been memorable. In these pages I will mention my colleagues from whom I learned so much, my friends who made my time as a PhD student so enjoyable.

Christoph and Srirang, thank you for your guidance and supervision. I thoroughly enjoyed our conversations and discussions, where we often taught each other mathematics, photoacoustics or deep learning. I am grateful for all the kind words and encouragement, but most of all for the freedom and trust that you gave me to advance the research in my own directions. Both of you established your own research group in the past four years: I hope that despite your additional tasks you will find the time to guide future PhD students and do the research that you both love so much.

Stephan and Wiendelt, because of you my Wednesday morning routine generally consisted of coffee, a group meeting, more coffee (with cake), espresso (with cake), followed by a second group meeting, rounding off with lunch. The combination of a welcoming atmosphere and interesting topics in mathematics and biomedical photonics made these mornings quite enjoyable. Thank you.

Markus and Johannes, thanks for a great stay in Innsbruck in the Autumn of 2019. I enjoyed the long discussions in Markus' office a lot. I learned much about inverse problems, and also quite a bit about Tyrolean culture, Austrian politics and a mathematical sense of humour.

Although my research was mostly academic, I felt its societal impact via the joint project with PA-Imaging. Tim, it was a great pleasure to have bi-weekly discussions on physics, mathematics and software implementation, where we established many creative solutions. You helped me to enrich my academic problems by addressing the real-world issues that you experienced in the PAM2 system.

It was interesting to work in three research groups, all with a different group dynamic and atmosphere. I thank my AA colleagues for the relaxed espresso group meetings and the SACS cluster for the diverse presentations on the mathematical topics researched on our floor. The board game evenings with

my SACS colleagues were a lot of fun. I thank the M3I group for expanding my knowledge in the medical area, and for the fun drinks and BBQs. Last, but definitely not least, I thank the BMPI group: I cannot imagine there is a more active group at the UT. There are too many sushi, pubquiz, movie, karaoke and other drinking nights to even start mentioning them.

A special thanks goes out to my 'roadtrip buddies' Colin, David, Gerwin, Maura and Sjoukje and my 'citytrip buddy' Manu. Besides going on road trips in Sweden and California and planning two city trips that never came, I enjoyed the times at work where you taught me a lot about physics, biomedical optics, biomedicine and mathematics.

I am glad my office mates Manu and David were always there to help me keep up the amount of complaining about students, supervisors, the university or basically anything that was not perfect at some time. A big thanks to Maura for putting up with this complaining, but also for doing most of the lab work for me (and letting me confiscate her GPU).

Bram and Pascal, it was a pleasure to supervise you during your final MSc project. You dove in the topics of photoacoustics, neural networks and inverse problems with great enthusiasm and helped me form my own research questions.

Finally, I am thankful for the support that I received from all my friends and family. I appreciate that you were so interested in my research, but more importantly, I am very happy with the distractions that created a nice work-life balance. Thanks for all the weekends and evenings filled with food, coffee, beer, parties, music and board games!

My biggest gratitude goes out to Evelien, who is always there for me. Thank you so much for listening to all my struggles and (double) petty problems. In the past four years, you have helped me much more than you know. I am very happy to be with you!