Computational Compensation for Model Imperfections in Photoacoustic Computed Tomography

Thesis by
Peng Hu

In Partial Fulfillment of the Requirements for
the Degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2023
(Defended May 19, 2023)

Peng Hu
ORCID: 0000-0002-2933-1239

For my wife Lizhen, my daughter Jocelyn, and my parents.

ACKNOWLEDGEMENTS

I am grateful for the invaluable guidance and mentorship provided by my advisor, Prof. Lihong Wang, throughout the past few years. Under his supervision, I have been privileged to work in a lab where I have been encouraged and inspired to tackle meaningful and complex problems in the field of medical imaging. Additionally, I would like to extend my appreciation to my dissertation committee members, Profs. Changhuei Yang, Katherine L. Bouman, and Timothy E. Colonius, for their kind support of my research. Furthermore, I would like to thank Profs. Methodius G. Tuuli and Molly J. Stout for their guidance and support in my clinical research on preterm birth prediction.

I am also grateful for the generous support I have received from my colleagues throughout my research journey. I extend my thanks to Drs. Li Lin, Rui Cao, Yang Zhang, and Yide Zhang for their invaluable contributions to my studies. Their critical efforts have greatly enhanced the quality of my research. I would also like to express my appreciation to Dr. Lei Li for his support during the initial stages of my research. Furthermore, I extend my gratitude to Dr. Peinan Zhao, Dr. Yuan Qu, Xin Tong, and Yilin Luo for their invaluable assistance in my studies. I would like to acknowledge and thank Dr. Junjie Yao, Dr. Junhui Shi, Dr. Liren Zhu, Dr. Terence T. W. Wong, Dr. Xiaohua Feng, Dr. Shuai Na, David Garrett, Karteek Dhara, and Manxiu Cui for their inspiring discussions. Lastly, I want to express my sincere appreciation to Dr. Konstantin Maslov, Jessica Chubiz, Dr. Geng Ku, and Catherine L. Pichotta for their support and assistance in the lab.

I would like to express my heartfelt gratitude to my family for their unwavering love and unconditional support. They have been a constant source of strength and comfort in my life. I would like to express my special gratitude to Lizhen Chen. Dear Zhenzhen, since the moment I met you, you have illuminated my life like the sun, moon, and stars. Your support and influence have been immeasurable, especially during the most challenging moments of my life. I want to convey my deepest appreciation for everything you have done for me. Being with you has not only brought me joy and happiness but has also transformed me into a better person. May this dissertation serve as a testament to our enduring journey together.

ABSTRACT

Photoacoustic computed tomography (PACT) images biological tissues' optical absorption through detection of photon-absorption-induced ultrasonic waves. Various systems have been proposed for PACT and they are described by different mathematical models to reconstruct from detected ultrasonic signals the photon-absorption-induced initial pressure, the main contrast in PACT. Accurate image reconstruction has high requirements for the system and the mathematical model, which is often imperfect in practice due to multiple factors, e.g., limited transducer bandwidth, finite transducer element size, sparse spatial sampling, partial-view detection, and tissue motion. The focus of this dissertation is on using computational methods to compensate for these model imperfections.

First, for a human breast imaging system based on a full-ring transducer array, we incorporate the limited transducer bandwidth into the model for spatiotemporal analysis to clarify the aliasing due to sparse spatial sampling and propose (1) two methods (radius-dependent spatiotemporal antialiasing and location-dependent spatiotemporal antialiasing) to mitigate these artifacts. Second, for an isotropic-resolution 3D PACT system formed by four arc arrays, we consider both the limited transducer bandwidth and the finite transducer element size and (2) compress the system matrix through singular value decomposition and fast Fourier transform for its efficient explicit expression. Enabled by this expression, we then propose (3) fast sparsely sampling functional imaging by incorporating a densely sampled prior image into the system matrix, which maintains the critical linearity while mitigating artifacts, and (4) intra-image nonrigid motion correction by incorporating the motion as subdomain translations into the system matrix and reconstructing the translations together with the image iteratively. Finally, for a single-shot 3D PACT system based on a single ultrasonic transducer, we propose (5) a fast implementation of the forward model by connecting traditional PACT with virtual detector responses through fast Fourier transform, and we iteratively reconstruct the image from signals with extremely compressed sensing and partial-view detection.

All these proposed methods enable image reconstruction or significantly improve image quality in numerical simulations, phantom experiments, and *in vivo* experiments. Although they are demonstrated only for certain PACT systems, they are directly applicable to other systems and can be extended to other tomographic imaging modalities such as X-ray computed tomography (X-ray CT) and magnetic resonance imaging (MRI).

PUBLISHED CONTENT AND CONTRIBUTIONS

1. **P. Hu**, X. Tong, L. Lin, and L. V. Wang, "Data-driven system matrix manipulation enabling fast functional imaging and intra-image nonrigid motion correction in tomography," *Submitted*.

Peng Hu conceived and designed the study, implemented the image reconstruction and processing algorithms, analyzed the data, and wrote the manuscript.

2. Y. Zhang†, **P. Hu**†, L. Li, R. Cao, A. Khadria, K. Maslov, X. Tong, Y. Zeng, L. Jiang, Q. Zhou, and L. V. Wang, "Single-shot 3D photoacoustic tomography using a single-element detector for ultrafast imaging of hemodynamics." bioRxiv, p. 2023.03.14.532661, 15-Mar-2023. Under review by *Nat. Biomed. Eng.* DOI: 10.1101/2023.03.14.532661

Peng Hu developed the 3D reconstruction algorithm, processed and analyzed the data, and wrote the manuscript.

3. **P. Hu**†, P. Zhao†, Y. Qu, K. Maslov, J. Chubiz, M. G. Tuuli, M. J. Stout, and L. V. Wang, "Quantification of cervical elasticity during pregnancy based on transvaginal ultrasound imaging and stress measurement." arXiv, 09-Mar-2023. DOI: 10.48550/arXiv.2303.05697

Peng Hu conceived and designed the study, implemented the system and image and signal processing algorithms, analyzed the data, and wrote the manuscript.

4. **P. Hu**, L. Li, and L. V. Wang, "Location-dependent spatiotemporal antialiasing in photoacoustic computed tomography," *IEEE Trans. Med. Imaging*, vol. 42, no. 4, pp. 1210–1224, Apr. 2023. DOI: 10.1109/TMI.2022.3225565

Peng Hu conceived and designed the study, implemented the image reconstruction and processing algorithms, analyzed the data, and wrote the manuscript.

5. L. Lin†, **P. Hu**†, X. Tong†, S. Na†, R. Cao, X. Yuan, D. C. Garrett, J. Shi, K. Maslov, and L. V. Wang, "High-speed three-dimensional photoacoustic computed tomography for preclinical research and clinical translation," *Nat. Commun.*, vol. 12, no. 1, p. 882, Feb. 2021. DOI: 10.1038/s41467-021-21232-1

Peng Hu developed the reconstruction algorithm, analyzed the data, and wrote the manuscript.

6. **P. Hu**, L. Li, L. Lin, and L. V. Wang, "Spatiotemporal antialiasing in photoacoustic computed tomography," *IEEE Trans. Med. Imaging*, vol. 39, no. 11, pp. 3535–3547, Nov. 2020. DOI: 10.1109/TMI.2020.2998509

Peng Hu conceived and designed the study, implemented the image reconstruction and processing algorithms, analyzed the data, and wrote the manuscript.

7. 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," *Nat. Commun.*, vol. 9, no. 1, p. 2352, Jun. 2018. DOI: 10.1038/s41467-018-04576-z

Peng Hu developed the software system and the reconstruction algorithm, analyzed the data, and wrote the manuscript.

†Equal contributions.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iv
ABSTRACT v
PUBLISHED CONTENT AND CONTRIBUTIONS vii
TABLE OF CONTENTS ix
LIST OF ILLUSTRATIONS xi
LIST OF TABLES xix
ABBREVIATIONS xx
Chapter 1 PACT AND ITS MODEL IMPERFECTIONS 1
Chapter 2 SPATIOTEMPORAL ANTIALIASING 5
2.1 Abstract 5
2.2 Background 6
2.2.1 Human breast imaging for tumor detection 6
2.2.2 Spatiotemporal antialiasing 9
2.3 Radius-dependent spatiotemporal antialiasing 11
2.3.1 Theory 11
2.3.2 Numerical simulations 19
2.3.3 *In vivo* experiments 22
2.4 Location-dependent spatiotemporal antialiasing 24
2.4.1 Theory 24
2.4.2 Numerical simulations 35
2.4.3 *In vivo* experiments 38
2.5 Discussion 39
Chapter 3 EFFICIENT EXPLICIT SYSTEM MATRIX EXPRESSION 41
3.1 Abstract 41
3.2 Background 42
3.2.1 An isotropic-resolution 3D PACT system 42
3.2.2 The need for an efficient explicit system matrix expression 44
3.3 Theory 45
3.3.1 The fast forward operator 45
3.3.2 Point source response measurement 51
3.3.3 Discretization of the forward operator 53
3.3.4 Iterative reconstruction 55
3.4 Numerical simulations 56
3.4.1 Efficiency and accuracy 56
3.4.2 Regularized iterative reconstruction 60
3.5 Discussion 64
Chapter 4 FAST FUNCTIONAL IMAGING WITH SPARSE SAMPLING 66
4.1 Abstract 66
4.2 Background 66
4.3 Theory 67
4.3.1 A hybrid method for fast functional imaging 67
4.3.2 Functional signal extraction 69

4.4 Numerical simulations 71
4.4.1 Antialiasing and linearity 71
4.4.2 Functional imaging 73
4.5 *In vivo* experiments 76
4.6 Discussion 81
Chapter 5 INTRA-IMAGE NONRIGID MOTION CORRECTION 82
5.1 Abstract 82
5.2 Background 82
5.3 Theory 83
5.4 Numerical simulations 88
5.5 *In vivo* experiments 93
5.6 Discussion 94
Chapter 6 SINGLE-ELEMENT SINGLE-SHOT 3D PACT 96
6.1 Abstract 96
6.2 Background 96
6.3 Theory 100
6.3.1 Forward operator 100
6.3.2 Image reconstruction 105
6.4 Phantom experiments 105
6.5 *In vivo* experiments 108
6.6 Discussion 112
Chapter 7 DISCUSSION 114
BIBLIOGRAPHY 115
Appendix A: Mathematical symbols 125

LIST OF ILLUSTRATIONS

Number *Page*

Figure 1 Representations of the SBH-PACT system. **a** Perspective cut-away view of the system with data acquisition components removed. **b** Perspective view of the system with patient bed and optical components removed. DAQ, data acquisition system; Pre-amp, pre-amplifier circuits. 7

Figure 2 PACT images and X-ray mammograms of a cancerous human breast. **a** A cross section of the PACT image of a cancerous human breast with the tumor circled by a green-dashed ellipse. PA, photoacoustic. **b** A closeup of the subset enclosed by a red-dotted box. **c1**-**c2** X-ray mammograms of the breast. RCC, right cranial-caudal; RML, right medio-lateral. **d** MAP of the PACT image along the elevational direction that is color-encoded by the elevational distance from the nipple. **e** MAP of the thick slice in the sagittal plane marked by white dashed lines in **d**. **f** Automatic tumor detection on a vessel density map. The tumor is indicated by a green-dashed ellipse. The gray-scale background is the MAP of the tissue image elevationally separated from the nipple. 8

Figure 3 Analysis of the spatial aliasing for a circular geometry. **a** A full-ring transducer array of radius R (red circle), where a detection element location $\mathbf{r}$ and a source point location $\mathbf{r}'$ are marked. The locations $\mathbf{r}$ and $\mathbf{r}'$ are also seen as vectors from the origin to these locations. Vectors $-\mathbf{r}'$ and $\mathbf{r}-\mathbf{r}'$ form an angle β, while the extension of line segment $\mathbf{r}-\mathbf{r}'$ forms an angle α' with the tangential dotted line that is perpendicular to $\mathbf{r}$. The angle formed by vectors $-\mathbf{r}$ and $\mathbf{r}'-\mathbf{r}$ can be expressed as $\alpha'-\frac{\pi}{2}$. This graph is used to analyze the aliasing in SS. **b** A full-ring transducer array with a detection element location $\mathbf{r}$, two reconstruction locations $\mathbf{r}''$ and $\mathbf{r}_1''$, and a source point location $\mathbf{r}'$ marked. Extensions of the line segments $\mathbf{r}-\mathbf{r}''$, $\mathbf{r}-\mathbf{r}_1''$, and $\mathbf{r}-\mathbf{r}'$ form angles α'', α_1'', and α', respectively, with the tangential dotted line that is perpendicular to $\mathbf{r}$. Vectors $\mathbf{r}''$ and $\mathbf{r}'$ form an angle $\gamma=\arccos\left(\frac{r''}{R}\right)+\arccos\left(\frac{r'}{R}\right)$, where $r''=\|\mathbf{r}''\|$ and $r'=\|\mathbf{r}'\|$. Points $\mathbf{r}''$ and $\mathbf{r}_1''$ are on the same circle centered at $\mathbf{r}$. This graph is used to analyze aliasing in IR. **c** Regions in the field of view representing different types of aliasing. In S_2 (green circle), which contains all source points and reconstruction locations, UBP reconstruction yields no aliasing. In S_1 (blue circle), aliasing does not exist in SS but may exist in UBP reconstruction. In S_0 (red circle), aliasing may exist in SS. 13

Figure 4 Different combinations of source locations and reconstruction locations subject to spatial aliasing in SS and IR. Three regions S_0, S_1, and S_2 are defined in Eqs. (8), (16), and (22), respectively. The first line radiating from the origin O represents the range of source locations for SS, while the second line radiating from the tip of the first line represents the range of reconstruction locations for IR. A solid line means no aliasing, while a dotted line means aliasing. **a** Spatial aliasing in UBP. The innermost two-way Nyquist zone S_2 is an aliasing-free region. **b** Spatial aliasing in

UBP with spatial interpolation. Spatial interpolation removes spatial aliasing in three cases of IR, making the one-way Nyquist zone S_1 an aliasing-free region. The dotted lines representing the three cases in **a** are changed to blue-solid lines in **b**. **c** Spatial aliasing in UBP with temporal filtering and spatial interpolation. Temporal filtering extends the one-way Nyquist zone S_1 in **b** to S_1' in **c**, and the original S_1 is marked as a blue-dashed circle for reference. Spatial interpolation further makes S_1' an aliasing-free region. 17

Figure 5 Spatial interpolation and temporal filtering's effects on IR in numerical simulations for two complex phantoms. **a1** Ground truth of a complex initial pressure p_0 distribution confined to S_1. **a2**–**a4** Reconstructions of the object in **a1** using **a2** UBP, **a3** UBP with SI, and **a4** UBP with TF and SI, respectively. SI, spatial interpolation; TF, temporal filtering. The artifacts in the red-boxed region are caused by spatial aliasing in IR, and they are mainly mitigated by SI. **b1** Comparison of the STDs in the ROIs A–C. **b2** and **b3** Comparisons of the profiles of lines P and Q, respectively, for the three methods. **c1** Ground truth of a complex initial pressure p_0 distribution beyond S_1. **c2**–**c4** Reconstructions of the object in **c1** using **c2** UBP, **c3** UBP with SI, and **c4** UBP with TF and SI, respectively. The artifacts in the red-boxed region are caused by spatial aliasing in SS and IR, and the artifacts are mitigated by TF and SI. **d1** Comparison of the STDs in the ROIs A–C. **d2** and **d3** Comparisons of the profiles of lines P and Q, respectively, for the three methods. The FWHM of the main lobe at Q was increased from 0.35 mm to 0.48 mm by temporal filtering, while the amplitude was changed from 0.90 to 0.56. 21

Figure 6 Spatial interpolation and temporal filtering's effects on IR in an *in vivo* human breast image. **a** Reconstructed image using UBP without either spatial interpolation or temporal filtering, and a closeup subset in the yellow-boxed region. Boundaries of S_1 and S_2 are shown as white-dashed and white-solid circles, respectively. **b** and **c** Reconstructions of the same region as **a** using **b** UBP with SI, and **c** UBP with TF and SI, respectively. SI, spatial interpolation; TF, temporal filtering. **d** and **e** Comparisons of the profiles of lines P and Q, respectively, for the three methods. The FWHM of the main lobe at Q was increased from 0.44 mm to 0.55 mm by temporal filtering, while the amplitude was changed from 0.64 to 0.38. 23

Figure 7 Location-dependent spatiotemporal analysis. **a** No point sources outside the subdomain. A full-ring transducer array (red circle), an image subdomain (rectangle with gray interior and blue boundary, denoted as D_{sub}, centered at $\mathbf{r}'_{\mathrm{c,sub}}$, and of size $l_x \times l_y$), two adjacent element locations $\mathbf{r}$ and $\mathbf{r}_{\mathrm{adj}}$, a source point $\mathbf{r}'$ inside D_{sub}, and a reconstruction location $\mathbf{r}''$ inside D_{sub}. There exists a hyperbola with $\mathbf{r}$ and $\mathbf{r}_{\mathrm{adj}}$ as the foci, and with one branch (blue dotted curve) crossing $\mathbf{r}'$. One of the branch's intersection points with the boundary of D_{sub} is denoted as $\hat{\mathbf{r}}'$. This graph is used in the spatiotemporal analysis for D_{sub} without point sources outside. **b** One or multiple point sources outside the subdomain. A full-ring transducer array, an image subdomain D_{sub} centered at $\mathbf{r}'_{\mathrm{c,sub}}$, two adjacent element locations $\mathbf{r}$ and $\mathbf{r}_{\mathrm{adj}}$, and a reconstruction location $\mathbf{r}''$ inside D_{sub}. We have a single source point ($\mathbf{r}'$) or multiple source points ($\mathbf{r}'_1, \mathbf{r}'_2, \mathbf{r}'_3$, ...) outside D_{sub}. This graph is used in the spatiotemporal analysis for D_{sub} with a single or multiple point sources outside. **c** Multiple point

sources outside multiple subdomains. A full-ring transducer array, two image subdomains D_1 and D_2 (centered at $\mathbf{r}'_{c,1}$ and $\mathbf{r}'_{c,2}$, respectively), and a group of source points: $\mathbf{r}'_1, \mathbf{r}'_2, \mathbf{r}'_3, \ldots$ 28
Figure 8 Workflow of the location-dependent spatiotemporal antialiasing for PACT. 34
Figure 9 Applying spatiotemporal antialiasing to the image reconstruction of a numerical phantom with complex blood vessel structures. **a1**–**c1** A numerical phantom consisting of complex blood vessel structures, and its closeup subsets. **a2**–**c2**, **a3**–**c3**, and **a4**–**c4** Images of the complex numerical phantom reconstructed using UBP, UBP with RDTF and spatial interpolation (SI), and UBP with LDTF and SI, respectively, and their closeup subsets. **d** Comparisons of the values along the line L marked in **b2** for the three methods. The FWHM of the dominant lobe is 0.93 mm for UBP with RDTF and SI, and 0.73 mm for UBP with LDTF and SI. The amplitudes are 1.15 and 1.31, respectively. **e** Comparisons of the STDs of pixel values in regions A and B marked in **b2** and **c2**, respectively, for the three methods. 36
Figure 10 Applying spatiotemporal antialiasing to human breast imaging *in vivo*. **a1**–**a3** Images of a human breast cross section *in vivo* reconstructed using **a1** UBP, **a2** UBP with RDTF and spatial interpolation (SI), and **a3** UBP with LDTF and SI. Two subdomains in the red-dashed box and the yellow-dashed box, respectively, are picked for comparisons of the three methods. **b1**-**c1**, **b2**-**c2**, and **b3**-**c3** Closeup images of the two subdomains for the three methods, respectively. Lines L1 and L2 are picked for comparisons. **d** and **e** Values along lines L1 and L2, respectively, for the three methods. 38
Figure 11 Representations and photographs of the 3D-PACT system and the leaf skeleton images. **a** Perspective view of the system. DAQ, data acquisition system. **b** Cut-away view of the system with the imaging platform removed. **c** Photograph of the system with a close-up view of the ultrasonic arrays. **d** Photograph and maximum amplitude projection (MAP) images of a leaf skeleton. PA, photoacoustic. 44
Figure 12 Compression of the forward operator based on SVD. **a** A point source (a red dot), the n-th transducer element (a black rectangular centered at $\mathbf{r}_n$), and the element's local coordinate system with axes $\hat{\mathbf{x}}_n$, $\hat{\mathbf{y}}_n$, and $\hat{\mathbf{z}}_n$. **b** Four point sources A, B, C, and D (red dots) in the local coordinate system of the n-th transducer element. Points A, D, and $\mathbf{r}_n$ are on the same line. **c** The responses of the transducer element to the signals from the four point sources and the PCC between every two responses. **d** Expression of the four responses using linear combinations (coefficients visualized with bars) of three temporal singular functions shown as red, gray, and blue curves, respectively, based on SVD. **e**-**f** Independent responses (**e**) of a transducer element to 50 point sources with decreasing distances to the element, and the temporally-shifted form (**f**), which aligns the nonzero signals in time. PA, photoacoustic. **g** White-noise responses of the same size as those in **e** and **f**. **h**-**i** Normalized singular values and proportions of the variances unexplained, respectively, in the SVDs of the signals in **e**, **f**, and **g**. 49
Figure 13 Accuracy of the fast operator in forward simulation. **a** A numerical phantom formed by three rectangular cuboids (each of size $2.6 \times 2.6 \times 10$ mm^3) intersecting at their centers. Voxel values are 1 inside the phantom and 0 outside. Three lines (L1,

L2, and L3) are picked for image-domain comparisons **Figure 14**. **b** A virtual 2D array (blue-solid arcs), four subdomains of size $1 \times 1 \times 1$ cm^3 (D_1, D_2, D_3, and D_4), and the domain occupied by rotations of the four subdomains around the array axis (black-dotted circles and arcs). **c** Computation times (t_{fast} and t_{slow}) of the fast and slow operators, respectively, for the forward simulations with the numerical phantom in D_1 (36 repetitions). **d1–d4** Relative errors of the simulated signals for all virtual elements ($n'_{\text{loc}} = 1,2,\dots,396, n'_{\text{ele}} = 1,2,\dots,128$), with the numerical phantom in D_1, D_2, D_3, and D_4, respectively. **e1–e4** Compares of the signals ($\hat{p}^{\text{slow}}$ and $\hat{p}^{\text{fast}}$) simulated by the fast and slow operators for the 64-th element ($n'_{\text{ele}} = 64$) in the first virtual arc array ($n'_{\text{loc}} = 1$), with the phantom in D_1, D_2, D_3, and D_4, respectively. 58

Figure 14 Image-domain accuracy of the fast forward operator. **a1** Relative error of the reconstructed image in the subdomain D_1 with 1 to 256 iterations. **a2–a4** Values on the lines L1, L2, and L3 (shown in **Figure 13a**), respectively, in the reconstructed images with $n_{\text{iter}} = 1, 4, 8, 64, 256$. **a5** Comparison between the ground truth ($p_{0,\text{L1}}$, $p_{0,\text{L2}}$, and $p_{0,\text{L3}}$) and 256-iteration reconstructed ($\hat{p}_{0,\text{L1}}$, $\hat{p}_{0,\text{L2}}$, and $\hat{p}_{0,\text{L3}}$) values along the three lines. **b1–b5**, **c1–c5**, and **d1–d5** The same analysis for reconstructions in subdomains D_2, D_3, and D_4, respectively. 60

Figure 15 Iterative reconstruction for noisy data. **a** Relative errors of the images of the numerical phantom in D_1 reconstructed iteratively with nonnegativity constraint, for SNRs of 2.4, 1.2, 0.8, and 0.6. **b1–b4** Relative errors of the images reconstructed iteratively with TV regularization, for SNRs of 2.4, 1.2, 0.8, and 0.6, respectively, for different choices of regularization parameters (mm). **c** A summary of the plots in **b1–b4** with the best choices of regularization parameters (enclosed in black-solid boxes). **d1–d3**, **e1–e3**, and **f1–f3** Values on the lines L1, L2, and L3 in the images reconstructed with 8, 32, and 256, iterations, respectively. 62

Figure 16 Regularized iterative reconstruction of mouse brain images with sparse sampling. 64

Figure 17 Workflow of the hybrid method for image reconstruction from sparsely sampled raw data and a prior image. 69

Figure 18 UBP, regularized iterative method, and the proposed hybrid method for sparse-sampling imaging, and their linearity tests. **a** Images reconstructed by the three methods (first three columns) from signals detected at sparsely distributed elements (red-bounded blue curves in the fourth column) for $4N_{\text{loc}} = 76, 40, 28, 20, 16, 12$. Examples of maintained features and suppressed artifacts are indicated by white-solid and white-dotted arrows, respectively. PA, photoacoustic. **b** SSIMs between the reconstructed images with $4N_{\text{loc}} = 40, 28, 20, 16, 12$ and those with $4N_{\text{loc}} = 76$ for the three methods. **c** Linearity tests of the three methods for $4N_{\text{loc}} = 12$. Scale bar, 5 mm. 72

Figure 19 Numerical phantoms for functional imaging. **a-b** The background and functional numerical phantoms, respectively, for functional imaging simulation. **c** A virtual array formed by 12 arc arrays, shown as arcs with red boundaries. **d-e** Modulation factors of the background and functional phantoms, respectively. 74

Figure 20 Functional images extracted using the regularized-correlation-based method with $\lambda_f = 1.6$ from ground-truth images and images reconstructed with UBP, the regularized iterative method, and the hybrid method for $A_f = 0.18, 0.06, 0.02$. The first three rows show both the 3D functional and background images and the last three rows show the MAPs of the functional images along the z-axis. In the first and fourth rows, the true functional regions and examples of false positive regions are indicated by white-solid and white-dotted arrows, respectively. 75

Figure 21 Sparse-sampling mouse brain functional imaging *in vivo*. **a** A densely sampled image of a mouse brain reconstructed by UBP (left column, $4N_{loc} = 396$) and sparsely sampled images of the mouse brain reconstructed using UBP (first row), the regularized iterative method (second row), and the hybrid method (third row), respectively, for $4N_{loc} = 40, 20, 12$. Examples of suppressed artifacts and maintained features are indicated by white-dotted and white-solid arrows, respectively. **b** Electrical stimulation to the mouse's right front paw: five cycles, each with 12-s stimulation on and 12-s off. **c** Functional images obtained from the images reconstructed using UBP (first row, $\lambda_f = 0.32$), the regularized iterative method (second row, $\lambda_f = 0.08$), and the hybrid method (third row, $\lambda_f = 0.32$), respectively, for $4N_{loc} = 40, 20, 12$. The true functional regions in all images are indicated by white-solid arrows, and examples of false positive regions are indicated by white-dotted arrows. 78

Figure 22 Mouse brain functional images *in vivo* extracted using the regularized-correlation-based method with $\lambda_f = 0.02, 0.08, 0.32, 1.28, 5.12$ from images reconstructed with UBP and the hybrid method ($4N_{loc} = 12$). The first two rows show both the 3D functional and background images and the last two rows show the MAPs of the functional images along the z-axis. In all images, the true functional regions are indicated by white-solid arrows, and the other high-amplitude functional regions are false positives. Images for $\lambda_f = 0.32$ are highlighted by red-solid boundaries. 79

Figure 23 Mouse brain functional images *in vivo* extracted using the regularized-correlation-based method with $\lambda_f = 0.02, 0.08, 0.32, 1.28, 5.12$ from images reconstructed through the regularized iterative method ($4N_{loc} = 40, 20, 12$). The first three rows show both the 3D functional and background images and the last three rows show the MAPs of the functional images along the z-axis. In all images, the true functional regions are indicated by white-solid arrows, and the other high-amplitude functional regions are false positives. Images for $\lambda_f = 0.08$ are highlighted by red-solid boundaries. 80

Figure 24 Workflow of the intra-image nonrigid motion correction based on data-driven manipulation of the system matrix. 88

Figure 25 Intra-image nonrigid correction of motions induced by tissue translation and deformation. **a** MAP of an image (a human breast image subset with a volume of 44.8×44.8×14.4 mm^3) along the z-axis. PA, photoacoustic. Scale bar, 1 cm. **b-c** Depictions of the intra-image motions induced by tissue translation and deformation, respectively. The image translates (deforms) from a rectangular cuboid with blue-dotted edges to one with red-solid edges as the four-arc array rotates, with the passed virtual element locations marked as blue dots, the latest element locations marked as

red curves, and the rotation direction indicated by black arrows. **d-e** MAPs of the images reconstructed from the signals with motions induced by tissue translation ($A_{\mathrm{tra}} = 0.6$ mm) and deformation ($A_{\mathrm{def}} = 0.06$), respectively, without and with motion correction for $N_{\mathrm{per}} = 0.5,1,2,3$. **f-g** Closed-up subsets of the MAPs enclosed in blue-dashed boxes (in **d** and **f)** and red-doted boxes (in **e** and **g**), respectively, with the ground truths. 91

Figure 26 Intra-image nonrigid motion correction for human breast imaging *in vivo*. **a1–a3** MAP of a human breast image with a volume of $6 \times 6 \times 3$ cm^3 reconstructed without motion correction (**a1**), and two closed-up subsets marked by blue-dashed (**a2**) and red-dotted (**a3**) boxes, respectively. **b1–b3** MAP of the image reconstructed with motion correction from the same signals as for **a1**, and the closed-up subsets. Examples of suppressed artifacts and enhanced features are highlighted in **a2-b2** and **a3-b3**, respectively, by blue-dotted and red-solid arrows. **c1**–**c3** and **d1**–**d3** The second experiment of the same human breast with the images reconstructed without (**c1–c3**) and with (**d1–d3**) motion correction, respectively. Examples of enhanced features are indicated by blue-solid and red-solid arrows in **a2-b2** and **a3-b3**, respectively. For the first-column images, scale bar, 1 cm; for the second and third columns, scale bar, 6 mm. 93

Figure 27 Four sets of signals ($N_{\mathrm{loc}} = 99$, $N_{\mathrm{ele}} = 4 \times 256$) from a human breast *in vivo* with heartbeat-induced motions and reconstructed images. **a-b** Signals in the first two experiments. The reconstructed images are shown in **Figure 26**. Temporal shifts in **a** caused by the quasiperiodic motion are indicated by white arrows. **c1**–**c3** and **d1**–**d3** Signals in the third and fourth experiments, respectively, with the images reconstructed without (**c2-d2**) and with (**c3-d3**) motion correction. Examples of the improvements due to motion correction are indicated by red-solid (feature enhanced) and red-dotted (artifacts suppressed) arrows. 94

Figure 28 PACTER system. **a-b** Schematic of the PACTER system in the calibration (**a**) and imaging (**b**) procedures. HWP, half-wave plate; PBS, polarizing beamsplitter; BT, beam trap; DAQ, data acquisition system. The differences between the two modes are highlighted in the black dotted boxes. **c** Schematic of the single-element ultrasonic transducer fabricated on the ER. **d** 1D (*t*) PACTER signals detected by the ultrasonic transducer at time instances t_0, t_1, and t_N. **e** Reconstruction of a 4D (*xyzt*) image of human palmar vessels from the signals in **d**. Norm., normalized. Scale bar, 1 mm. 100

Figure 29 Single-shot 3D reconstruction in PACTER. **a** Illustration of the calibration procedure of PACTER. Focused laser beams for calibration are shown in green. Calibration pixels are highlighted as orange dots. The calibration step size is 0.1 mm. The calibration pixels (80×80) become 6,400 virtual transducers. $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$ are the positions of three calibrated virtual transducers. **b** Illustration of PACTER of human palmar vessels. The homogenized beam for widefield illumination is shown in green. $\mathbf{r}_1'$ and $\mathbf{r}_2'$ are the positions of two source points in the vessels. Blue and green spheres denote the PA waves generated by the source points. The calibrated virtual transducers capture the PA signals from $\mathbf{r}_1'$ and $\mathbf{r}_2'$ with different delays, indicated by the thick blue and green lines. **c** PACTER signals, $k_1(t)$, $k_2(t)$, $k_3(t)$, of the

calibrated virtual transducers corresponding to $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$, respectively. **d** PACTER signal from the widefield imaging consists of PA signals from $\mathbf{r}_1'$ and $\mathbf{r}_2'$, which are essentially $k_1(t)$, $k_2(t)$, $k_3(t)$ delayed according to the distance between the calibrated virtual transducer and the source point. 104

Figure 30 Spatiotemporal characterization of PACTER. **a** 3D PACTER image of a curved black wire. Norm., normalized. **b** Snapshots of 4D PACTER showing bovine blood flushing through an S-shaped tube. **c** 3D PACTER images of three bars printed with black ink on a transparent film. In each image, the object was placed at a different *z* position. **d** Reconstructed versus real *z* positions of the objects in **c**. The measurement results are plotted as means ± standard errors of the means (n = 1,980). The blue curve represents a linear fit. **e** 3D PACTER image of two crossing human hairs in agarose. **f** Maximum *z*- (top) and *y*-projections (bottom) of the 3D volume in **e**. The *z* positions of the object are color-encoded. **g** Profiles along the yellow dashed lines in **f** denoted by gray dots. The blue curves represent two-term Gaussian fits. Black arrows denote the minimum distances that can separate the two hairs. **h** 3D PACTER image of bovine blood flushing through a tube. The white arrow indicates the flushing direction. **i** PA amplitudes along the tube in **h** versus time, when the blood flushes through the tube at different speeds. **j** Speeds of the blood flow quantified from the reconstructed images versus the real speeds in **i**. The measurement results are plotted as means ± standard errors of the means (n = 74). The blue curve represents a linear fit. Scale bars, 1 mm. 107

Figure 31 PACTER of mouse hemodynamics *in vivo*. **a** Schematic of the mouse imaging experiment. **b**-**c** 3D PACTER images of the abdominal vasculature of Mouse 1 (**b**) and Mouse 2 (**c**). Norm., normalized. **d**-**e** Cross-sectional 2D images corresponding to the yellow rectangle in **b** (**d**) and the magenta rectangle in **c** (**e**) at four different time instances from the 4D PACTER datasets. t_0 = 0.28 s, t_1 = 0.26 s. White solid curves represent the Gaussian fits of the vessels' profile denoted by the yellow (**d**) and magenta (**e**) dashed lines. Differences from the first image are highlighted. **f**-**g** PA amplitudes along the yellow dashed line (1D images) in **d** (**f**) and the magenta dashed line in **e** (**g**) versus time, where the time instances in **d** and **e** are labeled with vertical gray lines. **h**-**i** Center positions (blue solid curves) and widths (orange dash-dotted curves) of the vessels versus time, based on the fits in **d** (**h**) and **e** (**i**). The shaded areas denote the standard deviations (n = 5). **j**-**k** Fourier transforms of the center positions and widths of the vessels in **h** (**j**) and **i** (**k**), showing the respiratory frequency from the vessel center positions only (**j**) or both the vessel center positions and widths (**k**). Scale bars, 1 mm. 109

Figure 32 PACTER of human hand hemodynamics *in vivo*. **a** Schematic of the human hand imaging experiment. **b**-**c** 3D PACTER images of the thenar vasculature of Participant 1 (**b**) and Participant 2 (**c**). Norm., normalized. **d**-**e** Maximum amplitude projections of the 3D volumes from the 4D PACTER datasets along the *z* axis in **b** (**d**) and **c** (**e**) at the time instances before, during, and after cuffing. t_0 = 17.44 s, t_1 = 19.02 s. The solid lines flank the vessels under investigation. Differences from the images during cuffing are highlighted. **f**-**g** PA amplitudes along the vessels (1D images) in **d** (**f**) and **e** (**g**) versus time, where the time instances in **d** and **e** are labeled with vertical gray

lines. The blue and orange arrows indicate peak responses in the occlusion and recovery phases, respectively. **h** Positions (solid circles) of the blood front along the blood vessel during the occlusion (left) and recovery (right) phases in **f**. The blue curve is an exponential fit with an occlusion rate of 1.3 ± 0.1 m/s, and the orange curve is a linear fit showing the blood flow speed of 16.5 ± 2.8 m/s. **i** Comparison between the durations of the occlusion and recovery phases in **f**. ***$P < 0.001$, calculated by the two-sample *t*-test. **j** Positions (solid circles) of the blood front along the blood vessel during the occlusion (left) and recovery (right) phases in **g**. The blue curve is an exponential fit with an occlusion rate of 2.4 ± 0.3 m/s, and the orange curve is a linear fit showing the blood flow speed of 32.9 ± 6.5 m/s. **k** Comparison between the durations of the occlusion and recovery phases in **g**. ***$P < 0.001$, calculated by the two-sample *t*-test. **l** Comparison between the blood flow speeds during recovery in **f** and **g**. *$P < 0.05$. Scale bars, 1 mm. 111

LIST OF TABLES

Number *Page*

Table 1 Model imperfections and compensation purposes in different chapters of this dissertation. 4

Table 2 Symbols for the fast forward operator and image reconstruction. 125

Table 3 Additional symbols for fast functional imaging with sparse sampling. 128

Table 4 Additional symbols for intra-image nonrigid motion correction. 130

ABBREVIATIONS

1D: One-dimensional

2D: Two-dimensional

3D: Three-dimensional

4D: Four-dimensional

DAQ: Data acquisition system

EIR: Electrical impulse response

FFT: Fast Fourier transform

FISTA: Fast iterative shrinkage-thresholding algorithm

FOV: Field of view

FWHM: Full width at half maximum

IR: Image reconstruction

LDTF: Location-dependent temporal filtering

MAP: Maximum amplitude projection

MRI: Magnetic resonance imaging

PA: Photoacoustic

PACT: Photoacoustic computed tomography

PACTER: Photoacoustic computed tomography through an ergodic relay

PCCR: Regularized Pearson correlation coefficient

RDTF: Radius-dependent temporal filtering

SBH-PACT: Single-breath-hold photoacoustic computed tomography

SIR: Spatial impulse response

SNR: Signal-to-noise ratio

SS: Spatial sampling

SVD: Singular value decomposition

TV: Total variation

UBP: Universal back-projection

X-ray CT: X-ray computed tomography

Chapter 1

PACT AND ITS MODEL IMPERFECTIONS

Photoacoustic (PA) computed tomography (PACT) is a noninvasive hybrid imaging modality that combines the optical contrast of diffuse optical tomography with the high spatial resolution of ultrasonography. When a short-pulsed laser irradiates biological tissues, the light energy is partially absorbed by the tissue and converted into heat. Subsequently, a pressure rise induced by transient thermoelastic expansion propagates as a wideband ultrasonic wave (referred to as PA wave) through the biological tissue. The PA waves are then detected by ultrasonic transducers placed around the tissue surface to reconstruct the optical absorption distribution in the tissue [1]. The high sensitivity to optical absorption in biological tissues provides rich contrast mechanisms relating to various intrinsic and extrinsic chromophores, enabling structural, functional, and molecular imaging [2]. In addition to the rich optical contrast, the conversion from light to ultrasound also provides this imaging technique with ultrasonically defined spatial resolution at depths beyond the ~1-mm optical diffusion regime, which limits the penetration capability of ballistic optical imaging modalities [3].

This hybrid nature grants PACT unique advantages and makes it complementary to other mainstream imaging modalities in clinical practice: (1) Compared with pure optical imaging techniques such as fluorescence imaging in humans [4], PACT can sustain high spatial resolution in deep tissue. (2) Compared with ultrasonic imaging, PACT has rich intrinsic and extrinsic optical contrasts [3]; (3) Compared with X-ray computed tomography (X-ray CT) and positron emission tomography, PACT is free of radioactive material and uses

nonionizing illumination; (4) Compared with magnetic resonance imaging (MRI), laser-based PACT provides faster imaging with higher spatial resolution at a lower cost.

In a homogeneous medium, a photoacoustic wave can be expressed as [1], [5]

$$p(\mathbf{r},t)=\frac{1}{4\pi c^2}\frac{\partial}{\partial t}\left(\frac{1}{ct}\int_V \mathrm{d}\mathbf{r}' p_0(\mathbf{r}')\delta\left(t-\frac{\|\mathbf{r}'-\mathbf{r}\|}{c}\right)\right). \tag{1}$$

Here, $p(\mathbf{r},t)$ is the pressure at location $\mathbf{r}$ and time t, c is the speed of sound, V is the volumetric space occupied by the tissue, and $p_0(\mathbf{r}')$ is the initial pressure at $\mathbf{r}'$. For convenience in the following discussion, we rewrite Eq. (1) as

$$p(\mathbf{r},t)=\frac{1}{4\pi c^2}\int_V \frac{p_0(\mathbf{r}')}{\|\mathbf{r}'-\mathbf{r}\|}\frac{\partial}{\partial t}\delta\left(t-\frac{\|\mathbf{r}'-\mathbf{r}\|}{c}\right)\mathrm{d}\mathbf{r}'. \tag{2}$$

Discretizing Eq. (2) in space, we obtain

$$p(\mathbf{r}_n,t)=\frac{1}{4\pi c^2}\sum_{m=1}^{M} v_m \frac{p_0(\mathbf{r}_m')}{\|\mathbf{r}_m'-\mathbf{r}_n\|}\frac{\partial}{\partial t}\delta\left(t-\frac{\|\mathbf{r}_m'-\mathbf{r}_n\|}{c}\right),$$
$$n=1,2,\dots,N. \tag{3}$$

Here, we assume M source points distributed at $\mathbf{r}_m', m=1,2,\dots,M$, and N point detection elements distributed at $\mathbf{r}_n, n=1,2,\dots,N$. The term v_m is the volume of the m-th source point. For three ideal detection geometries (planar, spherical, and cylindrical surfaces), an image mapping the initial pressure $p_0(\mathbf{r}'')$ can be reconstructed through the universal back-projection (UBP) formula [6]:

$$p_0(\mathbf{r}'')=\int_{\Omega_0} b\left(\mathbf{r},t=\frac{\|\mathbf{r}''-\mathbf{r}\|}{c}\right)\frac{\mathrm{d}\Omega}{\Omega_0}, \tag{4}$$

where $b(\mathbf{r},t)=2p(\mathbf{r},t)-2t\frac{\partial p(\mathbf{r},t)}{\partial t}$ is the back-projection term, $\mathrm{d}\Omega=\frac{\mathrm{d}S}{\|\mathbf{r}''-\mathbf{r}\|^2}\frac{\mathbf{n}_S(\mathbf{r})\cdot(\mathbf{r}''-\mathbf{r})}{\|\mathbf{r}''-\mathbf{r}\|}$ is the solid angle for detection element at $\mathbf{r}$ with respect to the reconstruction location $\mathbf{r}''$, $\mathrm{d}S$ is the detection element surface area, and $\mathbf{n}_S(\mathbf{r})$ is the ingoing normal vector. The total solid angle is denoted as Ω_0.

In a homogenous medium, using an ideal detection geometry consisting of point detectors with unlimited bandwidth, we can detect the PA signals from an object and

reconstruct the initial pressure accurately. However, in practice, the detection system and its mathematical model are imperfect due to multiple factors, e.g., limited transducer bandwidth, finite transducer element size, sparse spatial sampling, partial-view detection, tissue motion, and acoustic inhomogeneity. Although hardware improvements are preferred options to break these limitations, they are not always feasible and may be very expensive. Numerical computation is a low-cost alternative to improve image quality. Numerous methods have been proposed to compensate for these model imperfections but have their limitations.

In this dissertation, we assume a homogeneous medium in all studies and propose computational methods to compensate for imperfections brought by limited transducer bandwidth, finite transducer element size, sparse spatial sampling, partial-view detection, and tissue motion.

In Chapter 2, we first introduce a human breast imaging system based on a full-ring transducer array. To mitigate the artifacts due to sparse sampling for better tumor detection, we incorporate the limited bandwidth of the transducer into the model and perform spatiotemporal analysis to classify the sources of aliasing. Based on the analysis, we propose two antialiasing methods: radius-dependent spatiotemporal antialiasing and location-dependent spatiotemporal antialiasing. In Chapter 3, we first introduce an isotropic-resolution 3D PACT system with four arc arrays, which has a massive system matrix. Then we propose an efficient explicit expression of the massive system matrix based on singular value decomposition and fast Fourier transform, which achieves $\times 42$ acceleration from the original explicit expression based on point source response and with negligible error. In Chapter 4, we focus on brain functional imaging using the four-arc 3D PACT system. We improve the temporal resolution for functional imaging by numerically reducing the requirement of sampling locations of the four-arc array for each image. This requirement reduction is achieved by incorporating a densely sampled prior image into the system matrix, which maintains the critical linearity while mitigating artifacts. In Chapter 5, we solve the problem of intra-image nonrigid motion correction for the four-arc 3D PACT system. We divide the whole image domain into multiple subdomains and approximate the nonrigid motion of the whole domain by translations of the subdomains. By incorporating these

translations into the forward model, we reconstruct them together with the image using Gauss-Seidel-type iterations. In Chapter 6, we introduce a single-shot 3D PACT system based on a single ultrasonic transducer, which has extremely compressed sensing and partial-view detection. We propose a fast implementation of the forward operator by connecting traditional PACT with virtual detector responses through fast Fourier transform. Based on the operator, we reconstruct the image from the compressed signals through regularized iteration. We demonstrate all the proposed methods with numerical simulations, phantom experiments, and *in vivo* experiments. All the studies in Chapters 2–6 are summarized in **Table 1**.

Table 1 Model imperfections and compensation purposes in different chapters of this dissertation.

Imperfection \ Compensation purpose	Fewer artifacts	Better true features	Finer spatial resolution	Finer temporal resolution
Limited transducer bandwidth	2	2, 3, 4, 5	3, 4, 5	
Finite transducer element size		3, 4, 5	3, 4, 5	
Sparse spatial sampling	2, 4	2, 4		4
Partial-view detection	6	6	6	6
Tissue motion	5	5		
Device-level signal compression	6	6	6	6

An underlined chapter number indicates that the imperfection-purpose combination is a main problem in this chapter.

C h a p t e r 2

SPATIOTEMPORAL ANTIALIASING

P. Hu, L. Li, and L. V. Wang, "Location-dependent spatiotemporal antialiasing in photoacoustic computed tomography," *IEEE Trans. Med. Imaging*, vol. 42, no. 4, pp. 1210–1224, Apr. 2023. DOI: 10.1109/TMI.2022.3225565

P. Hu, L. Li, L. Lin, and L. V. Wang, "Spatiotemporal antialiasing in photoacoustic computed tomography," *IEEE Trans. Med. Imaging*, vol. 39, no. 11, pp. 3535–3547, Nov. 2020. DOI: 10.1109/TMI.2020.2998509

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," *Nat. Commun.*, vol. 9, no. 1, p. 2352, Jun. 2018. DOI: 10.1038/s41467-018-04576-z

2.1 Abstract

We have developed a single-breath-hold photoacoustic computed tomography (SBH-PACT) system to reveal detailed angiographic structures in human breasts. By scanning the entire breast within a single breath hold (~15 s), a volumetric image can be acquired and subsequently reconstructed with negligible breathing-induced motion artifacts. SBH-PACT reveals tumors by visualizing higher blood vessel densities associated with tumors at a high spatial resolution, showing early promise for high sensitivity in radiographically dense breasts. However, the quantification of blood vessel density can be severely affected by artifacts caused by sparse spatial sampling violating the spatial Nyquist criterion, which happens in off-center regions for the proposed system. To mitigate these artifacts, we first clarified the source of spatial aliasing through spatiotemporal analysis. We demonstrated that the combination of spatial interpolation and radius-dependent temporal filtering, named radius-dependent spatiotemporal antialiasing, can effectively mitigate artifacts caused by aliasing in either image reconstruction or spatial sampling. Further, we refined the spatiotemporal analysis to clarify the spatial Nyquist criterion's different requirements in

different image subdomains and proposed a location-dependent spatiotemporal antialiasing method. Both methods mitigate aliasing artifacts in numerical simulations and *in vivo* experiments, and the location-dependent version has better performance in maintaining image resolution.

2.2 Background

2.2.1 Human breast imaging for tumor detection

Breast cancer is the second most common cancer to affect women in the U.S. and is the second-ranked cause of cancer-related deaths. About 1 in 8 (12%) women in the U.S. will develop invasive breast cancer during their lifetime [7]. Multiple large prospective clinical trials have demonstrated the importance of early detection in improving breast cancer survival [8]–[10]. While mammography is currently the gold standard used for breast cancer screening, it utilizes ionizing radiation and has lower sensitivity in women with dense breasts [11], [12]. Ultrasonography has been used as an adjunct to mammography but suffers from speckle artifacts and low specificity [13], [14]. Magnetic resonance imaging (MRI) poses a large financial burden and requires the use of intravenous contrast agents that can cause allergy [15], kidney damage [16], and permanent deposition in the central nervous system [17]. Diffuse optical tomography has been investigated to provide functional optical contrast. However, the spatial resolution of the current prototypes limits their clinical use [18], [19]. Overall, each modality has notable advantages and limitations. Photoacoustic computed tomography (PACT) is a promising complementary modality that overcomes many of these limitations.

We reported a significant advancement in breast PACT technology, a single-breath-hold PACT (SBH-PACT) system [20] with advantages: (1) Combining 1064-nm light illumination and a 2.25-MHz unfocused ultrasonic transducer array, SBH-PACT achieved up to 4 cm *in vivo* imaging depth and a 255-μm in-plane resolution (approximately four times finer than that of contrast-enhanced MRI [21]). (2) Equipped with one-to-one mapped signal amplification and data acquisition circuits, SBH-PACT can obtain an entire 2D cross-sectional breast image with a single laser pulse, or obtain a volumetric 3D image of the entire

breast by fast elevational scanning within a single breath-hold (~15 s). The 10-Hz 2D frame rate, currently limited by the laser repetition rate, enables SBH-PACT to observe biological dynamics in a cross-section associated with respiration and heartbeats without motion artifacts. (3) A full-ring 512-element ultrasonic transducer array enables SBH-PACT for full-view fidelity in 2D imaging planes and delivers high image quality. (4) Capitalizing on the optimized illumination method and signal amplification, SBH-PACT achieves sufficient noise-equivalent sensitivity to clearly reveal detailed angiographic structures both inside and outside breast tumors without the use of exogenous contrast agents. We visualize the system in **Figure 1**.

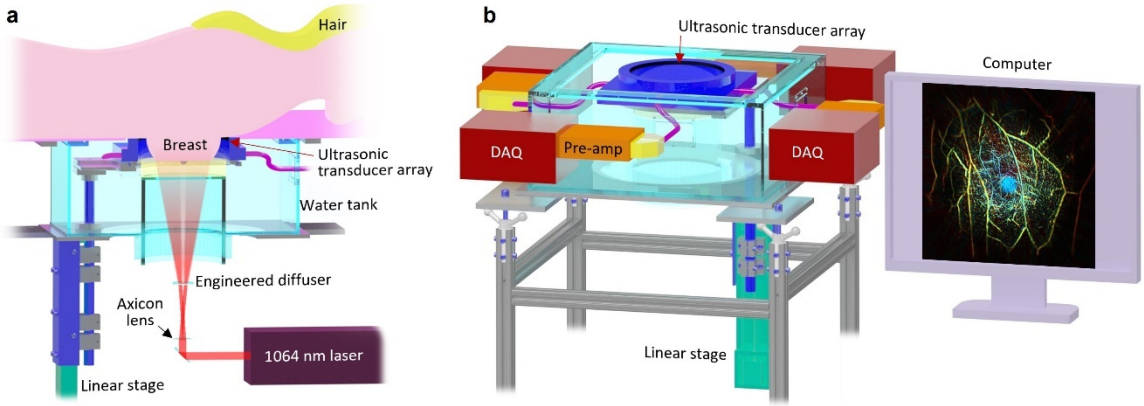

Figure 1 Representations of the SBH-PACT system. **a** Perspective cut-away view of the system with data acquisition components removed. **b** Perspective view of the system with patient bed and optical components removed. DAQ, data acquisition system; Pre-amp, pre-amplifier circuits.

In the pilot study [20], SBH-PACT was used to image one healthy volunteer and seven breast cancer patients. SBH-PACT identified eight of the nine breast tumors by delineation of angiographic anatomy. These tumors were subsequently verified by ultrasound-guided biopsy. A coronal-plane slice of an image reconstructed using 3D universal back-projection (UBP) for a breast cancer patient is shown in **Figure 2a** with its red-boxed closed-up subset shown in **Figure 2b**. The tumor, appearing as a group of dense blood vessels, is indicated by a green-dashed ellipse. Two X-ray mammograms of the breast are shown in **Figure 2c1** and **c2**, respectively, with the tumor indicated by two yellow-dashed ellipses. To compress the information in the 3D PACT image, we perform maximum amplitude projection (MAP) on the image along the direction normal to the coronal plane to obtain the maximum amplitudes

and the maximum-amplitude depths. We assign different colors to these depths to form a 2D depth-resolved color-encoded image and multiply it pixel by pixel with the maximum amplitudes to form the image shown in **Figure 2d** with the tumor indicated by a white-dashed ellipse. Further, we obtain the MAP of the image slices without the nipple, from which we quantify the vessel density by counting the number of vessels in 2×2 mm^2 squares. The blood vessel density map and the nipple-excluded MAP are shown in **Figure 2f**. The tumor is automatically identified through thresholding of the blood vessel density and is labeled by a green-dashed ellipse in **Figure 2f**, which matches well with the tumor locations shown in mammograms.

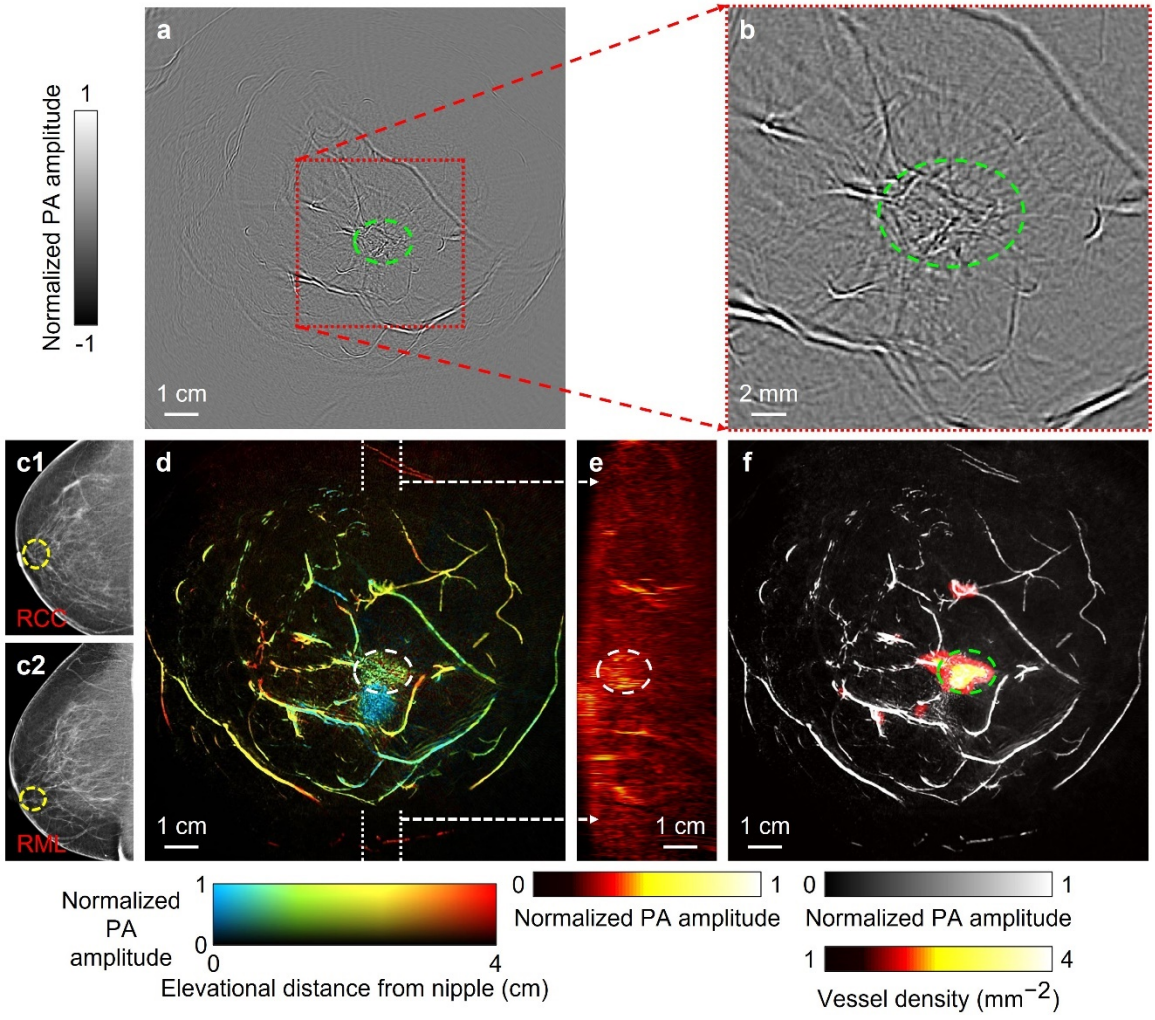

Figure 2 PACT images and X-ray mammograms of a cancerous human breast. **a** A cross section of the PACT image of a cancerous human breast with the tumor circled by a green-dashed ellipse. PA,

photoacoustic. **b** A closeup of the subset enclosed by a red-dotted box. **c1**-**c2** X-ray mammograms of the breast. RCC, right cranial-caudal; RML, right medio-lateral. **d** MAP of the PACT image along the elevational direction that is color-encoded by the elevational distance from the nipple. **e** MAP of the thick slice in the sagittal plane marked by white dashed lines in **d**. **f** Automatic tumor detection on a vessel density map. The tumor is indicated by a green-dashed ellipse. The gray-scale background is the MAP of the tissue image elevationally separated from the nipple.

2.2.2 Spatiotemporal antialiasing

As shown in our SBH-PACT study, reliably detecting angiogenesis-induced blood vessel structures is critical for tumor detection. However, these blood vessels can be contaminated by artifacts due to sparse spatial sampling, especially in regions close to the transducer elements.

In PACT, the ultrasonic transducer array should provide dense spatial sampling (SS) around the object to satisfy the Nyquist sampling theorem [22], [23]. The SS interval on the tissue surface should be less than half of the lowest detectable acoustic wavelength. Otherwise, artifacts may appear in image reconstruction (IR), a problem we call spatial aliasing. In practice, due to the high cost of a transducer array with a large number of elements or limited scanning time, spatially sparse sampling is common.

Various methods have been proposed to mitigate artifacts caused by spatial aliasing. In the image domain, total variation (TV) regularization has been used in model-based iterative methods to mitigate noise and aliasing artifacts [24]–[28]. In PACT, because the image domain is identical to the object domain, we will use them interchangeably henceforth. TV regularization shows high performance for piecewise smooth images [29], [30]. However, for PACT images with rich blood vessel structures, TV regularization tends to suppress vessels of small diameters. A regularization strategy specifically suited for vessel structures is needed. Deep learning has been proved effective in processing images with complex structures [31]–[36] and has shown an advantage in maintaining vessel structures [31]. However, a neural network is often system dependent and not universally applicable across different imaging systems or detection geometries. In the signal domain, temporal filtering and spatial interpolation have been used for antialiasing [37]. However, it is still a challenge to find a balance between mitigating aliasing artifacts and maintaining

image resolution. Another method proposed by Cai *et al*. [38] mitigates aliasing artifacts by connecting the image domain and the signal domain. It identifies potential sources of aliasing signals in the image domain, maps the sources to the signal domain, suppresses all the signals in the mapped region, and uses the remaining signals for image reconstruction [38]. This method performs well if there exist only a few dominant sources of aliasing signals. As sources of aliasing signals increase, this method may cause substantial information loss.

In this dissertation, we analyze spatial aliasing in PACT using UBP reconstruction [6]. We use a circular geometry (a full-ring ultrasonic transducer array or its scanning equivalent) with point elements as an example for analysis. In addition, we discuss only acoustically homogeneous media. Starting from the reconstruction at a source point, we identify, for the first time, two types of spatial aliasing: aliasing in SS and aliasing in IR. Then we demonstrate that aliasing in IR can be eliminated through spatial interpolation while aliasing in SS can be mitigated through radius-dependent temporal filtering. We name the combination of radius-dependent temporal filtering and spatial interpolation as radius-dependent spatiotemporal antialiasing.

Further, to mitigate aliasing artifacts without compromising image resolution, we perform detailed spatiotemporal analysis for image subdomains. We first reconstruct an image using the UBP method [6]. Applying a threshold to the reconstructed image, we identify the dominant sources of aliasing signals. Then we divide the whole image domain into multiple subdomains. We apply spatiotemporal analysis to source points, transducer locations, and each subdomain's reconstruction locations [37], revealing the spatial aliasing effects on the subdomain in detail. Next, we apply temporal filtering and spatial interpolation to signals so that the filtered signals satisfy the spatial Nyquist criterion specifically for this subdomain. We use the filtered signals to form an image in this subdomain. Repeating this process for all subdomains, we mitigate the aliasing artifacts for the whole image. We call this method location-dependent spatiotemporal antialiasing. Through numerical simulations, we demonstrate that the proposed method effectively mitigates aliasing artifacts with

minimal effects on the image resolution. We further validate this method through *in vivo* human breast imaging.

2.3 Radius-dependent spatiotemporal antialiasing

2.3.1 Theory

The response of an ultrasonic transducer can be described by the equation

$$\hat{p}(\mathbf{r}_n, t) = p(\mathbf{r}_n, t) *_t h_e(t), n = 1,2, \dots, N. \tag{5}$$

Here, $\hat{p}(\mathbf{r}_n, t)$ is the response of the n-th point detection element at time t, $h_e(t)$ is the ultrasonic transducer's electrical impulse response (EIR), and $*_t$ denotes temporal convolution. Substituting Eq. (3) into Eq. (5), we obtain

$$\hat{p}(\mathbf{r}_n, t) = \frac{1}{4\pi c^2} \sum_{m=1}^{M} v_m \frac{p_0(\mathbf{r}'_m)}{\|\mathbf{r}'_m - \mathbf{r}_n\|} h'_e\left(t - \frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c}\right). \tag{6}$$

The term $h'_e\left(t - \frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c}\right)$ is a function of both time and space, where the first prime denotes the temporal derivative. The following discussion is based on the spatiotemporal analysis of this term. When acoustic signals are digitized by a data acquisition system, an antialiasing filter for a sufficiently high temporal sampling rate avoids temporal aliasing. Thus, for simplicity, the time variable is assumed to be continuous here. The spatial variables are discretized, allowing for further discussion of SS.

We replace the true pressure $p(\mathbf{r}, t)$ with the detected pressure $\hat{p}(\mathbf{r}_n, t)$, leading to a discretized form of Eq. (4):

$$\hat{p}_0(\mathbf{r}'') \approx \sum_{n=1}^{N} w_n \hat{b}\left(\mathbf{r}_n, t = \frac{\|\mathbf{r}'' - \mathbf{r}_n\|}{c}\right). \tag{7}$$

Here, $\hat{p}_0(\mathbf{r}'')$ is the reconstructed initial pressure and $\hat{b}(\mathbf{r}_n, t) = 2\hat{p}(\mathbf{r}_n, t) - 2t\frac{\partial \hat{p}(\mathbf{r}_n, t)}{\partial t}$ is the back-projection term computed from the detected pressure. The weights $w_n, n = 1,2, \dots, N$ come from $\frac{\mathrm{d}\Omega}{\Omega_0}$ in Eq. (4). Also, it needs to be noted that UBP is approximately

valid for a ring array because it is not one of the three geometries required for accurate UBP reconstruction. This approximation is an important simplification for spatiotemporal analysis.

We assume that the full-ring transducer array with a radius of R has N evenly distributed detection elements, shown as the red circle in **Figure 3a**. The center O of the circle is the origin of a coordinate system for IR. The upper cutoff frequency of the ultrasonic transducer is f_c (the estimation of f_c is discussed in [37]), and the corresponding lower cutoff wavelength $\lambda_c = \frac{c}{f_c}$. The acquired signals were first filtered by a third-order lowpass Butterworth filter and a sinc filter (both with a cutoff frequency of f_c). Thus, the frequency components with frequencies higher than f_c were removed. We define S_0, called the detection zone here, as

$$S_0 = \{\mathbf{r}' | \|\mathbf{r}'\| \leq R\}. \tag{8}$$

We first analyze aliasing in SS. When the detection element location $\mathbf{r}$ varies discretely, the step size along the perimeter is $\frac{2\pi R}{N}$. The tangential direction is marked by a dotted line (**Figure 3a**), which is perpendicular to vector $\mathbf{r}$. We consider a source point at $\mathbf{r}'$, and extend the line segment $\mathbf{r} - \mathbf{r}'$ as a dashed line (**Figure 3a**). Vectors $-\mathbf{r}'$ and $\mathbf{r} - \mathbf{r}'$ form an angle β, while vector $\mathbf{r} - \mathbf{r}'$ forms an angle α' with the tangential dotted line. Then the angle formed by vectors $-\mathbf{r}$ and $\mathbf{r}' - \mathbf{r}$ can be expressed as $\alpha' - \frac{\pi}{2}$. The local sampling step size of $\|\mathbf{r} - \mathbf{r}'\|$ is approximately

$$\frac{2\pi R}{N} \cos \alpha', \tag{9}$$

whose absolute value means the length of the local sampling step size, while the sign means the sampling direction. From Eq. (6), at a given time t, and with the lower cutoff wavelength λ_c, we can express the Nyquist criterion as

$$\frac{2\pi R |\cos \alpha'|}{N} < \frac{\lambda_c}{2}. \tag{10}$$

To transform this inequality to a constraint for the source point location $\mathbf{r}'$, we use the Law of Sines:

$$\frac{R}{\sin\beta} = \frac{r'}{\sin\left(\alpha' - \frac{\pi}{2}\right)} = \frac{r'}{-\cos\alpha'}. \tag{11}$$

Here $r' = \|\mathbf{r}'\|$. Using Eq. (11), we transform Expression (9) to

$$\frac{2\pi R\cos\alpha'}{N} = \frac{-2\pi r'\sin\beta}{N}. \tag{12}$$

Combining Inequality (10) and Eq. (12), we obtain

$$r' < \frac{N\lambda_c}{4\pi|\sin\beta|}, \tag{13}$$

which must be satisfied for any $\beta \in [0,2\pi)$.

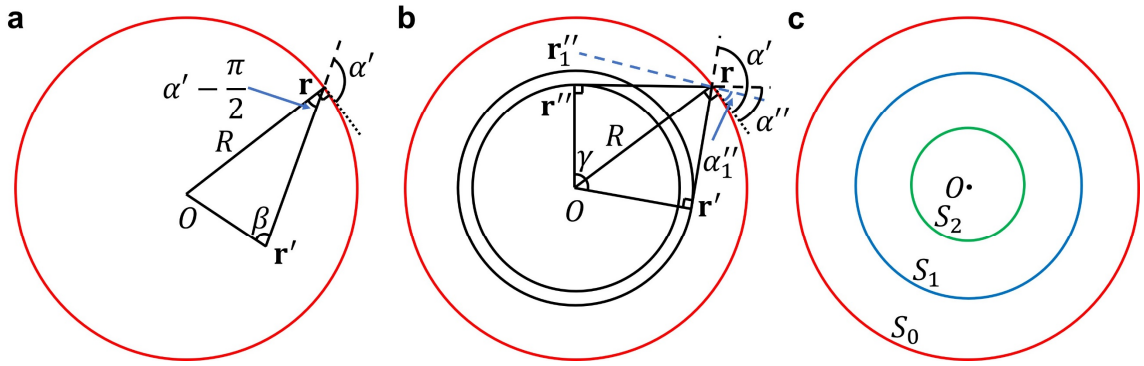

Figure 3 Analysis of the spatial aliasing for a circular geometry. **a** A full-ring transducer array of radius R (red circle), where a detection element location $\mathbf{r}$ and a source point location $\mathbf{r}'$ are marked. The locations $\mathbf{r}$ and $\mathbf{r}'$ are also seen as vectors from the origin to these locations. Vectors $-\mathbf{r}'$ and $\mathbf{r} - \mathbf{r}'$ form an angle β, while the extension of line segment $\mathbf{r} - \mathbf{r}'$ forms an angle α' with the tangential dotted line that is perpendicular to $\mathbf{r}$. The angle formed by vectors $-\mathbf{r}$ and $\mathbf{r}' - \mathbf{r}$ can be expressed as $\alpha' - \frac{\pi}{2}$. This graph is used to analyze the aliasing in SS. **b** A full-ring transducer array with a detection element location $\mathbf{r}$, two reconstruction locations $\mathbf{r}''$ and $\mathbf{r}_1''$, and a source point location $\mathbf{r}'$ marked. Extensions of the line segments $\mathbf{r} - \mathbf{r}''$, $\mathbf{r} - \mathbf{r}_1''$, and $\mathbf{r} - \mathbf{r}'$ form angles α'', α_1'', and α', respectively, with the tangential dotted line that is perpendicular to $\mathbf{r}$. Vectors $\mathbf{r}''$ and $\mathbf{r}'$ form an angle $\gamma = \arccos\left(\frac{r''}{R}\right) + \arccos\left(\frac{r'}{R}\right)$, where $r'' = \|\mathbf{r}''\|$ and $r' = \|\mathbf{r}'\|$. Points $\mathbf{r}''$ and $\mathbf{r}_1''$ are on the same circle centered at $\mathbf{r}$. This graph is used to analyze aliasing in IR. **c** Regions in the field of view representing different types of aliasing. In S_2 (green circle), which contains all source points and reconstruction locations, UBP reconstruction yields no aliasing. In S_1 (blue circle), aliasing does not exist in SS but may exist in UBP reconstruction. In S_0 (red circle), aliasing may exist in SS.

When $\beta = \frac{\pi}{2}$ or $\frac{3\pi}{2}$, Eq. (13) leads to the smallest upper limit of r':

$$r' < \frac{N\lambda_c}{4\pi}, \tag{14}$$

and the length of the local step size maximizes to $\frac{2\pi r'}{N}$. Obviously, if r' satisfies Eq. (14), the length of the local step size of $\|\mathbf{r}-\mathbf{r}'\|$ given by Eq. (12) also satisfies the Nyquist criterion:

$$\frac{2\pi R|\cos\alpha'|}{N}=\frac{2\pi r'|\sin\beta|}{N}\le\frac{2\pi r'}{N}<\frac{\lambda_c}{2}. \tag{15}$$

We define the region S_1, called the one-way Nyquist zone here, as

$$S_1=\left\{\mathbf{r}'\middle|\|\mathbf{r}'\|<\frac{N\lambda_c}{4\pi}\right\}. \tag{16}$$

Equivalently, we can consider the boundary of S_1 as a virtual detection surface, where the sampling spacing is scaled down from the actual detection spacing by $\frac{R_1}{R}$ with R_1 being the radius of S_1. For any source point inside S_1, there is no spatial aliasing during SS because the sampling spacing is less than half of the lower cutoff wavelength, which agrees with the result in Xu *et al.* [22].

Next, we analyze the spatial aliasing in IR. Substituting Eq. (6) into Eq. (7), we obtain

$$\hat{p}_0(\mathbf{r}'')\approx\frac{1}{2\pi c^2}\sum_{n=1}^{N}w_n\sum_{m=1}^{M}v_m\frac{p_0(\mathbf{r}_m')}{\|\mathbf{r}_m'-\mathbf{r}_n\|}$$
$$\left(1-\left(t+\frac{\|\mathbf{r}_m'-\mathbf{r}_n\|}{c}\right)\frac{\partial}{\partial t}\right)h_e'\left(\frac{\|\mathbf{r}''-\mathbf{r}_n\|}{c}-\frac{\|\mathbf{r}_m'-\mathbf{r}_n\|}{c}\right). \tag{17}$$

Here, we use the differential operator

$$\left(1-\left(t+\frac{\|\mathbf{r}_m'-\mathbf{r}_n\|}{c}\right)\frac{\partial}{\partial t}\right)h_e'(t)=h_e'(t)-\left(t+\frac{\|\mathbf{r}_m'-\mathbf{r}_n\|}{c}\right)h_e''(t). \tag{18}$$

In Eq. (17), we need to analyze only the expression $\left(1-\left(t+\frac{\|\mathbf{r}_m'-\mathbf{r}_n\|}{c}\right)\frac{\partial}{\partial t}\right)h_e'\left(\frac{\|\mathbf{r}''-\mathbf{r}_n\|}{c}-\frac{\|\mathbf{r}_m'-\mathbf{r}_n\|}{c}\right)$, which consists of two terms: $h_e'\left(\frac{\|\mathbf{r}''-\mathbf{r}_n\|}{c}-\frac{\|\mathbf{r}_m'-\mathbf{r}_n\|}{c}\right)$ and $\frac{\|\mathbf{r}''-\mathbf{r}_n\|}{c}h_e''\left(\frac{\|\mathbf{r}''-\mathbf{r}_n\|}{c}-\frac{\|\mathbf{r}_m'-\mathbf{r}_n\|}{c}\right)$. The difference between the spectra of $\frac{\|\mathbf{r}''-\mathbf{r}_n\|}{c}h_e''\left(\frac{\|\mathbf{r}''-\mathbf{r}_n\|}{c}-\frac{\|\mathbf{r}_m'-\mathbf{r}_n\|}{c}\right)$ and $h_e''\left(\frac{\|\mathbf{r}''-\mathbf{r}_n\|}{c}-\frac{\|\mathbf{r}_m'-\mathbf{r}_n\|}{c}\right)$ is negligible [37]. Thus we only need to analyze the spatial aliasing in $h_e'\left(\frac{\|\mathbf{r}''-\mathbf{r}_n\|}{c}-\frac{\|\mathbf{r}_m'-\mathbf{r}_n\|}{c}\right)$ and $h_e''\left(\frac{\|\mathbf{r}''-\mathbf{r}_n\|}{c}-\frac{\|\mathbf{r}_m'-\mathbf{r}_n\|}{c}\right)$.

If $h'_e(t)$ has an upper cutoff frequency f_c, $h''_e(t)$ will have the same upper cutoff frequency [37]. Given a reconstruction location $\mathbf{r}''$ and a source point location $\mathbf{r}'$ (**Figure 3b**), we need to analyze the sampling step size of $\|\mathbf{r}'' - \mathbf{r}\| - \|\mathbf{r}' - \mathbf{r}\|$ while the detection element location $\mathbf{r}$ varies. In fact, the lengths of the step sizes of both $\|\mathbf{r}'' - \mathbf{r}\|$ and $\|\mathbf{r}' - \mathbf{r}\|$ reach maxima when $\mathbf{r}'' - \mathbf{r}$ and $\mathbf{r}' - \mathbf{r}$ are perpendicular to $\mathbf{r}''$ and $\mathbf{r}'$, respectively. If the angle γ between vectors $\mathbf{r}''$ and $\mathbf{r}'$ satisfies $\gamma = \arccos\left(\frac{r''}{R}\right) + \arccos\left(\frac{r'}{R}\right)$, where $r'' = \|\mathbf{r}''\|$ and $r' = \|\mathbf{r}'\|$, then the lengths of the step sizes of $\|\mathbf{r}'' - \mathbf{r}\|$ and $\|\mathbf{r}' - \mathbf{r}\|$ achieve maxima of $\frac{2\pi r''}{N}$ and $\frac{2\pi r'}{N}$, respectively, with $\mathbf{r}$ at the same location, as shown in **Figure 3b**. In addition, as $\mathbf{r}$ passes this location clockwise, $\|\mathbf{r}'' - \mathbf{r}\|$ increases while $\|\mathbf{r}' - \mathbf{r}\|$ decreases. Thus, the length of the step size of $\|\mathbf{r}'' - \mathbf{r}\| - \|\mathbf{r}' - \mathbf{r}\|$ achieves its maximum of

$$\frac{2\pi r''}{N} + \frac{2\pi r'}{N} = \frac{2\pi(r'' + r')}{N}. \tag{19}$$

The Nyquist criterion requires that

$$\frac{2\pi(r'' + r')}{N} < \frac{\lambda_c}{2}, \tag{20}$$

which is equivalent to

$$r'' + r' < \frac{N\lambda_c}{4\pi}. \tag{21}$$

One may interpret this condition as follows. The physical propagation of the photoacoustic wave in the object to the detectors is succeeded by a time-reversal propagation for the IR. The combined region encompasses a disc with a radius of $r'' + r'$. On the perimeter of this disc, the Nyquist sampling criterion requires that the sampling spacing be less than half of the lower cutoff wavelength, i.e., Eq. (20).

We denote S_2, referred to as the two-way Nyquist zone here, as

$$S_2 = \left\{\mathbf{r}' \,\middle|\, \|\mathbf{r}'\| < \frac{N\lambda_c}{8\pi}\right\}. \tag{22}$$

Again, we can consider the boundary of S_2 as a virtual detection surface, where the sampling spacing is scaled down from the actual detection spacing by $\frac{R_2}{R}$ with R_2 being the

radius of S_2. If the source points are inside S_2 and we reconstruct at points within S_2, then $r' < \frac{N\lambda_c}{8\pi}$ and $r'' < \frac{N\lambda_c}{8\pi}$, respectively, and Inequality (21) is satisfied. Thus, there is no spatial aliasing during reconstruction, and we call S_2 an aliasing-free region. It needs to be pointed out that, due to the finite duration of the transducer's temporal response, the function $\left(1-\left(t+\frac{\|\mathbf{r}'_m-\mathbf{r}_n\|}{c}\right)\frac{\partial}{\partial t}\right)h'_e(t)$ has nonzero value only when t is within a finite interval, denoted as T_e. The broader the bandwidth of $h_e(t)$, the shorter T_e. When $\frac{\|\mathbf{r}''-\mathbf{r}\|}{c} - \frac{\|\mathbf{r}'-\mathbf{r}\|}{c}$ is out of T_e, signals from source point $\mathbf{r}'$ that are detected by the element at $\mathbf{r}$ have no contribution to the reconstruction at $\mathbf{r}''$.

In the following discussion, we assume that $\frac{\|\mathbf{r}''-\mathbf{r}\|}{c} - \frac{\|\mathbf{r}'-\mathbf{r}\|}{c}$ belongs to T_e. Even with source points inside S_2, we may still have aliasing when reconstruction locations are outside S_2 but inside S_1. To demonstrate this, we assume that both $\mathbf{r}''$ and $\mathbf{r}'$ are on the boundary of S_2, and that the length of the step size of $\|\mathbf{r}''-\mathbf{r}\| - \|\mathbf{r}'-\mathbf{r}\|$ achieves the maximum value of $\frac{2\pi R(\cos\alpha''-\cos\alpha')}{N} = \frac{\lambda_c}{2}$. Here, α'' and α' denote the angles formed by the line segments $\mathbf{r}-\mathbf{r}''$ and $\mathbf{r}-\mathbf{r}'$, respectively, with the tangential dotted line that is perpendicular to $\mathbf{r}$. We move the reconstruction location $\mathbf{r}''$ to a new position $\mathbf{r}''_1$ outside S_2 but inside S_1, as shown in **Figure 3b**. We keep the distance $\|\mathbf{r}''_1-\mathbf{r}\| = \|\mathbf{r}''-\mathbf{r}\|$ constant. Thus, $\frac{\|\mathbf{r}''_1-\mathbf{r}\|}{c} - \frac{\|\mathbf{r}'-\mathbf{r}\|}{c} = \frac{\|\mathbf{r}''-\mathbf{r}\|}{c} - \frac{\|\mathbf{r}'-\mathbf{r}\|}{c}$ still belongs to T_e. As $\mathbf{r}''$ moves to $\mathbf{r}''_1$, the angle α'' decreases to α''_1. Both α'' and α''_1 belong to $\left(0,\frac{\pi}{2}\right)$, then we have $\cos\alpha''_1 > \cos\alpha''$. Thus, for the local step size of $\|\mathbf{r}''_1-\mathbf{r}\| - \|\mathbf{r}'-\mathbf{r}\|$, we have the estimation $\frac{2\pi R(\cos\alpha''_1-\cos\alpha')}{N} > \frac{2\pi R(\cos\alpha''-\cos\alpha')}{N} = \frac{\lambda_c}{2}$, which means that spatial aliasing appears in reconstruction. Switching the source and reconstruction locations, we can repeat the analysis and draw a similar conclusion: with source points inside S_1 but outside S_2, we may have aliasing when reconstruction locations are inside S_2.

We visualize the relative sizes of the three regions S_0, S_1, and S_2 in **Figure 3c**. Spatial aliasing in SS does not appear for objects inside S_1, but appears for objects outside S_1.

Spatial aliasing in IR does not appear for objects and reconstruction locations inside S_2, but appears for other combinations of objects and reconstruction locations. A detailed classification is shown in **Figure 4a**.

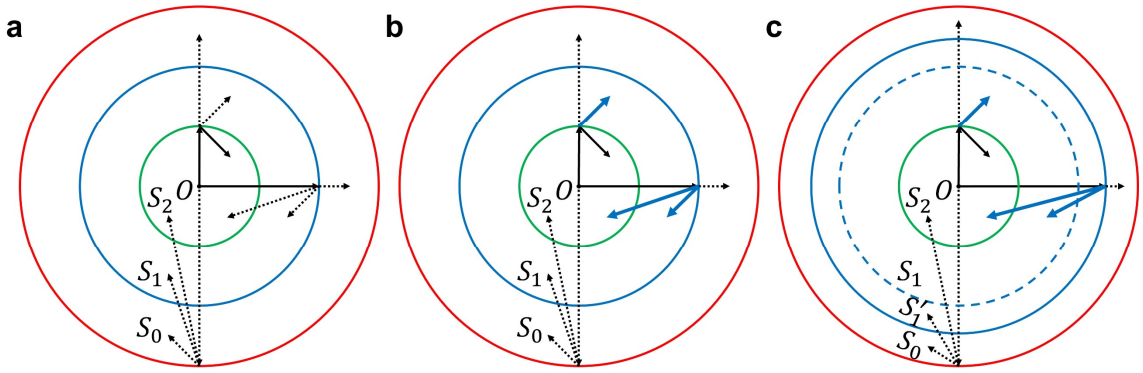

Figure 4 Different combinations of source locations and reconstruction locations subject to spatial aliasing in SS and IR. Three regions S_0, S_1, and S_2 are defined in Eqs. (8), (16), and (22), respectively. The first line radiating from the origin O represents the range of source locations for SS, while the second line radiating from the tip of the first line represents the range of reconstruction locations for IR. A solid line means no aliasing, while a dotted line means aliasing. **a** Spatial aliasing in UBP. The innermost two-way Nyquist zone S_2 is an aliasing-free region. **b** Spatial aliasing in UBP with spatial interpolation. Spatial interpolation removes spatial aliasing in three cases of IR, making the one-way Nyquist zone S_1 an aliasing-free region. The dotted lines representing the three cases in **a** are changed to blue-solid lines in **b**. **c** Spatial aliasing in UBP with temporal filtering and spatial interpolation. Temporal filtering extends the one-way Nyquist zone S_1 in **b** to S_1' in **c**, and the original S_1 is marked as a blue-dashed circle for reference. Spatial interpolation further makes S_1' an aliasing-free region.

Spatial aliasing solely in IR but not in SS can be removed by spatial interpolation. In fact, without spatial aliasing in SS, spatially continuous signals can be accurately recovered from spatially discrete signals through Whittaker–Shannon interpolation. Then, in theory, no aliasing would occur in reconstructing the image using the spatially continuous signals. In practice, the number of detection elements is numerically increased. To clarify the process, at any given time t, we define

$$f_R(\theta) = \hat{p}(\mathbf{r}, t), \tag{23}$$

where $\mathbf{r} = (R\cos\theta, R\sin\theta), \theta \in [0, 2\pi)$. The function $f_R(\theta)$ is sampled at $\theta_n = \frac{2\pi n}{N}, n = 0,1,2,\dots,N-1$. For objects inside the region S_1, SS has no aliasing. Thus, the function $f_R(\theta)$ can be well recovered from $f_R(\theta_n), n = 0,1,2,\dots,N-1$ through spatial Whittaker–Shannon interpolation. To extend the region S_2, we can numerically double the number of

detection elements $N' = 2N$ based on the interpolation. Substituting N' for N in Eq. (22), we obtain a larger region:

$$S_2' = \left\{\mathbf{r}' \middle| \|\mathbf{r}'\| < \frac{N'\lambda_c}{8\pi}\right\} = \left\{\mathbf{r}' \middle| \|\mathbf{r}'\| < \frac{N\lambda_c}{4\pi}\right\} = S_1. \tag{24}$$

From the above discussion about Eq. (22), we know that with source and reconstruction locations inside S_2', IR has no aliasing. From $S_1 = S_2'$, we can see that spatial interpolation successfully removes spatial aliasing in IR. **Figure 4a** is now replaced by **Figure 4b**. For source points outside the region S_1, SS has aliasing, and spatial interpolation cannot recover the lost information, which is mitigated by the next method.

Spatial aliasing in SS can be eliminated by temporal lowpass filtering. We consider a region larger than S_1:

$$S_1' = \{\mathbf{r}''| \|\mathbf{r}''\| < r'\} \text{ with } r' > \frac{N\lambda_c}{4\pi}. \tag{25}$$

We have already shown that source points in this region can produce spatial aliasing during SS. To avoid this concern, before spatial interpolation and reconstruction, we process the signals using a lowpass filter with upper cutoff frequency $f_c' = \frac{Nc}{4\pi r'}$. Replacing λ_c in Eq. (16) with $\lambda_c' = \frac{c}{f_c'}$, we extend the region S_1 to

$$S_1' = \left\{\mathbf{r}'' \middle| \|\mathbf{r}''\| < \frac{N\lambda_c'}{4\pi}\right\} = \{\mathbf{r}''| \|\mathbf{r}''\| < r'\}. \tag{26}$$

Based on the above discussion about Eq. (16), for source points inside S_1', using spatial interpolation, we can reconstruct any points inside S_1' without aliasing artifacts. Thus, we extend the one-way Nyquist zone through temporal lowpass filtering at the expense of spatial resolution and replace **Figure 4b** with **Figure 4c**.

An ideal antialiasing method should extend the region S_1' in **Figure 4c** to the whole region S_0. However, lowpass filtering removes the high-frequency signals, and blurs the reconstructed images. Directly extending S_1' to S_0 would greatly compromise the image resolution. As a balance between spatial antialiasing and high resolution, to reconstruct the image at $\mathbf{r}' \in S_0$, we design the lowpass filter based on its distance to the center $r' = \|\mathbf{r}'\|$.

If $r' < \frac{N\lambda_c}{4\pi}$, we apply spatial interpolation, then perform reconstruction. If $r' \geq \frac{N\lambda_c}{4\pi}$, we first filter the signals with upper cutoff frequency $f_c' = \frac{Nc}{4\pi r'}$, then perform spatial interpolation and reconstruction. We call this method radius-dependent spatiotemporal antialiasing.

2.3.2 Numerical simulations

To demonstrate the proposed method, we perform numerical simulations using a full-ring transducer array of radius $R = 30$ mm. The frequency range of the transducer is from 0.1 MHz to 4.5 MHz. We set the number of detection elements $N = 512$ and the speed of sound $c = 1.5\ \text{mm} \cdot \mu\text{s}^{-1}$. The radius of the one-way Nyquist zone S_1 is thus $r = \frac{Nc}{4\pi f_c} \approx 13.6$ mm.

We simulated two cases with complex structures. In the first case, the object is completely within S_1, as shown in **Figure 5a1**. The reconstructions of the object in **Figure 5a1** using UBP, UBP with spatial interpolation, and UBP with temporal filtering and spatial interpolation are shown in **Figure 5a2**–**a4**, respectively. region of interest (ROIs) A–C ($1.2 \times 1.2\ \text{mm}^2$) were chosen at locations with zero initial pressure. The standard deviations (STDs) were calculated and compared, as shown in **Figure 5b1**. Profiles of lines P and Q are shown in **Figure 5b2** and **b3**, respectively. In the second case, the object is beyond S_1, and covers most of the area inside the full-ring transducer array, as shown in **Figure 5c1**. The reconstructions of the object in **Figure 5c1** using the three methods are shown in **Figure 5c2**–**c4**, respectively. The STDs of ROIs A–C are compared in **Figure 5d1**, while the profiles of lines P and Q are shown in **Figure 5d2** and **d3**, respectively. Although spatial interpolation mitigates the aliasing artifacts in both **Figure 5a2** and **c2**, visible artifacts remain in the red-boxed region in **Figure 5c3** but not in **Figure 5a3**. It can also be seen in **Figure 5b1** and **d1** that spatial interpolation has more obvious antialiasing effects on **Figure 5a2** than on **Figure 5c2**, while temporal filtering's effect on **Figure 5c2** is more obvious than on **Figure 5a2**. In fact, the aliasing artifacts in **Figure 5a2** are solely from the IR, while those in **Figure 5c2** are from both the SS and the IR. Thus, spatial

interpolation works well in antialiasing for **Figure 5a2**, for which temporal filtering's smoothing effect slightly helps; but not as well for **Figure 5c2** due to the spatial aliasing in SS, for which temporal filtering is necessary. In general, the aliasing artifacts are mitigated by spatial interpolation and further diminished by temporal filtering, as shown in both **Figure 5b1** and **d1**. The antialiasing methods maintain the image resolution well inside S_1, as shown in **Figure 5b2**, **b3**, and **d2**. Due to spatial aliasing in SS for objects outside S_1, the profile of line Q is affected by spatial interpolation. Adding temporal filtering further smooths the profile. The full width at half maximum (FWHM) of the main lobe in line Q's profile was increased by temporal filtering while the amplitude was reduced, as shown in **Figure 5d3**.

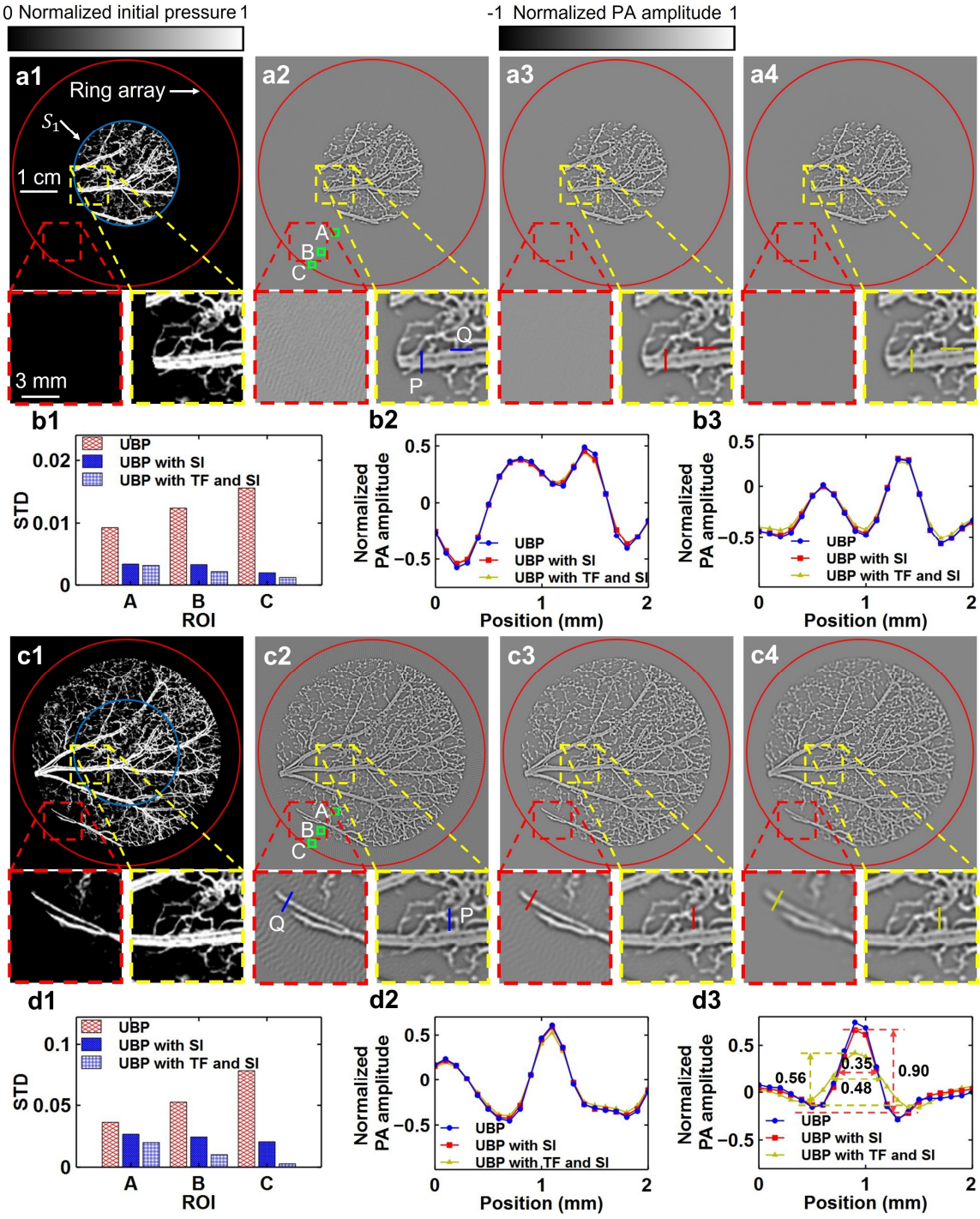

Figure 5 Spatial interpolation and temporal filtering's effects on IR in numerical simulations for two complex phantoms. **a1** Ground truth of a complex initial pressure p_0 distribution confined to S_1. **a2**–**a4** Reconstructions of the object in **a1** using **a2** UBP, **a3** UBP with SI, and **a4** UBP with TF and SI, respectively. SI, spatial interpolation; TF, temporal filtering. The artifacts in the red-boxed region are caused by spatial aliasing in IR, and they are mainly mitigated by SI. **b1** Comparison of the STDs in the ROIs A–C. **b2** and **b3** Comparisons of the profiles of lines P and Q, respectively, for the three

methods. **c1** Ground truth of a complex initial pressure p_0 distribution beyond S_1. **c2**–**c4** Reconstructions of the object in **c1** using **c2** UBP, **c3** UBP with SI, and **c4** UBP with TF and SI, respectively. The artifacts in the red-boxed region are caused by spatial aliasing in SS and IR, and the artifacts are mitigated by TF and SI. **d1** Comparison of the STDs in the ROIs A–C. **d2** and **d3** Comparisons of the profiles of lines P and Q, respectively, for the three methods. The FWHM of the main lobe at Q was increased from 0.35 mm to 0.48 mm by temporal filtering, while the amplitude was changed from 0.90 to 0.56.

2.3.3 *In vivo* experiments

Finally, we applied spatial interpolation and temporal filtering methods to human breast imaging. The imaging system, as shown in **Figure 1**, employed a 512-element full-ring ultrasonic transducer array (Imasonic, Inc., 110-mm radius, 2.25-MHz central frequency, 95% one-way bandwidth). Based on point source measurements, the cutoff frequency is estimated to be $f_c \approx 3.80$ MHz [37]. The acquired signals were filtered by a third-order lowpass Butterworth filter and a sinc filter (both with cutoff frequency 3.80 MHz). Thus, the one-way Nyquist zone S_1 has a radius $r = \frac{Nc}{4\pi f_c} \approx 16.0\,\mathrm{mm}$, while the two-way Nyquist zone S_2 has a radius of $8.0\,\mathrm{mm}$. Here we use the speed of sound $c = 1.49\,\mathrm{mm} \cdot \mu\mathrm{s}^{-1}$.

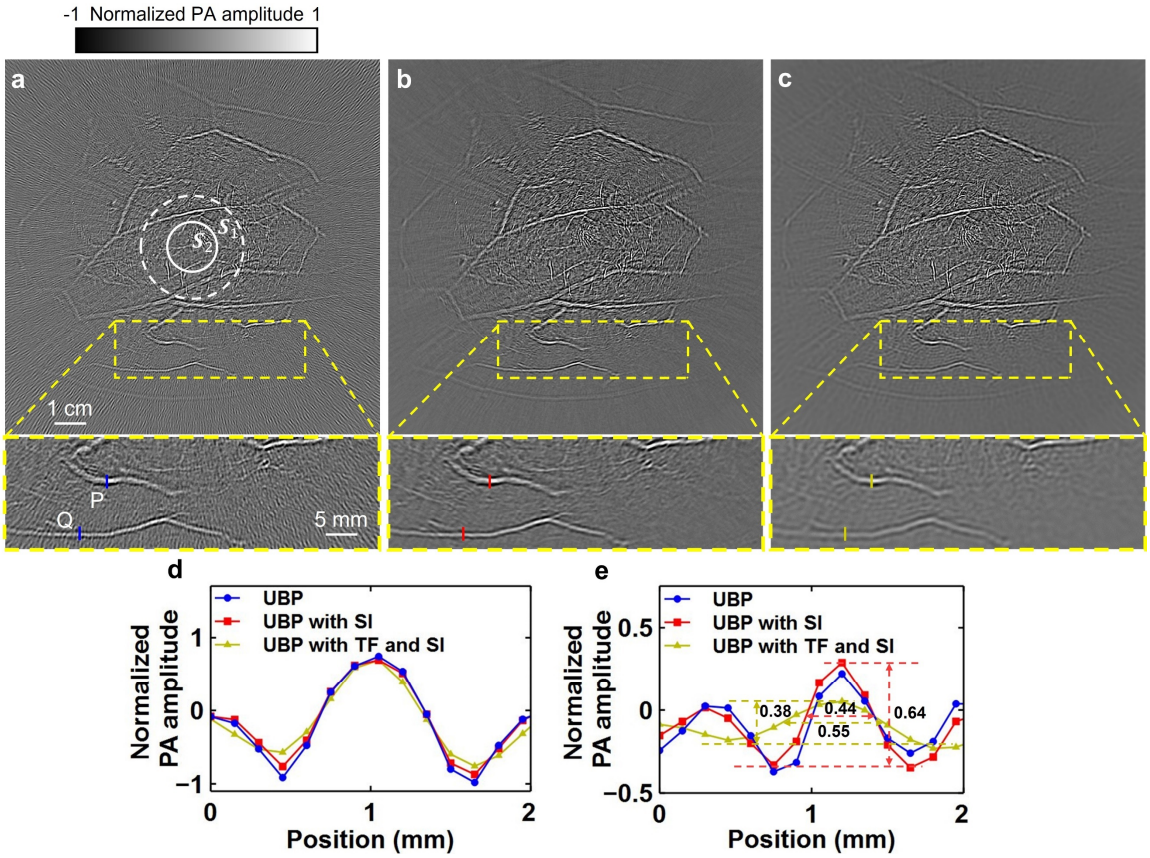

Figure 6 Spatial interpolation and temporal filtering's effects on IR in an *in vivo* human breast image. **a** Reconstructed image using UBP without either spatial interpolation or temporal filtering, and a closeup subset in the yellow-boxed region. Boundaries of S_1 and S_2 are shown as white-dashed and white-solid circles, respectively. **b** and **c** Reconstructions of the same region as **a** using **b** UBP with SI, and **c** UBP with TF and SI, respectively. SI, spatial interpolation; TF, temporal filtering. **d** and **e** Comparisons of the profiles of lines P and Q, respectively, for the three methods. The FWHM of the main lobe at Q was increased from 0.44 mm to 0.55 mm by temporal filtering, while the amplitude was changed from 0.64 to 0.38.

Using UBP, we reconstructed a cross-sectional image of a breast, shown in **Figure 6a**. The aliasing artifacts are obvious in the peripheral regions, as shown in **Figure 6a**'s closeup subset in a yellow-boxed region. After spatial interpolation of the raw data, the reconstructed image is shown in **Figure 6b**. Applying temporal filtering and spatial interpolation, we obtained **Figure 6c**. As can be seen from these subsets, the image quality is improved by spatial interpolation, and the aliasing artifacts are further mitigated by temporal filtering. For comparison, the profiles of lines P and Q for the three images are shown in **Figure 6d** and **e**, respectively. As shown by the numerical simulation, image resolution outside S_1 is compromised by both spatial interpolation and temporal filtering.

Temporal filtering smooths the profiles, as shown in **Figure 6d** and **e**. Quantitatively, as shown in **Figure 6e**, temporal filtering increases the FWHM of the main lobe of line Q's profile and reduces the amplitude.

2.4 Location-dependent spatiotemporal antialiasing

The radius-dependent spatiotemporal antialiasing mitigates aliasing artifacts but compromises the image resolution outside of the one-way Nyquist zone. To avoid the blurring effect, we propose a location-dependent spatiotemporal antialiasing method.

2.4.1 Theory

As demonstrated in [37], spatial aliasing in PACT has two sources: spatial sampling and image reconstruction. Spatial aliasing in spatial sampling and image reconstruction can be explained by analyzing the step sizes of $\frac{\|\mathbf{r}'_m-\mathbf{r}_n\|}{c}$ (in Eq. (6)) and $\frac{\|\mathbf{r}''-\mathbf{r}_n\|}{c}-\frac{\|\mathbf{r}'_m-\mathbf{r}_n\|}{c}$ (in Eq. (17)), respectively, as n varies [37]. Here, we divide the whole image domain D into subdomains; then we analyze these two terms and develop antialiasing strategies for each subdomain.

For simplicity, we focus on 2D image reconstruction and consider only rectangular subdomains. For a subdomain D_{sub} of size $l_x \times l_y$ centered at $\mathbf{r}'_{\text{c,sub}}$, we shift the time t to t' according to $t' = t - \frac{\|\mathbf{r}'_{\text{c,sub}}-\mathbf{r}_n\|}{c}$; instead of analyzing $\hat{p}(\mathbf{r}_n, t)$, we analyze

$$\hat{p}_{D_{\text{sub}}}(\mathbf{r}_n, t') = \hat{p}\left(\mathbf{r}_n, t' + \frac{\left\|\mathbf{r}'_{\text{c,sub}} - \mathbf{r}_n\right\|}{c}\right), n = 1,2,\dots,N, \tag{27}$$

which corresponds to a temporal recentering of signals based on the transducer elements' distances to the subdomain center $\mathbf{r}'_{\text{c,sub}}$. After the recentering, the signals originating from $\mathbf{r}'_{\text{c,sub}}$ arrive at all detectors at time 0, and the exact range of interest for t' is dynamically determined for each subdomain. This recentering is essential in antialiasing to minimize temporal filtering and thus image blurring. Substituting Eq. (6) into Eq. (27) yields

$$\hat{p}_{D_{\mathrm{sub}}}(\mathbf{r}_n, t') = \frac{1}{4\pi c^2}\sum_{m=1}^{M} v_m \frac{p_0(\mathbf{r}'_m)}{\|\mathbf{r}'_m - \mathbf{r}_n\|} h'_{\mathrm{e}}\left(t' - \left(\frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c} - \frac{\left\|\mathbf{r}'_{\mathrm{c,sub}} - \mathbf{r}_n\right\|}{c}\right)\right),$$
$$n = 1,2,\ldots,N. \tag{28}$$

Given the subdomain D_{sub}, we categorize our analysis into four cases with increasing complexities: without point sources outside D_{sub}, with a single point source outside D_{sub}, with multiple point sources outside D_{sub}, and with general sources outside D_{sub}. In all cases, we discuss image reconstruction only in D_{sub}.

A Without point sources outside the subdomain

In the first case, without point sources outside the subdomain D_{sub}, we perform spatiotemporal analysis only for source points and reconstruction locations inside D_{sub}. Let $\mathbf{r}'$ and $\mathbf{r}''$ be a source point and a reconstruction location, respectively, in D_{sub}. Let $\mathbf{r}$ and $\mathbf{r}_{\mathrm{adj}}$ be two adjacent element locations, as shown in **Figure 7a**.

First, we analyze spatial aliasing in spatial sampling based on Eq. (28). We define

$$t'_{\mathbf{r}} = \frac{\|\mathbf{r}' - \mathbf{r}\|}{c} - \frac{\left\|\mathbf{r}'_{\mathrm{c,sub}} - \mathbf{r}\right\|}{c}, \tag{29}$$

and its step size when the element location $\mathbf{r}$ changes to $\mathbf{r}_{\mathrm{adj}}$:

$$\tau\left(\mathbf{r}, \mathbf{r}_{\mathrm{adj}}, \mathbf{r}'_{\mathrm{c,sub}}, \mathbf{r}'\right) = \left|t'_{\mathbf{r}_{\mathrm{adj}}} - t'_{\mathbf{r}}\right| = \left|\begin{array}{l}\left(\frac{\left\|\mathbf{r}' - \mathbf{r}_{\mathrm{adj}}\right\|}{c} - \frac{\|\mathbf{r}' - \mathbf{r}\|}{c}\right) - \\ \left(\frac{\left\|\mathbf{r}'_{\mathrm{c,sub}} - \mathbf{r}_{\mathrm{adj}}\right\|}{c} - \frac{\left\|\mathbf{r}'_{\mathrm{c,sub}} - \mathbf{r}\right\|}{c}\right)\end{array}\right|,$$
$$\mathbf{r}' \in D_{\mathrm{sub}}. \tag{30}$$

For any $\mathbf{r}' \in D_{\mathrm{sub}}$ with $\|\mathbf{r}' - \mathbf{r}\| \neq \left\|\mathbf{r}' - \mathbf{r}_{\mathrm{adj}}\right\|$, there exists a branch (blue-dotted curve in **Figure 7a**) of a hyperbola crossing $\mathbf{r}'$ and with $\mathbf{r}$ and $\mathbf{r}_{\mathrm{adj}}$ as the foci. We denote either one of the intersection points (using the other one leads to the same result) between the branch and the boundary of D_{sub} as $\hat{\mathbf{r}}'$, as shown in **Figure 7a**. For convenience, we denote the boundary of D_{sub} as $\partial D_{\mathrm{sub}}$ in the following discussions. Based on one of the hyperbola's

definitions, we have $\frac{\|\mathbf{r}'-\mathbf{r}_{\mathrm{adj}}\|}{c}-\frac{\|\mathbf{r}'-\mathbf{r}\|}{c}=\frac{\|\hat{\mathbf{r}}'-\mathbf{r}_{\mathrm{adj}}\|}{c}-\frac{\|\hat{\mathbf{r}}'-\mathbf{r}\|}{c}$. For any $\mathbf{r}'\in D_{\mathrm{sub}}$ with $\|\mathbf{r}'-\mathbf{r}\|=\|\mathbf{r}'-\mathbf{r}_{\mathrm{adj}}\|$, $\mathbf{r}'$ will be on the perpendicular bisector (black-dashed line in **Figure 7a**) of the line segment with $\mathbf{r}$ and $\mathbf{r}_{\mathrm{adj}}$ as endpoints. Here, we define

$$\tau_{D_{\mathrm{sub}}}(\mathbf{r},\mathbf{r}_{\mathrm{adj}})=\max_{\mathbf{r}'\in\partial D_{\mathrm{sub}}}\tau(\mathbf{r},\mathbf{r}_{\mathrm{adj}},\mathbf{r}'_{\mathrm{c,sub}},\mathbf{r}'). \tag{31}$$

We choose $\hat{\mathbf{r}}'$ as one of the intersection points between the perpendicular bisector and $\partial D_{\mathrm{sub}}$. Thus, for any $\mathbf{r}'\in D_{\mathrm{sub}}$, replacing $\mathbf{r}'$ with $\hat{\mathbf{r}}'$ in $\tau(\mathbf{r},\mathbf{r}_{\mathrm{adj}},\mathbf{r}'_{\mathrm{c,sub}},\mathbf{r}')$ does not change its value, yielding

$$\tau(\mathbf{r},\mathbf{r}_{\mathrm{adj}},\mathbf{r}'_{\mathrm{c,sub}},\mathbf{r}')=\tau(\mathbf{r},\mathbf{r}_{\mathrm{adj}},\mathbf{r}'_{\mathrm{c,sub}},\hat{\mathbf{r}}')\le\tau_{D_{\mathrm{sub}}}(\mathbf{r},\mathbf{r}_{\mathrm{adj}}),\quad \mathbf{r}'\in D_{\mathrm{sub}}. \tag{32}$$

Through Eq. (32), we simplify the estimation of the upper limit of $\tau(\mathbf{r},\mathbf{r}_{\mathrm{adj}},\mathbf{r}'_{\mathrm{c,sub}},\mathbf{r}')$ from searching $\mathbf{r}'$ in D_{sub} to searching $\mathbf{r}'$ on $\partial D_{\mathrm{sub}}$, which reduces the computation cost by one dimension. Due to spatiotemporal coupling, as shown in the term $t-\left(\frac{\|\mathbf{r}'_m-\mathbf{r}_n\|}{c}-\frac{\|\mathbf{r}'_{\mathrm{c,sub}}-\mathbf{r}_n\|}{c}\right)$ in Eq. (28), the upper cutoff frequency $f_{\mathrm{c,SS}}$ for spatial sampling must meet the Nyquist criterion:

$$f_{\mathrm{c,SS}}<\frac{1}{2\tau_{D_{\mathrm{sub}}}(\mathbf{r},\mathbf{r}_{\mathrm{adj}})}. \tag{33}$$

Denoting

$$f_{\mathrm{c},D_{\mathrm{sub}}}(\mathbf{r}_n)=\min_{\mathbf{r}_{n'}\text{ is adjacent to }\mathbf{r}_n}\frac{1}{2\tau_{D_{\mathrm{sub}}}(\mathbf{r}_n,\mathbf{r}_{n'})},n,n'=1,2,\dots,N, \tag{34}$$

we can remove aliasing in the spatial sampling by processing signals of the element located at $\mathbf{r}_n$ using a lowpass filter with an upper cutoff frequency $f_{\mathrm{c},D_{\mathrm{sub}}}(\mathbf{r}_n),n=1,2,\dots,N$. The filter is implemented as a third-order lowpass Butterworth filter combined with a sinc filter with the same upper cutoff frequency.

Next, we analyze spatial aliasing in the image reconstruction based on Eq. (17). For this analysis, we first estimate the upper limit of the step size of $\frac{\|\mathbf{r}''-\mathbf{r}\|}{c}-\frac{\|\mathbf{r}'-\mathbf{r}\|}{c}$ between two adjacent element locations $\mathbf{r}$ and $\mathbf{r}_{\mathrm{adj}}$:

$$\tau\left(\mathbf{r},\mathbf{r}_{\text{adj}},\mathbf{r}',\mathbf{r}''\right)=\left|\begin{array}{c}\left(\frac{\left\|\mathbf{r}''-\mathbf{r}_{\text{adj}}\right\|}{c}-\frac{\|\mathbf{r}''-\mathbf{r}\|}{c}\right)-\\ \left(\frac{\left\|\mathbf{r}'-\mathbf{r}_{\text{adj}}\right\|}{c}-\frac{\|\mathbf{r}'-\mathbf{r}\|}{c}\right)\end{array}\right|\mathbf{r}',\mathbf{r}''\in D_{\text{sub}}. \tag{35}$$

Based on Eq. (32), we use the triangular inequality to obtain

$$\begin{aligned}\tau\left(\mathbf{r},\mathbf{r}_{\text{adj}},\mathbf{r}',\mathbf{r}''\right)&\leq\tau\left(\mathbf{r},\mathbf{r}_{\text{adj}},\mathbf{r}'_{\text{c,sub}},\mathbf{r}'\right)+\tau\left(\mathbf{r},\mathbf{r}_{\text{adj}},\mathbf{r}'_{\text{c,sub}},\mathbf{r}''\right)\\&\leq 2\tau_{D_{\text{sub}}}\left(\mathbf{r},\mathbf{r}_{\text{adj}}\right),\mathbf{r}',\mathbf{r}''\in D_{\text{sub}}.\end{aligned} \tag{36}$$

From Eq. (32) and Inequality (36) as well as Eq. (34), we conclude that removing aliasing in the image reconstruction can be accomplished by additional lowpass filtering with an upper cutoff frequency $\frac{f_{\text{c},D_{\text{sub}}}(\mathbf{r}_n)}{2}, n=1,2,\dots,N$. We observe that this cutoff frequency is half the value required for removing aliasing in the spatial sampling. This observation agrees with our previous finding [37], where the whole image domain was globally analyzed. This filtering would further compromise the spatial resolution.

Fortunately, the additional filtering is avoided by spatial interpolation if aliasing in the spatial sampling is removed first. We denote the denser element locations after the spatial interpolation as $\mathbf{r}_{\beta,n}, n=1,2,\dots,\beta N$ with β being an integer. Note that the denser element locations coincide with the physical element locations at $\mathbf{r}_n=\mathbf{r}_{\beta,\beta(n-1)+1}, n=1,2,\dots,N$. For each t', the recentered signals ($\hat{p}_{D_{\text{sub}}}(\mathbf{r}_n,t')$) from all the transducer elements form a vector of length N. We apply fast Fourier transform (FFT) to the vector and pad zeros following the highest frequency components to form a new vector of length βN. Then we apply inverse FFT to the new vector to finish the spatial interpolation. Updating Eq. (34) with the denser locations, we obtain

$$\begin{aligned}f_{\text{c},D_{\text{sub}},\beta}\left(\mathbf{r}_{\beta,n}\right)=&\min_{\mathbf{r}_{\beta,n'}\text{ is adjacent to }\mathbf{r}_{\beta,n}}\frac{1}{2\tau_{D_{\text{sub}}}\left(\mathbf{r}_{\beta,n},\mathbf{r}_{\beta,n'}\right)},\\&n,n'=1,2,\dots,\beta N.\end{aligned} \tag{37}$$

From Eqs. (30) and (31), we see that the larger the value of β, the closer the adjacent element locations, the smaller the value of $\tau_{D_{\text{sub}}}\left(\mathbf{r}_{\beta,n},\mathbf{r}_{\beta,n'}\right)$, and the larger the value of $f_{\text{c},D_{\text{sub}},\beta}\left(\mathbf{r}_{\beta,n}\right)$. According to Inequality (36) and Eq. (37), we can filter the signals of the

n-th channel with upper cutoff frequency $\frac{f_{c,D_{sub},\beta}(\mathbf{r}_{\beta,n})}{2}$ to remove aliasing in the image reconstruction. To avoid compromising spatial resolution further, we choose β such that

$$\frac{f_{c,D_{sub},\beta}(\mathbf{r}_{\beta,\beta(n-1)+1})}{2} \geq f_{c,D_{sub}}(\mathbf{r}_n), n = 1,2,\dots,N. \quad (38)$$

Thus, after the first temporal filtering with the upper cutoff frequency of $f_{c,D_{sub}}(\mathbf{r}_n)$ to remove aliasing in the spatial sampling and the spatial interpolation with factor β, we no longer need to perform additional temporal filtering for image reconstruction. Because a general subdomain is off-centered in the image domain, the spatial interpolation is applied to the recentered signals $\hat{p}_{D_{sub}}(\mathbf{r}_n, t')$ instead of the original signals $\hat{p}(\mathbf{r}_n, t)$. In our previous study for the whole image domain [37], we have $\beta = 2$. In this research, the minimal β satisfying Inequality (38) is obtained through numerical computations for each subdomain.

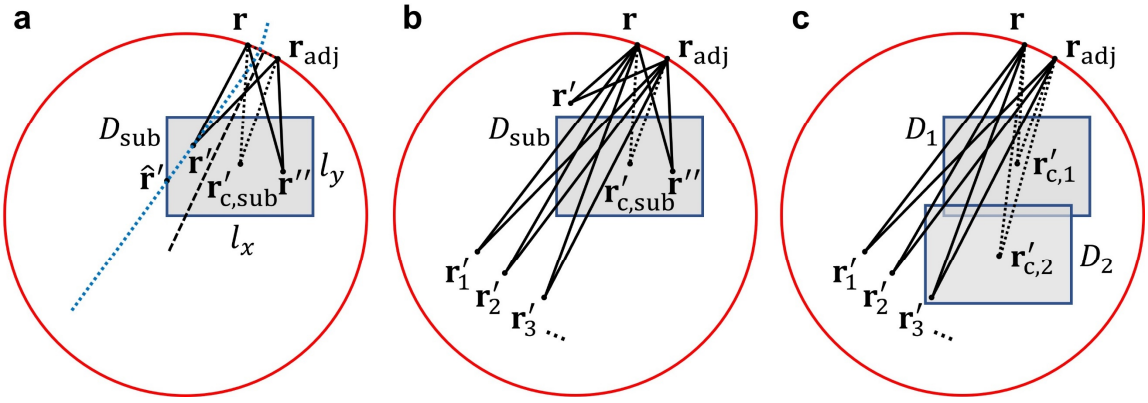

Figure 7 Location-dependent spatiotemporal analysis. **a** No point sources outside the subdomain. A full-ring transducer array (red circle), an image subdomain (rectangle with gray interior and blue boundary, denoted as D_{sub}, centered at $\mathbf{r}'_{c,sub}$, and of size $l_x \times l_y$), two adjacent element locations $\mathbf{r}$ and $\mathbf{r}_{adj}$, a source point $\mathbf{r}'$ inside D_{sub}, and a reconstruction location $\mathbf{r}''$ inside D_{sub}. There exists a hyperbola with $\mathbf{r}$ and $\mathbf{r}_{adj}$ as the foci, and with one branch (blue dotted curve) crossing $\mathbf{r}'$. One of the branch's intersection points with the boundary of D_{sub} is denoted as $\hat{\mathbf{r}}'$. This graph is used in the spatiotemporal analysis for D_{sub} without point sources outside. **b** One or multiple point sources outside the subdomain. A full-ring transducer array, an image subdomain D_{sub} centered at $\mathbf{r}'_{c,sub}$, two adjacent element locations $\mathbf{r}$ and $\mathbf{r}_{adj}$, and a reconstruction location $\mathbf{r}''$ inside D_{sub}. We have a single source point ($\mathbf{r}'$) or multiple source points ($\mathbf{r}'_1, \mathbf{r}'_2, \mathbf{r}'_3$, ...) outside D_{sub}. This graph is used in the spatiotemporal analysis for D_{sub} with a single or multiple point sources outside. **c** Multiple point sources outside multiple subdomains. A full-ring transducer array, two image subdomains D_1 and D_2 (centered at $\mathbf{r}'_{c,1}$ and $\mathbf{r}'_{c,2}$, respectively), and a group of source points: $\mathbf{r}'_1, \mathbf{r}'_2, \mathbf{r}'_3$, ...

In summary, to implement spatiotemporal antialiasing in D_{sub}, we first apply location-dependent temporal filtering (LDTF) to the recentered signals $\hat{p}_{D_{\mathrm{sub}}}(\mathbf{r}_n, t')$ of each element with an upper cutoff frequency $f_{\mathrm{c},D_{\mathrm{sub}}}(\mathbf{r}_n)$, and obtain $\hat{p}_{D_{\mathrm{sub}},\mathrm{LDTF}}(\mathbf{r}_n, t'), n = 1,2,\dots,N$. Then we apply spatial interpolation with a factor β to the filtered signals and obtain $\tilde{p}_{D_{\mathrm{sub}}}(\mathbf{r}_{\beta,n}, t'), n = 1,2,\dots,\beta N$. After reversing the temporal recentering of the signals based on the relation

$$\tilde{p}(\mathbf{r}_{\beta,n}, t) = \tilde{p}_{D_{\mathrm{sub}}}\left(\mathbf{r}_{\beta,n}, t - \frac{\left\|\mathbf{r}'_{\mathrm{c,sub}} - \mathbf{r}_{\beta,n}\right\|}{c}\right), n = 1,2,\dots,\beta N, t \geq 0, \tag{39}$$

we reconstruct the image in D_{sub} using Eq. (7) for the denser element locations.

B With a single point source outside the subdomain

We increase the complexity of our theory by adding a single point source at $\mathbf{r}'$ outside the subdomain D_{sub}, as shown in **Figure 7b**. For the single source point $\mathbf{r}'$, Eq. (28) reduces to

$$\hat{p}_{D_{\mathrm{sub}}}(\mathbf{r}_n, t') = \frac{v p_0(\mathbf{r}')}{4\pi c^2 \|\mathbf{r}' - \mathbf{r}_n\|} = h'_{\mathrm{e}}\left(t' - \left(\frac{\|\mathbf{r}' - \mathbf{r}_n\|}{c} - \frac{\left\|\mathbf{r}'_{\mathrm{c,sub}} - \mathbf{r}_n\right\|}{c}\right)\right), \quad n = 1,2,\dots,N, \tag{40}$$

where $p_0(\mathbf{r}')$ is the initial pressure at $\mathbf{r}' \in D \backslash D_{\mathrm{sub}}$ and v is the volume of the source point. Similarly, Eq. (17) reduces to

$$\hat{p}_0(\mathbf{r}'') \approx \frac{v p_0(\mathbf{r}')}{2\pi c^2} \sum_{n=1}^{N} \frac{w_n}{\|\mathbf{r}' - \mathbf{r}_n\|}\left(1 - \left(t + \frac{\|\mathbf{r}' - \mathbf{r}_n\|}{c}\right)\frac{\partial}{\partial t}\right) h'_{\mathrm{e}}\left(\frac{\|\mathbf{r}'' - \mathbf{r}_n\|}{c} - \frac{\|\mathbf{r}' - \mathbf{r}_n\|}{c}\right), \mathbf{r}'' \in D_{\mathrm{sub}}. \tag{41}$$

Here, we confine the image reconstruction in D_{sub} by letting $\mathbf{r}'' \in D_{\mathrm{sub}}$.

We first analyze spatial aliasing in spatial sampling for signals from the source point $\mathbf{r}'$. Based on Eq. (40), the spatial aliasing is determined by the step size of $\frac{\|\mathbf{r}' - \mathbf{r}_n\|}{c} - \frac{\left\|\mathbf{r}'_{\mathrm{c,sub}} - \mathbf{r}_n\right\|}{c}$ as n varies. Using Eq. (30), we express the step size as $\tau(\mathbf{r}_n, \mathbf{r}_{n'}, \mathbf{r}'_{\mathrm{c,sub}}, \mathbf{r}')$ for

adjacent element locations $\mathbf{r}_n$ and $\mathbf{r}_{n'}$, and we define the upper cutoff frequency as

$$f_{\mathrm{c,SS},D_{\mathrm{sub}},\mathrm{OS}}(\mathbf{r}_n,\mathbf{r}') = \min_{\mathbf{r}_{n'}\text{ is adjacent to }\mathbf{r}_n} \frac{1}{2\tau\left(\mathbf{r}_n,\mathbf{r}_{n'},\mathbf{r}'_{\mathrm{c,sub}},\mathbf{r}'\right)}, \quad n,n' = 1,2,\dots,N. \tag{42}$$

To remove aliasing in spatial sampling (SS in the subscript) for signals from the source point $\mathbf{r}'$ outside the subdomain (OS in the subscript), we apply lowpass filtering to $\hat{p}_{D_{\mathrm{sub}}}(\mathbf{r}_n,t')$ with the above upper cutoff frequency. We assume that the value of $h'_{\mathrm{e}}(t)$ is nonzero only for t in an interval $[0,T_{\mathrm{e}}]$ ([0, 1.8 μs] for this research), which is often small for the transducers used in PACT. To minimize unwanted smoothing of signals, we filter $\hat{p}_{D_{\mathrm{sub}}}(\mathbf{r}_n,t')$ only for $t' - \left(\frac{\|\mathbf{r}'-\mathbf{r}_n\|}{c} - \frac{\|\mathbf{r}'_{\mathrm{c,sub}}-\mathbf{r}_n\|}{c}\right) \in [0,T_{\mathrm{e}}]$.

Then we analyze spatial aliasing in the image reconstruction for signals from the source point $\mathbf{r}'$ based on Eq. (41). We estimate the upper limit of the step size of $\frac{\|\mathbf{r}''-\mathbf{r}_n\|}{c} - \frac{\|\mathbf{r}'-\mathbf{r}_n\|}{c}$ as n varies using

$$\begin{aligned}\tau\left(\mathbf{r},\mathbf{r}_{\mathrm{adj}},\mathbf{r}',\mathbf{r}''\right) &\le \tau\left(\mathbf{r},\mathbf{r}_{\mathrm{adj}},\mathbf{r}',\mathbf{r}'_{\mathrm{c,sub}}\right) + \tau\left(\mathbf{r},\mathbf{r}_{\mathrm{adj}},\mathbf{r}'_{\mathrm{c,sub}},\mathbf{r}''\right)\\ &\le \tau\left(\mathbf{r},\mathbf{r}_{\mathrm{adj}},\mathbf{r}',\mathbf{r}'_{\mathrm{c,sub}}\right) + \tau_{D_{\mathrm{sub}}}\left(\mathbf{r},\mathbf{r}_{\mathrm{adj}}\right),\\ &\mathbf{r}' \in D\backslash D_{\mathrm{sub}}, \mathbf{r}'' \in D_{\mathrm{sub}}.\end{aligned} \tag{43}$$

To remove aliasing in the image reconstruction (IR in the following subscript) for signals from the source point $\mathbf{r}'$, we can apply a second lowpass filtering to $\hat{p}_{D_{\mathrm{sub}}}(\mathbf{r}_n,t')$ with an upper cutoff frequency

$$f_{\mathrm{c,IR},D_{\mathrm{sub}},\mathrm{OS}}(\mathbf{r}_n,\mathbf{r}') = \min_{\mathbf{r}_{n'}\text{ is adjacent to }\mathbf{r}_n} \frac{1}{2\left(\tau\left(\mathbf{r}_n,\mathbf{r}_{n'},\mathbf{r}',\mathbf{r}'_{\mathrm{c,sub}}\right) + \tau_{D_{\mathrm{sub}}}(\mathbf{r}_n,\mathbf{r}_{n'})\right)}, \quad n,n' = 1,2,\dots,N. \tag{44}$$

To avoid the second lowpass filtering, after the first lowpass filtering, we apply spatial interpolation to the filtered signals and obtain the interpolated signals at virtual locations $\mathbf{r}_{\beta,n}, n = 1,2,\dots,\beta N$. Updating Eq. (44) with these virtual locations, we obtain

$$f_{\mathrm{c,IR},D_{\mathrm{sub}},\mathrm{OS},\beta}\left(\mathbf{r}_{\beta,n},\mathbf{r}'\right) =$$

$$\min_{\mathbf{r}_{\beta,n'} \text{ is adjacent to } \mathbf{r}_{\beta,n}} \frac{1}{2\left(\tau\left(\mathbf{r}_{\beta,n}, \mathbf{r}_{\beta,n'}, \mathbf{r}', \mathbf{r}'_{\mathrm{c,sub}}\right) + \tau_{D_{\mathrm{sub}}}\left(\mathbf{r}_{\beta,n}, \mathbf{r}_{\beta,n'}\right)\right)},$$
$$n, n' = 1,2, \dots, \beta N. \tag{45}$$

Here, we still have the relation $\mathbf{r}_n = \mathbf{r}_{\beta,\beta(n-1)+1}, n = 1,2, \dots, N$. We choose a minimal β such that

$$f_{\mathrm{c,IR},D_{\mathrm{sub}},\mathrm{OS},\beta}\left(\mathbf{r}_{\beta,\beta(n-1)+1}, \mathbf{r}'\right) \geq f_{\mathrm{c,SS},D_{\mathrm{sub}},\mathrm{OS}}(\mathbf{r}_n, \mathbf{r}'), n = 1,2, \dots, N. \tag{46}$$

The factor β is obtained through numerical computations.

In practice, we have signals from both the subdomain D_{sub} and the source point $\mathbf{r}'$. First, we rewrite Eq. (42) (for the source point $\mathbf{r}'$), which applies to only the time domain subsets, as

$$f_{\mathrm{c,SS},D_{\mathrm{sub}},\mathrm{OS}}(\mathbf{r}_n, t', \mathbf{r}') = \begin{cases} f_{\mathrm{c,SS},D_{\mathrm{sub}},\mathrm{OS}}(\mathbf{r}_n, \mathbf{r}'), & t' - t'_0 \in T_{\mathrm{e}} \\ f_{\mathrm{c,IS}}, & \text{else} \end{cases},$$
$$n = 1,2, \dots, N, \tag{47}$$

which applies to the whole time domain. Here, $t'_0 = \frac{\|\mathbf{r}' - \mathbf{r}_n\|}{c} - \frac{\|\mathbf{r}'_{\mathrm{c,sub}} - \mathbf{r}_n\|}{c}$ denotes the recentered first arrival time from $\mathbf{r}'$ to $\mathbf{r}_n$, and $f_{\mathrm{c,IS}}$ denotes the upper cutoff frequency of the imaging system (IS in the subscript). Next, we combine the upper cutoff frequencies in Eq. (34) (for the subdomain D_{sub}) and Eq. (47) to yield the following upper cutoff frequency of the recentered signal $\hat{p}_{D_{\mathrm{sub}}}(\mathbf{r}_n, t')$:

$$f_{\mathrm{c},D_{\mathrm{sub}},\mathbf{r}'}(\mathbf{r}_n, t') = \min\left\{f_{\mathrm{c},D_{\mathrm{sub}}}(\mathbf{r}_n), f_{\mathrm{c,SS},D_{\mathrm{sub}},\mathrm{OS}}(\mathbf{r}_n, t', \mathbf{r}')\right\}. \tag{48}$$

For the n-th element at time t', by applying lowpass filtering with the above upper cutoff frequency, we remove aliasing in spatial sampling for signals from both the subdomain D_{sub} and the source point $\mathbf{r}'$.

This LDTF process is computationally intensive if implemented directly. For fast reconstruction, we give an efficient implementation of the LDTF through precomputation and interpolation. Before the reconstruction of any subdomain, we process the original signals $\hat{p}(\mathbf{r}_n, t)$ of each element using lowpass filters with upper cutoff frequencies of $f_{\mathrm{c},k}, k = 1,2, \dots, K+1$ satisfying $0 < f_{\mathrm{c},1} < f_{\mathrm{c},2} < \cdots < f_{\mathrm{c},K} < f_{\mathrm{c},K+1} = f_{\mathrm{c}}$. Here, a lowpass filter with an upper cutoff frequency means a third-order lowpass Butterworth

filter followed by a sinc filter with the same upper cutoff frequency. We denote the filtered signals as $\hat{p}_{f_{\mathrm{c},k}}(\mathbf{r}_n, t), k = 1,2,\dots,K+1$. For reconstruction of a subdomain D_{sub}, we recenter the filtered signals based on Eq. (27) and obtain

$$\hat{p}_{D_{\mathrm{sub}},f_{\mathrm{c},k}}(\mathbf{r}_n, t') = \hat{p}_{f_{\mathrm{c},k}}\left(\mathbf{r}_n, t' + \frac{\left\|\mathbf{r}'_{\mathrm{c,sub}} - \mathbf{r}_n\right\|}{c}\right), \quad n = 1,2,\dots,N, k = 1,2,\dots,K+1. \tag{49}$$

For a general cutoff frequency $f > 0$, we obtain the filtered signals through the following linear interpolation:

$$\hat{p}_{D_{\mathrm{sub}},f}(\mathbf{r}_n, t') = \begin{cases} \dfrac{f}{f_{\mathrm{c},1}}\hat{p}_{D_{\mathrm{sub}},f_{\mathrm{c},1}}(\mathbf{r}_n, t'), & 0 < f \le f_{\mathrm{c},1} \\ \dfrac{f_{\mathrm{c},k+1} - f}{f_{\mathrm{c},k+1} - f_{\mathrm{c},k}}\hat{p}_{D_{\mathrm{sub}},f_{\mathrm{c},k}}(\mathbf{r}_n, t') + \dfrac{f - f_{\mathrm{c},k}}{f_{\mathrm{c},k+1} - f_{\mathrm{c},k}}\hat{p}_{D_{\mathrm{sub}},f_{\mathrm{c},k+1}}(\mathbf{r}_n, t'), & f_{\mathrm{c},k} < f \le f_{\mathrm{c},k+1}, \\ & k = 1,2,\dots,K \\ \hat{p}_{D_{\mathrm{sub}},f_{\mathrm{c}}}(\mathbf{r}_n, t'), & f > f_{\mathrm{c},K+1} = f_{\mathrm{c}} \end{cases}$$
$$n = 1,2,\dots,N, f > 0. \tag{50}$$

In practice, we let the upper cutoff frequencies $f_{\mathrm{c},k}, k = 1,2,\dots,K+1$ be dense enough so that further increasing their density has minor effects on the reconstructed images. We substitute $f_{\mathrm{c},D_{\mathrm{sub}},\mathbf{r}'}(\mathbf{r}_n, t')$ for f in Eq. (50) to obtain the filtered signals $\hat{p}_{D_{\mathrm{sub}},\mathrm{LDTF},\mathbf{r}'}(\mathbf{r}_n, t')$.

Further, we apply to the filtered signals a spatial interpolation with the minimal factor β satisfying Inequalities (38) and (46), and reverse the recentering of the interpolated signals. We finish the reconstruction in the subdomain D_{sub} based on Eq. (7) for the virtual locations and denote the reconstructed image in D_{sub} as $\hat{p}_{0,D_{\mathrm{sub}},\mathbf{r}'}(\mathbf{r}''), \mathbf{r}'' \in D_{\mathrm{sub}}$.

C With multiple point sources outside the subdomain

We further extend our theory to include multiple point sources outside the subdomain D_{sub}. We denote the set of source points outside D_{sub} as $G = \{\mathbf{r}'_1, \mathbf{r}'_2, \mathbf{r}'_3, \dots\}$, as shown in **Figure 7b**, and update the upper cutoff frequency as follows:

$$f_{\mathrm{c},D_{\mathrm{sub}},G}(\mathbf{r}_n,t') = \min\left\{f_{\mathrm{c},D_{\mathrm{sub}}}(\mathbf{r}_n), \min_{\mathbf{r}'\in G} f_{\mathrm{c,SS},D_{\mathrm{sub}},\mathrm{OS}}(\mathbf{r}_n,t',\mathbf{r}')\right\},$$
$$n = 1,2,\dots,N. \tag{51}$$

To remove spatial aliasing in the spatial sampling for signals from the subdomain D_{sub} and the source points in G, we apply lowpass filtering with the above upper cutoff frequency to the recentered signal $\hat{p}_{D_{\mathrm{sub}}}(\mathbf{r}_n,t')$ of the n-th element at time t', and obtain $\hat{p}_{D_{\mathrm{sub}},\mathrm{LDTF},G}(\mathbf{r}_n,t')$ through Eq. (50). Then we find the minimal β such that Inequalities (38) and (46) are satisfied for all $n = 1,2,\dots,N$ and $\mathbf{r}' \in G$. Spatially interpolating the filtered signals with the factor β, reversing the recentering of the interpolated signals, and using Eq. (7) for the denser locations, we reconstruct the image for the subdomain D_{sub}, denoted as $\hat{p}_{0,D_{\mathrm{sub}},G}(\mathbf{r}''), \mathbf{r}'' \in D_{\mathrm{sub}}$.

D With general sources outside the subdomain

In the fourth case, we consider general sources outside the subdomain D_{sub}. A direct method for LDTF with general sources is selecting all voxels outside D_{sub} as source points (grouped as G) and using Eq. (51) to obtain the upper cutoff frequencies for lowpass filtering. However, this direct method causes severe blurring in the reconstructed images due to unwanted filtering. To minimize unwanted filtering during spatiotemporal antialiasing, we select multiple sets of sparsely distributed source points in the image domain, denoted as $G_1, G_2, \dots, G_J$. We repeat the process in the previous case for each G_j and obtain an image of the subdomain D_{sub}, denoted as $\hat{p}_{0,D_{\mathrm{sub}},G_j}(\mathbf{r}''), \mathbf{r}'' \in D_{\mathrm{sub}}, j = 1,2,\dots,J$. The final image for the subdomain D_{sub} is obtained through averaging

$$\hat{p}_{0,D_{\mathrm{sub}}}(\mathbf{r}'') = \frac{1}{J}\sum_{j=1}^{J} \hat{p}_{0,D_{\mathrm{sub}},G_j}(\mathbf{r}''), \mathbf{r}'' \in D_{\mathrm{sub}}. \tag{52}$$

For $\mathbf{r}'' \in D\backslash D_{\mathrm{sub}}$, we define $\hat{p}_{0,D_{\mathrm{sub}}}(\mathbf{r}'')$ as zero.

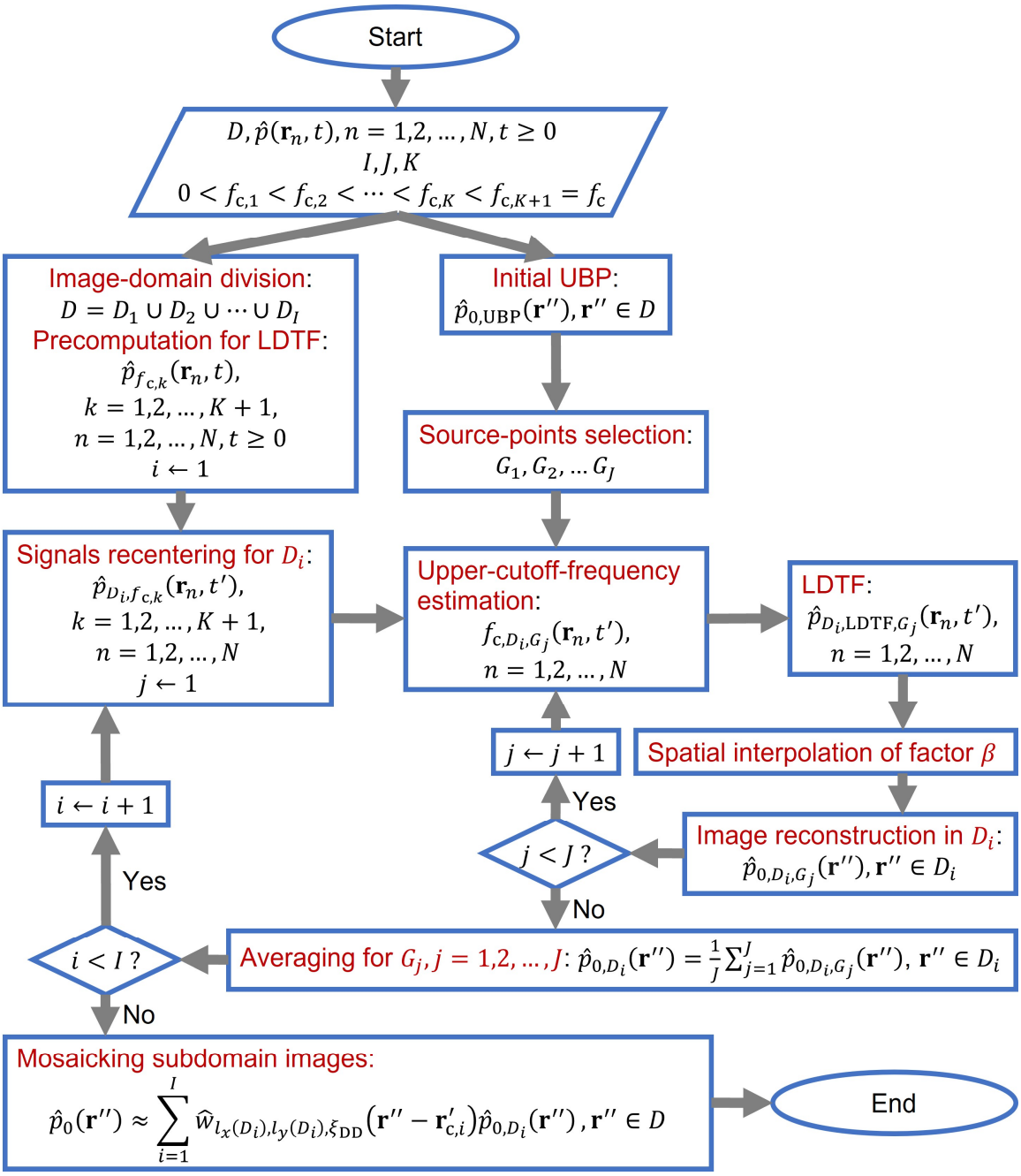

Figure 8 Workflow of the location-dependent spatiotemporal antialiasing for PACT.

We clarify the reconstruction of a single subdomain through the analysis in the above four cases. To reconstruct a whole image, we divide the whole image domain D into multiple subdomains $D_1, D_2, \dots, D_I$. In **Figure 7c**, we depict two subdomains with a group of outside point sources. To form the whole image, these subdomains must satisfy

$$D = D_1 \cup D_2 \cup \ldots \cup D_I. \tag{53}$$

To mitigate artifacts caused by pixel-value mismatch on subdomain boundaries, we overlap adjacent subdomains by a length of ξ_{DD}. Then, for each subdomain D_i, we repeat the process described in the fourth case above to obtain $\hat{p}_{0,D_i}(\mathbf{r}''), i = 1,2, \ldots I$. Finally, we mosaic these subdomain images to form the whole image:

$$\hat{p}_0(\mathbf{r}'') \approx \sum_{i=1}^{I} \widehat{w}_{l_x(D_i),l_y(D_i),\xi_{\mathrm{DD}}}\left(\mathbf{r}'' - \mathbf{r}'_{\mathrm{c},i}\right)\hat{p}_{0,D_i}(\mathbf{r}''), \mathbf{r}'' \in D. \tag{54}$$

Here, $l_x(D_i)$ and $l_y(D_i)$ denote the sizes of the rectangle D_i in x-axis and y-axis directions, respectively, and $\mathbf{r}'_{\mathrm{c},i}$ is the center of D_i. We obtain the normalized weight function $\widehat{w}$ from

$$\widehat{w}_{l_x(D_i),l_y(D_i),\xi_{\mathrm{DD}}}\left(\mathbf{r}'' - \mathbf{r}'_{\mathrm{c},i}\right) = \frac{w_{l_x(D_i),l_y(D_i),\xi_{\mathrm{DD}}}\left(\mathbf{r}'' - \mathbf{r}'_{\mathrm{c},i}\right)}{\sum_{i'=1}^{I} w_{l_x(D_{i'}),l_y(D_{i'}),\xi_{\mathrm{DD}}}\left(\mathbf{r}'' - \mathbf{r}'_{\mathrm{c},i'}\right)},$$
$$i = 1,2, \ldots, I, \tag{55}$$

where $l_x(D_i)$ and $l_y(D_i)$ denote the sizes of the rectangle D_i in x-axis and y-axis directions, respectively, and $\mathbf{r}'_{\mathrm{c},i}$ is the center of D_i. The weight function w is defined as

$$w_{l_x,l_y,\xi}(\mathbf{r}) = w_{l_x,l_y,\xi}(x,y) = w_{l_x,\xi}(x)w_{l_y,\xi}(y), \tag{56}$$

where the 1D weight function is of the form

$$w_{l,\xi}(x) = \begin{cases} 1, & |x| \le \frac{l}{2} \\ 1 - \frac{2|x| - l}{\xi}, & \frac{l}{2} < |x| \le \frac{l+\xi}{2} \\ 0, & |x| > \frac{l+\xi}{2} \end{cases}. \tag{57}$$

In summary, we have the general workflow of the location-dependent spatiotemporal antialiasing for PACT, shown in **Figure 8**.

2.4.2 Numerical simulations

We validate the proposed method using a numerical phantom shown in **Figure 9a1**, which consists of blood vessels with complex structures. Two subsets of the complex numerical phantom are shown in **Figure 9b1** and **c1**, respectively. Applying the parameter-tuning strategy [39] to the complex numerical phantom, we choose $(l_{\mathrm{SP}}, \alpha, J, l_{\mathrm{DD}}, \xi_{\mathrm{DD}}, T_{\mathrm{o}})$ to be

(3.6 mm, 0.08, 36, 18 mm, 1.8 mm, 1.2 μs). Reconstructions of the complex numerical phantom using UBP, UBP with radius-dependent temporal filtering (RDTF, our previous method in [37]) and spatial interpolation, and UBP with LDTF and spatial interpolation are shown in **Figure 9a2**–**a4**, respectively. The subdomain images reconstructed by the three methods are shown in **Figure 9b2-c2**, **b3-c3**, and **b4-c4**, respectively. Comparing these subdomain images, we see that LDTF with spatial interpolation more effectively mitigates aliasing artifacts than RDTF with spatial interpolation. Moreover, using numerical simulations of the complex numerical phantom, we demonstrate the advantage of using the location-dependent parameter $T_{\mathrm{b}}(D_i)$ over using a constant parameter T_{o} across all subdomains for temporal filtering [39].

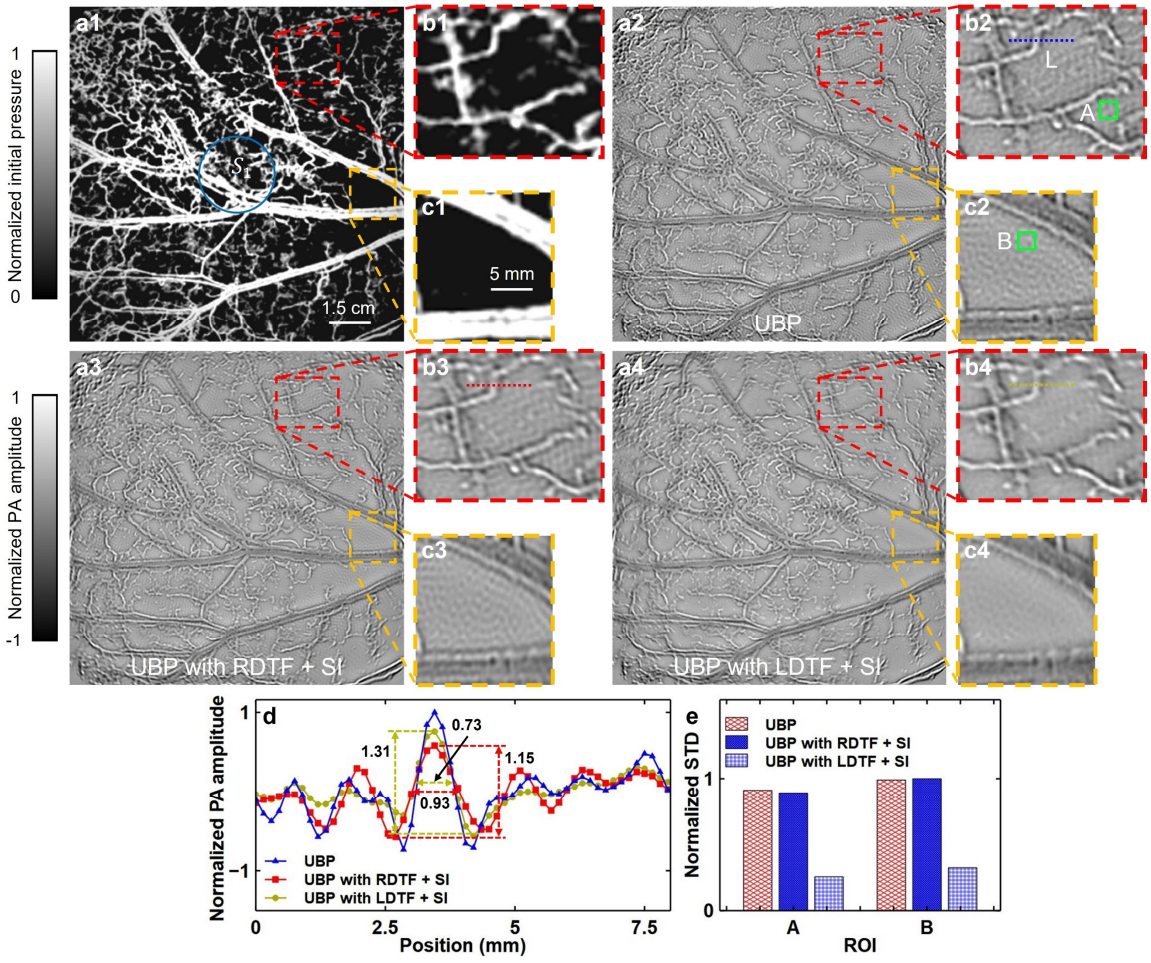

Figure 9 Applying spatiotemporal antialiasing to the image reconstruction of a numerical phantom with complex blood vessel structures. **a1**–**c1** A numerical phantom consisting of complex blood

vessel structures, and its closeup subsets. **a2**–**c2**, **a3**–**c3**, and **a4**–**c4** Images of the complex numerical phantom reconstructed using UBP, UBP with RDTF and spatial interpolation (SI), and UBP with LDTF and SI, respectively, and their closeup subsets. **d** Comparisons of the values along the line L marked in **b2** for the three methods. The FWHM of the dominant lobe is 0.93 mm for UBP with RDTF and SI, and 0.73 mm for UBP with LDTF and SI. The amplitudes are 1.15 and 1.31, respectively. **e** Comparisons of the STDs of pixel values in regions A and B marked in **b2** and **c2**, respectively, for the three methods.

For quantitative comparisons, we draw a line L (marked in **Figure 9b2**) in the red-boxed subdomain and pick two small regions A and B (marked in **Figure 9b2** and **c2**, respectively) in different subdomains for the three methods. The values along the line L are shown in **Figure 9d** while the STDs of the pixel values in A and B are compared in **Figure 9e**. In **Figure 9d**, we still see that RDTF with spatial interpolation blurs the image and introduces new artifacts; whereas LDTF with spatial interpolation blurs the image to a smaller degree and mitigates the aliasing artifacts. The blurring effect of LDTF with spatial interpolation is more obvious in the complex numerical phantom than in the simple numerical phantom. In fact, compared with the simple numerical phantom, more source points are selected for LDTF in the complex numerical phantom, which results in the filtering of more signals and more blurring of the image. Thus, for images with complex structures, we find a balance between mitigating aliasing artifacts and maintaining image resolution by tuning the parameters. Because of this balance, aliasing artifacts still appear in **Figure 9b4**. In **Figure 9e**, we still observe that LDTF with spatial interpolation is better than RDTF with spatial interpolation in mitigating aliasing artifacts.

Due to the more intricate temporal filtering for antialiasing, the proposed method has a significantly higher computation cost. On a computer with Windows 10 Home and Intel® Core™ i7-6700 CPU @ 3.40 GHz, the reconstructions of the simple numerical phantom through UBP, UBP with RDTF and spatial interpolation, and UBP with LDTF and spatial interpolation (single-thread implementations) take 17.6 s, 98.8 s, and 682.7 s (average values for 10 repetitions), respectively. For the complex numerical phantom, the computation times of the first two methods do not change but the third one takes 1345.5 s due to the differences in reconstruction parameters. All methods can be accelerated through GPU. For example, using an NVIDIA GeForce GTX 1050 Ti graphics card, we reduce the

computation time of the UBP method from 17.6 s to 0.54 s. Although not demonstrated in this study, an efficient GPU acceleration of the proposed method is preferred for faster parameter tuning and image reconstruction in future studies.

2.4.3 *In vivo* experiments

Finally, we apply the proposed LDTF with spatial interpolation to human breast imaging *in vivo*. The cutoff frequency has been estimated to be $f_{c,IS} \approx 3.80$ MHz. Here we use the speed of sound $c = 1.49\ \text{mm} \cdot \mu\text{s}^{-1}$. A cross-sectional image of a breast *in vivo* is reconstructed using UBP and shown in **Figure 10a1**. The result of UBP with RDTF and spatial interpolation is shown in **Figure 10a2**.

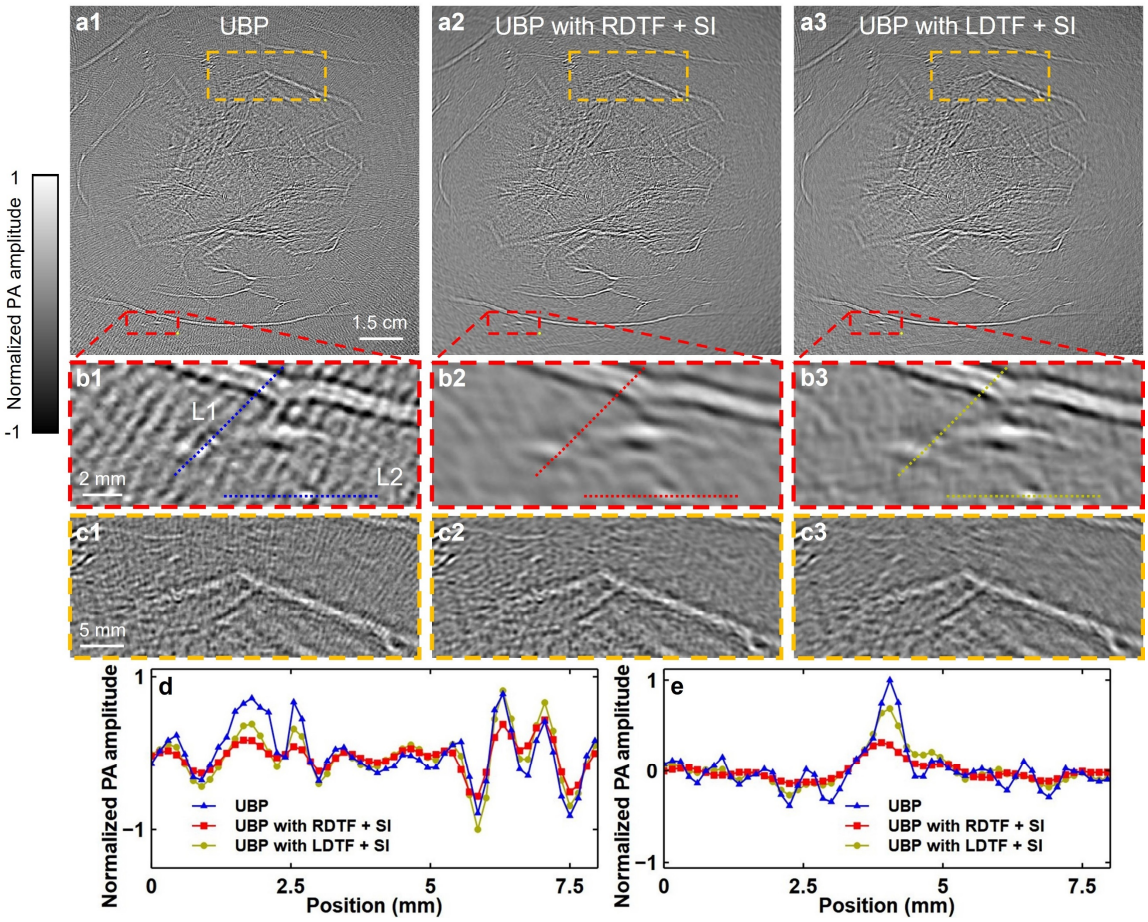

Figure 10 Applying spatiotemporal antialiasing to human breast imaging *in vivo*. **a1**–**a3** Images of a human breast cross section *in vivo* reconstructed using **a1** UBP, **a2** UBP with RDTF and spatial interpolation (SI), and **a3** UBP with LDTF and SI. Two subdomains in the red-dashed box and the

yellow-dashed box, respectively, are picked for comparisons of the three methods. **b1**-**c1**, **b2**-**c2**, and **b3**-**c3** Closeup images of the two subdomains for the three methods, respectively. Lines L1 and L2 are picked for comparisons. **d** and **e** Values along lines L1 and L2, respectively, for the three methods.

We use $(l_{\mathrm{SP}}, \alpha, J, l_{\mathrm{DD}}, \xi_{\mathrm{DD}}, T_{\mathrm{o}})$ = (1.8 mm, 0.04, 36, 18 mm, 1.8 mm, 1.2 μs) for the reconstruction of the breast image. The result of UBP with LDTF and spatial interpolation is shown in **Figure 10a3**. For better comparisons, we select two subdomains in a red-dashed box and a yellow-dashed box, respectively, as shown in **Figure 10a1**–**a3**. Closeup images of the subdomains are shown in **Figure 10b1**-**c1**, **b2**-**c2**, and **b3**-**c3** for the three methods, respectively. Comparing **Figure 10b1**–**b3**, we see that both RDTF with spatial interpolation and LDTF with spatial interpolation mitigate aliasing artifacts. However, RDTF with spatial interpolation compromises image resolution. From **Figure 10c1**–**c3**, we see that LDTF with spatial interpolation is more effective than RDTF with spatial interpolation in mitigating the aliasing artifacts. For quantitative comparisons, we pick two lines L1 and L2 in the red-boxed subdomain (**Figure 10b1**). Pixel values along these two lines, respectively, for the three methods are shown in **Figure 10d** and **e**, which further validate that LDTF with spatial interpolation is more effective than RDTF with spatial interpolation in both mitigating aliasing artifacts and maintaining image resolution.

2.5 Discussion

In the radius-dependent spatiotemporal antialiasing method, we clarified the source of spatial aliasing in PACT through spatiotemporal analysis. Then we classified the aliasing into two categories: aliasing in SS, and aliasing in IR. Spatial interpolation maintains the resolution in the one-way Nyquist zone S_1 while mitigating the artifacts caused by aliasing in IR. It extends the aliasing-free region from S_2 to S_1. For objects outside S_1, spatial interpolation is inaccurate due to spatial aliasing in SS, thus compromises the resolution. Adding radius-dependent temporal filtering does not affect the resolution inside S_1. For objects outside S_1, temporal filtering suppresses high-frequency signals to satisfy the temporal Nyquist sampling requirement. Although reducing the spatial resolution in the affected regions, temporal filtering mitigates aliasing in SS and makes the spatial interpolation accurate, thus further extends the aliasing-free region.

Then, we performed a location-dependent refinement of the spatiotemporal analysis and proposed location-dependent spatiotemporal antialiasing. To apply this method, we first divide the image domain into subdomains and select multiple groups of source points with maximum amplitudes from an initial image reconstructed using UBP. Then for each subdomain and each group of source points, we temporally filter the signals from the source points that overlap with signals from the subdomain. We recentered signals for this subdomain, apply spatial interpolation to the recentered signals, and use them to reconstruct the image in the subdomain. In this process, doing temporal filtering only for signals from source points with high amplitudes is essential for mitigating the dominant aliasing artifacts while minimizing unwanted blurring of the image. Location-dependent recentering of signals before spatial interpolation is essential for protecting signals from the subdomain of interest during spatial interpolation and maintaining image resolution. The proposed method outperforms radius-dependent antialiasing in mitigating aliasing artifacts and maintaining image resolution. We validated both methods through numerical simulations and *in vivo* experiments.

In practice, the radius-dependent antialiasing method can be easily implemented for a given system for the initial test and the location-dependent antialiasing method can be further implemented for better performance.

Both methods are applicable to any other detection geometry: 1D arrays for 2D imaging and 2D arrays for 3D imaging. They can also be used for signal preprocessing and combined with another method for further image reconstruction (e.g., model-based iterative methods and deep learning methods). Due to the high similarities in system matrices, both methods are directly applicable to CT and radial-sampling MRI.

C h a p t e r 3

EFFICIENT EXPLICIT SYSTEM MATRIX EXPRESSION

P. Hu, X. Tong, L. Lin, and L. V. Wang, "Data-driven system matrix manipulation enabling fast functional imaging and intra-image nonrigid motion correction in tomography," *Submitted*.

L. Lin†, **P. Hu**†, X. Tong†, S. Na†, R. Cao, X. Yuan, D. C. Garrett, J. Shi, K. Maslov, and L. V. Wang, "High-speed three-dimensional photoacoustic computed tomography for preclinical research and clinical translation," *Nat. Commun.*, vol. 12, no. 1, p. 882, Feb. 2021. DOI: 10.1038/s41467-021-21232-1

3.1 Abstract

We developed a three-dimensional photoacoustic computed tomography (3D-PACT) system that features large imaging depth, scalable field of view with isotropic spatial resolution, high imaging speed, and superior image quality. 3D-PACT allows for multipurpose imaging to reveal detailed angiographic information in biological tissues ranging from the mouse brain to the human breast. The system has a massive system matrix, which may be degraded by sparse sampling in mouse brain functional imaging and tissue motion in human breast imaging. Due to the high computation cost of system matrix manipulation, various existing techniques improve the image quality without correcting the system matrix and have limitations. Here, we compress the system matrix to improve computational efficiency (e.g., 42 times) using singular value decomposition and fast Fourier transform. This compression delivers an explicit expression of the system matrix that allows for efficient matrix slicing and data-driven system matrix manipulation. Both the limited bandwidth and finite size of the transducer elements are incorporated into the system matrix.

3.2 Background

3.2.1 An isotropic-resolution 3D PACT system

Employing different light illumination and detection schemes, several PACT systems with a three-dimensional (3D) field of view (FOV) have been developed for preclinical studies [40]–[43] and clinical practice [44]–[50]. While these systems have advanced PACT performance, key limitations remain unaddressed. Specifically, the current systems' limitations mainly arise from their shallow imaging depth, slow imaging speed, and/or compromised image quality due to the limited noise-equivalent sensitivity [51] and FOV. Here, we define high image quality as a sufficient contrast-to-noise ratio to reveal detailed structures on the order of the inherent spatial resolution (i.e., high clarity) within an FOV large and deep enough to cover the target region.

To provide a large view aperture (i.e., view angle) [52], some researchers utilized spherical detection matrices, which, however, were designed and implemented differently [40]–[42], [44]–[48]. For example, Deán-Ben *et al*. recently reported a PACT system with 256 transducer elements integrated on a partial cap for small animal imaging [41]. This design can generate a small volumetric image with a single laser shot, but its well-resolved FOV is theoretically limited to a few millimeters (e.g., diameter < 4 mm) according to the spatial Nyquist sampling criterion [37]. Subsequently, the authors scanned the system helically around a mouse body to densely sample the surface tissue and generated an elegant superficial image. However, without scanning along the radial direction, the mouse's internal organs were not imaged clearly. The limited-view aperture (i.e., < 2π steradian solid angle) further compromised the image quality when a higher imaging speed was required for brain functional imaging [40].

For human breast imaging, Matsumoto *et al*. developed a PACT system using a sparse hemispherical detector array that scanned in a spiral pattern on a plane [44]. The dense sampling produced high-quality vascular images near the skin surface. Similar to the aforementioned system, however, the well-resolved FOV of the detector array could only recover a small breast volume and the system could not image the deep tissue well without

an elevational scan. In addition, the illumination wavelengths they used suffer from higher scattering in the breast and a lower safety limit for maximum permissible exposure. Accordingly, the imaging depth of the system was limited to the superficial region of the breast. More recent work published by Oraevsky *et al*. [45] and Schoustra *et al*. [46] employ scanned arc ultrasonic arrays for human breast imaging. Both studies showed neither cross-sectional breast images at different depths nor 3D-rendered views, presumably due to the limited imaging depth. Similarly, the sparse spatial sampling limited the well-resolved FOV in breasts, thus compromising the image quality while also requiring a long scanning time of several minutes. In addition, the scanning of the laser beam varied the light energy distribution [46], which violates the assumption of 3D image reconstruction algorithms [6], [53] that light energy distribution remains consistent during scanning. Other than the systems that utilized spherical detection matrices, researchers also explored 3D imaging by detecting with a planar sensor [49] or scanning a linear array [50]. Though the systems were simpler, the spatial resolution was anisotropic and the image quality was reduced due to the limited-view aperture.

Here, we developed a PACT system that produces volumetric images of biological tissues with superior imaging depth, image quality, and speed (3D-PACT). Rather than being dedicatedly used for a single application, we have demonstrated its *in vivo* imaging versatility in the rat brain and human breast, representing two extreme imaging scales. In the rat brain, we imaged angiographic anatomies and functions from the cortex to the Circle of Willis. In the human breast, the system can generate a volumetric image with penetration of up to 4 cm by scanning the breast within a single breath-hold of 10 seconds. Our *in vivo* imaging results distinguish 3D-PACT as a cutting-edge photoacoustic imaging system that meets all the following conditions: (1) deep penetration and scalable FOV to accommodate imaging objects from rat brain to human breast, (2) isotropically high spatial resolution within the FOV in 3D space, (3) high noise-equivalent sensitivity to reveal small structures deep in tissues, (4) high imaging speed to minimize motion artifacts and enable functional studies, and (5) minimal limited-view artifacts. Accordingly, 3D-PACT could provide high-quality images with functional optical contrast and high imaging speed to directly benefit both

preclinical research and clinical practice, which have been highlighted by increasing interest from clinical researchers and practitioners. The proposed system and samples images are shown in **Figure 11**.

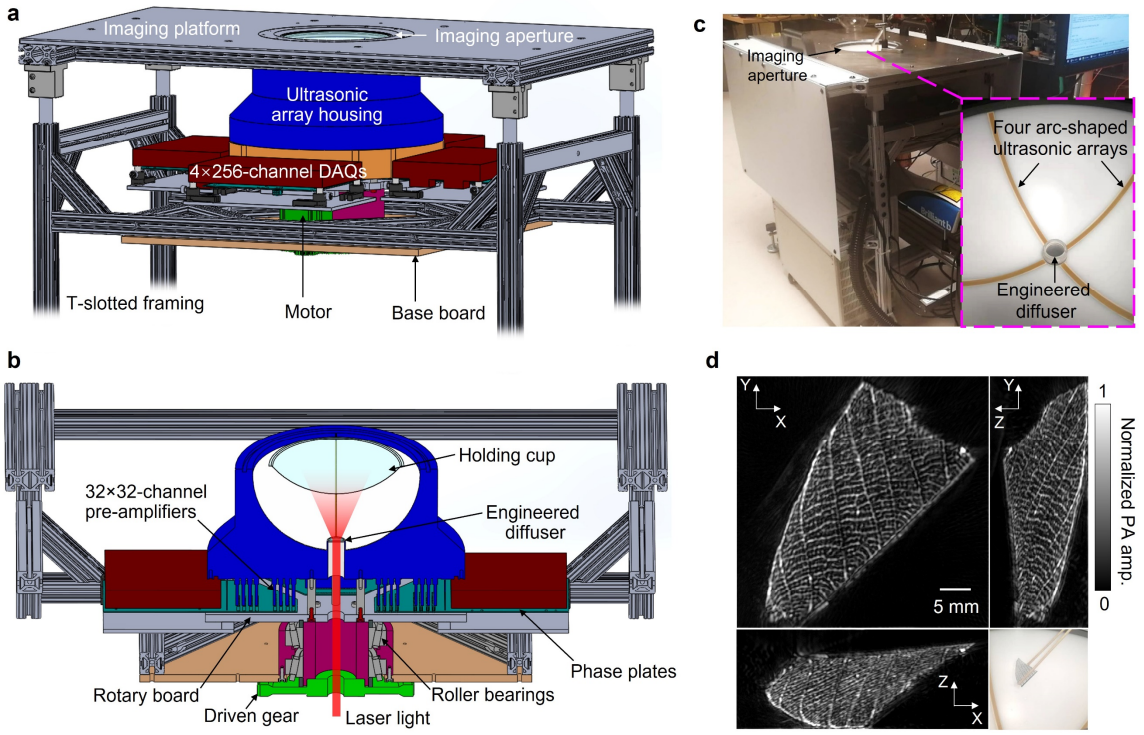

Figure 11 Representations and photographs of the 3D-PACT system and the leaf skeleton images. **a** Perspective view of the system. DAQ, data acquisition system. **b** Cut-away view of the system with the imaging platform removed. **c** Photograph of the system with a close-up view of the ultrasonic arrays. **d** Photograph and maximum amplitude projection (MAP) images of a leaf skeleton. PA, photoacoustic.

3.2.2 The need for an efficient explicit system matrix expression

The proposed 3D-PACT system can be modeled by a system matrix, and an efficient explicit expression of the system matrix is critical for certain applications. The proposed 3D-PACT supports mouse brain functional imaging and isotropic-resolution whole breast imaging. However, it has a relatively low volumetric imaging rate (0.5 Hz) for functional imaging because of the 20 rotating locations of the four-arc array and a 10-Hz laser repetition rate. Also, human breast imaging suffers from heart-beat-induced motion as the four-arc array rotates across 400 locations for a complete acquisition in 10 s. Advanced numerical methods

for fast functional imaging with sparse sampling and intra-image nonrigid motion correction will directly boost the performance of this system and may also be beneficial to other tomographic imaging systems.

Numerous methods have been proposed to compensate for system-matrix imperfections from image-domain [31], [54]–[58], signal-domain [37], [59]–[63], and cross-domain [34], [39], [64]–[66] perspectives. However, due to the large size of each system matrix, these methods tend not to manipulate or correct the system matrix directly and have limitations. For PACT, wave-equation-solver-based methods [25], [26] are efficient for forward simulation and iterative reconstruction but the system matrix is implicitly expressed by a series of linear operations. Thus, performing computations for a subset of the image domain and a subset of transducer elements may still require us to perform computations for a much larger domain enclosing the subsets of the image domain and transducer elements, which is highly inefficient.

To perform data-driven manipulation of the system matrix, we need to express it explicitly for efficient slicing. A point-source-response-based method [24] has been proposed for explicit system matrix expression, which is convenient for slicing but highly inefficient for large-scale simulation. To improve the computational efficiency while maintaining the slicing efficiency, we compress the explicit system matrix through singular value decomposition (SVD) and fast Fourier transform (FFT). In the following sections, we derive the efficient explicit expression of the system matrix and demonstrate its acceleration factor and accuracy through numerical simulations.

3.3 Theory

3.3.1 The fast forward operator

Assume that we have N finite-size ultrasonic transducer elements. For the n-th transducer element at $\mathbf{r}_n$ with detection surface S_n, the average pressure on the detection surface at time t is expressed as

$$\tilde{p}(\mathbf{r}_n, t) = \frac{1}{A_n} \int_{S_n} p(\mathbf{r}, t) \mathrm{d}a_n(\mathbf{r}) . \tag{58}$$

Here, A_n denotes the area of the surface S_n. We denote the spatial impulse response (SIR) in Eq. (58) as [24]

$$h_{\mathrm{s},n}(\mathbf{r}', t) = \int_{S_n} \frac{\delta\left(t - \frac{\|\mathbf{r}' - \mathbf{r}\|}{c}\right)}{2\pi c \|\mathbf{r}' - \mathbf{r}\|} \mathrm{d}a_n(\mathbf{r}) . \tag{59}$$

Then Eq. (58) becomes

$$\begin{aligned} \tilde{p}(\mathbf{r}_n, t) &= \frac{1}{2c} \int_V p_0(\mathbf{r}') \frac{1}{A_n} \frac{\partial}{\partial t} \left[\int_{S_n} \frac{\delta\left(t - \frac{\|\mathbf{r}' - \mathbf{r}\|}{c}\right)}{2\pi c \|\mathbf{r}' - \mathbf{r}\|} \mathrm{d}a_n(\mathbf{r}) \right] \mathrm{d}\mathbf{r}' \\ &= \frac{1}{2c} \int_V p_0(\mathbf{r}') \frac{1}{A_n} \frac{\partial}{\partial t} h_{\mathrm{s},n}(\mathbf{r}', t) \mathrm{d}\mathbf{r}' . \end{aligned} \tag{60}$$

We denote the electrical impulse response (EIR) of the n-th transducer as $h_{\mathrm{e},n}(t)$ and express the transducer's response using temporal convolution $*_t$ as

$$\hat{p}(\mathbf{r}_n, t) = \tilde{p}(\mathbf{r}_n, t) *_t h_{\mathrm{e},n}(t). \tag{61}$$

Substituting Eq. (60) into Eq. (61) yields

$$\begin{aligned} \hat{p}(\mathbf{r}_n, t) &= \frac{1}{2c} \int_V p_0(\mathbf{r}') \frac{1}{A_n} \frac{\partial}{\partial t} h_{\mathrm{s},n}(\mathbf{r}', t) \mathrm{d}\mathbf{r}' *_t h_{\mathrm{e},n}(t) \\ &= \int_V p_0(\mathbf{r}') \frac{h'_{\mathrm{e},n}(t) *_t h_{\mathrm{s},n}(\mathbf{r}', t)}{2cA_n} \mathrm{d}\mathbf{r}' \\ &= \int_V p_0(\mathbf{r}') h_n(\mathbf{r}', t) \mathrm{d}\mathbf{r}' . \end{aligned} \tag{62}$$

Here, the prime in $h'_{\mathrm{e},n}(t)$ denotes the time derivative, and h_n denotes the point source response per unit initial pressure per unit infinitesimal tissue volume received by a finite-size transducer element:

$$h_n(\mathbf{r}', t) = \frac{h'_{\mathrm{e},n}(t) *_t h_{\mathrm{s},n}(\mathbf{r}', t)}{2cA_n}. \tag{63}$$

For convenience in the following discussion, we temporally shift $h_{\mathrm{s},n}(\mathbf{r}', t)$ and $h_n(\mathbf{r}', t)$ for $\mathbf{r}'$ such that time 0 aligns with the onset of the nonzero signal received by the center of the n-th transducer element $\mathbf{r}_n$:

$$\hat{h}_{\mathrm{s},n}(\mathbf{r}',t)=h_{\mathrm{s},n}\left(\mathbf{r}',t+\frac{\|\mathbf{r}'-\mathbf{r}_n\|}{c}\right)=\int_{S_n}\frac{\delta\left(t+\frac{\|\mathbf{r}'-\mathbf{r}_n\|}{c}-\frac{\|\mathbf{r}'-\mathbf{r}\|}{c}\right)}{2\pi c\|\mathbf{r}'-\mathbf{r}\|}\mathrm{d}a_n(\mathbf{r}) \quad (64)$$

and

$$\hat{h}_n(\mathbf{r}',t)=h_n\left(\mathbf{r}',t+\frac{\|\mathbf{r}'-\mathbf{r}_n\|}{c}\right)=\frac{h'_{\mathrm{e},n}(t)*_t h_{\mathrm{s},n}\left(\mathbf{r}',t+\frac{\|\mathbf{r}'-\mathbf{r}_n\|}{c}\right)}{2cA_n}$$
$$=\frac{h'_{\mathrm{e},n}(t)*_t \hat{h}_{\mathrm{s},n}(\mathbf{r}',t)}{2cA_n}. \quad (65)$$

Next, we express the SIR and point source response in the local coordinates of the transducer elements. Each transducer element used in this research has a flat rectangular detection surface with a length a of 0.7 mm and a width b of 0.6 mm, yielding $A_n = ab = 0.42\ \mathrm{mm}^2, n = 1,2,\dots,N$. These transducer elements also have the same EIR: $h_{\mathrm{e},n}(t) = h_{\mathrm{e}}(t), n = 1,2,\dots,N$, which is measured experimentally based on the following Eq. (80). We define the local coordinates of the n-th transducer element using the center the detection surface as the origin, the length direction as the x-axis, the width direction as the y-axis, and the normal direction (toward the detection region) as the z-axis. Here, we choose the x-axis and y-axis to let the coordinates satisfy the right-hand rule. We express the three axes of the local coordinates as three vectors of unit length: $\hat{\mathbf{x}}_n$, $\hat{\mathbf{y}}_n$, and $\hat{\mathbf{z}}_n$ (shown in **Figure 12a**), which form an orthonormal matrix

$$\mathbf{A}_n = (\hat{\mathbf{x}}_n^{\mathrm{T}}, \hat{\mathbf{y}}_n^{\mathrm{T}}, \hat{\mathbf{z}}_n^{\mathrm{T}}). \quad (66)$$

Locations $\mathbf{r}'$, $\mathbf{r}_n$, and $\mathbf{r}$ in the global coordinates correspond to the locations $\mathbf{r}'_{\mathrm{lcl}}$, $\mathbf{0}$, and $\mathbf{r}_{\mathrm{lcl}}$ in the local coordinates of the n-th transducer element. Coordinate transformations yield $\mathbf{r}'_{\mathrm{lcl}} = (\mathbf{r}' - \mathbf{r}_n)\mathbf{A}_n$ and $\mathbf{r}_{\mathrm{lcl}} = (\mathbf{r} - \mathbf{r}_n)\mathbf{A}_n$. These global and local coordinates satisfy $\|\mathbf{r}' - \mathbf{r}_n\| = \|\mathbf{r}'_{\mathrm{lcl}}\|$ and $\|\mathbf{r}' - \mathbf{r}\| = \|\mathbf{r}'_{\mathrm{lcl}} - \mathbf{r}_{\mathrm{lcl}}\|$ due to the orthonormality of the transformations. We denote the detection surface (S_n in the global coordinates) in the local coordinates as $S_{n,\mathrm{lcl}}$, and denote the local Leibniz notation as $\mathrm{d}a_{n,\mathrm{lcl}}(\mathbf{r}_{\mathrm{lcl}}) = \mathrm{d}a_n(\mathbf{r})$. Thus, in the local coordinates of the n-th transducer element, we express Eq. (64) as

$$\hat{h}_{\mathrm{s},n}(\mathbf{r}',t)=\int_{S_{n,\mathrm{lcl}}}\frac{\delta\left(t+\frac{\|\mathbf{r}'_{\mathrm{lcl}}\|}{c}-\frac{\|\mathbf{r}'_{\mathrm{lcl}}-\mathbf{r}_{\mathrm{lcl}}\|}{c}\right)}{2\pi c\|\mathbf{r}'_{\mathrm{lcl}}-\mathbf{r}_{\mathrm{lcl}}\|}\mathrm{d}a_{n,\mathrm{lcl}}(\mathbf{r}_{\mathrm{lcl}}),$$
$$\mathbf{r}'_{\mathrm{lcl}}=(\mathbf{r}'-\mathbf{r}_n)\mathbf{A}_{\mathbf{r}_n},n=1,2,\dots,N. \quad (67)$$

All transducer elements are geometrically identical and have the same local coordinates, meaning $S_{n,\mathrm{lcl}}=S_{n',\mathrm{lcl}}$ and $\mathrm{d}a_{n,\mathrm{lcl}}(\mathbf{r}_{\mathrm{lcl}})=\mathrm{d}a_{n',\mathrm{lcl}}(\mathbf{r}_{\mathrm{lcl}})$ for $n,n'\in\{1,2,\dots,N\}$. We define $S_{\mathrm{lcl}}=S_{1,\mathrm{lcl}}$, $\mathrm{d}a_{\mathrm{lcl}}(\mathbf{r}_{\mathrm{lcl}})=\mathrm{d}a_{1,\mathrm{lcl}}(\mathbf{r}_{\mathrm{lcl}})$, and rewrite Eq. (67) as

$$\hat{h}_{\mathrm{s,lcl}}(\mathbf{r}'_{\mathrm{lcl}},t)=\int_{S_{\mathrm{lcl}}}\frac{\delta\left(t+\frac{\|\mathbf{r}'_{\mathrm{lcl}}\|}{c}-\frac{\|\mathbf{r}'_{\mathrm{lcl}}-\mathbf{r}_{\mathrm{lcl}}\|}{c}\right)}{2\pi c\|\mathbf{r}'_{\mathrm{lcl}}-\mathbf{r}_{\mathrm{lcl}}\|}\mathrm{d}a_{\mathrm{lcl}}(\mathbf{r}_{\mathrm{lcl}}), \quad (68)$$

which is now independent of the transducer element index n. Replacing $h'_{\mathrm{e},n}(t)$ and $\hat{h}_{\mathrm{s},n}(\mathbf{r}',t)$ with $h'_{\mathrm{e}}(t)$ and $\hat{h}_{\mathrm{s,lcl}}(\mathbf{r}'_{\mathrm{lcl}},t)$, respectively, in Eq. (65), we define

$$\hat{h}_{\mathrm{lcl}}(\mathbf{r}'_{\mathrm{lcl}},t)=\frac{h'_{\mathrm{e}}(t)*_t\hat{h}_{\mathrm{s,lcl}}(\mathbf{r}'_{\mathrm{lcl}},t)}{2cab}. \quad (69)$$

Thus, we need to calculate only the values of $\hat{h}_{\mathrm{s,lcl}}(\mathbf{r}'_{\mathrm{lcl}},t)$ and $\hat{h}_{\mathrm{lcl}}(\mathbf{r}'_{\mathrm{lcl}},t)$, then obtain the values of $\hat{h}_{\mathrm{s},n}(\mathbf{r}',t)$ and $\hat{h}_n(\mathbf{r}',t)$ through coordinate transformation:

$$\hat{h}_{\mathrm{s},n}(\mathbf{r}',t)=\hat{h}_{\mathrm{s,lcl}}(\mathbf{r}'_{\mathrm{lcl}},t) \quad (70)$$

and

$$\hat{h}_n(\mathbf{r}',t)=\frac{h'_{\mathrm{e},n}(t)*_t\hat{h}_{\mathrm{s},n}(\mathbf{r}',t)}{2cA_n}=\frac{h'_{\mathrm{e}}(t)*_t\hat{h}_{\mathrm{s,lcl}}(\mathbf{r}'_{\mathrm{lcl}},t)}{2cab}=\hat{h}_{\mathrm{lcl}}(\mathbf{r}'_{\mathrm{lcl}},t), \quad (71)$$

respectively, with $\mathbf{r}'_{\mathrm{lcl}}=(\mathbf{r}'-\mathbf{r}_n)\mathbf{A}_n$ for $n=1,2,\dots,N$. Through these relations, we express both the SIR and the point source response in the local coordinates.

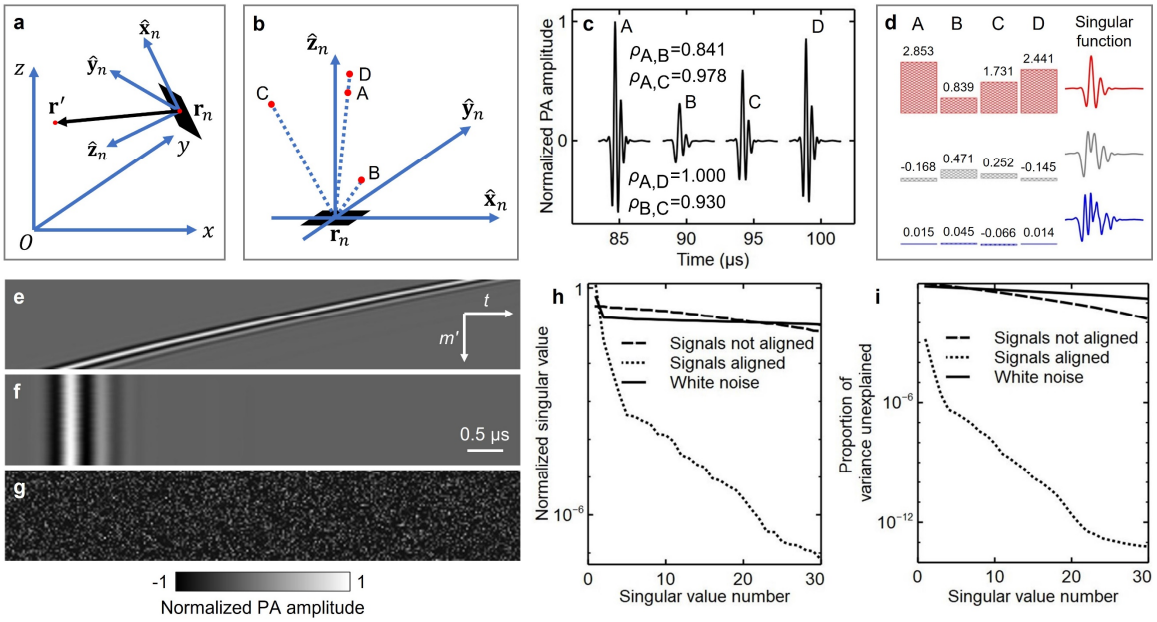

Figure 12 Compression of the forward operator based on SVD. **a** A point source (a red dot), the n-th transducer element (a black rectangular centered at $\mathbf{r}_n$), and the element's local coordinate system with axes $\hat{\mathbf{x}}_n$, $\hat{\mathbf{y}}_n$, and $\hat{\mathbf{z}}_n$. **b** Four point sources A, B, C, and D (red dots) in the local coordinate system of the n-th transducer element. Points A, D, and $\mathbf{r}_n$ are on the same line. **c** The responses of the transducer element to the signals from the four point sources and the PCC between every two responses. **d** Expression of the four responses using linear combinations (coefficients visualized with bars) of three temporal singular functions shown as red, gray, and blue curves, respectively, based on SVD. **e-f** Independent responses (**e**) of a transducer element to 50 point sources with decreasing distances to the element, and the temporally-shifted form (**f**), which aligns the nonzero signals in time. PA, photoacoustic. **g** White-noise responses of the same size as those in **e** and **f**. **h-i** Normalized singular values and proportions of the variances unexplained, respectively, in the SVDs of the signals in **e**, **f**, and **g**.

We visualize the signals detected by a transducer element by picking four point sources, labeled as A, B, C, and D, respectively, in the local coordinate system of the n-th transducer element (**Figure 12b**), with points A, D, and the element center $\mathbf{r}_n$ on the same line. The element's responses to the signals from the point sources are shown in **Figure 12c**. We let $\rho_{\mathrm{A,B}}$ denote the Pearson correlation coefficient (PCC) between the responses corresponding to A and B. A direct implementation of the forward operator based on Eq. (62) is computationally intensive. Although $\rho_{\mathrm{A,D}} = 1$, due to the effects of SIR, $\rho_{\mathrm{A,B}}$, $\rho_{\mathrm{A,C}}$, and $\rho_{\mathrm{B,C}}$ are less than 1, indicating the signals from points A, B, and C are not shift-invariant; thus, an efficient temporal convolution with one kernel function cannot yield the detected signals accurately.

To accelerate the forward operator, we decouple the spatial and temporal dimensions of $\hat{h}_{\mathrm{lcl}}(\mathbf{r}'_{\mathrm{lcl}}, t)$ using SVD while keeping only the dominant components:

$$\hat{h}_{\mathrm{lcl}}(\mathbf{r}'_{\mathrm{lcl}}, t) \approx \sum_{k=1}^{K} \hat{h}_{\mathrm{lcl},k}(\mathbf{r}'_{\mathrm{lcl}})\eta_k(t)\,. \tag{72}$$

Here, $\hat{h}_{\mathrm{lcl},k}(\mathbf{r}'_{\mathrm{lcl}})$ and η_k denote the k-th spatial singular function and the k-th temporal singular function, respectively; and we use the first K terms to approximate the whole series. Combining Eqs. (65), (71), and (72), we obtain

$$\begin{aligned} h_n(\mathbf{r}', t) &= \hat{h}_n\left(\mathbf{r}', t - \frac{\|\mathbf{r}' - \mathbf{r}_n\|}{c}\right) = \hat{h}_{\mathrm{lcl}}\left((\mathbf{r}' - \mathbf{r}_n)\mathbf{A}_n, t - \frac{\|\mathbf{r}' - \mathbf{r}_n\|}{c}\right) \\ &\approx \sum_{k=1}^{K} \hat{h}_{\mathrm{lcl},k}\big((\mathbf{r}' - \mathbf{r}_n)\mathbf{A}_n\big)\eta_k\left(t - \frac{\|\mathbf{r}' - \mathbf{r}_n\|}{c}\right). \end{aligned} \tag{73}$$

Substituting Eq. (73) into Eq. (62), we obtain

$$\begin{aligned} \hat{p}(\mathbf{r}_n, t) &\approx \int_V p_0(\mathbf{r}') \sum_{k=1}^{K} \hat{h}_{\mathrm{lcl},k}\big((\mathbf{r}' - \mathbf{r}_n)\mathbf{A}_n\big)\eta_k\left(t - \frac{\|\mathbf{r}' - \mathbf{r}_n\|}{c}\right)\mathrm{d}\mathbf{r}' \\ &= \sum_{k=1}^{K} \eta_k(t) *_t \int_V p_0(\mathbf{r}')\hat{h}_{\mathrm{lcl},k}\big((\mathbf{r}' - \mathbf{r}_n)\mathbf{A}_n\big)\delta\left(t - \frac{\|\mathbf{r}' - \mathbf{r}_n\|}{c}\right)\mathrm{d}\mathbf{r}', \\ & \qquad n = 1,2,\ldots,N, t \geq 0. \end{aligned} \tag{74}$$

As shown in this equation, we split the temporal variable from the spatial integrals, which allows for a fast implementation of the forward operator.

We apply Eq. (72) with $K = 3$ to the responses in **Figure 12c** for an initial demonstration. The three temporal singular functions $\eta_1(t), \eta_2(t), \eta_3(t)$ are shown as red, gray, and blue curves, respectively, in **Figure 12d**, and the values of spatial singular functions are visualized as bars. Then we explain the necessity of temporal shifting for alignment, described in Eqs. (64) and (65), for SVD. We select 50 point sources with decreasing distances to a transducer element and visualize the element's independent responses to them in **Figure 12e**. The temporally-shifted form of these responses based on Eq. (65) is shown in **Figure 12f**. We also add white-noise responses with the same size for comparison, as shown in **Figure 12g**. Performing SVD to the three sets of responses, we

observe different compression efficiencies from the perspectives of normalized singular value (**Figure 12h**) and proportion of the variance (**Figure 12i**). In both figures, we see that the compression efficiency of the original responses is similar to that of the white-noise responses, whereas the efficiency for the temporally-shifted responses is significantly higher (necessary for the compression).

3.3.2 Point source response measurement

The fast operator can be configured for a transducer with any detection surface. Here, we only discuss an ultrasonic transducer with a flat rectangular detection surface, which is used in our 3D imaging system [67]. We see from Eqs. (62) and (71) that the forward operator is based on the point source response $\hat{g}_{\mathrm{lcl}}(\mathbf{r}'_{\mathrm{lcl}}, t)$ in the local coordinates. The response is determined by the first derivative of EIR ($h'_{\mathrm{e}}(t)$) and the realigned SIR ($\hat{h}_{\mathrm{s,lcl}}(\mathbf{r}'_{\mathrm{lcl}}, t)$) in the local coordinates. In general, we can measure $\hat{h}_{\mathrm{lcl}}(\mathbf{r}'_{\mathrm{lcl}}, t)$ experimentally. In this research, for an ultrasonic transducer with a flat rectangular detection surface, we calculate $\hat{h}_{\mathrm{lcl}}(\mathbf{r}'_{\mathrm{lcl}}, t)$ efficiently using the far-field approximation of $\hat{h}_{\mathrm{s,lcl}}(\mathbf{r}'_{\mathrm{lcl}}, t)$ [24]. From a special case of this approximation, we derive a method to measure $h'_{\mathrm{e}}(t)$ experimentally.

In the first step of obtaining $\hat{h}_{\mathrm{lcl}}(\mathbf{r}'_{\mathrm{lcl}}, t)$, we introduce the calculation of $\hat{h}_{\mathrm{lcl}}(\mathbf{r}'_{\mathrm{lcl}}, t)$ based on the far-field approximation of $\hat{h}_{\mathrm{s,lcl}}(\mathbf{r}'_{\mathrm{lcl}}, t)$. In this research, the image domain is within the far-field regions of all transducer elements. Using the far-field approximation, we express the SIR in the frequency domain as [24]

$$F_t\left(h_{\mathrm{s},n}(\mathbf{r}', t)\right)(f) \approx \frac{ab \exp\left(-j2\pi f \frac{\|\mathbf{r}' - \mathbf{r}_n\|}{c}\right)}{2\pi c\|\mathbf{r}' - \mathbf{r}_n\|} \operatorname{sinc}\left(\pi f \frac{a|(\mathbf{r}' - \mathbf{r}_n)\hat{\mathbf{x}}_n^{\mathrm{T}}|}{c\|\mathbf{r}' - \mathbf{r}_n\|}\right) \operatorname{sinc}\left(\pi f \frac{b|(\mathbf{r}' - \mathbf{r}_n)\hat{\mathbf{y}}_n^{\mathrm{T}}|}{c\|\mathbf{r}' - \mathbf{r}_n\|}\right). \quad (75)$$

Here, F_t denotes the temporal FT, and j denotes the imaginary unit. Combining Eqs. (70) and (75), we apply the temporal Fourier transform (FT) to $\hat{h}_{\mathrm{s,lcl}}(\mathbf{r}'_{\mathrm{lcl}}, t)$ and obtain

$$F_t\left(\hat{h}_{\mathrm{s,lcl}}(\mathbf{r}'_{\mathrm{lcl}},t)\right)(f) = F_t\left(h_{\mathrm{s},n}\left(\mathbf{r}', t+\frac{\|\mathbf{r}'-\mathbf{r}_n\|}{c}\right)\right)(f)$$

$$= \exp\left(j2\pi f\frac{\|\mathbf{r}'-\mathbf{r}_n\|}{c}\right)F_t\left(h_{\mathrm{s},n}(\mathbf{r}',t)\right)(f)$$

$$\approx \frac{ab}{2\pi c\|\mathbf{r}'-\mathbf{r}_n\|}\mathrm{sinc}\left(\pi f\frac{a|(\mathbf{r}'-\mathbf{r}_n)\hat{\mathbf{x}}_n^{\mathrm{T}}|}{c\|\mathbf{r}'-\mathbf{r}_n\|}\right)\mathrm{sinc}\left(\pi f\frac{b|(\mathbf{r}'-\mathbf{r}_n)\hat{\mathbf{y}}_n^{\mathrm{T}}|}{c\|\mathbf{r}'-\mathbf{r}_n\|}\right)$$

$$= \frac{ab}{2\pi c\|\mathbf{r}'_{\mathrm{lcl}}\|}\mathrm{sinc}\left(\pi f\frac{a|x'_{\mathrm{lcl}}|}{c\|\mathbf{r}'_{\mathrm{lcl}}\|}\right)\mathrm{sinc}\left(\pi f\frac{b|y'_{\mathrm{lcl}}|}{c\|\mathbf{r}'_{\mathrm{lcl}}\|}\right). \quad (76)$$

Here, we use identities $\mathbf{r}'_{\mathrm{lcl}} = (\mathbf{r}'-\mathbf{r}_n)\mathbf{A}_n = (\mathbf{r}'-\mathbf{r}_n)(\hat{\mathbf{x}}_n^{\mathrm{T}},\hat{\mathbf{y}}_n^{\mathrm{T}},\hat{\mathbf{z}}_n^{\mathrm{T}}) = (x'_{\mathrm{lcl}},y'_{\mathrm{lcl}},z'_{\mathrm{lcl}})$ and $\|\mathbf{r}'-\mathbf{r}_n\| = \|\mathbf{r}'_{\mathrm{lcl}}\|$, and $\mathbf{r}'_{\mathrm{lcl}}$ must not be the origin in the local coordinates. Substituting Eq. (76) into the temporal FT of Eq. (69), we obtain

$$F_t\left(\hat{h}_{\mathrm{lcl}}(\mathbf{r}'_{\mathrm{lcl}},t)\right)(f) = \frac{1}{2cab}F_t\big(h'_{\mathrm{e}}(t)\big)(f)F_t\left(\hat{h}_{\mathrm{s,lcl}}(\mathbf{r}'_{\mathrm{lcl}},t)\right)(f)$$

$$\approx \frac{F_t\big(h'_{\mathrm{e}}(t)\big)(f)}{4\pi c^2\|\mathbf{r}'_{\mathrm{lcl}}\|}\mathrm{sinc}\left(\pi f\frac{a|x'_{\mathrm{lcl}}|}{c\|\mathbf{r}'_{\mathrm{lcl}}\|}\right)\mathrm{sinc}\left(\pi f\frac{b|y'_{\mathrm{lcl}}|}{c\|\mathbf{r}'_{\mathrm{lcl}}\|}\right). \quad (77)$$

Applying temporal inverse FT to Eq. (77), we obtain

$$\hat{h}_{\mathrm{lcl}}(\mathbf{r}'_{\mathrm{lcl}},t) \approx F_t^{-1}\left(\frac{F_t\big(h'_{\mathrm{e}}(t)\big)(f)}{4\pi c^2\|\mathbf{r}'_{\mathrm{lcl}}\|}\mathrm{sinc}\left(\pi f\frac{a|x'_{\mathrm{lcl}}|}{c\|\mathbf{r}'_{\mathrm{lcl}}\|}\right)\mathrm{sinc}\left(\pi f\frac{b|y'_{\mathrm{lcl}}|}{c\|\mathbf{r}'_{\mathrm{lcl}}\|}\right)\right). \quad (78)$$

In the second step of obtaining $\hat{h}_{\mathrm{lcl}}(\mathbf{r}'_{\mathrm{lcl}},t)$, we derive a method to measure $h'_{\mathrm{e}}(t)$ by analyzing a special case of Eq. (78). We constrain the location $\mathbf{r}'_{\mathrm{lcl}}$ on the axis of the transducer by letting $\mathbf{r}'_{\mathrm{lcl}} = (0,0,z'_{\mathrm{lcl}})$, which simplifies Eq. (78) to

$$\hat{h}_{\mathrm{lcl}}(\mathbf{r}'_{\mathrm{lcl}},t) \approx F_t^{-1}\left(\frac{F_t\big(h'_{\mathrm{e}}(t)\big)(f)}{4\pi c^2|z'_{\mathrm{lcl}}|}\right) = \frac{h'_{\mathrm{e}}(t)}{4\pi c^2|z'_{\mathrm{lcl}}|}. \quad (79)$$

Solving for $h'_{\mathrm{e}}(t)$ from Eq. (79), we obtain

$$h'_{\mathrm{e}}(t) \approx 4\pi c^2|z'_{\mathrm{lcl}}|\hat{h}_{\mathrm{lcl}}(\mathbf{r}'_{\mathrm{lcl}},t) = 4\pi c^2|z'_{\mathrm{lcl}}|\hat{h}_n(\mathbf{r}',t)$$

$$= 4\pi c^2|z'_{\mathrm{lcl}}|h_n\left((0,0,z'_{\mathrm{lcl}})\mathbf{A}_{\mathbf{r}_n}^{-1}+\mathbf{r}_n, t+\frac{|z'_{\mathrm{lcl}}|}{c}\right). \quad (80)$$

Here, we use the identity $\mathbf{r}' = \mathbf{r}'_{\mathrm{lcl}}\mathbf{A}_n^{-1} + \mathbf{r}_n = (0,0,z'_{\mathrm{lcl}})\mathbf{A}_n^{-1} + \mathbf{r}_n$. In practice, we repeat the measurement of the right-hand side of Eq. (80) and use the average to represent $h'_{\mathrm{e}}(t)$.

In summary, we measure $h'_{\mathrm{e}}(t)$ experimentally on the basis of Eq. (80) and substitute the measurement into Eq. (78) to obtain $\hat{h}_{\mathrm{lcl}}(\mathbf{r}'_{\mathrm{lcl}}, t)$. Further, we perform SVD to $\hat{h}_{\mathrm{lcl}}(\mathbf{r}'_{\mathrm{lcl}}, t)$ according to Eq. (72) and obtain singular functions $\hat{h}_{\mathrm{lcl},k}(\mathbf{r}'_{\mathrm{lcl}})$ and $\eta_k(t), k = 1,2,\dots,K$. We use these functions in Eq. (74) to implement the fast forward operator.

3.3.3 Discretization of the forward operator

We express the forward operator in two different forms in Eqs. (62) and (74), respectively. First, we discretize the forward operator implemented in Eq. (62). In the temporal domain, we choose points of interest $t_l = t_1 + (l-1)\tau, l = 1,2,\dots,L$, where t_1 is the initial time, τ is the sampling step size, and L is the number of time points. Then we discretize the temporal FT of $h'_{\mathrm{e}}(t)$ as $F_t\big(h'_{\mathrm{e}}(t)\big)_{l'} \approx F_l(h'_l)(l'), l' = 1,2,\dots,L$, where we define $h'_l = h'_{\mathrm{e}}(t_l)$, $F_t\big(h'_{\mathrm{e}}(t)\big)_{l'} = F_t\big(h'_{\mathrm{e}}(t)\big)(f_{l'})$, $l, l' \in \{1,2,\dots,L\}$, and let F_l represent the discrete FT with respect to l. The temporal frequencies $f_{l'}, l' = 1,2,\dots,L$ are selected according to the requirement of the discrete FT. We further denote the m_{lcl}-th location in the local coordinates as $\mathbf{r}'_{\mathrm{lcl},m_{\mathrm{lcl}}} = \big(x'_{\mathrm{lcl},m_{\mathrm{lcl}}}, y'_{\mathrm{lcl},m_{\mathrm{lcl}}}, z'_{\mathrm{lcl},m_{\mathrm{lcl}}}\big)$ and define $\hat{h}_{\mathrm{lcl},m_{\mathrm{lcl}},l} = \hat{h}_{\mathrm{lcl}}\big(\mathbf{r}'_{\mathrm{lcl},m_{\mathrm{lcl}}}, t_l\big), m_{\mathrm{lcl}} = 1,2,\dots,M_{\mathrm{lcl}}$, where M_{lcl} is the number of locations of interest in the local coordinates. Thus, we discretize Eq. (78) as

$$\hat{h}_{\mathrm{lcl},m_{\mathrm{lcl}},l} \approx \frac{1}{4\pi c^2\left\|\mathbf{r}'_{\mathrm{lcl},m_{\mathrm{lcl}}}\right\|} F_l^{-1}\left(F_l(h'_l)(l')\mathrm{sinc}\left(\pi f_{l'}\frac{a\left|x'_{\mathrm{lcl},m_{\mathrm{lcl}}}\right|}{c\left\|\mathbf{r}'_{\mathrm{lcl},m_{\mathrm{lcl}}}\right\|}\right)\mathrm{sinc}\left(\pi f_{l'}\frac{b\left|y'_{\mathrm{lcl},m_{\mathrm{lcl}}}\right|}{c\left\|\mathbf{r}'_{\mathrm{lcl},m_{\mathrm{lcl}}}\right\|}\right)\right)(l),$$
$$m_{\mathrm{lcl}} = 1,2,\dots,M_{\mathrm{lcl}}, l = 1,2,\dots,L. \quad (81)$$

Further, we discretize the point source response $h_n(\mathbf{r}', t)$ as $h_{n,m,l} = h_n(\mathbf{r}'_m, t_l), n = 1,2,\dots,N, m = 1,2,\dots,M, l = 1,2,\dots,L$. Here, M is the number of voxels (source points) in the image domain. On the basis of the relation $h_n(\mathbf{r}'_m, t_l) = \hat{h}_n\left(\mathbf{r}'_m, t_l - \frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c}\right) = \hat{h}_{\mathrm{lcl}}\left((\mathbf{r}'_m - \mathbf{r}_n)\mathbf{A}_n, t_l - \frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c}\right)$, we obtain the values of $h_{n,m,l}$ through spatiotemporal

interpolation of the values of $\hat{h}_{\mathrm{lcl},m_{\mathrm{lcl}},l}$. Denoting $\hat{p}_{n,l} = \hat{p}(\mathbf{r}_n, t_l)$ and $p_{0,m} = p_0(\mathbf{r}'_m)$, we discretize Eq. (62) as

$$\hat{p}_{n,l}^{\mathrm{slow}} = \sum_{m=1}^{M} v_m h_{n,m,l} p_{0,m}\,, n = 1,2,\dots,N, l = 1,2,\dots,L. \tag{82}$$

Here, $\hat{p}_{n,l}^{\mathrm{slow}}$ represents an approximation of $\hat{p}_{n,l}$ using this relatively slow operator, and v_m represents the volume of the m-th voxel. In practice, due to the finite duration of every point source response (shown in **Figure 12c**, **e**, and **f**), given a combination of m and n, we need to calculate only for l in a range of length $L' < L$. Here, L' is the effective length for nonzero values in the discretized point source responses, and the computational complexity of the discrete forward operator in Eq. (82) is $O(NML')$.

Next, we discretize the forward operator to a form with lower computational complexity from Eq. (74). We denote $\bar{h}_{\mathrm{lcl},k,m_{\mathrm{lcl}}} = \hat{h}_{\mathrm{lcl},k}\left(\mathbf{r}'_{\mathrm{lcl},m_{\mathrm{lcl}}}\right)$ and $\eta_{k,l} = \eta_k(t_l)$, $k = 1,2,\dots,K$, $m_{\mathrm{lcl}} = 1,2,\dots,M_{\mathrm{lcl}}$, $l = 1,2,\dots,L$. After obtaining the array $\hat{h}_{\mathrm{lcl},m_{\mathrm{lcl}},l}$ through Eq. (81), we estimate $\hat{h}_{\mathrm{lcl},k,m_{\mathrm{lcl}}}$ and $\eta_{k,l}$ through SVD:

$$\hat{h}_{\mathrm{lcl},m_{\mathrm{lcl}},l} \approx \sum_{k=1}^{K} \bar{h}_{\mathrm{lcl},k,m_{\mathrm{lcl}}} \eta_{k,l}\,, m_{\mathrm{lcl}} = 1,2,\dots,M_{\mathrm{lcl}}, l = 1,2,\dots,L. \tag{83}$$

We denote $\hat{h}_{\mathrm{lcl},k,n,m} = \hat{h}_{\mathrm{lcl},k}\left((\mathbf{r}'_m - \mathbf{r}_n)\mathbf{A}_n\right)$, which is obtained from the values of $\bar{h}_{\mathrm{lcl},k,m_{\mathrm{lcl}}}$ through spatial interpolation. Moreover, we express $\delta\left(t - \frac{\|\mathbf{r}' - \mathbf{r}_n\|}{c}\right)$ in the discrete form as

$$\frac{1}{\tau}\left(\left(t_{l_{n,m}+1} - \frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c}\right)\delta_{l,l_{n,m}} + \left(\frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c} - t_{l_{n,m}}\right)\delta_{l,l_{n,m}+1}\right),$$
$$n = 1,2,\dots,N, m = 1,2,\dots,M, l = 1,2,\dots,L. \tag{84}$$

Here, $l_{n,m}$ denotes the temporal index such that $t_{l_{n,m}} \le \frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c} < t_{l_{n,m}+1}$, and we use the Kronecker delta function

$$\delta_{l,l'} = \begin{cases} 0, l \neq l', \\ 1, l = l'. \end{cases} \tag{85}$$

We discretize the forward operator in Eq. (74) as

$$\hat{p}_{n,l}^{\text{fast}} = \sum_{k=1}^{K} \eta_{k,l} *_l \sum_{m=1}^{M} v_m \hat{h}_{\text{lcl},k,n,m} p_{0,m} \frac{1}{\tau} \left[\begin{array}{l} \left(t_{l_{n,m}+1} - \dfrac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c} \right) \delta_{l,l_{n,m}} \\ + \left(\dfrac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c} - t_{l_{n,m}} \right) \delta_{l,l_{n,m}+1} \end{array} \right],$$

$$n = 1,2,\dots,N, l = 1,2,\dots,L. \tag{86}$$

Here, $\hat{p}_{n,l}^{\text{fast}}$ represents an approximation of $\hat{p}_{n,l}$ using the fast operator, and $*_l$ denotes the discrete convolution with respect to l. In practice, we choose a value of L so that $\log_2 L$ is an integer, and we implement the discrete convolution using the temporal FFT. Thus, the computational complexity of the discrete forward operator in Eq. (86) is $O(NMK) + O(N(L\log_2 L)K)$.

3.3.4 Iterative reconstruction

We denote the discretized initial pressure and transducer response as column vectors $\mathbf{p}_0$ and $\hat{\mathbf{p}}$, respectively, and express the forward model as

$$\hat{\mathbf{p}} = \mathbf{H}\mathbf{p}_0, \tag{87}$$

where $\mathbf{H}$ is the discretized system matrix of shape (NL, M). If $\mathbf{H}$ is full column rank (columns of the matrix are linearly independent), we can reconstruct an image by solving the optimization problem

$$\hat{\mathbf{p}}_0 = \underset{\mathbf{p}_0 \in \mathbb{R}^M}{\arg\min} \|\mathbf{H}\mathbf{p}_0 - \hat{\mathbf{p}}\|^2 \tag{88}$$

using FISTA with a constant step size [68]. We obtain the Lipschitz constant (the largest eigenvalue) of $2\mathbf{H}^{\mathrm{T}}\mathbf{H}$ through the power method with 32 iterations and use it in FISTA. Further, we can add nonnegativity of the initial pressure as a constraint to facilitate the reconstruction:

$$\hat{\mathbf{p}}_0 = \underset{\mathbf{p}_0 \in \mathbb{R}^M, \mathbf{p}_0 \geq \mathbf{0}}{\arg\min} \|\mathbf{H}\mathbf{p}_0 - \hat{\mathbf{p}}\|^2 . \tag{89}$$

If $\mathbf{H}$ is not full column rank or the noise level is high, we reconstruct the image by solving the regularized optimization problem

$$\hat{\mathbf{p}}_0 = \underset{\mathbf{p}_0 \in \mathbb{R}^M, \mathbf{p}_0 \geq \mathbf{0}}{\arg\min} \gamma_{\mathrm{c}} \|\mathbf{H}\mathbf{p}_0 - \hat{\mathbf{p}}\|^2 + \lambda |\mathbf{p}_0|_{\mathrm{TV}} \tag{90}$$

using FISTA with a constant step size [68]. Here, λ denotes the regularization parameter

and γ_c represents the system-specific measurement calibration factor. In practice, instead of dealing with $\hat{\mathbf{p}}$ directly, we obtain $\gamma_c\hat{\mathbf{p}}_0$ from $\gamma_c\hat{\mathbf{p}}$ (reading of the data acquisition system) by solving the problem $\gamma_c\hat{\mathbf{p}}_0 = \underset{\gamma_c\mathbf{p}_0\in\mathbb{R}^M,\gamma_c\mathbf{p}_0\geq\mathbf{0}}{\arg\min} \|\mathbf{H}(\gamma_c\mathbf{p}_0) - \gamma_c\hat{\mathbf{p}}\|^2 + \lambda|\gamma_c\mathbf{p}_0|_{\mathrm{TV}}$, which is equivalent to Eq. (90). The normalized values of $\gamma_c\hat{\mathbf{p}}_0$ (the same as those of $\hat{\mathbf{p}}_0$) are used for further analysis, and we ignore the effect of γ_c in this study. The term $|\mathbf{p}_0|_{\mathrm{TV}}$ in Eq. (90) denotes the total variation (TV) norm, defined as

$$|\mathbf{p}_0|_{\mathrm{TV}} = \sum_{\substack{2\leq m_1\leq M_1\\ 2\leq m_2\leq M_2\\ 2\leq m_3\leq M_3}} \sqrt{\begin{matrix}\left(\frac{\mathbf{p}_{0,m_1,m_2,m_3} - \mathbf{p}_{0,m_1-1,m_2,m_3}}{v_{m,1}}\right)^2 + \\ \left(\frac{\mathbf{p}_{0,m_1,m_2,m_3} - \mathbf{p}_{0,m_1,m_2-1,m_3}}{v_{m,2}}\right)^2 + \\ \left(\frac{\mathbf{p}_{0,m_1,m_2,m_3} - \mathbf{p}_{0,m_1,m_2,m_3-1}}{v_{m,3}}\right)^2\end{matrix}}. \quad (91)$$

In this definition, $v_{m,1}$, $v_{m,2}$, and $v_{m,3}$ represent the voxel sizes along the x-axis, y-axis, and z-axis, respectively; M_1, M_2, and M_3 ($M = M_1M_2M_3$) denote the number of voxels in the first, second, and third dimensions, respectively; and we reshape the column vector $\mathbf{p}_0$ to a 3D array $\mathbf{p}_{0,M_1\times M_2\times M_3}$ to express 3D information. Given a set of parameters for FISTA, the solution to Eq. (88) is linearly dependent on $\hat{\mathbf{p}}$. The solutions to Eqs. (89) and (90) are nonlinearly dependent on $\hat{\mathbf{p}}$ due to the nonnegative constraint ($\mathbf{p}_0 \geq \mathbf{0}$) and the TV regularization. Symbols for the forward operator and image reconstruction are summarized in **Table 2** in Appendix A.

3.4 Numerical simulations

3.4.1 Efficiency and accuracy

To quantify the efficiency and accuracy of fast forward operator based on Eq. (86), we perform numerical simulations by placing a numerical phantom of size $1 \times 1 \times 1$ cm^3 (**Figure 13a**, with voxel values of 1 inside the solid and 0 elsewhere) in four image subdomains of the system (D_1, D_2, D_3, and D_4 shown in **Figure 13b**). The virtual 2D array formed by the rotation of the four arc arrays is marked by blue arcs and the rotation of the

four subdomains around the same axis covers a domain (marked by black-dotted circles and arcs) enclosing all the image domains discussed in this study. We also implement the slow operator according to Eq. (82) for comparison.

We first quantify the efficiency of the fast operator by performing forward simulations with the numerical phantom in the subdomain D_1 using both operators with parameters $M = 50 \times 50 \times 50$ (voxel size of $0.2 \times 0.2 \times 0.2$ mm^3), $N = 396 \times 128$ (the four-arc array rotated for 99 locations and each arc downsized to 128 elements), $L = 4096$, $L' = 151$, and $K = 3$. We ran the single-thread-CPU version of each operator 36 times on a server with Ubuntu 20.04.6 LTS and Intel(R) Xeon(R) Gold 6248R CPU @ 3.00GHz. Denoting the computation times of both operators as t_{fast} and t_{slow}, respectively, we obtain $\text{mean}(t_{\text{fast}}) = 6.3$ min, $\text{std}(t_{\text{fast}}) = 0.4$ min, $\text{mean}(t_{\text{slow}}) = 267$ min, and $\text{std}(t_{\text{slow}}) = 19$ min for the 36 simulations and compare the values in **Figure 13c**. Further, we perform a Welch's *t*-test between $42t_{\text{fast}}$ and t_{slow}, showing insignificant difference (p-value = 0.59). Thus, the fast operator with $K = 3$ has approximately 42 times the speed of the slow operator, and the acceleration ratio can be further improved by reducing K.

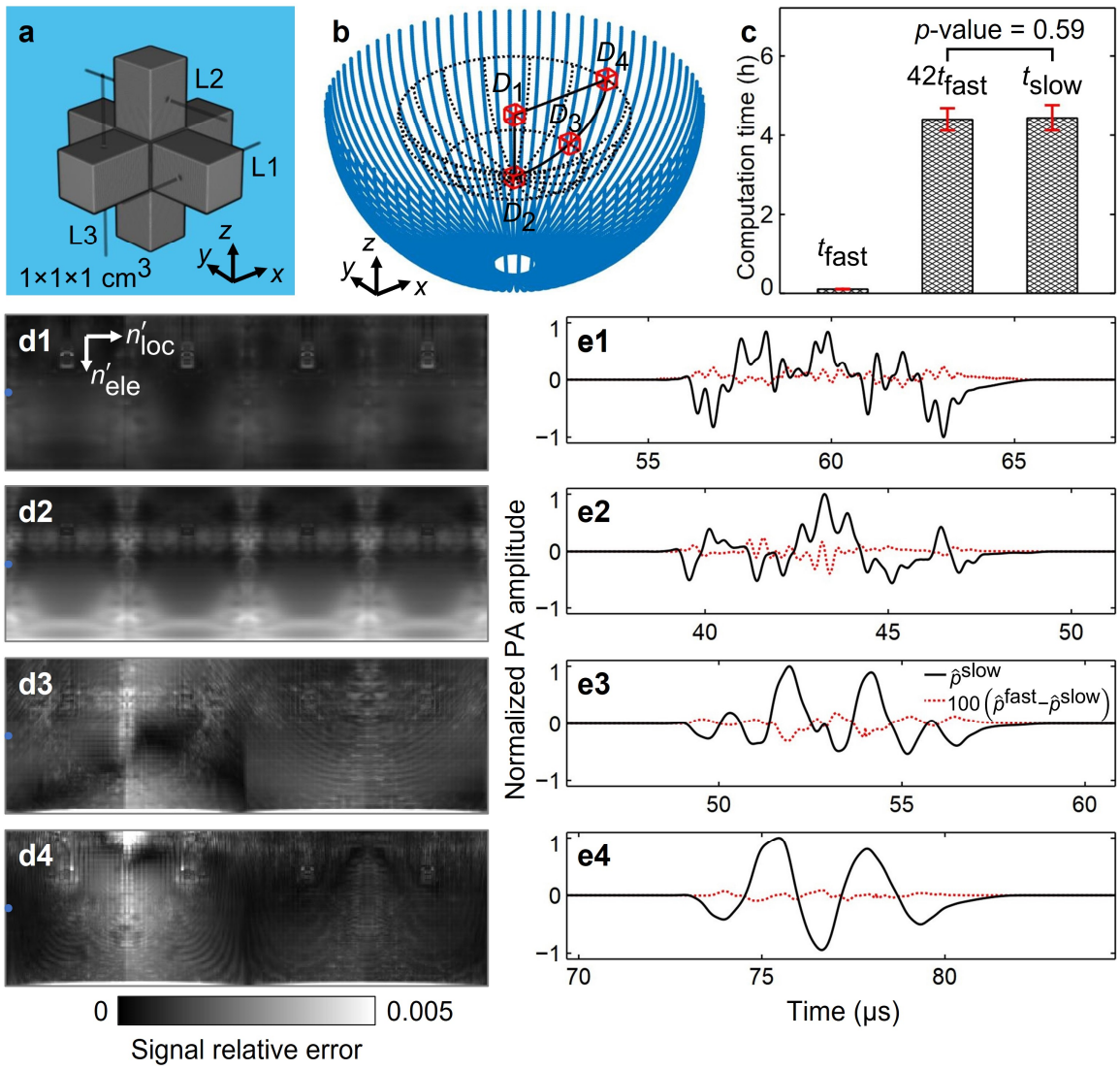

Figure 13 Accuracy of the fast operator in forward simulation. **a** A numerical phantom formed by three rectangular cuboids (each of size $2.6 \times 2.6 \times 10$ mm^3) intersecting at their centers. Voxel values are 1 inside the phantom and 0 outside. Three lines (L1, L2, and L3) are picked for image-domain comparisons **Figure 14**. **b** A virtual 2D array (blue-solid arcs), four subdomains of size $1 \times 1 \times 1$ cm^3 (D_1, D_2, D_3, and D_4), and the domain occupied by rotations of the four subdomains around the array axis (black-dotted circles and arcs). **c** Computation times (t_{fast} and t_{slow}) of the fast and slow operators, respectively, for the forward simulations with the numerical phantom in D_1 (36 repetitions). **d1–d4** Relative errors of the simulated signals for all virtual elements ($n'_{\text{loc}} = 1,2,\ldots,396, n'_{\text{ele}} = 1,2,\ldots,128$), with the numerical phantom in D_1, D_2, D_3, and D_4, respectively. **e1–e4** Compares of the signals ($\hat{p}^{\text{slow}}$ and $\hat{p}^{\text{fast}}$) simulated by the fast and slow operators for the 64-th element ($n'_{\text{ele}} = 64$) in the first virtual arc array ($n'_{\text{loc}} = 1$), with the phantom in D_1, D_2, D_3, and D_4, respectively.

Then, we demonstrate the accuracy of the fast operator by performing forward

simulations for the numerical phantom at the four subdomains (all the other parameters are the same as above). We regard the results from the slow operator as ground truth ($\hat{p}_{\mathrm{gt},n,l} = \hat{p}_{n,l}^{\mathrm{slow}}$) and define the relative error of the fast-operator results ($\hat{p}_{n,l}^{\mathrm{fast}}$) as $e_n = \frac{\sqrt{\sum_{l=1}^{L}\left(\hat{p}_{n,l}^{\mathrm{fast}} - \hat{p}_{\mathrm{gt},n,l}\right)^2}}{\sqrt{\sum_{l=1}^{L}\hat{p}_{\mathrm{gt},n,l}^2}}, n = 1,2,\dots,N, l = 1,2,\dots,L$. The values of e_n for the four subdomains (D_1, D_2, D_3, and D_4) are compared in **Figure 13d1–d4**, respectively, with n'_{loc} the index of the virtual arc array and n'_{ele} the index of the element in each array, $n'_{\mathrm{loc}} = 1,2,\dots,396, n'_{\mathrm{ele}} = 1,2,\dots,128$. These relative errors are generally small and all of them are below 0.005, which is enough for this study. For the 64-th element ($n'_{\mathrm{ele}} = 64$) on the first virtual array ($n'_{\mathrm{loc}} = 1$), marked as blue dots in **Figure 13d1–d4**, we compare the signals generated by the slow and fast operators, denoted as $\hat{p}^{\mathrm{slow}}$ and $\hat{p}^{\mathrm{fast}}$, respectively, in **Figure 13e1–e4**, which show the fast operator's high accuracy across the time domain.

Most importantly, we quantify the accuracy of the fast forward operator in the image domain by performing forward simulations using the slow operator (**H** implemented through Eq. (82), regarded as ground truth) and image reconstructions using iterations with nonnegativity constraints (Eq. (89) with **H** implemented through Eq. (86) and accelerated by two NVIDIA A100 GPUs), with the numerical phantom (**Figure 13a**) in the four subdomains D_1, D_2, D_3, and D_4, respectively. We define the reconstructed image $\hat{\mathbf{p}}_0$'s relative error to the ground-truth image $\mathbf{p}_{0,\mathrm{gt}}$ as $\frac{\|\hat{\mathbf{p}}_0 - \mathbf{p}_{0,\mathrm{gt}}\|}{\|\mathbf{p}_{0,\mathrm{gt}}\|}$. For the reconstructions with the numerical phantom in D_1, the relative error reduces to 0.003 as the iteration number n_{iter} reaches 256, as shown in **Figure 14a1**. We plot the values in the reconstructed images along three lines (L1, L2, and L3 shown in **Figure 13a**), in **Figure 14a2–a4**, respectively, for $n_{\mathrm{iter}} = 1, 4, 8, 64, 256$. Further, we denote the ground-truth (256-iteration reconstructed) values along the lines as $p_{0,\mathrm{L1}}$, $p_{0,\mathrm{L2}}$, and $p_{0,\mathrm{L3}}$ ($\hat{p}_{0,\mathrm{L1}}$, $\hat{p}_{0,\mathrm{L2}}$, and $\hat{p}_{0,\mathrm{L3}}$), respectively, and compare them in **Figure 14a5**. All the plots show the high accuracy of the fast forward operator in the image domain. Repeating the analysis in subdomains D_2, D_3, and D_4 further confirms the accuracy, as shown in **Figure 14b1–b5**, **c1–c5**, and **d1–**

d5, respectively.

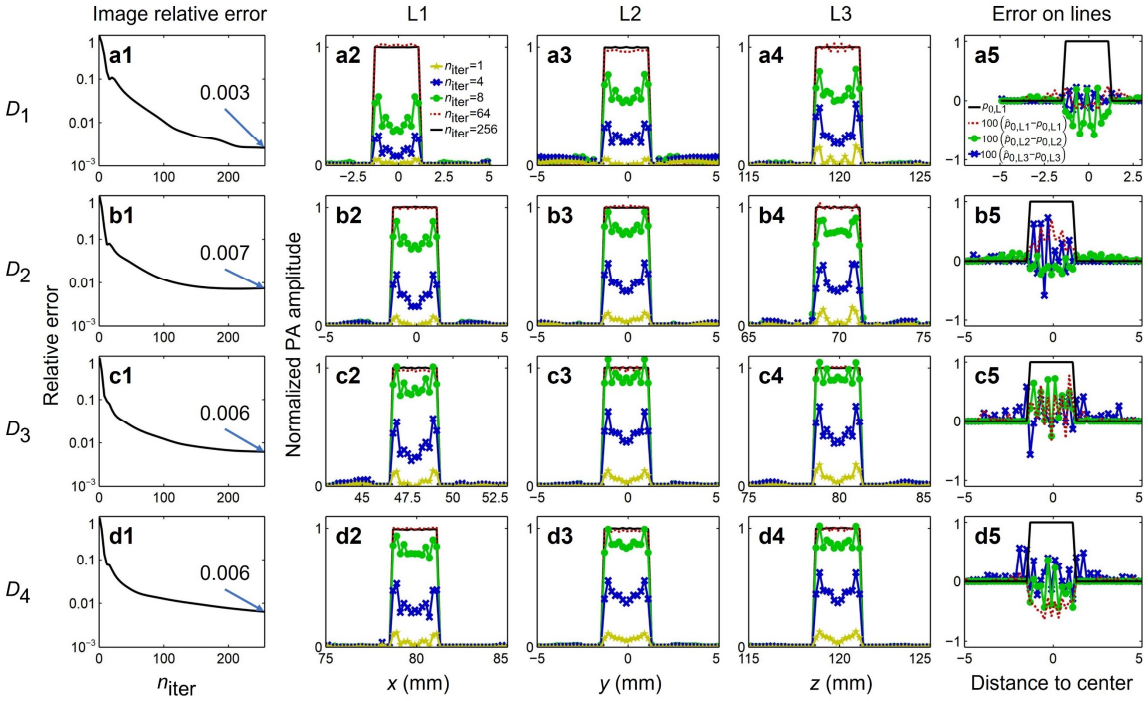

Figure 14 Image-domain accuracy of the fast forward operator. **a1** Relative error of the reconstructed image in the subdomain D_1 with 1 to 256 iterations. **a2–a4** Values on the lines L1, L2, and L3 (shown in **Figure 13a**), respectively, in the reconstructed images with $n_{\text{iter}} = 1, 4, 8, 64, 256$. **a5** Comparison between the ground truth ($p_{0,\text{L1}}$, $p_{0,\text{L2}}$, and $p_{0,\text{L3}}$) and 256-iteration reconstructed ($\hat{p}_{0,\text{L1}}$, $\hat{p}_{0,\text{L2}}$, and $\hat{p}_{0,\text{L3}}$) values along the three lines. **b1–b5**, **c1–c5**, and **d1–d5** The same analysis for reconstructions in subdomains D_2, D_3, and D_4, respectively.

3.4.2 Regularized iterative reconstruction

After validating the accuracy of the fast forward operator in both the signal domain and image domain, we use its GPU-accelerated implementation for both forward simulation and image reconstruction in the following analyses.

First, we perform a forward simulation with the numerical phantom in the subdomain D_1 (**Figure 13a-b**), add white noise with different amplitudes to the signals, and reconstruct the images through iterations with nonnegativity constraint (Eq. (89)) and iterations with TV regularization (Eq. (90)), respectively. We define the amplitude of the signal as half of the difference between its maximum and minimum values, the amplitude of noise as its

standard deviation, and their division as signal-to-noise ratio (SNR). We still use the relative error $\frac{\|\widehat{\mathbf{p}}_0-\mathbf{p}_{0,\mathrm{gt}}\|}{\|\mathbf{p}_{0,\mathrm{gt}}\|}$ between the reconstructed image $\widehat{\mathbf{p}}_0$ and ground truth $\mathbf{p}_{0,\mathrm{gt}}$ to quantify the accuracy. For SNRs of 2.4, 1.2, 0.8, and 0.6, the relative errors of the images reconstructed without TV regularization are shown in **Figure 15a**, which means that the reconstructions are unsuccessful. For iterations with TV regularization, the relative errors for regularization parameters of 5.0×10^5, 1.0×10^6, 2.0×10^6, and 4.0×10^6 (mm) are shown in **Figure 15b1–b4** for the four SNRs, respectively, which demonstrate that TV regularization stabilizes the iterations and the best choice of regularization parameter (enclosed in black-solid boxes) is noise level dependent. Plots with the best regularization parameters in **Figure 15b1–b4** are summarized in **Figure 15c**, which satisfies that the higher the signal noise level the poorer the image quality. Values on the lines L1, L2, and L3 (**Figure 13a**) in the reconstructed images after 8, 32, and 256 iterations are shown in **Figure 15d1–d3**, **e1–e3**, and **f1–f3**, respectively. The fact that 8 iterations already reconstruct a large portion of the image inspires us to perform only 8 iterations in certain steps of the motion correction discussed below, where the general structure is more important than accurate amplitude. The fact that 256 iterations bring minor improvements to 32 iterations inspires us to perform only 32 iterations for image reconstruction in functional imaging and the final step of motion correction as discussed below.

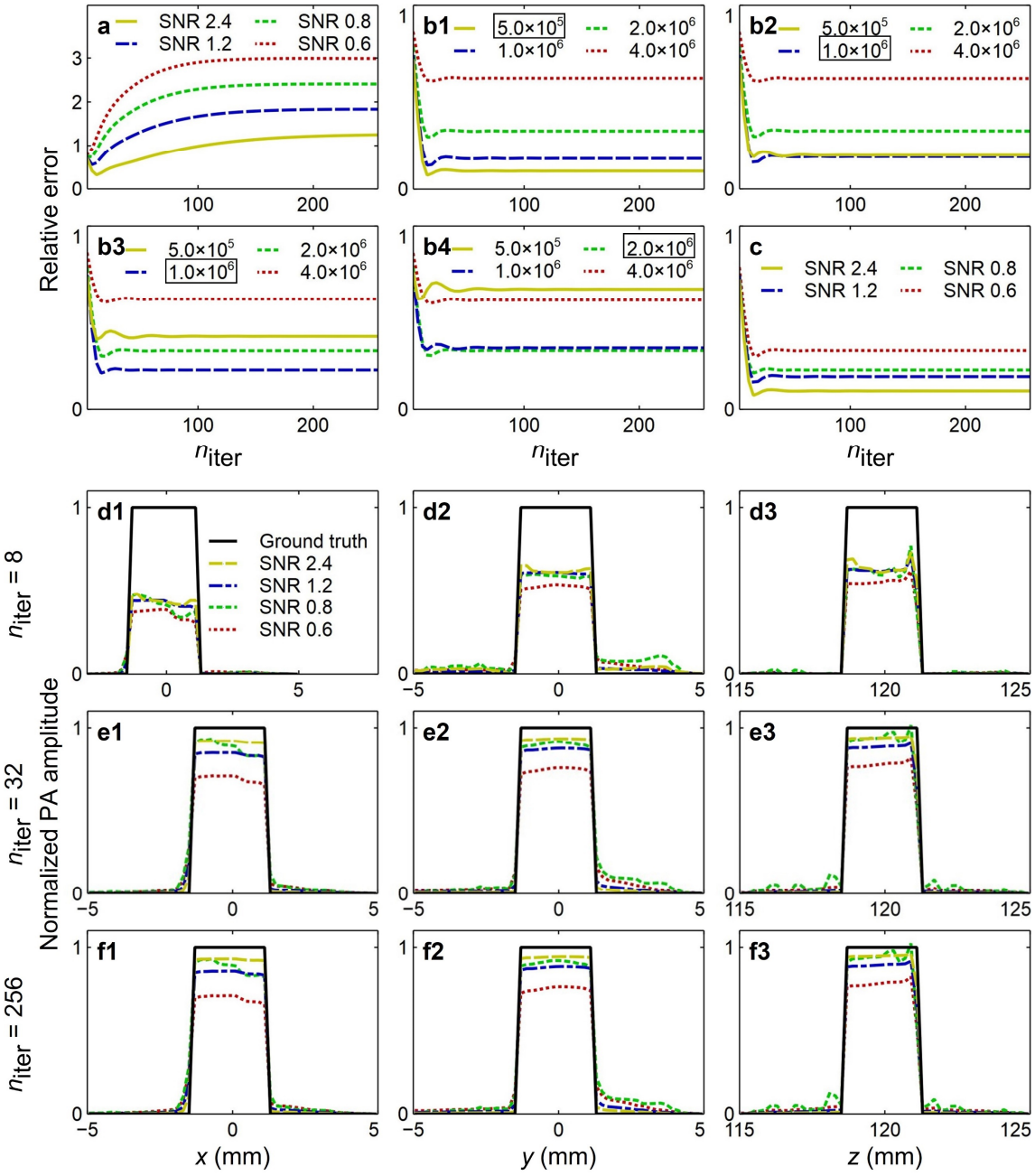

Figure 15 Iterative reconstruction for noisy data. **a** Relative errors of the images of the numerical phantom in D_1 reconstructed iteratively with nonnegativity constraint, for SNRs of 2.4, 1.2, 0.8, and 0.6. **b1**–**b4** Relative errors of the images reconstructed iteratively with TV regularization, for SNRs of 2.4, 1.2, 0.8, and 0.6, respectively, for different choices of regularization parameters (mm). **c** A summary of the plots in **b1**–**b4** with the best choices of regularization parameters (enclosed in black-solid boxes). **d1**–**d3**, **e1**–**e3**, and **f1**–**f3** Values on the lines L1, L2, and L3 in the images reconstructed with 8, 32, and 256, iterations, respectively.

Then we use a regularized iterative method to reconstruct a mouse brain numerical phantom with signals from the phantom detected by virtual arrays with different numbers of arcs: $4N_{\mathrm{loc}} = 76,40,28,20,16,12$, as shown in the fourth column in **Figure 18a**. Maximum amplitude projections (MAPs) of the images (along the z-axis) reconstructed by regularized iterative method with different regularization parameters for $4N_{\mathrm{loc}} = 76, 40, 28, 20, 16, 12$ are shown in **Figure 16**. Here, instead of selecting the regularization parameters quantitatively according to the relative error as in **Figure 15b1–b4**, we select them qualitatively by balancing mitigating artifacts and maintaining image resolution, which mimics the reality where the ground truth is unknown. The best images are indicated by red-solid boxes in **Figure 16**.

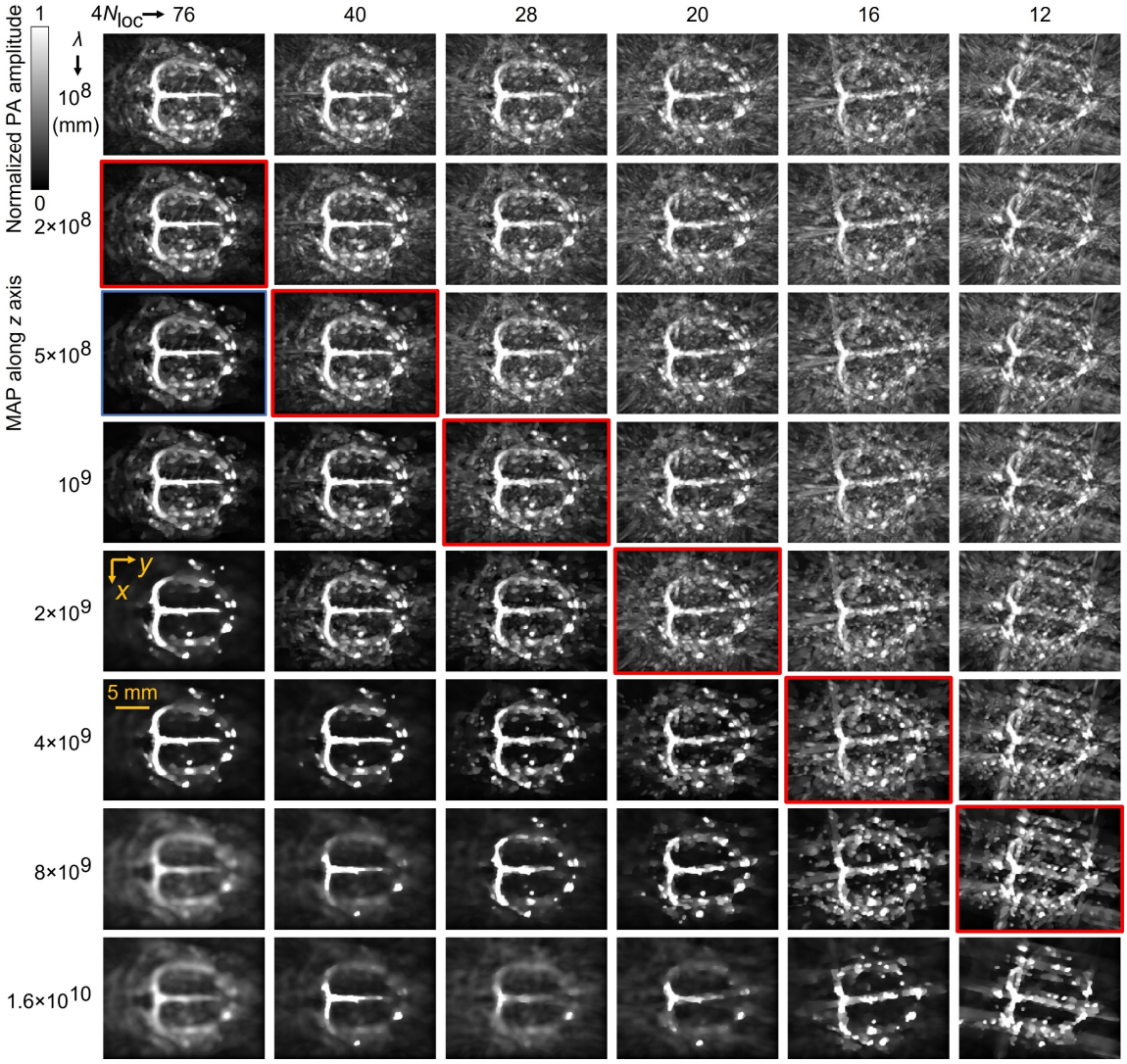

Figure 16 Regularized iterative reconstruction of mouse brain images with sparse sampling.

3.5 Discussion

We incorporate both EIR (limited bandwidth) and SIR (finite element size) into a system matrix of the 3D-PACT system explicitly expressed as point source responses and compress the matrix through SVD and FFT for efficient computation and slicing. Through numerical simulations, we validate that the compression results in ×42 significant acceleration with negligible error. This efficient explicit expression helps to improve image quality in accelerated iterative reconstructions. Also, the efficient system matrix slicing enables a method for fast functional imaging with sparse sampling and a method for intra-image

nonrigid motion correction, which are discussed in the following two chapters, respectively.

The theory is developed for homogeneous mediums, and further study may extend it to inhomogeneous mediums. For example, we can use a wave-equation-based forward operator to generate the point source responses and compress them for fast image reconstruction and/or data-driven manipulation. The compression technique can be used for any detection geometry and any detection surface of a transducer: 1D arrays for 2D imaging and 2D arrays for 3D imaging. It is also applicable to other tomographic imaging modalities, such as radial-sample MRI due to the high similarity in system matrices.

C h a p t e r 4

FAST FUNCTIONAL IMAGING WITH SPARSE SAMPLING

P. Hu, X. Tong, L. Lin, and L. V. Wang, "Data-driven system matrix manipulation enabling fast functional imaging and intra-image nonrigid motion correction in tomography," *Submitted*.

4.1 Abstract

Tomographic imaging modalities are described by large system matrices. Sparse sampling degrades the system matrix and image quality. Various existing techniques improve the image quality without correcting the system matrices. Here, we compress the system matrices to improve computational efficiency (e.g., 42 times) using singular value decomposition and fast Fourier transform. Enabled by the efficiency, we propose fast sparsely sampling functional imaging by incorporating a densely sampled prior image into the system matrix, which maintains the critical linearity while mitigating artifacts. We demonstrate the method in 3D photoacoustic computed tomography (PACT) with significantly improved image qualities and clarify their applicability to X-ray computed tomography (X-ray CT) and magnetic resonance imaging (MRI) due to the similarities in system matrices.

4.2 Background

To achieve high temporal resolution in tomographic functional imaging, the spatial sampling density is often sacrificed, which introduces artifacts in the reconstructed image [54], [69], [70] and may affect the functional signal extraction.

For sparse sampling functional imaging, traditional methods [37], [39], [54], [57], [58] mitigate artifacts in images but their performances drop sharply as the sampling density reduces. Deep neural networks (DNNs) [31], [34], [59], [60], [62], [66], [71] show high performance in mitigating artifacts but tend to generate false image features when the

sampling density is low, and they require imaging-modality- and device-dependent datasets, which are not always available. Moreover, most of the methods introduce nonlinearity while mitigating artifacts, which disrupts the functional signals that are often much weaker than background signals.

In this study, using the four-arc PACT system as an example, we compress the system matrices using singular value decomposition (SVD) and fast Fourier transform (FFT), which enables efficient system matrix slicing and manipulation. Then we incorporate a prior image into the system matrix to reduce unknown variables in image reconstruction. Special configurations in the method maintain linearity in image reconstruction while substantially mitigating artifacts, which is critical for weak functional signal extraction.

4.3 Theory

4.3.1 A hybrid method for fast functional imaging

We first acquire signals from an object with transducers at enough locations (N, dense sampling). Then we continuously acquire multiple sets of signals from the object with transducers at a smaller number of locations (N', sparse sampling) for fast functional imaging. We assume that the object does not move as a whole during functional imaging.

For image reconstruction, we denote the densely sampled signals and a set of sparsely sampled signals as $\widehat{\mathbf{p}}_{NL\times 1}$ and $\widehat{\mathbf{p}}^{S}_{N'L\times 1}$, respectively. We first obtain an image $\widehat{\mathbf{p}}_{0,M\times 1}$ from the densely sampled signals $\widehat{\mathbf{p}}_{NL\times 1}$ by solving the regularized optimization problem

$$\widehat{\mathbf{p}}_{0,M\times 1} = \underset{\mathbf{p}_{0,M\times 1}\in\mathbb{R}^{M},\mathbf{p}_{0,M\times 1}\geq\mathbf{0}}{\operatorname{argmin}} \gamma_{\mathrm{c}}\left\|\mathbf{H}_{NL\times M}\mathbf{p}_{0,M\times 1} - \widehat{\mathbf{p}}_{NL\times 1}\right\|^{2} + \lambda\left|\mathbf{p}_{0,M\times 1}\right|_{\mathrm{TV}}. \qquad (92)$$

Here, γ_{c} represents the system-specific measurement calibration factor, meaning $\gamma_{\mathrm{c}}\widehat{\mathbf{p}}_{NL\times 1}$ is the reading of the data acquisition system; $\mathbf{H}_{NL\times M}$ is the dense-sampling system matrix; $\left|\mathbf{p}_{0,M\times 1}\right|_{\mathrm{TV}}$ denotes $\mathbf{p}_{0,M\times 1}$'s total variation (TV) norm, defined in Eq. (91); and λ means the regularization parameter. To obtain an image from $\widehat{\mathbf{p}}^{S}_{N'L\times 1}$ directly, we can use $\widehat{\mathbf{p}}^{S}_{N'L\times 1}$ and the sparse-sampling system matrix $\mathbf{H}^{\mathrm{S}}_{N'L\times M}$ to replace $\widehat{\mathbf{p}}_{NL\times 1}$ and $\mathbf{H}_{NL\times M}$, respectively, in Eq. (92) and solve it. However, the nonlinearity introduced by the nonnegativity

constraint and TV regularization may disrupt the functional signals. To maintain the linearity under sparse sampling, we treat $\widehat{\mathbf{p}}_{0,M\times 1}$ as a prior image and apply a smooth modulation to it to extract the dominant sources of aliasing artifacts by solving the optimization problem

$$\widehat{\boldsymbol{\mu}}_{0,M'\times 1} = \underset{\mathbf{p}_{0,M'\times 1}\in\mathbb{R}^{M'}}{\operatorname{argmin}} \left\| \mathbf{H}^{\mathrm{s}}_{N'L\times M}\left(\widehat{\mathbf{p}}_{0,M\times 1}\odot\left(\mathbf{U}_{M\times M'}\boldsymbol{\mu}_{0,M'\times 1}\right)\right) - \widehat{\mathbf{p}}^{s}_{N'L\times 1}\right\|^2 \qquad (93)$$

using a fast iterative shrinkage-thresholding algorithm (FISTA) with a constant step size [68]. Here, $\boldsymbol{\mu}_{0,M'\times 1}$ is a modulation image in the form of a column vector of size $M' = M'_1M'_2M'_3$ ($\ll M$), the symbol $\odot$ denotes the element-wise product, and $\mathbf{U}_{M\times M'}$ is an upsampling operator transferring $\boldsymbol{\mu}_{0,M'\times 1}$ to a smooth modulation image in the column-vector form of size M. To implement $\mathbf{U}_{M\times M'}$, we reshape the vector $\boldsymbol{\mu}_{0,M'\times 1}$ into a 3D array $\boldsymbol{\mu}_{0,M'_1\times M'_2\times M'_3}$, apply trilinear interpolation to the 3D array to obtain $\boldsymbol{\mu}_{0,M_1\times M_2\times M_3}$, and flatten it to obtain $\mathbf{U}_{M\times M'}\boldsymbol{\mu}_{0,M'\times 1}$. The array $\boldsymbol{\mu}_{0,M_1\times M_2\times M_3}$ is a smooth array determined only by M' independent values. Thus, the expression $\widehat{\mathbf{p}}_{0,M\times 1}\odot\left(\mathbf{U}_{M\times M'}\boldsymbol{\mu}_{0,M'\times 1}\right)$ represents a smooth modulation of the prior image $\widehat{\mathbf{p}}_{0,M\times 1}$. By solving the optimization problem in Eq. (93) we obtain $\widehat{\boldsymbol{\mu}}_{0,M'\times 1}$ and represent the dominant sources causing aliasing artifacts by $\widehat{\mathbf{p}}_{0,M\times 1}\odot\left(\mathbf{U}_{M\times M'}\widehat{\boldsymbol{\mu}}_{0,M'\times 1}\right)$, which correspond to the dominant signals causing aliasing artifacts $\mathbf{H}^{\mathrm{s}}_{N'L\times M}\left(\widehat{\mathbf{p}}_{0,M\times 1}\odot\left(\mathbf{U}_{M\times M'}\widehat{\boldsymbol{\mu}}_{0,M'\times 1}\right)\right)$. It needs to be noted that the implementation of the forward operator based on Eq. (86) allows for efficient slicing of the system matrix, such as the matrix $\mathbf{H}^{\mathrm{s}}_{N'L\times M}$, a slicing with respect to transducer element indices. Also, for a given set of parameters in FISTA, the solution to Eq. (93), $\widehat{\boldsymbol{\mu}}_{0,M'\times 1}$, depends linearly on $\widehat{\mathbf{p}}^{s}_{N'L\times 1}$.

Then, we apply the dense-sampling system matrix $\mathbf{H}_{NL\times M}$ to $\widehat{\mathbf{p}}_{0,M\times 1}\odot\left(\mathbf{U}_{M\times M'}\widehat{\boldsymbol{\mu}}_{0,M'\times 1}\right)$ to simulate the densely sampled signals $\mathbf{H}_{NL\times M}\left(\widehat{\mathbf{p}}_{0,M\times 1}\odot\left(\mathbf{U}_{M\times M'}\widehat{\boldsymbol{\mu}}_{0,M'\times 1}\right)\right)$, from which we use the universal back-projection (UBP) method to reconstruct an image, denoted as $\widehat{\mathbf{p}}^{\mathrm{m}}_0$. We further remove the aliasing artifacts caused by signals from $\widehat{\mathbf{p}}_{0,M\times 1}\odot\left(\mathbf{U}_{M\times M'}\widehat{\boldsymbol{\mu}}_{0,M'\times 1}\right)$ by subtracting these signals from $\widehat{\mathbf{p}}^{s}_{N'L\times 1}$ to obtain the residual signals $\widehat{\mathbf{p}}^{s}_{N'L\times 1} - \mathbf{H}^{\mathrm{s}}_{N'L\times M}\left(\widehat{\mathbf{p}}_{0,M\times 1}\odot\left(\mathbf{U}_{M\times M'}\widehat{\boldsymbol{\mu}}_{0,M'\times 1}\right)\right)$ and using them with

the UBP method to reconstruct an image, denoted as $\hat{\mathbf{p}}_0^{\mathrm{r}}$. Due to the linearity of the UBP method, both images $\hat{\mathbf{p}}_0^{\mathrm{m}}$ and $\hat{\mathbf{p}}_0^{\mathrm{r}}$ are linearly dependent on $\hat{\boldsymbol{\mu}}_{0,M'\times 1}$ and thus on $\hat{\mathbf{p}}_{N'L\times 1}^{s}$. We combine these two images to obtain the final image

$$\hat{\mathbf{p}}_0^{\mathrm{h}} = \hat{\mathbf{p}}_0^{\mathrm{m}} + \hat{\mathbf{p}}_0^{\mathrm{r}}, \tag{94}$$

for $\hat{\mathbf{p}}_{N'L\times 1}^{s}$. We summarize the workflow of the hybrid method in **Figure 17**. We repeat the process for other sets of sparsely sampled raw data to obtain the images for further functional signal extraction. Additional symbols for fast functional imaging are shown in **Table 3** in Appendix A.

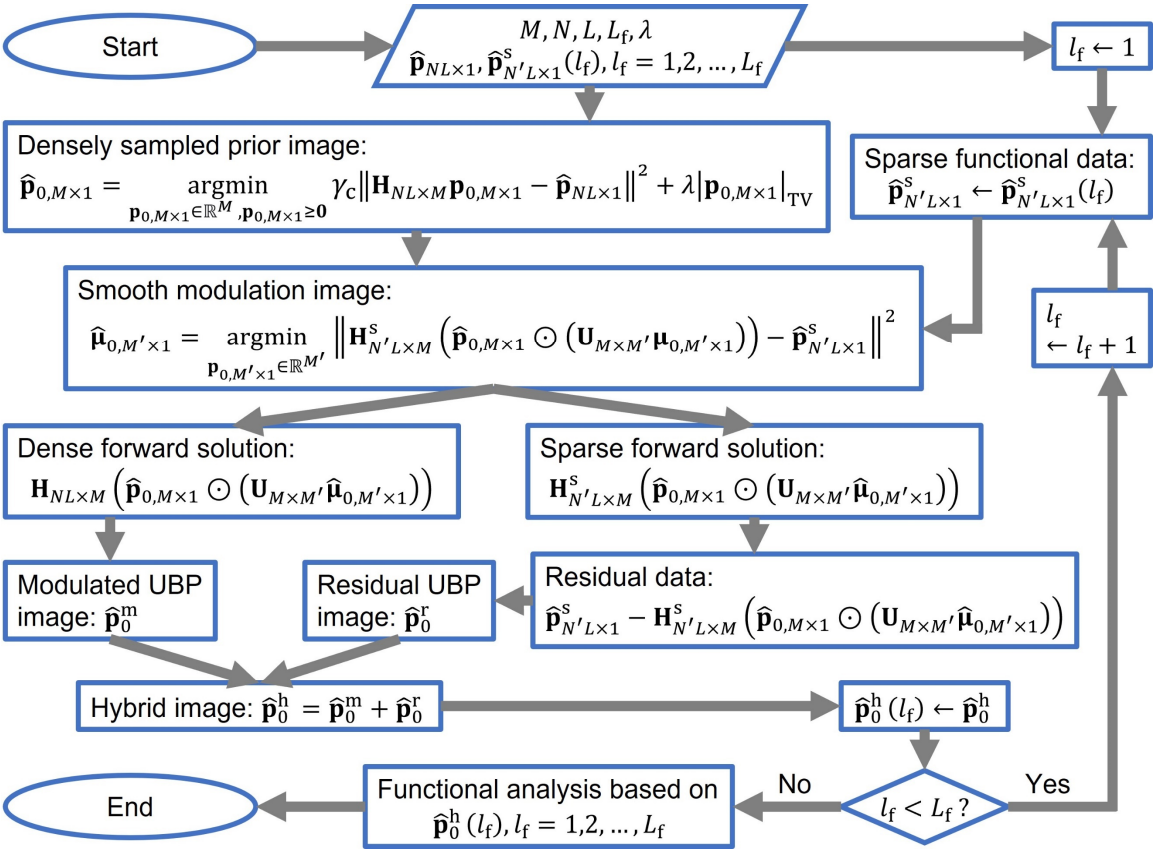

Figure 17 Workflow of the hybrid method for image reconstruction from sparsely sampled raw data and a prior image.

4.3.2 Functional signal extraction

Given a set of reconstructed images $\hat{\mathbf{p}}_{0,l_{\mathrm{f}}}$, l_{f}=1,2, ... ,L_{f} in functional imaging, we propose a method to extract functional signals from them through regularized correlation. To explain

the process, we define $\underset{j}{\text{sum}}\, a_j = \sum_{j=1}^{J} a_j$, $\underset{j}{\text{mean}}\, a_j = \frac{\underset{j}{\text{sum}}\, a_j}{J}$, and $\underset{j}{\text{norm}}\, a_j = \sqrt{\sum_{j=1}^{J} a_j^2}$ for a series of numbers $a_j, j = 1,2, \dots, J$. We assume that the functional signal has a profile $\alpha_{\mathrm{f},l_\mathrm{f}}, l_\mathrm{f}=1,2, \dots, L_\mathrm{f}$, normalized to $\bar{\alpha}_{\mathrm{f},l_\mathrm{f}} = \left(\alpha_{\mathrm{f},l_\mathrm{f}} - \underset{l_\mathrm{f}'}{\text{mean}}\, \alpha_{\mathrm{f},l_\mathrm{f}'}\right)\left(\underset{l_\mathrm{f}''}{\text{norm}}\left(\alpha_{\mathrm{f},l_\mathrm{f}''} - \underset{l_\mathrm{f}'}{\text{mean}}\, \alpha_{\mathrm{f},l_\mathrm{f}'}\right)\right)^{-1}$. The m-th voxel has a value of $\hat{p}_{0,l_\mathrm{f},m}$ in the l_f-th image $\hat{\mathbf{p}}_{0,l_\mathrm{f}}$. We can quantify the functional amplitude at the m-th voxel through the Pearson correlation coefficient (PCC) between $\hat{p}_{0,l_\mathrm{f},m}$ and $\alpha_{\mathrm{f},l_\mathrm{f}}$:

$$\underset{l_\mathrm{f}}{\mathrm{PCC}}\left(\hat{p}_{0,l_\mathrm{f},m}, \alpha_{\mathrm{f},l_\mathrm{f}}\right) = \frac{\underset{l_\mathrm{f}}{\text{sum}}\left(\bar{\alpha}_{\mathrm{f},l_\mathrm{f}}\left(\hat{p}_{0,l_\mathrm{f},m} - \underset{l_\mathrm{f}'}{\text{mean}}\, \hat{p}_{0,l_\mathrm{f}',m}\right)\right)}{\underset{l_\mathrm{f}}{\text{norm}}\left(\hat{p}_{0,l_\mathrm{f},m} - \underset{l_\mathrm{f}'}{\text{mean}}\, \hat{p}_{0,l_\mathrm{f}',m}\right)}, m = 1,2, \dots, M, \tag{95}$$

which is not robust for very-low-amplitude regions. To improve the robustness, we add a regularization term to the denominator and obtain the regularized PCC (PCCR):

$$\underset{l_\mathrm{f}}{\mathrm{PCCR}_{\lambda_\mathrm{f}}}\left(\hat{p}_{0,l_\mathrm{f},m}, \alpha_{\mathrm{f},l_\mathrm{f}}\right) = \frac{\underset{l_\mathrm{f}}{\text{sum}}\left(\bar{\alpha}_{\mathrm{f},l_\mathrm{f}}\left(\hat{p}_{0,l_\mathrm{f},m} - \underset{l_\mathrm{f}'}{\text{mean}}\, \hat{p}_{0,l_\mathrm{f}',m}\right)\right)}{\underset{l_\mathrm{f}}{\text{norm}}\left(\hat{p}_{0,l_\mathrm{f},m} - \underset{l_\mathrm{f}'}{\text{mean}}\, \hat{p}_{0,l_\mathrm{f}',m}\right) + \lambda_\mathrm{f}\, \underset{m'}{\text{mean}}\, \underset{l_\mathrm{f}}{\text{norm}}\left(\hat{p}_{0,l_\mathrm{f},m'} - \underset{l_\mathrm{f}'}{\text{mean}}\, \hat{p}_{0,l_\mathrm{f}',m'}\right)}, \quad m = 1,2, \dots, M. \tag{96}$$

A 3D functional image is formed by assigning the regularized correlation to each voxel.

The assumed functional signal profile $\alpha_{\mathrm{f},l_\mathrm{f}}$ is directly available in the following numerical simulations (Eq. (98)). For mouse brain functional imaging *in vivo*, we first let $\alpha_{\mathrm{f},l_\mathrm{f}}$ be a sinusoidal profile synchronized with the paw stimulation pattern and apply it to the UBP reconstructed images with $4N_{\mathrm{loc}} = 76$ to obtain the functional region. Then we extract the true functional profile from this region and apply it to images with $4N_{\mathrm{loc}} = 40, 20, 12$.

4.4 Numerical simulations

4.4.1 Antialiasing and linearity

We first demonstrate the performance of the hybrid method by using it to reconstruct images from signals of a numerical phantom acquired by virtual arrays with different numbers of arcs: $4N_{\mathrm{loc}} = 76, 40, 28, 20, 16, 12$. Here N_{loc} is the number of rotating locations of the four-arc array (with 128 transducer elements in each arc) in a virtual array. Images of the numerical phantom reconstructed using UBP, the regularized iterative method (Eq. (90)), and the hybrid method (Eq. (94) with a prior image obtained by performing a smooth modulation to the numerical phantom) are shown in the first three columns in **Figure 18a**. The used virtual arrays are shown as blue arcs with red boundaries in the fourth column in **Figure 18a**. The best regularization parameters (obtained through parameter tunning) for different values of $4N_{\mathrm{loc}}$ are used in the iterative method in **Figure 18a**. We see that, as $4N_{\mathrm{loc}}$ decreases from 76 to 12, the artifacts in the images reconstructed using UBP become more abundant. The regularized iterative method mitigates the relatively weak artifacts in all images but failed to suppress the strong artifacts such as in the images with $4N_{\mathrm{loc}} = 16, 12$. In contrast, with the help of the prior image, the hybrid method significantly mitigates the artifacts. Quantitatively, for each method, we calculate the structural similarity index measures (SSIMs) between the images with $4N_{\mathrm{loc}} = 40, 28, 20, 16, 12$ and that with $4N_{\mathrm{loc}} = 76$, and compare the values in **Figure 18b**. The hybrid method performs the best in mitigating artifacts and maintaining true features.

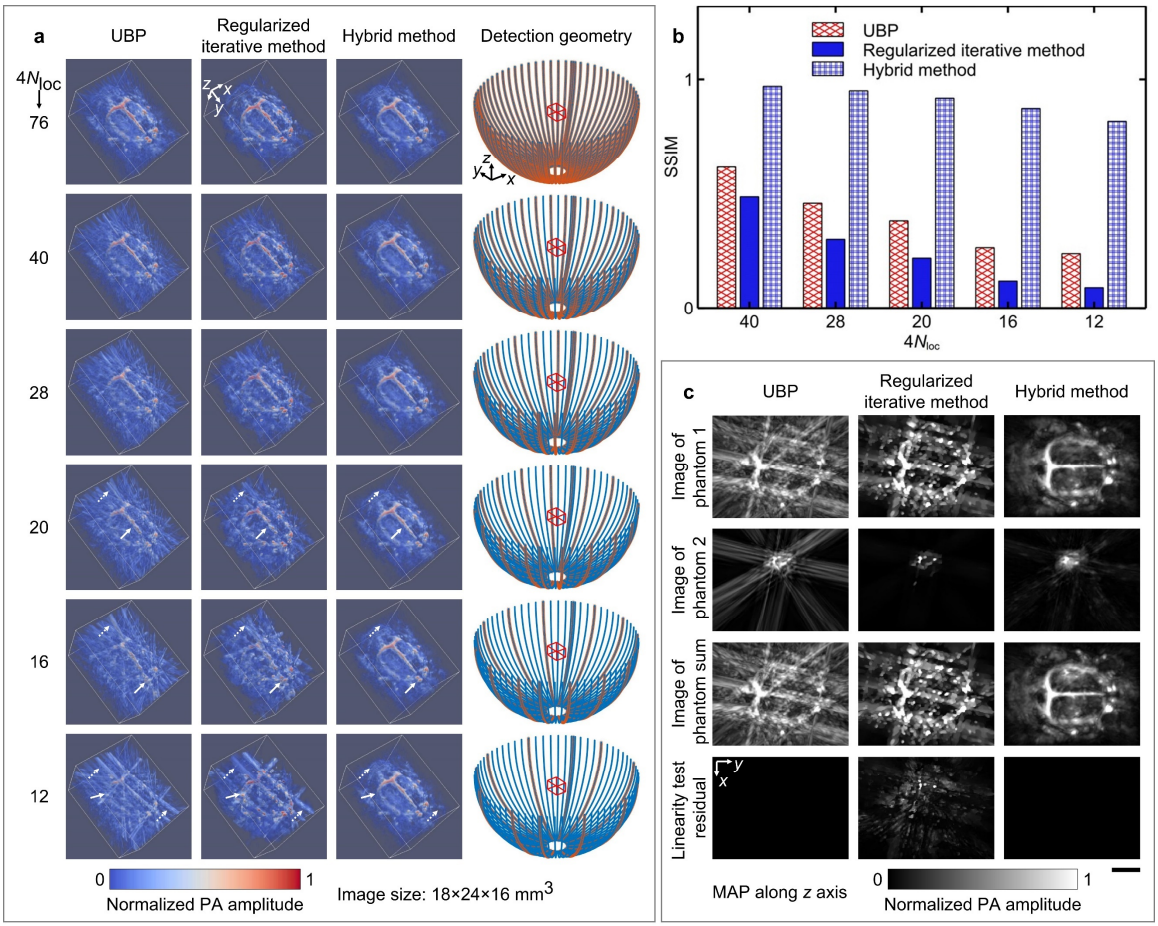

Figure 18 UBP, regularized iterative method, and the proposed hybrid method for sparse-sampling imaging, and their linearity tests. **a** Images reconstructed by the three methods (first three columns) from signals detected at sparsely distributed elements (red-bounded blue curves in the fourth column) for $4N_{\mathrm{loc}} = 76, 40, 28, 20, 16, 12$. Examples of maintained features and suppressed artifacts are indicated by white-solid and white-dotted arrows, respectively. PA, photoacoustic. **b** SSIMs between the reconstructed images with $4N_{\mathrm{loc}} = 40, 28, 20, 16, 12$ and those with $4N_{\mathrm{loc}} = 76$ for the three methods. **c** Linearity tests of the three methods for $4N_{\mathrm{loc}} = 12$. Scale bar, 5 mm.

Additionally, we test the linearity of each method by reconstructing two numerical phantoms and their summation. We summarize the maximum amplitude projections (MAPs) in the linearity tests with $4N_{\mathrm{loc}} = 12$ in **Figure 18c**, which shows that UBP and the hybrid method are linear, but the regularized iterative method is nonlinear. In the hybrid method, the prior image quality does not affect the linearity but affects the reconstructed image quality.

4.4.2 Functional imaging

Due to the requirement of a prior image, the hybrid method's practical value mainly lies in fast functional imaging with sparse sampling. We perform numerical simulations of functional imaging with $4N_{\mathrm{loc}} = 12$. We obtain numerical phantoms for functional imaging using

$$\mathbf{p}_{0,l_\mathrm{f}} = \alpha_{\mathrm{b},l_\mathrm{f}}\mathbf{p}_{0,\mathrm{b}} + \alpha_{\mathrm{f},l_\mathrm{f}}\mathbf{p}_{0,\mathrm{f}}, l_\mathrm{f}=1,2,\ldots,L_\mathrm{f}. \tag{97}$$

Here, $\mathbf{p}_{0,l_\mathrm{f}}$ (**Figure 19a**, voxel size $0.1 \times 0.1 \times 0.1$ mm^3) is the l_f-th numerical phantom; $\mathbf{p}_{0,\mathrm{b}}$ (**Figure 19b**) is the background phantom obtained in imaging with dense sampling and $\mathbf{p}_{0,\mathrm{f}}$ is the functional phantom obtained from by smoothing, downsampling, and zero padding $\mathbf{p}_{0,\mathrm{b}}$; $\alpha_{\mathrm{b},l_\mathrm{f}}$ and $\alpha_{\mathrm{f},l_\mathrm{f}}$ are modulation factors of the two phantoms, respectively; and L_f is the number of numerical phantoms for functional imaging. The way $\mathbf{p}_{0,\mathrm{b}}$ is obtained guarantees that the mean value of nonzero voxels in $\mathbf{p}_{0,\mathrm{b}}$ approximately equals that in $\mathbf{p}_{0,\mathrm{f}}$. For simulations in this study, we use a virtual array formed by 12 arc arrays (**Figure 19c**), let $L_\mathrm{f} = 36$, and let $\alpha_{\mathrm{b},l_\mathrm{f}} \sim N(1,0.1)$, $l_\mathrm{f}=1,2,\ldots,L_\mathrm{f}$ (**Figure 19d**), an amplitude similar to the image relative difference we observed in mouse brain functional imaging. Also, we let

$$\alpha_{\mathrm{f},l_\mathrm{f}} = \frac{A_\mathrm{f}}{2}\left(\sin\frac{6\pi(l_\mathrm{f}-1)}{L_\mathrm{f}} + 1\right), l_\mathrm{f}=1,2,\ldots,L_\mathrm{f}, \tag{98}$$

where A_f is the functional amplitude. The values of $\alpha_{\mathrm{f},l_\mathrm{f}}$ with $A_\mathrm{f} = 0.18, 0.06, 0.02$ are shown in **Figure 19d** and used in the simulations.

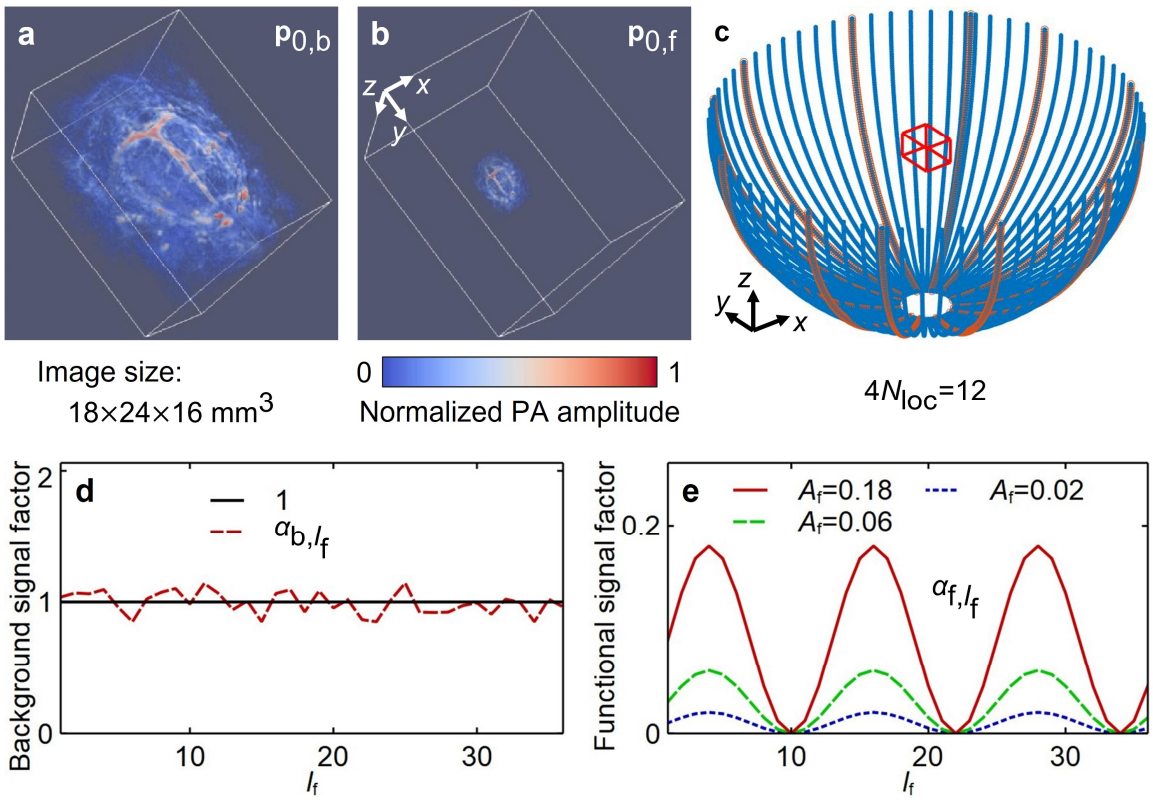

Figure 19 Numerical phantoms for functional imaging. **a-b** The background and functional numerical phantoms, respectively, for functional imaging simulation. **c** A virtual array formed by 12 arc arrays, shown as arcs with red boundaries. **d-e** Modulation factors of the background and functional phantoms, respectively.

We perform forward simulations, image reconstructions (UBP, the regularized iterative method in Eq. (90), and the hybrid method in Eq. (94)), and functional signal extractions (Eq. (96)) with different λ_f. The functional images extracted from four sets of images (ground-truth images and images reconstructed with three methods) using the regularized-correlation-based method with $\lambda_f = 1.6$ are shown in **Figure 20**. We observe that artifacts in the UBP-reconstructed images cause artifacts in the functional images, the regularized iterative method mitigates artifacts in functional images but also compromises the true functional region, and the proposed hybrid method is superior to both UBP and the regularized iterative method. The results with other values of λ_f also support this observation.

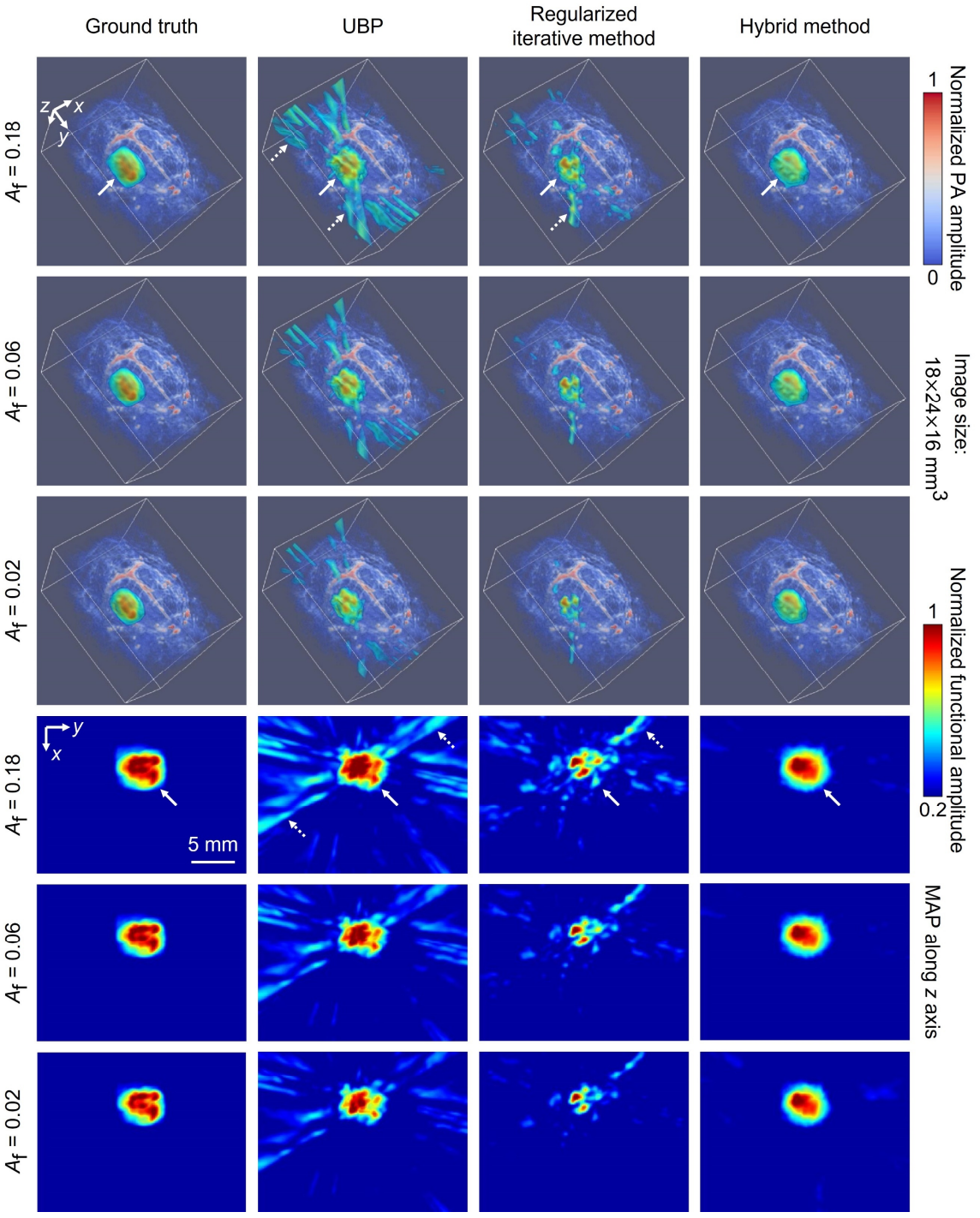

Figure 20 Functional images extracted using the regularized-correlation-based method with $\lambda_f = 1.6$ from ground-truth images and images reconstructed with UBP, the regularized iterative method, and the hybrid method for $A_f = 0.18, 0.06, 0.02$. The first three rows show both the 3D functional and background images and the last three rows show the MAPs of the functional images along the z-axis. In the first and fourth rows, the true functional regions and examples of false positive regions are indicated by white-solid and white-dotted arrows, respectively.

4.5 *In vivo* experiments

We apply UBP, the regularized iterative method, and the hybrid method to mouse brain functional imaging *in vivo* using the four-arc system. We first obtain a prior image of a mouse brain through dense sampling ($4N_{\mathrm{loc}} = 396$), then electrically stimulate its right front paw and continuously acquire signals from the mouse brain through sparse sampling ($4N_{\mathrm{loc}} = 76$, 2 s per image). We use subsets of the sparsely sampled signals ($4N_{\mathrm{loc}} = 40, 20, 12$) to demonstrate the performance of the hybrid method. For one set of sparsely sampled signals, the images reconstructed using UBP, the regularized iterative method, and the hybrid method for $4N_{\mathrm{loc}} = 40, 20, 12$ are shown in **Figure 21a**. We observe that the iterative method mitigates the artifacts (e.g., those indicated by white-dotted arrows) but compromises low-amplitude features (e.g., those indicated by white-solid arrows for $4N_{\mathrm{loc}} = 20$). In contrast, the hybrid method maintains low-amplitude features while substantially mitigating the artifacts, resulting in images more similar to the densely sampled image. Electrical stimulation of the mouse's right front paw occurs in five cycles, each with 12-s stimulation on and 12-s off, as shown in **Figure 21b**.

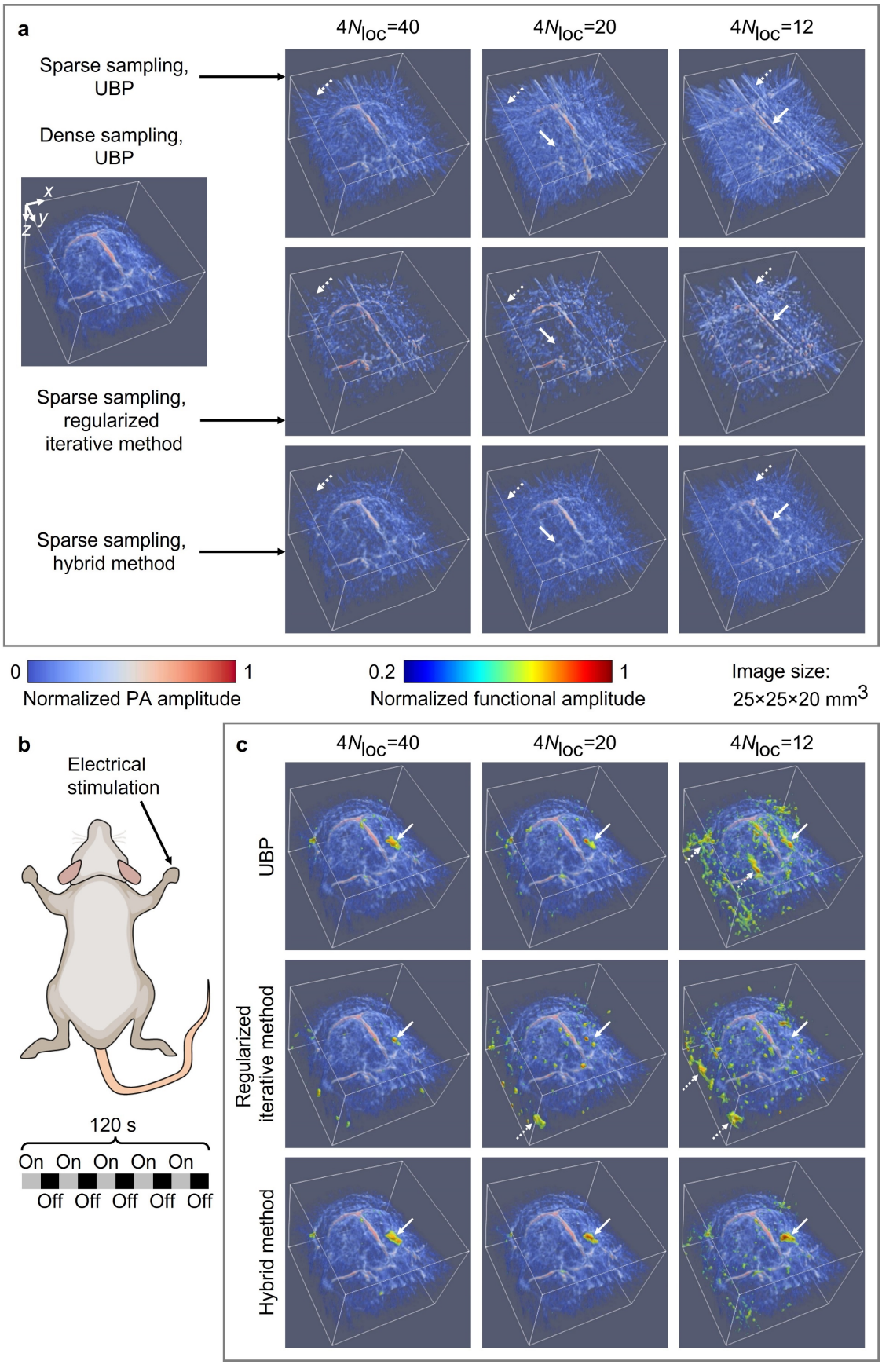
a
4Nloc=40
4Nloc=20
4Nloc=12
Sparse sampling, UBP
Dense sampling, UBP
x
y
z
Sparse sampling, regularized iterative method
Sparse sampling, hybrid method
0
1
Normalized PA amplitude
0.2
1
Normalized functional amplitude
Image size: 25×25×20 mm3
b
Electrical stimulation
120 s
On On On On On
Off Off Off Off Off
c
4Nloc=40
4Nloc=20
4Nloc=12
UBP
Regularized iterative method
Hybrid method

Figure 21 Sparse-sampling mouse brain functional imaging *in vivo*. **a** A densely sampled image of a mouse brain reconstructed by UBP (left column, $4N_{\mathrm{loc}} = 396$) and sparsely sampled images of the mouse brain reconstructed using UBP (first row), the regularized iterative method (second row), and the hybrid method (third row), respectively, for $4N_{\mathrm{loc}} = 40, 20, 12$. Examples of suppressed artifacts and maintained features are indicated by white-dotted and white-solid arrows, respectively. **b** Electrical stimulation to the mouse's right front paw: five cycles, each with 12-s stimulation on and 12-s off. **c** Functional images obtained from the images reconstructed using UBP (first row, $\lambda_{\mathrm{f}} = 0.32$), the regularized iterative method (second row, $\lambda_{\mathrm{f}} = 0.08$), and the hybrid method (third row, $\lambda_{\mathrm{f}} = 0.32$), respectively, for $4N_{\mathrm{loc}} = 40, 20, 12$. The true functional regions in all images are indicated by white-solid arrows, and examples of false positive regions are indicated by white-dotted arrows.

To find the best regularization parameter (λ_{f}) in the regularized correlation (Eq. (96)), we obtain functional images from the images reconstructed through the three methods for $4N_{\mathrm{loc}} = 40, 20, 12$ using $\lambda_{\mathrm{f}} = 0.02, 0.08, 0.32, 1.28, 5.12$. For UBP and the hybrid method, $\lambda_{\mathrm{f}} = 0.32$ is the best choice to maintain the true functional region and suppress false positive regions, and the functional images for $4N_{\mathrm{loc}} = 12$ and all choices of λ_{f} are shown in **Figure 22**. For the regularized iterative method, we compare the images for $4N_{\mathrm{loc}} = 40, 20, 12$ in **Figure 23** and observe that $\lambda_{\mathrm{f}} = 0.08$ is the best choice. We summarize the obtained functional images with best values of λ_{f} in **Figure 21c**. Results from UBP and the hybrid method match well for $4N_{\mathrm{loc}} = 40$. The hybrid method is slightly (significantly) better than UBP for $4N_{\mathrm{loc}} = 20$ ($4N_{\mathrm{loc}} = 12$). Due to the violation of linearity, the regularized iterative method compromises the true functional region: leading to its shrinkage for $4N_{\mathrm{loc}} = 40, 20$ and its decimation altogether for $4N_{\mathrm{loc}} = 12$. In summary, the proposed hybrid method enables fast functional imaging with highly sparse sampling.

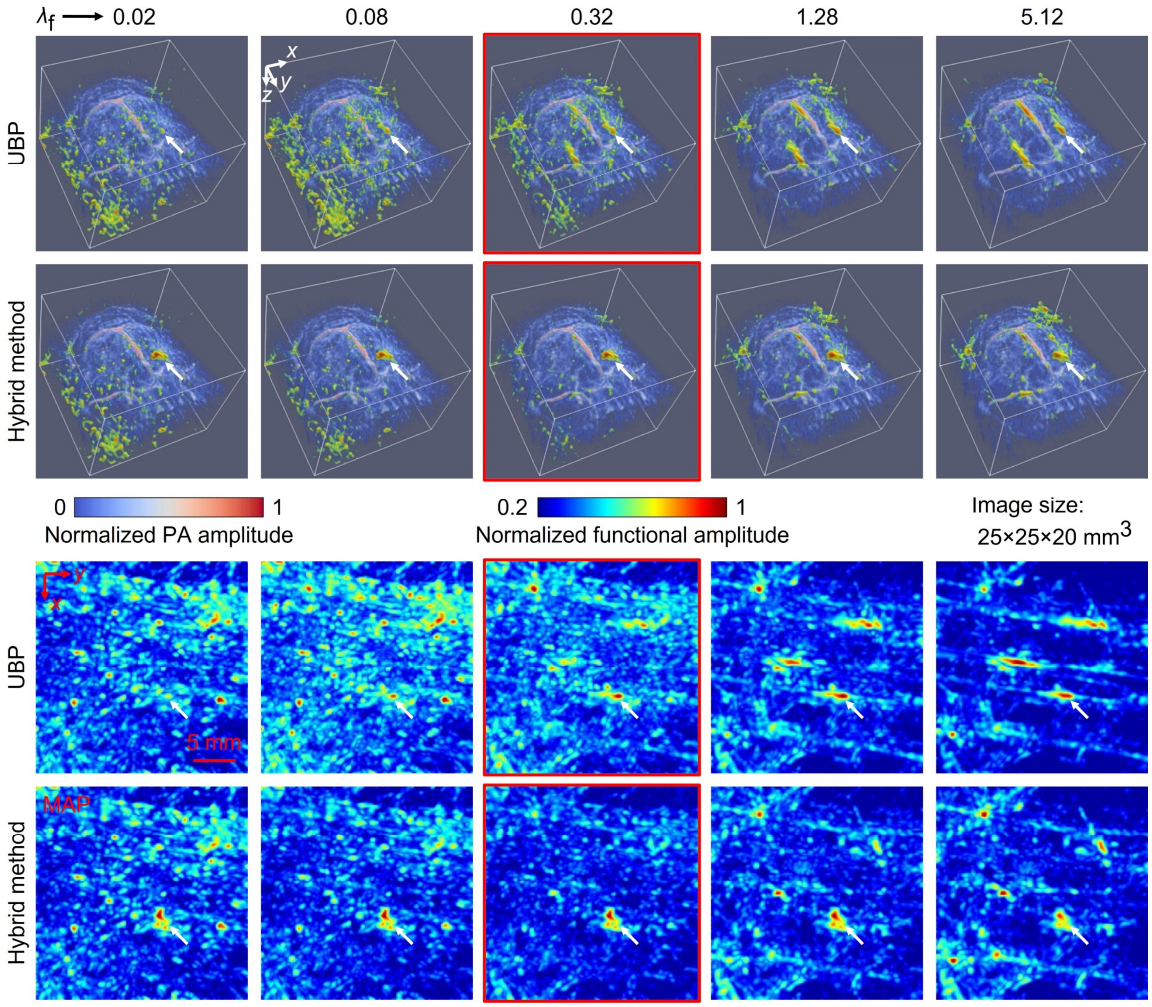

Figure 22 Mouse brain functional images *in vivo* extracted using the regularized-correlation-based method with $\lambda_f = 0.02, 0.08, 0.32, 1.28, 5.12$ from images reconstructed with UBP and the hybrid method ($4N_{loc} = 12$). The first two rows show both the 3D functional and background images and the last two rows show the MAPs of the functional images along the z-axis. In all images, the true functional regions are indicated by white-solid arrows, and the other high-amplitude functional regions are false positives. Images for $\lambda_f = 0.32$ are highlighted by red-solid boundaries.

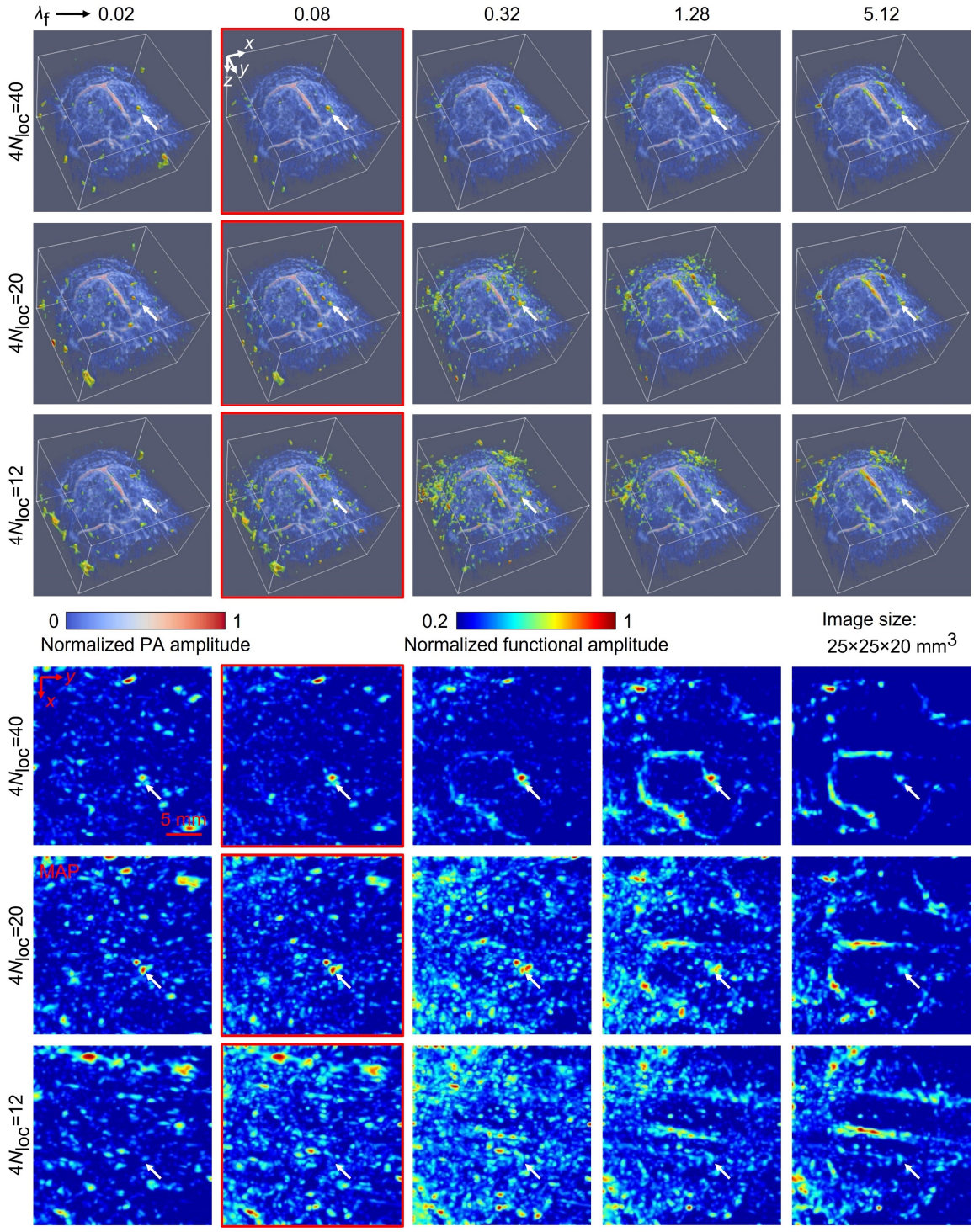

Figure 23 Mouse brain functional images *in vivo* extracted using the regularized-correlation-based method with $\lambda_f = 0.02, 0.08, 0.32, 1.28, 5.12$ from images reconstructed through the regularized iterative method ($4N_{loc} = 40, 20, 12$). The first three rows show both the 3D functional and background images and the last three rows show the MAPs of the functional images along the z-axis. In all images, the true functional regions are indicated by white-solid arrows, and the other

high-amplitude functional regions are false positives. Images for $\lambda_f = 0.08$ are highlighted by red-solid boundaries.

4.6 Discussion

In both numerical simulations and mouse brain functional imaging *in vivo*, the hybrid method substantially mitigates artifacts in the reconstructed images and reduces false positive regions in the functional image, and its linearity is important for maintaining the true functional region. Due to its high robustness, the method can accelerate or enhance the performance of an existing system and reduce the cost of a future system for functional imaging.

Although we demonstrate the proposed methods only using PACT, they are applicable to CT and MRI. System matrices in 3D PACT and CT correspond to sphere [24] and line [72] integrals, respectively, and the latter can be transformed using Grangeat's method [73] into plane integrals, which are locally equivalent to sphere integrals. System matrices in 2D PACT and CT [74] correspond to circle (reduced from a sphere) and line integrals, respectively, which are locally equivalent. MRI is more complex due to its high flexibility in k-space sampling. For radial sampling MRI [64], [75], the acquired signals can be transformed to integrals on lines and planes, respectively, for 2D and 3D imaging using the Fourier slice theorem. For other sampling patterns, further analysis may disclose proper transformations to obtain integrals that are locally equivalent to those in PACT. It needs to be noted that the proposed method relies on efficient slicing and manipulation of the system matrix, which is directly available for CT where the system matrix is highly sparse and is achieved in PACT through system matrix compression based on SVD and FFT as shown in Chapter 3. For radial sampling MRI, we can transform the k-space signals to line or plane integrals and compress them through SVD and FFT. For general MRI, further study may transform the k-space signals to compressible integrals. In summary, through certain transformations, system matrices in CT, MRI, and PACT share similar local structures, which allows the knowledge in this study to be transferred to CT and MRI.

C h a p t e r 5

INTRA-IMAGE NONRIGID MOTION CORRECTION

P. Hu, X. Tong, L. Lin, and L. V. Wang, "Data-driven system matrix manipulation enabling fast functional imaging and intra-image nonrigid motion correction in tomography," *Submitted*.

5.1 Abstract

Tomographic imaging modalities are described by large system matrices. Tissue motion degrades the system matrix and image quality. Various existing techniques improve the image quality without correcting the system matrices. Here, we compress the system matrices to improve computational efficiency (e.g., 42 times) using singular value decomposition and fast Fourier transform. Enabled by the efficiency, we propose intra-image nonrigid motion correction by incorporating the motion as subdomain translations into the system matrix and reconstructing the translations together with the image iteratively. We demonstrate the method in 3D photoacoustic computed tomography (PACT) with significantly improved image qualities and clarify their applicability to X-ray computed tomography (X-ray CT) and magnetic resonance imaging (MRI) due to the similarities in system matrices.

5.2 Background

In tomographic imaging (e.g., X-ray CT, MRI, and PACT), tissue motions such as heart beating [76]–[78], breathing [79]–[81], abdominal movement [82], [83], and fetal movement [84], cause complex geometric errors in each system matrix, which introduce artifacts in the reconstructed image and compromise valuable image features.

For intra-image nonrigid motion correction, gating- and binning-based methods [61], [63]–[65], [78] are commonly used. However, they require repeated data acquisition, which is time-consuming and infeasible for unrepeated motions. DNNs have also been used for

motion correction [55], [56], however, they need specific training datasets that are not universally available, and it is challenging to reject falsely generated features in DNNs. Two system-matrix-level methods [85], [86] have been proposed for motion correction. In the first method [85], [87], the authors approximate general nonrigid motions with localized linear translations, identify possible motion paths from multichannel navigator data, and estimate the motion at each pixel using localized gradient-entropy metric in the image domain. However, quantifying localized motion from only navigator data is not robust, especially when the motion amplitude and noise level increase. In the second method [86], the authors express breathing- and heartbeat-induced motions in basis functions by performing singular value decomposition (SVD) and resolve these motions in imaging. However, for general motions, especially unrepeated motions, the method's performance is unknown. The high computation cost of SVD in the method also restricts its application to 3D imaging.

In this study, using the four-arc PACT system as an example, we compress the system matrices using SVD and fast Fourier transform (FFT), which enables efficient system matrix slicing and manipulation. For intra-image nonrigid motion correction, we approximate the motion with localized linear translations. Starting from an initially reconstructed motion-blurred image, we first estimate the translation of each subdomain of the object by minimizing the difference between the simulated signals from the subdomain and the detected signals. With the estimated translations, we update the system matrix and reconstruct the image again. We iterate the correction-reconstruction process to obtain the final image. This method does not require repeated data acquisition and is effective for unrepeated motions.

5.3 Theory

Implementing the system matrix **H** using Eq. (86) allows for efficient slicing of **H** with respect to image voxel indices and transducer element indices. We propose a method for intra-image nonrigid motion correction through data-driven slicing and manipulation of **H**. In this research, we move a transducer array with N_{ele} elements across N_{loc} locations to form a virtual 2D transducer array for 3D imaging (with the number of transducer elements $N =$

$N_{\mathrm{loc}}N_{\mathrm{ele}}$). The motion of a source point in the image domain corresponds to temporal shifts of the signals from the source point in the signal domain. To encode the motion information in the forward operator, we modify Eq. (74) to

$$\hat{p}_{\zeta}\left(\mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}},t\right)\approx\sum_{k=1}^{K}\eta_k(t)*_t$$
$$\int_V p_0(\mathbf{r}')\hat{h}_{\mathrm{lcl},k}\left(\left(\mathbf{r}'-\mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}}\right)\mathbf{A}_{\mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}}}\right)\delta\left(t-\zeta\left(\mathbf{r}',\mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}}\right)-\frac{\left\|\mathbf{r}'-\mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}}\right\|}{c}\right)\mathrm{d}\mathbf{r}',$$
$$n_{\mathrm{loc}}=1,2,\ldots,N_{\mathrm{loc}},n_{\mathrm{ele}}=1,2,\ldots,N_{\mathrm{ele}},t\geq 0. \tag{99}$$

Here, we define $\mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}}=\mathbf{r}_{n_{\mathrm{loc}}+(n_{\mathrm{ele}}-1)N_{\mathrm{loc}}}$ and let $\zeta\left(\mathbf{r}',\mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}}\right)$ be the motion-induced temporal shift of the signals from the source point at $\mathbf{r}'$ that are detected by the n_{ele}-th element when the transducer array is at the n_{loc}-th location. It needs to be noted that, due to motion, a source point does not stay at $\mathbf{r}'$, and we use $\mathbf{r}'$ to represent the source point's average location.

Next, we mathematically connect the temporal shifts $\zeta\left(\mathbf{r}',\mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}}\right)$ with the motions. For efficiency and robustness, instead of modeling the motions of all source points in the image domain D, we model the motions of subdomains $D_1,D_2,\ldots,D_{M_{\mathrm{sub}}}$ occupying the image domain:

$$D=D_1\cup D_2\cup\ldots\cup D_{M_{\mathrm{sub}}}. \tag{100}$$

For simplicity, we let all the subdomains be rectangular cuboids of the same size. We denote the center of the m_{sub}-th subdomain $D_{m_{\mathrm{sub}}}$ as $\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}$ and use the motions of $\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}$ to represent the motions of $D_{m_{\mathrm{sub}}},m_{\mathrm{sub}}=1,2,\ldots,M_{\mathrm{sub}}$. In this research, the motions of the tissues are generally small, and we observe small deformations and rotations of each subdomain. Because of this observation, we ignore the deformations and rotations of each subdomain and only discretize the translations with step sizes a_x, a_y, and a_z along the x-axis, y-axis, and z-axis, respectively. We represent the motion-induced translation steps of the domain center from $\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}$ as $k_x\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}},n_{\mathrm{loc}}\right)$, $k_y\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}},n_{\mathrm{loc}}\right)$, and $k_z\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}},n_{\mathrm{loc}}\right)$ along the x-axis, y-axis, and z-axis, respectively, when the transducer array is at the n_{loc}-th location. In summary, we express the temporal shifts caused by these translations as

$$\zeta_{\mathbf{k}}\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, \mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}}\right) = \frac{\left\|\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}} + \mathbf{k}\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right) \odot \mathbf{a} - \mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}}\right\|}{c} - \frac{\left\|\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}} - \mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}}\right\|}{c},$$
$$m_{\mathrm{sub}} = 1,2,\dots,M_{\mathrm{sub}}, n_{\mathrm{loc}} = 1,2,\dots,N_{\mathrm{loc}}, n_{\mathrm{ele}} = 1,2,\dots,N_{\mathrm{ele}}. \quad (101)$$

Here, we define a constant vector $\mathbf{a} = (a_x, a_y, a_z)$ and variable vectors $\mathbf{k}\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right) = \left(k_x\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right), k_y\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right), k_z\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right)\right) \in \mathbb{z}^{\mathbf{3}}$, and we group the variable vectors into a set $\mathbf{k} = \left\{\mathbf{k}\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right) \in \mathbb{z}^{\mathbf{3}} | m_{\mathrm{sub}} = 1,2,\dots,M_{\mathrm{sub}}, n_{\mathrm{loc}} = 1,2,\dots,N_{\mathrm{loc}}\right\}$. Then we obtain the values of $\zeta\left(\mathbf{r}', \mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}}\right)$ from the values of $\zeta_{\mathbf{k}}\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, \mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}}\right)$ through spatial interpolation in the image domain, which has high accuracy due to the spatial smoothness of the motions of tissues. Thus, we mathematically connect the temporal shifts $\zeta\left(\mathbf{r}', \mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}}\right)$ with the motions, described by $\mathbf{k}$.

Furthermore, we discretize the motion-incorporated forward operator in Eq. (99) and propose a motion correction method through Gauss-Seidel-type iterations. We replace $\zeta\left(\mathbf{r}', \mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}}\right)$ with $\zeta_{\mathbf{k}}\left(\mathbf{r}', \mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}}\right)$ in Eq. (99) to emphasize these temporal shifts' dependency on $\mathbf{k}$. Then we discretize the forward operator to

$$\hat{p}_{\mathbf{k},n,l} \approx \sum_{k=1}^{K} \eta_{k,l} *_l \sum_{m=1}^{M} v_m \hat{h}_{\mathrm{lcl},k,n,m} p_{0,m} \frac{1}{\tau}\left[\begin{array}{l}\left(t_{l_{n,m}+1} - \zeta_{\mathbf{k}}(\mathbf{r}'_m, \mathbf{r}_n) - \frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c}\right)\delta_{l,l_{n,m}} \\ + \left(\zeta_{\mathbf{k}}(\mathbf{r}'_m, \mathbf{r}_n) + \frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c} - t_{l_{n,m}}\right)\delta_{l,l_{n,m}+1}\end{array}\right],$$
$$n = 1,2,\dots,N, l = 1,2,\dots,L, \quad (102)$$

where, $l_{n,m}$ denotes the temporal index such that $t_{l_{n,m}} \le \zeta_{\mathbf{k}}(\mathbf{r}'_m, \mathbf{r}_n) + \frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c} < t_{l_{n,m}+1}$. Accordingly, we express the forward operator in matrix form as

$$\hat{\mathbf{p}} = \mathbf{H}_{\mathbf{k}}\mathbf{p}_0, \quad (103)$$

through which we express the data-driven motion correction as a dual-objective optimization problem

$$\left(\hat{\mathbf{p}}_0, \hat{\mathbf{k}}\right) = \underset{\mathbf{p}_0 \in \mathbb{R}^M, \mathbf{p}_0 \ge \mathbf{0}, \mathbf{k} \in (\mathbb{Z}^3)^{M_{\mathrm{sub}} N_{\mathrm{loc}}}}{\operatorname{argmin}} \gamma_{\mathrm{c}} \|\mathbf{H}_{\mathbf{k}}\mathbf{p}_0 - \hat{\mathbf{p}}\|^2 + \lambda |\mathbf{p}_0|_{\mathrm{TV}}. \quad (104)$$

Here, $|\mathbf{p}_0|_{\mathrm{TV}}$ denotes $\mathbf{p}_0$'s total variation (TV) norm, defined in Eq. (91). It is challenging to solve this problem directly. Therefore, we simplify it into a convex optimization problem and a combinatorial optimization problem. The convex optimization problem, expressed as

$$\widehat{\mathbf{p}}_0 = \underset{\mathbf{p}_0 \in \mathbb{R}^M, \mathbf{p}_0 \geq \mathbf{0}}{\operatorname{argmin}} \gamma_{\mathrm{c}} \|\mathbf{H}_{\mathbf{k}} \mathbf{p}_0 - \widehat{\mathbf{p}}\|^2 + \lambda |\mathbf{p}_0|_{\mathrm{TV}}, \tag{105}$$

can be solved by the FISTA. The combinatorial optimization problem, expressed as

$$\hat{\mathbf{k}} = \underset{\mathbf{k} \in (\mathbb{Z}^3)^{M_{\mathrm{sub}} N_{\mathrm{loc}}}}{\operatorname{argmin}} \|\mathbf{H}_{\mathbf{k}} \mathbf{p}_0 - \widehat{\mathbf{p}}\|^2, \tag{106}$$

is still challenging.

Instead of solving the problem in Eq. (106) directly, we solve each vector $\mathbf{k}\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right)$ independently without changing other vectors:

$$\hat{\mathbf{k}}\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right) = \underset{\mathbf{k}\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right) \in \mathbb{Z}^3}{\operatorname{argmin}} \|\mathbf{H}_{\mathbf{k}} \mathbf{p}_0 - \widehat{\mathbf{p}}\|^2,$$
$$m_{\mathrm{sub}} = 1,2,\ldots,M_{\mathrm{sub}}, n_{\mathrm{loc}} = 1,2,\ldots,N_{\mathrm{loc}}. \tag{107}$$

To further simplify Eq. (107), for each subdomain D_i, we decompose $\mathbf{p}_0$ into two images:

$$\mathbf{p}_0 = \mathbf{p}_{0,D_{m_{\mathrm{sub}}}} + \mathbf{p}_{0,D\setminus D_{m_{\mathrm{sub}}}}, m_{\mathrm{sub}} = 1,2,\ldots,M_{\mathrm{sub}}. \tag{108}$$

Here, $\mathbf{p}_{0,D_{m_{\mathrm{sub}}}}$ is zero outside the subdomain $D_{m_{\mathrm{sub}}}$ and the same as $\mathbf{p}_0$ inside $D_{m_{\mathrm{sub}}}$, whereas $\mathbf{p}_{0,D\setminus D_{m_{\mathrm{sub}}}}$ is zero inside the subdomain $D_{m_{\mathrm{sub}}}$ and the same as $\mathbf{p}_0$ outside $D_{m_{\mathrm{sub}}}$. Using this decomposition, we rewrite Eq. (107) as

$$\hat{\mathbf{k}}\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right) = \underset{\mathbf{k}\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right) \in \mathbb{Z}^3}{\operatorname{argmin}} \left\| \mathbf{H}_{\mathbf{k}} \left(\mathbf{p}_{0,D_{m_{\mathrm{sub}}}} + \mathbf{p}_{0,D\setminus D_{m_{\mathrm{sub}}}} \right) - \widehat{\mathbf{p}} \right\|^2$$
$$= \underset{\mathbf{k}\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right) \in \mathbb{Z}^3}{\operatorname{argmin}} \left\{ \begin{array}{c} \left\| \mathbf{H}_{\mathbf{k}} \mathbf{p}_{0,D_{m_{\mathrm{sub}}}} - \widehat{\mathbf{p}} \right\|^2 + 2\mathbf{p}_{0,D_{m_{\mathrm{sub}}}}^{\mathrm{T}} \mathbf{H}_{\mathbf{k}}^{\mathrm{T}} \mathbf{H}_{\mathbf{k}} \mathbf{p}_{0,D\setminus D_{m_{\mathrm{sub}}}} + \\ \left\| \mathbf{H}_{\mathbf{k}} \mathbf{p}_{0,D\setminus D_{m_{\mathrm{sub}}}} \right\|^2 - 2\widehat{\mathbf{p}}^{\mathrm{T}} \mathbf{H}_{\mathbf{k}} \mathbf{p}_{0,D\setminus D_{m_{\mathrm{sub}}}} \end{array} \right\},$$
$$m_{\mathrm{sub}} = 1,2,\ldots,M_{\mathrm{sub}}, n_{\mathrm{loc}} = 1,2,\ldots,N_{\mathrm{loc}}. \tag{109}$$

Given the initial pressure $\mathbf{p}_{0,D\setminus D_{m_{\mathrm{sub}}}}$, we apply forward simulation ($\mathbf{H}_{\mathbf{k}}$) and adjoint reconstruction ($\mathbf{H}_{\mathbf{k}}^{\mathrm{T}}$) to it to obtain $\mathbf{H}_{\mathbf{k}}^{\mathrm{T}} \mathbf{H}_{\mathbf{k}} \mathbf{p}_{0,D\setminus D_{m_{\mathrm{sub}}}}$, which has small amplitudes inside $D_{m_{\mathrm{sub}}}$ for the detection geometry used in this research. Thus, we have $2\mathbf{p}_{0,D_{m_{\mathrm{sub}}}}^{\mathrm{T}} \mathbf{H}_{\mathbf{k}}^{\mathrm{T}} \mathbf{H}_{\mathbf{k}} \mathbf{p}_{0,D\setminus D_{m_{\mathrm{sub}}}} \approx 0$. Additionally, although the values of $\mathbf{k}\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right)$ (the

motion inside $D_{m_{\mathrm{sub}}}$) affect $\mathbf{H_k p}_{0,D\setminus D_{m_{\mathrm{sub}}}}$ (the signals from the sources outside $D_{m_{\mathrm{sub}}}$) due to spatial interpolation, these effects are minor compared with the values' effects on $\mathbf{H_k p}_{0,D_{m_{\mathrm{sub}}}}$. Thus, we ignore $\left\|\mathbf{H_k p}_{0,D\setminus D_{m_{\mathrm{sub}}}}\right\|^2$ and $2\hat{\mathbf{p}}^{\mathrm{T}}\mathbf{H_k p}_{0,D\setminus D_{m_{\mathrm{sub}}}}$ in the optimization for $\mathbf{k}\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right)$. Through these approximations, we simplify Eq. (109) to

$$\hat{\mathbf{k}}\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right) \approx \underset{\mathbf{k}\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right)\in\mathbb{Z}^3}{\operatorname{argmin}} \left\|\mathbf{H_k p}_{0,D_{m_{\mathrm{sub}}}} - \hat{\mathbf{p}}\right\|^2 = \underset{\mathbf{k}\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right)\in\mathbb{Z}^3}{\operatorname{argmin}} \left\|\mathbf{H}_{\mathbf{k},n_{\mathrm{loc}}}\mathbf{p}_{0,D_{m_{\mathrm{sub}}}} - \hat{\mathbf{p}}_{n_{\mathrm{loc}}}\right\|^2,$$
$$m_{\mathrm{sub}} = 1,2,\dots,M_{\mathrm{sub}}, n_{\mathrm{loc}} = 1,2,\dots,N_{\mathrm{loc}}. \quad (110)$$

Here, $\hat{\mathbf{p}}_{n_{\mathrm{loc}}}$ (of shape $N_{\mathrm{ele}}L \times 1$) denotes the signals detected by the transducer array when it is at the n_{loc}-th location whereas signals for other locations are not affected by $\mathbf{k}\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right)$, and $\mathbf{H}_{\mathbf{k},n_{\mathrm{loc}}}$ (of shape $N_{\mathrm{ele}}L \times M$) denotes the corresponding system matrix slices. Additionally, our forward operator allows for forward simulation of any subdomain $D_{m_{\mathrm{sub}}}$ (with M_{SD} voxels), which is much smaller than the whole image domain D (with M voxels). Hence, obtaining $\hat{\mathbf{k}}\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right)$ using Eq. (110) is much more efficient than obtaining it through Eq. (107). We need to solve the optimization problems in Eqs. (105) and (110) in multiple iterations. Thus, instead of obtaining an accurate choice of $\mathbf{k}\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right)$ in each iteration, we further decompose the optimization problem in Eq. (110) into problems

$$\hat{k}_x\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right) = \underset{k_x\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right)\in\{-K_x,-K_x+1,\dots,K_x\}}{\operatorname{argmin}} \left\|\mathbf{H}_{\mathbf{k},n_{\mathrm{loc}}}\mathbf{p}_{0,D_{m_{\mathrm{sub}}}} - \hat{\mathbf{p}}_{n_{\mathrm{loc}}}\right\|^2, \quad (111)$$

$$\hat{k}_y\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right) = \underset{k_y\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right)\in\{-K_y,-K_y+1,\dots,K_y\}}{\operatorname{argmin}} \left\|\mathbf{H}_{\mathbf{k},n_{\mathrm{loc}}}\mathbf{p}_{0,D_{m_{\mathrm{sub}}}} - \hat{\mathbf{p}}_{n_{\mathrm{loc}}}\right\|^2, \quad (112)$$

and

$$\hat{k}_z\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right) = \underset{k_z\left(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}\right)\in\{-K_z,-K_z+1,\dots,K_z\}}{\operatorname{argmin}} \left\|\mathbf{H}_{\mathbf{k},n_{\mathrm{loc}}}\mathbf{p}_{0,D_{m_{\mathrm{sub}}}} - \hat{\mathbf{p}}_{n_{\mathrm{loc}}}\right\|^2, \quad (113)$$
$$m_{\mathrm{sub}} = 1,2,\dots,M_{\mathrm{sub}}, n_{\mathrm{loc}} = 1,2,\dots,N_{\mathrm{loc}},$$

with limited searching ranges determined by parameters K_x, K_y, and K_z, respectively. For each subdomain and each transducer array location, we solve the problems in Eqs. (111)–(113) by performing the forward simulation for the subdomain and the transducer array location for $2K_x + 1$, $2K_y$, and $2K_y$ times, respectively. In summary, the computational complexity of updating all values in $\mathbf{k}$ once is

$$N_{\text{loc}}M_{\text{sub}}\left[2\left(K_x + K_y + K_z\right) + 1\right]\left[O\left(N_{\text{ele}}M_{\text{SD}}K\right) + O\left(N_{\text{ele}}\left(L\log_2 L\right)K\right)\right]. \quad (114)$$

We summarize the workflow of the proposed motion correction method in **Figure 24** and the additional symbols in **Table 4**.

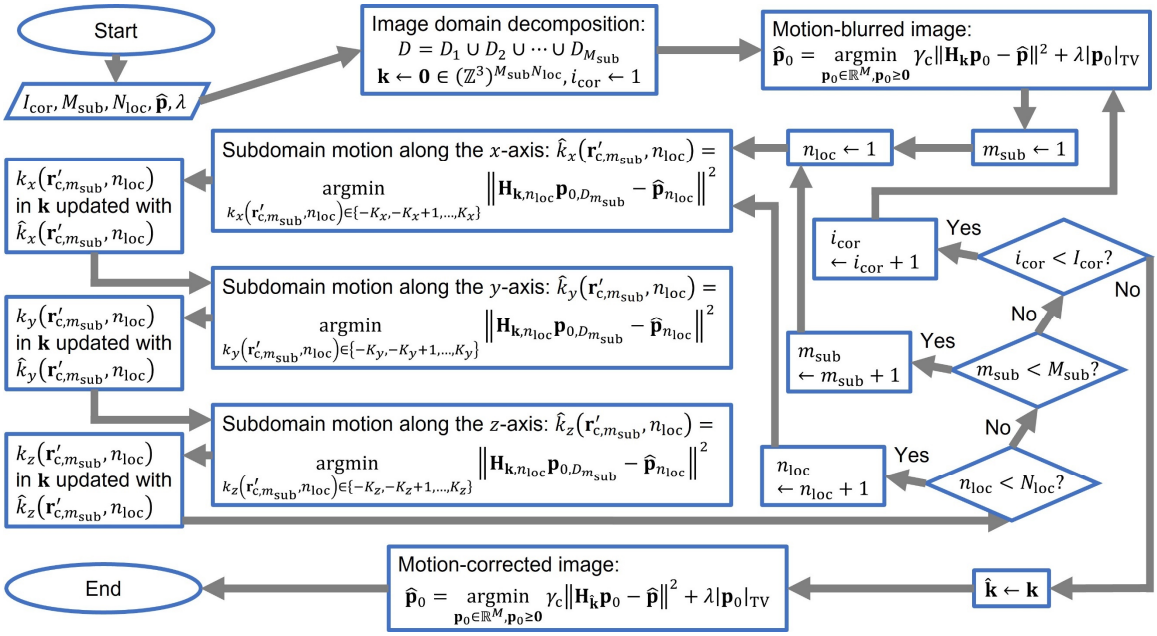

Figure 24 Workflow of the intra-image nonrigid motion correction based on data-driven manipulation of the system matrix.

5.4 Numerical simulations

To simulate translation- and deformation-induced motions, we let the values at $\left(x, y, z + z_{\text{tra},n_{\text{loc}}}\right)$ and $\left(\sqrt{\alpha_{\text{def},n_{\text{loc}}}}\left(x - \frac{x_0+x_1}{2}\right), \sqrt{\alpha_{\text{def},n_{\text{loc}}}}\left(y - \frac{y_0+y_1}{2}\right), \frac{1}{\alpha_{\text{def},n_{\text{loc}}}}(z - z_0)\right) + \left(\frac{x_0+x_1}{2}, \frac{y_0+y_1}{2}, z_0\right)$, respectively, in the n_{loc}-th numerical phantom be the value at

$(x, y, z) \in [x_0, x_1] \times [y_0, y_1] \times [z_0, z_1]$ in a predefined numerical phantom, with the translation distances defined as

$$z_{\mathrm{tra},n_{\mathrm{loc}}} = A_{\mathrm{tra}} \sin \frac{2\pi(n_{\mathrm{loc}} - 1)N_{\mathrm{per}}}{N_{\mathrm{loc}}}, n_{\mathrm{loc}} = 1, \dots, N_{\mathrm{loc}} \tag{115}$$

and deformation ratios defined as

$$\alpha_{\mathrm{def},n_{\mathrm{loc}}} = 1 + A_{\mathrm{def}} \sin \frac{2\pi(n_{\mathrm{loc}} - 1)N_{\mathrm{per}}}{N_{\mathrm{loc}}}, n_{\mathrm{loc}} = 1, \dots, N_{\mathrm{loc}}. \tag{116}$$

Here, A_{tra} and A_{def} are amplitudes of the translation distance and deformation ratio, respectively. In the forward simulations, signals from the n_{loc}-th numerical phantom are only detected by the transducer array at the n_{loc}-th location. We let $x_1 - x_0 = 44.8$ mm, $y_1 - y_0 = 44.8$ mm, $z_1 - z_0 = 14.4$ mm, $N_{\mathrm{loc}} = 72$, $N_{\mathrm{ele}} = 4 \times 72$, and $N_{\mathrm{per}} = 0.5, 1, 2, 3$ for all motion simulations, and use $A_{\mathrm{tra}} = 0.2, 0.4, 0.6, 0.8, 1.0, 1.2$ (mm) and $A_{\mathrm{def}} = 0.02, 0.04, 0.06, 0.08, 0.10, 0.12$, respectively, for translation and deformation simulations.

From a numerical phantom (**Figure 25a**) obtained in 3D imaging of a human breast, we simulate numerical phantoms with translation- and deformation-induced intra-image motions, whose patterns are defined by Eqs. (115) and (116), respectively. We depict the array rotation with tissue translation and deformation in **Figure 25b** and **c**, respectively. A 90° rotation of the four-arc array ($N_{\mathrm{loc}} = 72$, $N_{\mathrm{ele}} = 4 \times 72$) allows the transducer elements to detect signals needed to form a 3D image, during which tissue motions occur. The amplitudes of translation and deformation are controlled by A_{tra} and A_{def}, respectively, and the number of periods of the motion during a 90° rotation of the array is denoted as N_{per}.

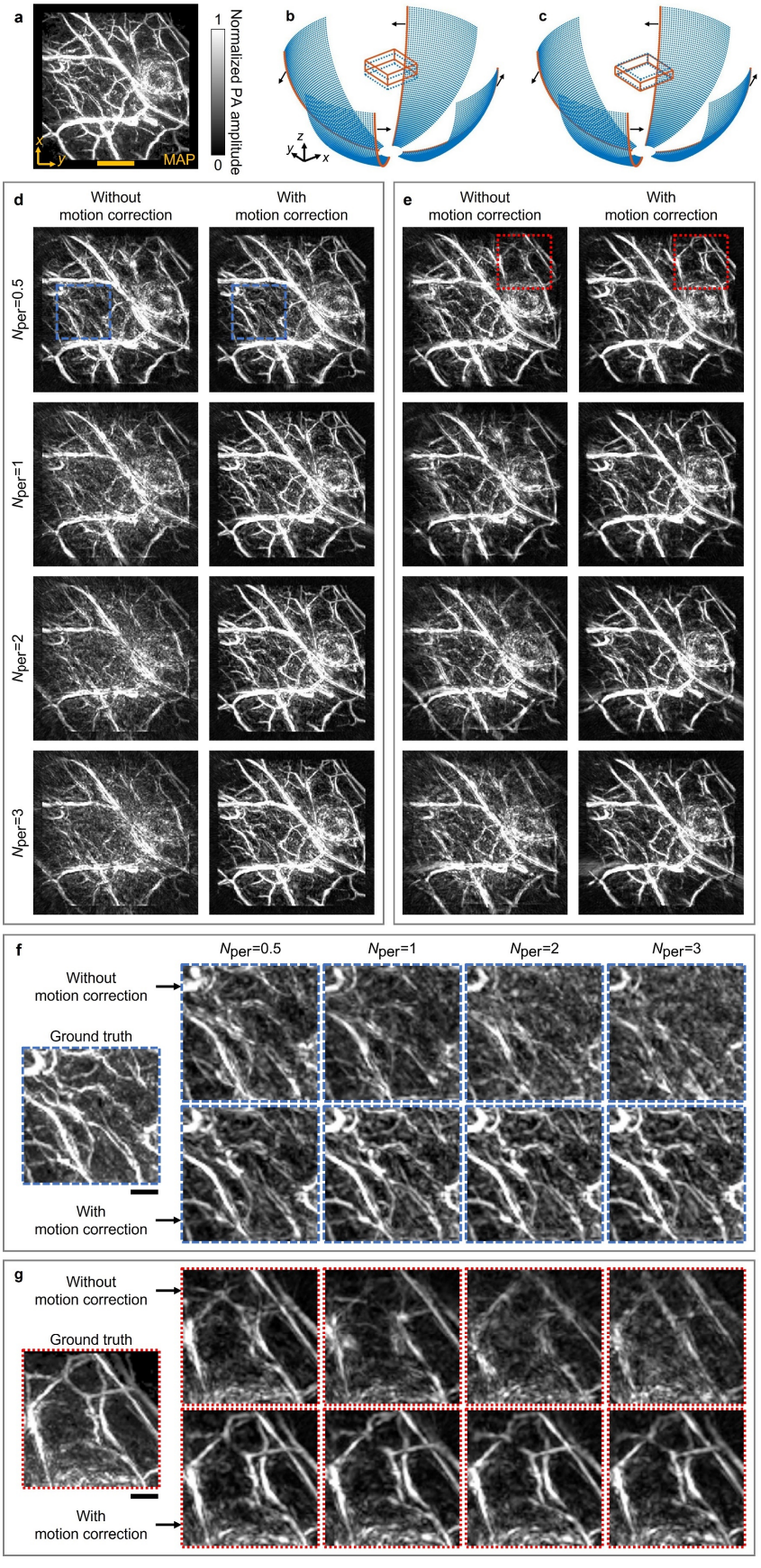
a
MAP
1
0
Normalized PA amplitude
b
c
d
Without motion correction
With motion correction
N_{per}=0.5
N_{per}=1
N_{per}=2
N_{per}=3
e
Without motion correction
With motion correction
f
N_{per}=0.5
N_{per}=1
N_{per}=2
N_{per}=3
Without motion correction
Ground truth
With motion correction
g
Without motion correction
Ground truth
With motion correction

Figure 25 Intra-image nonrigid correction of motions induced by tissue translation and deformation. **a** MAP of an image (a human breast image subset with a volume of $44.8\times44.8\times14.4$ mm^3) along the z-axis. PA, photoacoustic. Scale bar, 1 cm. **b-c** Depictions of the intra-image motions induced by tissue translation and deformation, respectively. The image translates (deforms) from a rectangular cuboid with blue-dotted edges to one with red-solid edges as the four-arc array rotates, with the passed virtual element locations marked as blue dots, the latest element locations marked as red curves, and the rotation direction indicated by black arrows. **d-e** MAPs of the images reconstructed from the signals with motions induced by tissue translation ($A_{\mathrm{tra}} = 0.6$ mm) and deformation ($A_{\mathrm{def}} = 0.06$), respectively, without and with motion correction for $N_{\mathrm{per}} = 0.5$,1,2,3. **f-g** Closed-up subsets of the MAPs enclosed in blue-dashed boxes (in **d** and **f)** and red-doted boxes (in **e** and **g**), respectively, with the ground truths.

We reconstruct images from the simulated signals without and with motion correction and compare their maximum amplitude projections (MAPs) along the z-axis (translation-induced motions with $N_{\mathrm{per}} = 0.5, 1, 2, 3$ and $A_{\mathrm{tra}} = 0.2, 0.4, 0.6, 0.8, 1.0, 1.2$ (mm) ; deformation-induced motions with $N_{\mathrm{per}} = 0.5, 1, 2, 3$ and $A_{\mathrm{def}} = 0.02, 0.04, 0.06, 0.08, 0.10, 0.12$). The proposed motion correction method improves image quality for every set of signals. We show those MAPs for $N_{\mathrm{per}} = 0.5, 1, 2, 3$ in **Figure 25d** (translation, $A_{\mathrm{tra}} = 0.6$ mm) and **Figure 25e** (deformation, $A_{\mathrm{def}} = 0.06$). Moreover, we pick subsets of the MAPs in blue-dashed box and red-dotted box in **Figure 25d** and **e**, respectively, and compare closeups of them in **Figure 25f** and **g** with ground-truth MAPs. Motion correction not only reduces the motion-induced artifacts but also reveals more true blood vessels, which match with those in the ground-truth images. It needs to be noted that although image-domain DNNs for motion correction are powerful in mitigating artifacts, they are less capable of revealing features not observable in the original images. In practice, for a certain region in the tissue, if most of its motions have amplitudes greater than the image resolution, the detected signals from this region will have severe mismatches, which causes feature loss in the reconstructed images. The fact that the proposed hybrid method reveals these features proves its high tolerance to tissue motion and highlights the importance of system matrix manipulation in motion correction.

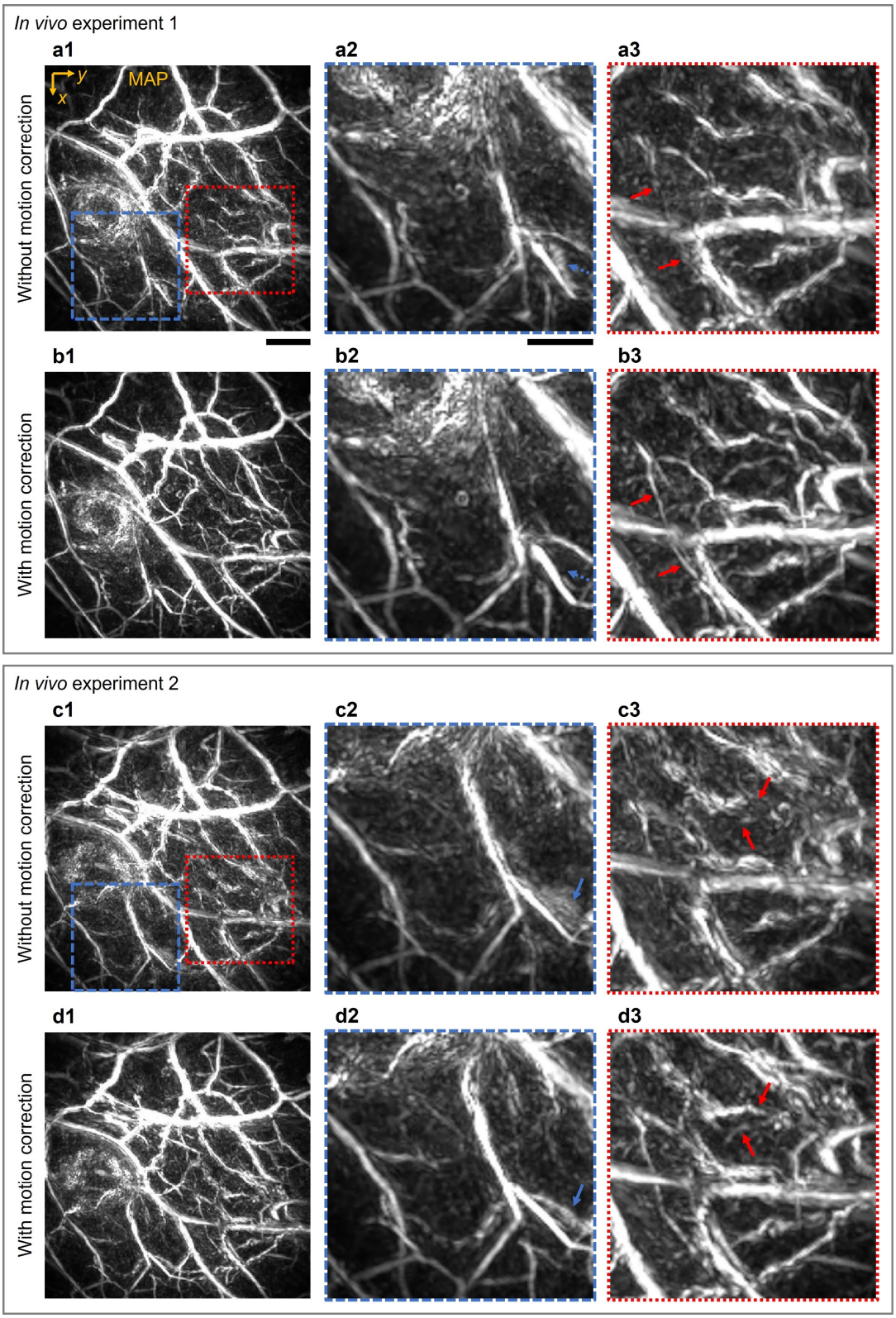
In vivo experiment 1
a1
a2
a3
MAP
y
x
Without motion correction
b1
b2
b3
With motion correction
In vivo experiment 2
c1
c2
c3
Without motion correction
d1
d2
d3
With motion correction
1
0
Normalized PA amplitude

Figure 26 Intra-image nonrigid motion correction for human breast imaging *in vivo*. **a1–a3** MAP of a human breast image with a volume of $6 \times 6 \times 3$ cm^3 reconstructed without motion correction (**a1**), and two closed-up subsets marked by blue-dashed (**a2**) and red-dotted (**a3**) boxes, respectively. **b1–b3** MAP of the image reconstructed with motion correction from the same signals as for **a1**, and the closed-up subsets. Examples of suppressed artifacts and enhanced features are highlighted in **a2**-**b2** and **a3**-**b3**, respectively, by blue-dotted and red-solid arrows. **c1**–**c3** and **d1**–**d3** The second experiment of the same human breast with the images reconstructed without (**c1–c3**) and with (**d1–d3**) motion correction, respectively. Examples of enhanced features are indicated by blue-solid and red-solid arrows in **a2**-**b2** and **a3**-**b3**, respectively. For the first-column images, scale bar, 1 cm; for the second and third columns, scale bar, 6 mm.

5.5 *In vivo* experiments

Furthermore, we use the proposed method to correct heartbeat-induced motions in human breast imaging ($N_{\text{loc}} = 99$, $N_{\text{ele}} = 4 \times 256$). We acquired signals from the left breast of a volunteer with breath holding (10 s) in 4 experiments. Samples of the signals in the first two experiments are shown in **Figure 27a** and **b**, respectively, from which we observe the signals' temporal shifts caused by heartbeat-induced motions. We reconstruct the human breast images in the first two experiments and compare MAPs of them in **Figure 26**. For the first experiment, the MAPs of images reconstructed without and with motion correction are shown in **Figure 26a1** and **a2**, respectively, with their closed-up subsets (blue-dashed and red-dotted boxes) shown in **Figure 26a2**-**a3** and **b2**-**b3**. We observe that motion correction mitigates motion-induced artifacts (e.g., the region indicated by blue-dotted arrows in **Figure 26a2**-**b2**) and reveals motion-compromised features (e.g., the blood vessel indicated by red-solid arrows in **Figure 26a3**-**b3**). MAPs of the images in the second experiment without and with motion correction are shown in **Figure 26c1**–**c3** and **d1**–**d3**, respectively. Still, motion correction recovers motion-compromised vessels, such as those indicated by blue-solid and red-solid arrows in **Figure 26c2**-**d2** and **c3**-**d3**, respectively. Importantly, the compromised features in **Figure 26a2**-**a3** are very different from those in **Figure 26c2**-**c3** due to motions' different effects in the two experiments, although the breast just deforms slightly between the two experiments. After motion correction, the features have high similarity as shown in **Figure 26b2**-**b3** and **Figure 26d2**-**d3**, which validates that the features recovered by motion correction are true. Image qualities in the las two experiments are also improved by motion correction, as shown in **Figure 27c1**–**c3** and **d1**–**d3**, respectively.

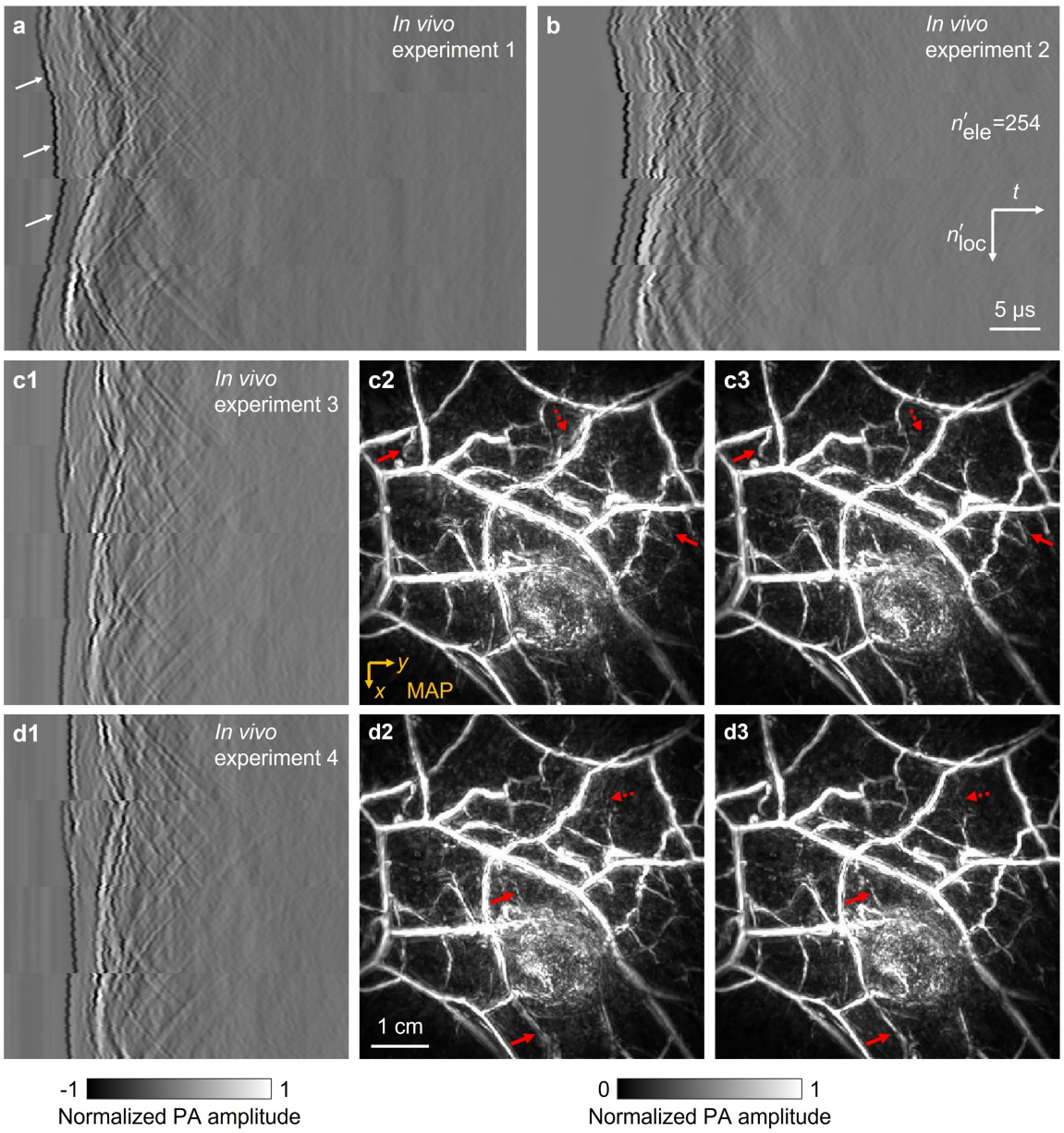

Figure 27 Four sets of signals ($N_{\mathrm{loc}} = 99$, $N_{\mathrm{ele}} = 4 \times 256$) from a human breast *in vivo* with heartbeat-induced motions and reconstructed images. **a**-**b** Signals in the first two experiments. The reconstructed images are shown in **Figure 26**. Temporal shifts in **a** caused by the quasiperiodic motion are indicated by white arrows. **c1**–**c3** and **d1**–**d3** Signals in the third and fourth experiments, respectively, with the images reconstructed without (**c2**-**d2**) and with (**c3**-**d3**) motion correction. Examples of the improvements due to motion correction are indicated by red-solid (feature enhanced) and red-dotted (artifacts suppressed) arrows.

5.6 Discussion

In both numerical simulations and human breast imaging *in vivo*, the motion correction

method successfully mitigates motion-induced artifacts and recovers motion-compromised features. Due to its low requirement in data acquisition and high robustness, the method can be broadly applied to brain, heart, chest, abdominal, and fetus imaging. Currently, the method reduces the effects of motions in a set of signals to form a single image. In a future study, we may perform finer manipulations of the system matrix to recover the complete motion profile from these signals: an important feature for heart imaging. Although we demonstrate the proposed method only using PACT, it is applicable to CT and MRI due to the high similarities in their system matrices as explained in Chapter 4.

Chapter 6

SINGLE-ELEMENT SINGLE-SHOT 3D PACT

Y. Zhang†, **P. Hu**†, L. Li, R. Cao, A. Khadria, K. Maslov, X. Tong, Y. Zeng, L. Jiang, Q. Zhou, and L. V. Wang, "Single-shot 3D photoacoustic tomography using a single-element detector for ultrafast imaging of hemodynamics." bioRxiv, p. 2023.03.14.532661, 15-Mar-2023. Under review by *Nat. Biomed. Eng.* DOI: 10.1101/2023.03.14.532661

6.1 Abstract

Imaging hemodynamics is crucial for the diagnosis, treatment, and prevention of vascular diseases. However, current imaging techniques are limited due to the use of ionizing radiation or contrast agents, short penetration depth, or complex and expensive data acquisition systems. Photoacoustic tomography shows promise as a solution to these issues. However, existing photoacoustic tomography methods collect signals either sequentially or through numerous detector elements, leading to either low imaging speed or high system complexity and cost. To address these issues, here we introduce a method to capture a 3D photoacoustic image of vasculature using a single laser pulse and a single-element detector that functions as 6,400 virtual ones. Our method enables ultrafast volumetric imaging of hemodynamics in the human body at up to 1 kHz and requires only a single calibration for different objects and for long-term operations. We demonstrate 3D imaging of hemodynamics at depth in humans and small animals, capturing the variability in blood flow speeds. This concept can inspire other imaging technologies and find applications such as home-care monitoring, biometrics, point-of-care testing, and wearable monitoring.

6.2 Background

Vascular diseases, including atherosclerosis, thrombosis, and aneurysms, pose serious health risks such as heart attack, stroke, and organ failure [88]. The early detection of these diseases through imaging of hemodynamics is crucial, enabling prompt intervention and treatment [89]. The efficacy of pharmacological therapies and surgical interventions, such as

angioplasty and stenting, can also be evaluated through monitoring changes in blood flow and velocity [90]. Additionally, assessing an individual's risk of developing vascular diseases is possible by measuring factors such as blood flow velocity, facilitating preventive measures. In biomedical research, imaging of hemodynamics can provide insights into the physiology and pathology of blood vessels, aiding in the development of new treatments for diseases such as hypertension, diabetes, and cancer. Overall, imaging of hemodynamics plays a vital role in disease diagnosis, treatment, and prevention, ultimately improving patient outcomes and advancing medical research [88].

There are several techniques available to image hemodynamics in the human body, each with its own strengths and limitations. Magnetic resonance imaging (MRI), computed tomography (CT) angiography, and positron emission tomography (PET) are all capable of producing high-resolution images of the vascular system and blood flow dynamics, but they require the use of ionizing radiation and the injection of contrast agents, which can have adverse health effects [91]–[93]. Moreover, relying on strong ionizing sources and numerous detector elements, these techniques are bulky and expensive, making them inaccessible to mobile clinics or small healthcare facilities. Optical imaging techniques, such as fluorescence imaging and optical coherence tomography (OCT), offer noninvasive visualization of hemodynamics, but their penetration depths are constrained by the optical diffusion limit (~1 mm) and do not have sufficient specificity to hemoglobin [94], [95]. Doppler ultrasound is another option, providing real-time measurement of blood flow velocity and direction. However, even with recent improvements in minimizing ultrasound probes, state-of-the-art ultrasound imaging techniques still require burdensome and costly data acquisition systems, e.g., Verasonics, due to hundreds to thousands of detector elements [96], [97].

PACT offers a promising solution to the limitations faced by other imaging techniques for hemodynamic imaging [98]–[100]. However, a fast PACT system typically requires numerous detection elements and the corresponding data acquisition systems, making it complex, expensive, and bulky. To overcome the challenges of complexity, cost, and size in PACT systems, researchers have been working on ways to reduce the number of detector elements needed to reconstruct a 3D image. One promising approach utilizes the principles

of compressive sensing and single-pixel imaging [101]–[104], which use acoustic scatterers to achieve photoacoustic (PA) or ultrasound computed tomography with just a single detector element [105]–[108]. However, these techniques are time-consuming, as they require a sequence of measurements with different mask configurations, limiting their speed. To address this issue, researchers have developed novel methods that take advantage of the spatiotemporal encoding of an ergodic relay (ER) or a chaotic cavity [109], [110]. These techniques can produce high-quality single-shot images while using fewer detector elements [111]–[115]. However, they have only been demonstrated for 2D imaging and require recalibration for different objects, which can be time-consuming. Additionally, they may not be suitable for long-term imaging in unstable environments due to their sensitivity to boundary conditions.

Here, we present photoacoustic computed tomography through an ergodic relay (PACTER), a PACT system that simultaneously addresses the challenges faced by other imaging techniques. PACTER provides a highly accessible and efficient solution, paving the way for noninvasive, label-free, and ultrafast 3D imaging of hemodynamics at depth in humans. With PACTER, a single-element detector encodes information equivalent to that of 6,400 virtual ones, enabling the reconstruction of a tomographic image of vasculature in 3D with just a single laser pulse. The system achieves ultrafast volumetric imaging at kilohertz rates, making it possible to capture fast hemodynamics in the human body in real time. We demonstrate PACTER's capability in monitoring vital signs in small animals and visualizing human hemodynamics in response to cuffing, capturing the variability in blood flow speeds. Because PACTER signals are unaffected by the boundary conditions of the object, the system only needs to be calibrated once and is suitable for long-term imaging in an unstable environment. PACTER's single-element detector design makes it convenient, affordable, and compact, thus translatable to clinical applications such as home-care monitoring [116], [117], biometrics, point-of-care testing [118], and noninvasive hemodynamic monitoring in intensive care units [119]. The single-element detector concept in PACTER can also be generalized to other imaging technologies, such as ultrasonography [120], sonar [121], and radar [122].

The PACTER system requires calibration only once prior to its utilization for a complete series of imaging. In the calibration procedure (**Figure 28a**), the laser beam was transmitted through the ER and focused on a uniform optical absorber placed on top of the ER. We used bovine blood as our calibration target. Using two motorized stages, we controlled the positions of a pair of mirrors to steer the focused laser beam across the field of view (FOV) in the *x*-*y* plane and recorded the PACTER signals at each scanning position. After calibration, the uniform optical absorber could be removed, and the system was ready for imaging. In the imaging procedure (**Figure 28b**), the focusing lens was replaced by a fly's eye homogenizer, which converted the incident laser beam into a widefield, homogenized illumination pattern that had the same shape and width as that of the calibration FOV. To acquire imaging data, the object was directly placed on top of the ER using ultrasound gel as the coupling medium, and we recorded the PACTER signal generated by the object following each laser pulse.

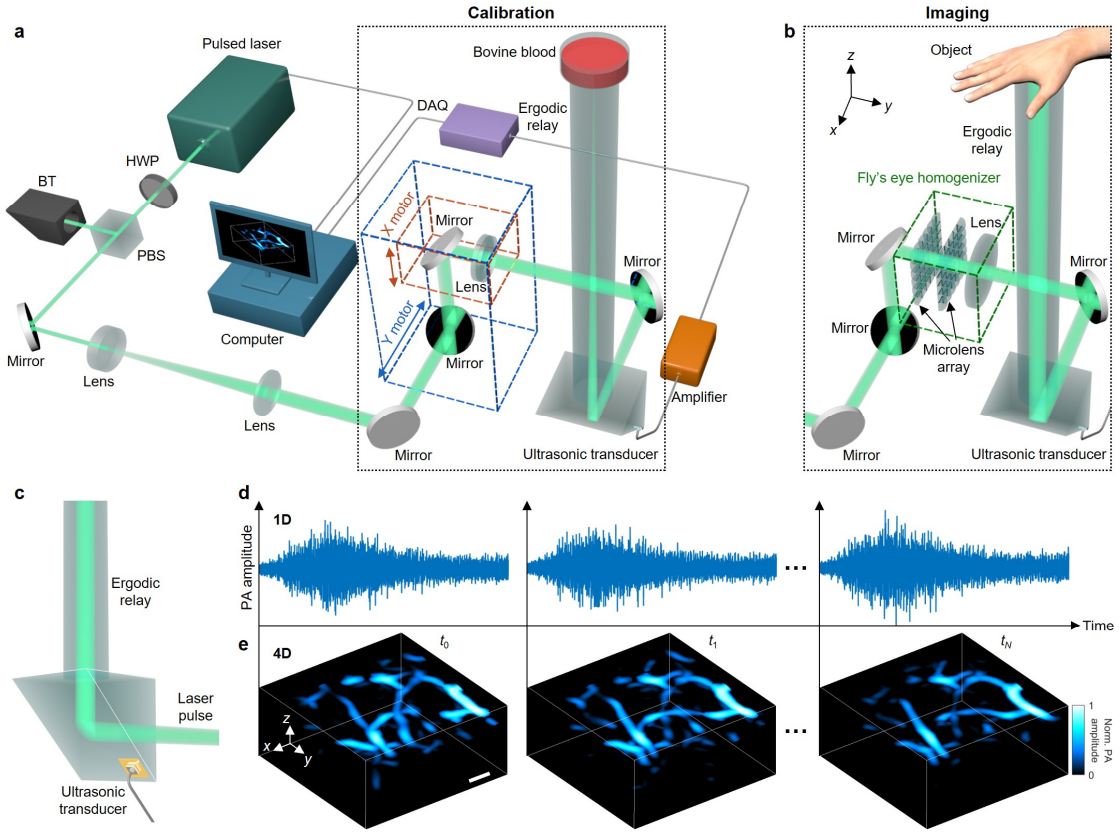

Figure 28 PACTER system. **a**-**b** Schematic of the PACTER system in the calibration (**a**) and imaging (**b**) procedures. HWP, half-wave plate; PBS, polarizing beamsplitter; BT, beam trap; DAQ, data acquisition system. The differences between the two modes are highlighted in the black dotted boxes. **c** Schematic of the single-element ultrasonic transducer fabricated on the ER. **d** 1D (*t*) PACTER signals detected by the ultrasonic transducer at time instances t_0, t_1, and t_N. **e** Reconstruction of a 4D (*xyzt*) image of human palmar vessels from the signals in **d**. Norm., normalized. Scale bar, 1 mm.

To enhance the detection sensitivity and improve the stability of the system, we fabricated a single-element ultrasonic transducer on the hypotenuse surface of the prism of the ER (**Figure 28c**). The transducer was based on a lead magnesium niobate–lead titanate (PMN-PT) single crystal, which achieved exceptional piezoelectric performance, such as high piezoelectric constant (d_{33}) and electromechanical coupling coefficient (k_t) [123]–[125]. Further, because one of the gold electrodes of the transducer was directly sputtered on the surface of the ER, the PACTER signals inside the ER could reach the transducer with the maximum transmission, and the transducer and the ER became a whole piece that facilitated stable data acquisition.

Conventionally, a single-element ultrasonic transducer can only acquire a 1D signal in the time domain (**Figure 28d**). However, with the ER, PACTER can use the single-element transducer to encode spatiotemporal information equivalent to those captured by 6,400 detectors, which can then be used to reconstruct a 3D map of the optical absorbers in the imaging volume. With a kilohertz laser repetition rate, PACTER can use the 1D (*t*) signals to generate a thousand 3D (*xyz*) volumes per second, leading to a high-speed 4D (*xyzt*) image of optical absorption in, e.g., human palmar vessels (**Figure 28e**).

6.3 Theory

6.3.1 Forward operator

We performed calibrations at pixels on a 2D plane and used these pixels as virtual ultrasonic transducers for 3D imaging. If non-zero initial pressure exists only on the calibration plane, the detected signal $s(t)$ at time t can be expressed as

$$s(t) = \sum_{n=1}^{N} p_n k_n(t)\,, t \geq 0, \tag{117}$$

where N is the number of calibrated virtual transducers, $k_n(t)$ is the normalized impulse response from the calibration at the n-th virtual transducer, and p_n is the root-mean-squared PA amplitude proportional to the initial pressure at the n-th virtual transducer.

For initial pressure in a 3D volume, we assume M source points located at $\mathbf{r}'_m, m = 1,2,\dots,M$, in an acoustically homogeneous 3D region attached to the calibration plane. The PA wave generated from the source point at $\mathbf{r}'_m$ propagates to the calibrated virtual transducer $\mathbf{r}_n$ with the speed of sound c after time $t_{m,n} = \frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c}$, which, through the ER, adds $p_{m,n}k_n\left(t - \frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c}\right)$ to the detected signal, with the PA amplitude $p_{m,n}$ quantified as $p_{m,n} = \frac{w(\theta_{m,n})p_{0,m}}{\|\mathbf{r}'_m - \mathbf{r}_n\|}$. Here, $\theta_{m,n}$ denotes the incidence angle satisfying $\cos\theta_{m,n} = \frac{(\mathbf{r}'_m - \mathbf{r}_n)\cdot\mathbf{n}}{\|\mathbf{r}'_m - \mathbf{r}_n\|}$ with $\mathbf{n}$ being the normal vector of the calibration plane; function $w(\theta_{m,n})$ describes a virtual transducer's angle-dependent sensitivity; and $p_{0,m}$ is proportional to the initial pressure at $\mathbf{r}'_m$. We replace $p_n k_n(t)$ in Eq. (117) with $p_{m,n}k_n\left(t - \frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c}\right)$ from all the M source points and obtain the detected wide-field PA signal [126]

$$s(t) = \sum_{n=1}^{N}\sum_{m=1}^{M} \frac{w(\theta_{m,n})p_{0,m}}{\|\mathbf{r}'_m - \mathbf{r}_n\|} k_n\left(t - \frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c}\right), t \geq 0. \tag{118}$$

Here, we define $k_n(t) = 0,\ n = 1,2\dots,N,\ t < 0$. For sufficiently small virtual ultrasonic transducers, we assume that

$$w(\theta_{m,n}) = \mathbf{1}_{[0,\theta_l]}(\theta_{m,n})\cos\theta_{m,n}\,. \tag{119}$$

Here, we use the indicator function

$$\mathbf{1}_A(x) = \begin{cases} 1, x \in A \\ 0, x \notin A \end{cases} \tag{120}$$

to rejection detections with incidence angles greater than the critical angle θ_l. Substituting Eq. (119) into Eq. (118) yields

$$s(t) = \sum_{n=1}^{N}\sum_{m=1}^{M} p_{0,m} \frac{\mathbf{1}_{[0,\theta_l]}(\theta_{m,n})\cos\theta_{m,n}}{\|\mathbf{r}'_m - \mathbf{r}_n\|} k_n\left(t - \frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c}\right), t \geq 0. \tag{121}$$

We let L be the number of time points after temporal discretization. Then the computational complexity of a forward model based on Eq. (121) is $O(MNL)$.

To accelerate the forward model in Eq. (121), we split the delay term $\frac{\|\mathbf{r}'_m-\mathbf{r}_n\|}{c}$ from function $k_n(t)$ through temporal convolution:

$$k_n\left(t-\frac{\|\mathbf{r}'_m-\mathbf{r}_n\|}{c}\right)=\delta\left(t-\frac{\|\mathbf{r}'_m-\mathbf{r}_n\|}{c}\right)*k_n(t),$$
$$m=1,2,\dots,M, n=1,2,\dots,N, t\geq 0. \quad (122)$$

Substituting Eq. (122) into Eq. (121), we obtain [126]

$$s(t)=\sum_{n=1}^{N}k_n(t)*\sum_{m=1}^{M}p_{0,m}\frac{\mathbf{1}_{[0,\theta_1]}(\theta_{m,n})\cos\theta_{m,n}}{\|\mathbf{r}'_m-\mathbf{r}_n\|}\delta\left(t-\frac{\|\mathbf{r}'_m-\mathbf{r}_n\|}{c}\right), t\geq 0. \quad (123)$$

The inner summation in Eq. (123) has a complexity of $O(MN)$ and each temporal convolution is implemented through three fast Fourier transforms (FFTs) with a complexity of $O(L\log_2 L)$. Thus, the forward model based on Eq. (123) has a computational complexity of $\max\{O(MN), O(NL\log_2 L)\}$, which is $\max\left\{\frac{1}{L}, \frac{\log_2 L}{M}\right\}$ times of $O(MNL)$. In this work, L = 65,536, $M = 80\times 80\times 120$, and $\max\left\{\frac{1}{L}, \frac{\log_2 L}{M}\right\}^{-1} = 48{,}000$, which indicates the factor of acceleration.

PACTER needs only a one-time universal calibration despite its more stringent requirement for 3D imaging. In the calibration procedure, the focused laser beam was scanned across the FOV in 80 by 80 steps with a step size of 0.1 mm (**Figure 29a**). To ensure that the PACTER signal acquired at each calibration pixel was distinct from others, we chose the step size to be about a half of the full width at half maximum (FWHM) of a line profile along the cross-correlation map, i.e., ~0.21 mm. Although the calibration signals were obtained by scanning the laser beam across a 2D plane, they could be used as $80\times 80 = 6{,}400$ virtual transducers for 3D reconstruction (**Figure 29b**) because (1) the PACTER signals were object-independent, and (2) the calibration signals were generated at the bottom of the 3D imaging volume. When source points in the 3D volume ($\mathbf{r}'_m$, m = 1, 2, …) were illuminated by a laser pulse, the PA signals they generated would propagate to the calibrated virtual transducers ($\mathbf{r}_n$, n = 1, 2, …) after time $t_{m,n} = \|\mathbf{r}'_m-\mathbf{r}_n\|/c$, where c is the speed of the sound in the medium. Then, these signals would follow the same acoustic path inside the ER to the ultrasonic transducer as that of the calibration signals. From the transducer's

perspective, compared with the calibration signal $k_n(t)$ acquired at $\mathbf{r}_n$ (**Figure 29c**), the signal from the source point $\mathbf{r}'_m$ relayed through $\mathbf{r}_n$ is proportional to $k_n(t)$ delayed by $t_{m,n}$, i.e., $k_n(t - \|\mathbf{r}'_m - \mathbf{r}_n\|/c)$ (**Figure 29d**). The signal is modulated by both $p_{0,m}$, the initial pressure at $\mathbf{r}'_m$, and a weighting factor dependent on the angle and distance. Accordingly, we developed an algorithm to reconstruct the initial pressure in the 3D volume. Because the reconstruction is prohibitively computationally intensive, we reformulated the forward model through temporal convolution and implemented it using the fast Fourier transformation, increasing the computational efficiency by 48,000 times.

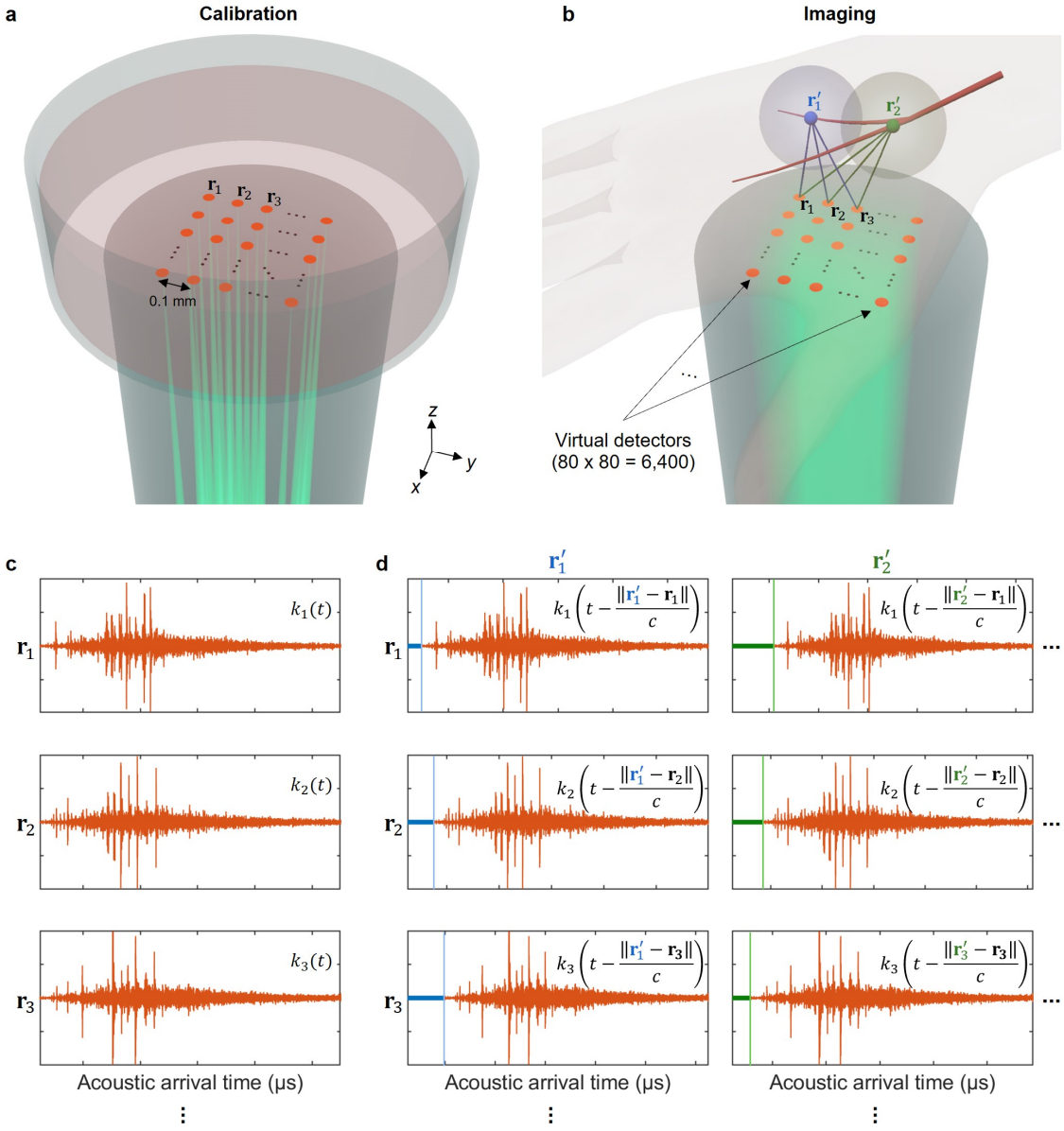

Figure 29 Single-shot 3D reconstruction in PACTER. **a** Illustration of the calibration procedure of PACTER. Focused laser beams for calibration are shown in green. Calibration pixels are highlighted as orange dots. The calibration step size is 0.1 mm. The calibration pixels (80 × 80) become 6,400 virtual transducers. $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$ are the positions of three calibrated virtual transducers. **b** Illustration of PACTER of human palmar vessels. The homogenized beam for widefield illumination is shown in green. $\mathbf{r}_1'$ and $\mathbf{r}_2'$ are the positions of two source points in the vessels. Blue and green spheres denote the PA waves generated by the source points. The calibrated virtual transducers capture the PA signals from $\mathbf{r}_1'$ and $\mathbf{r}_2'$ with different delays, indicated by the thick blue and green lines. **c** PACTER signals, $k_1(t)$, $k_2(t)$, $k_3(t)$, of the calibrated virtual transducers corresponding to $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$, respectively. **d** PACTER signal from the widefield imaging consists of PA signals from $\mathbf{r}_1'$ and $\mathbf{r}_2'$, which are essentially $k_1(t)$, $k_2(t)$, $k_3(t)$ delayed according to the distance between the calibrated virtual transducer and the source point.

6.3.2 Image reconstruction

Discretizing Eq. (123), we obtain the forward model

$$\mathbf{s} = \mathbf{H}\mathbf{p}_0, \tag{124}$$

where $\mathbf{s}$ represents a vector of length L, $\mathbf{p}_0$ denotes a vector of length $M(=M_1M_2M_3)$ which consists of all voxels in a 3D image of size $M_1 \times M_2 \times M_3$, and $\mathbf{H}$ is the system matrix of size $L \times M$. This forward model has a computational complexity of $\max\{O(MN), O(NL\log_2 L)\}$. To obtain an image from the signals $\mathbf{s}$, we invert the forward model by solving the regularized optimization problem

$$\widehat{\mathbf{p}}_0 = \underset{\mathbf{p}_0 \in \mathbb{R}^M, \mathbf{p}_0 \geq \mathbf{0}}{\arg\min} \|\mathbf{H}\mathbf{p}_0 - \mathbf{s}\|^2 + \lambda |\mathbf{p}_0|_{\mathrm{TV}}. \tag{125}$$

Here, $|\mathbf{p}_0|_{\mathrm{TV}}$ denotes the total variation (TV) of the 3D image corresponding to $\mathbf{p}_0$, and λ is the regularization parameter. Numerically, we solve this optimization problem through a fast iterative shrinkage-thresholding algorithm (FISTA) [68].

6.4 Phantom experiments

Using the signals acquired by the single detector, PACTER can image the 3D structure of a curved black wire with a single laser pulse (**Figure 30a**) and the 4D dynamics of bovine blood flushing through an S-shaped tube when the tube was illuminated by multiple laser pulses (**Figure 30b**). To evaluate whether the 3D volumes reconstructed by PACTER were correct measurements of the objects, we first compared the perspective views of the 3D volumes reconstructed by PACTER with the ground-truth projection images formed by raster-scanning the laser beam across the objects. Despite a lower spatial resolution compared with the ground truth, the comparison demonstrates that PACTER can correctly reconstruct the 3D objects in the lateral (*x*-*y*) directions. Second, we imaged a thin object in water, whose *z* positions were precisely controlled and measured by a linear translation stage. We imaged the object at multiple *z* positions, reconstructed the 3D volumes (**Figure 30c**), and compared the *z* positions in the reconstructed volumes with the real ones. As shown in **Figure 30d**, the reconstructed and real *z* positions were linearly related (R^2 = 1.000) with a

slope (0.993) close to 1, demonstrating that PACTER can accurately reconstruct the 3D objects in the axial (*z*) direction.

To quantify the resolution of PACTER, we imaged two human hairs embedded in an agarose block (**Figure 30e**). The hairs were intentionally positioned such that they were in close contact with each other, forming a crossing pattern that could be seen in both *z*- and *y*-projections (**Figure 30f**). Defining the spatial resolution as the minimum distance that can distinguish the peaks of the two hairs, we found the lateral and axial resolutions of PACTER to be 0.56 mm and 0.13 mm, respectively (**Figure 30g**). The anisotropic spatial resolutions along the lateral and axial directions were related to the image formation process in PACTER. The coarser lateral resolution was due to the acoustic impedance mismatch between the object and the ER and both resolutions are limited by the frequency-dependent acoustic attenuation.

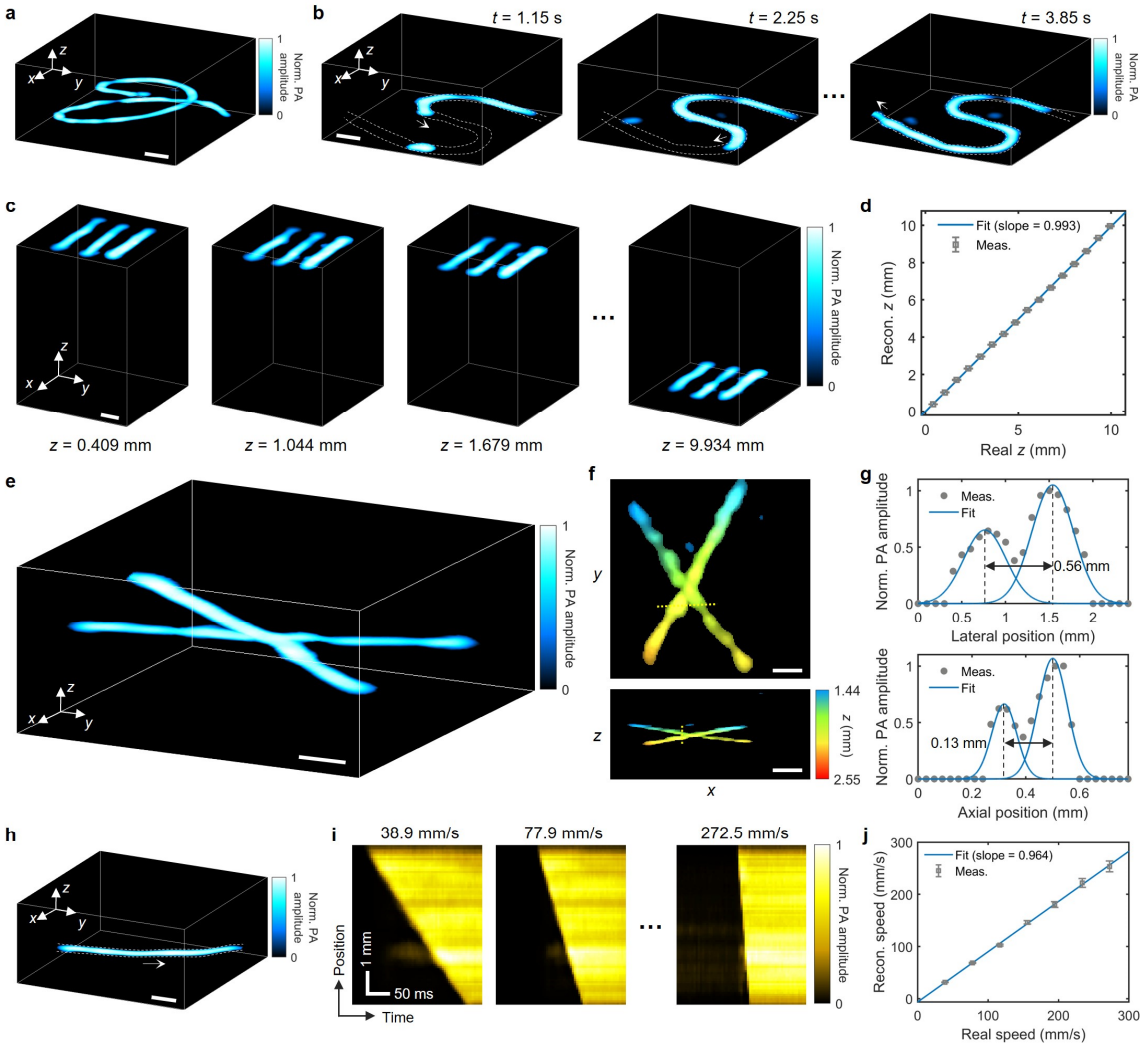

Figure 30 Spatiotemporal characterization of PACTER. **a** 3D PACTER image of a curved black wire. Norm., normalized. **b** Snapshots of 4D PACTER showing bovine blood flushing through an S-shaped tube. **c** 3D PACTER images of three bars printed with black ink on a transparent film. In each image, the object was placed at a different *z* position. **d** Reconstructed versus real *z* positions of the objects in **c**. The measurement results are plotted as means ± standard errors of the means (n = 1,980). The blue curve represents a linear fit. **e** 3D PACTER image of two crossing human hairs in agarose. **f** Maximum *z*- (top) and *y*-projections (bottom) of the 3D volume in **e**. The *z* positions of the object are color-encoded. **g** Profiles along the yellow dashed lines in **f** denoted by gray dots. The blue curves represent two-term Gaussian fits. Black arrows denote the minimum distances that can separate the two hairs. **h** 3D PACTER image of bovine blood flushing through a tube. The white arrow indicates the flushing direction. **i** PA amplitudes along the tube in **h** versus time, when the blood flushes through the tube at different speeds. **j** Speeds of the blood flow quantified from the reconstructed images versus the real speeds in **i**. The measurement results are plotted as means ± standard errors of the means (n = 74). The blue curve represents a linear fit. Scale bars, 1 mm.

To evaluate whether PACTER could be reliably used to image 4D dynamics, i.e., time-lapse movements of 3D objects, we captured 4D images of bovine blood flushing through a tube at different speeds precisely controlled by a syringe pump (**Figure 30h**). Based on the reconstructed 4D images, we plotted the PA amplitudes along the tube (1D images) over time (**Figure 30i**), calculated the speeds of the blood flow, and compared them with the real speeds set by the syringe pump (**Figure 30j**). A linear relationship ($R^2 = 0.999$) between the reconstructed and real speeds with a slope (0.964) close to 1 can be observed, demonstrating that PACTER is capable of 4D imaging, faithfully reconstructing the dynamics of 3D objects over time. Empowered by the imaging speed of up to a thousand volumes per second, PACTER could resolve the high-speed dynamics of the blood flushing through the tube at 272.5 mm/s in 3D, with a temporal resolution of 1 ms.

6.5 *In vivo* experiments

Enabled by the capability of noninvasive, label-free, and ultrafast 3D imaging, PACTER is expected to be suitable for monitoring hemodynamics *in vivo*. Here, we evaluated PACTER's capability in monitoring vital signs in small animals. We imaged the hemodynamics of the abdominal regions of mice (**Figure 31a**). With a single laser pulse, PACTER could reconstruct the abdominal vasculature in 3D (**Figure 31b-c**). When multiple laser pulses were used, PACTER revealed the 4D dynamics of the blood vessels. Based on the 4D PACTER datasets, we isolated individual blood vessels from the cross sections of the 3D volumes (**Figure 31d-e**) and visualized their motions and structural changes (**Figure 31f-g**).

By recording the time-lapse changes of the center positions and widths of the blood vessels, the respiratory motion could be tracked and identified (**Figure 31h-i**). Using Fourier analysis, we found that the center position of the blood vessel of Mouse 1 fluctuated periodically, exhibiting a respiratory frequency of 1.8 Hz (**Figure 31j**), whereas the width of the vessel was relatively stable. In comparison, a respiratory frequency of 1.4 Hz could be observed from both the center position and width of the blood vessel of Mouse 2 (**Figure 31k**). Further, when we imaged the third mouse, we observed a respiratory frequency of 1.9

Hz from the width, not the center position, of the blood vessel. The distinct 4D hemodynamics of the blood vessels from the three mice demonstrated that PACTER could be a practical tool in monitoring vital signs, such as breathing, in small animals.

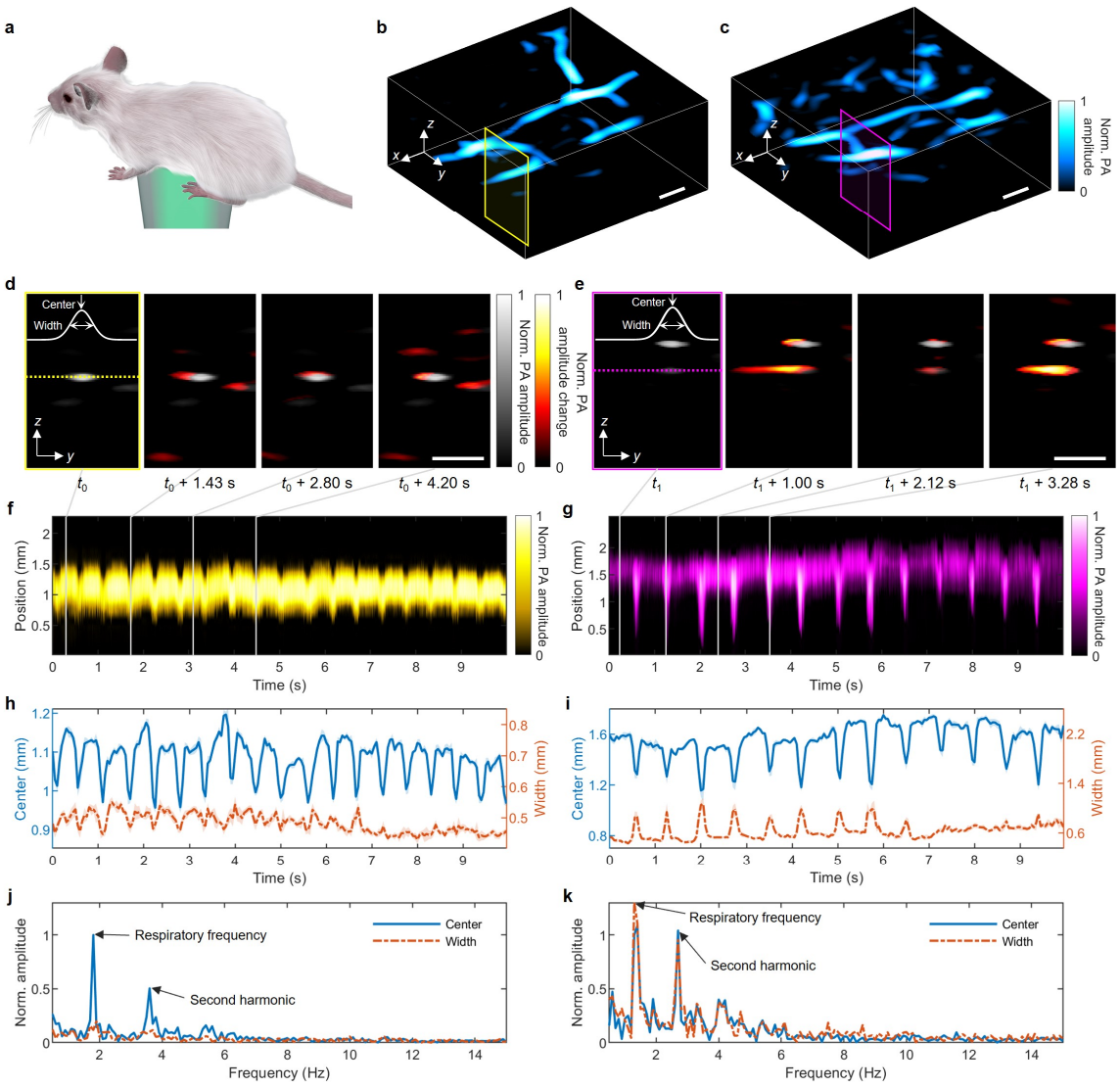

Figure 31 PACTER of mouse hemodynamics *in vivo*. **a** Schematic of the mouse imaging experiment. **b-c** 3D PACTER images of the abdominal vasculature of Mouse 1 (**b**) and Mouse 2 (**c**). Norm., normalized. **d-e** Cross-sectional 2D images corresponding to the yellow rectangle in **b** (**d**) and the magenta rectangle in **c** (**e**) at four different time instances from the 4D PACTER datasets. t_0 = 0.28 s, t_1 = 0.26 s. White solid curves represent the Gaussian fits of the vessels' profile denoted by the yellow (**d**) and magenta (**e**) dashed lines. Differences from the first image are highlighted. **f-g** PA amplitudes along the yellow dashed line (1D images) in **d** (**f**) and the magenta dashed line in **e** (**g**) versus time, where the time instances in **d** and **e** are labeled with vertical gray lines. **h-i** Center positions (blue solid curves) and widths (orange dash-dotted curves) of the vessels versus time, based

on the fits in **d** (**h**) and **e** (**i**). The shaded areas denote the standard deviations (n = 5). **j-k** Fourier transforms of the center positions and widths of the vessels in **h** (**j**) and **i** (**k**), showing the respiratory frequency from the vessel center positions only (**j**) or both the vessel center positions and widths (**k**). Scale bars, 1 mm.

To demonstrate PACTER's effectiveness in monitoring hemodynamics in humans, we imaged the hand vasculature of two participants. Different regions of the hand, e.g., fingers, proximal phalanx, and thenar regions, were imaged independently as the participants moved their hands to align those regions with the ER. In the following study, we focused on imaging the participants' thenar vasculature and their responses to cuffing, which was induced by a sphygmomanometer wrapped around the participants' upper arm (**Figure 32a**). Using PACTER, we imaged the thenar vasculature in 3D with single laser pulses (**Figure 32b-c**) and reconstructed the 4D dynamics of the blood vessels in response to cuffing. As shown in the maximum amplitude projections of the 4D datasets (**Figure 32d-e**), whereas some blood vessels exhibited a relatively stable PA amplitude throughout the experiment, the other vessels showed a decreased PA amplitude after cuffing due to the occlusion of blood flows; when the cuffing was released, the blood flows recovered, and the PA amplitude was rapidly restored (**Figure 32f-g**). The different hemodynamics of these two types of blood vessels in response to cuffing may indicate their distinct roles in the circulatory system [127]: the blood vessels with stable and changing PA amplitudes could be venous and arterial, respectively, agreeing with the observations reported in other cuffing-based studies [128]–[130]. With the capability to simultaneously image both arterial and venous blood *in vivo*, PACTER provides additional benefits over conventional pulse oximetry, which can only monitor arterial blood without spatial resolution [131].

Because PA amplitudes have 100% sensitivity to optical absorption [98], the 4D hemodynamics imaged by PACTER revealed the real-time changes in the blood vessels in response to cuffing, and the linear position of the blood front during the recovery phase could be used to measure the blood flow speed [130]. For Participant 1, the occlusion rate of the vessel was found to be 1.3 $\pm$ 0.1 m/s, significantly slower than the blood flow speed of 16.5 $\pm$ 2.8 m/s extracted from the recovery phase (**Figure 32h-i**). For Participant 2, the occlusion rate and the blood flow speed of the vessel were found to be 2.4 $\pm$ 0.3 m/s and 32.9 $\pm$ 6.5

m/s, respectively (**Figure 32j-k**), exhibiting a greater blood flow speed compared with Participant 1 (**Figure 32l**). Immediately after an imaging session, we asked Participant 1 to slightly move their hand and used PACTER to image a different area of the thenar region. Using the same analysis on a different blood vessel, the occlusion rate and the blood flow speed were found to be 0.6 ± 0.1 m/s and 22.4 ± 6.4 m/s, respectively. Taken together, we demonstrated that PACTER could monitor the hemodynamics in human, including the consistent responses of thenar vasculature to cuffing, and capture the variability in blood flow speeds.

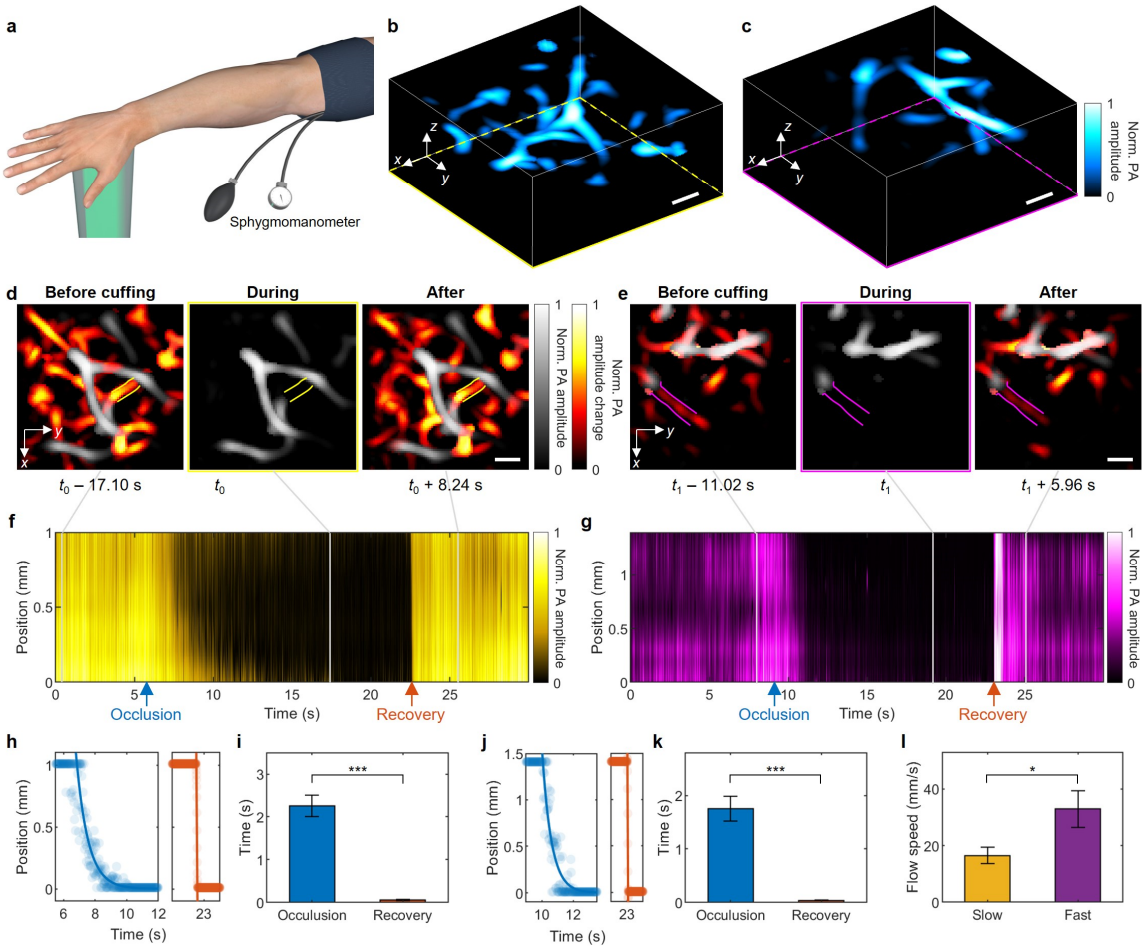

Figure 32 PACTER of human hand hemodynamics *in vivo*. **a** Schematic of the human hand imaging experiment. **b-c** 3D PACTER images of the thenar vasculature of Participant 1 (**b**) and Participant 2 (**c**). Norm., normalized. **d-e** Maximum amplitude projections of the 3D volumes from the 4D PACTER datasets along the *z* axis in **b** (**d**) and **c** (**e**) at the time instances before, during, and after cuffing. t_0 = 17.44 s, t_1 = 19.02 s. The solid lines flank the vessels under investigation. Differences from the images during cuffing are highlighted. **f-g** PA amplitudes along the vessels (1D images) in

d (**f**) and **e** (**g**) versus time, where the time instances in **d** and **e** are labeled with vertical gray lines. The blue and orange arrows indicate peak responses in the occlusion and recovery phases, respectively. **h** Positions (solid circles) of the blood front along the blood vessel during the occlusion (left) and recovery (right) phases in **f**. The blue curve is an exponential fit with an occlusion rate of 1.3 ± 0.1 m/s, and the orange curve is a linear fit showing the blood flow speed of 16.5 ± 2.8 m/s. **i** Comparison between the durations of the occlusion and recovery phases in **f**. ***$P < 0.001$, calculated by the two-sample *t*-test. **j** Positions (solid circles) of the blood front along the blood vessel during the occlusion (left) and recovery (right) phases in **g**. The blue curve is an exponential fit with an occlusion rate of 2.4 ± 0.3 m/s, and the orange curve is a linear fit showing the blood flow speed of 32.9 ± 6.5 m/s. **k** Comparison between the durations of the occlusion and recovery phases in **g**. ***$P < 0.001$, calculated by the two-sample *t*-test. **l** Comparison between the blood flow speeds during recovery in **f** and **g**. *$P < 0.05$. Scale bars, 1 mm.

6.6 Discussion

Although the current implementation of PACTER requires motorized stages for calibration, a pulsed laser for illumination, and a DAQ card for data acquisition, these requirements could be fulfilled using cheaper and more compact alternatives. Owing to the universal calibration capability, the ER in PACTER could be pre-calibrated, eliminating the need for motor-based calibration. The pulsed laser and the DAQ card could be replaced with cost-effective light-emitting diodes (LEDs) [132] and microcontrollers [133], respectively, which could further enhance the portability of the system. Additionally, mass production of the ER and the single-element ultrasonic transducer could substantially lower the cost of the system, making PACTER more accessible to users or researchers in low-resource settings, further lowering the barriers to clinical translation.

Due to the large dimensions of the ER compared with the acoustic wavelength, the PA waves need to propagate a long distance inside the ER; therefore, a slight change in the speed of sound due to temperature fluctuations [132] would cause large differences in the measured PACTER signal. To address this problem, we built a temperature stabilizing box to maintain the temperature of the ER, which stabilizes the temperature of the ER at a set temperature, e.g., 30ºC, at all times, guaranteeing a constant speed of sound throughout the experiments. In addition, the penetration depth of our PACTER system was limited to 3.6 mm *in vivo* due to the strong attenuation of 532 nm light by endogenous chromophores in biological tissue [134]. Changing the wavelength to 1064 nm could increase the penetration depth to several centimeters [20], [23]. Another limitation of PACTER is the relatively small FOV (8 mm ×

8 mm) constrained by the diameter of the glass rod in the ER. We believe that the design of the ER could be optimized further to achieve a larger FOV, enabling new applications such as vascular biometrics [135]. Finally, the spatial resolution of PACTER is currently limited by the acoustic impedance mismatch between the object and the ER, which could be addressed in the future by adding an impedance-matching layer on top of the ER.

In summary, PACTER, a noninvasive, label-free, and ultrafast imaging technique, enables 4D imaging of hemodynamics in humans using the 1D signal captured by a single detector, achieving an imaging speed of up to a thousand volumes per second. We have demonstrated PACTER's capability to visualize the 4D hemodynamics in humans and small animals. We have also shown the convenience of using PACTER to image different objects, including human hands and mouse abdomens, without the need for recalibration. PACTER's high imaging speed allows for immediate intervention in case of abnormal hemodynamic changes. Additionally, PACTER's low cost and compact form factor are ideal for point-of-care testing, facilitating quick and easy assessment of hemodynamic parameters at the bedside or in remote locations. We envision that PACTER will have profound impacts on a wide range of applications in biomedical research and clinical settings, including home care of diabetic foot ulcers [116] or carotid-artery disease [117], point-of-care screening for hypertension [118], and simultaneous oximetry of both arterial and venous blood in intensive care units [119]. Furthermore, PACTER's single-shot 3D imaging concept using a single-element detector can extend beyond optical imaging, aiding fields such as medical ultrasonography [120], underwater sonar [121], and airborne radar [122].

Chapter 7

DISCUSSION

In this dissertation, we propose computational methods in five topics for model imperfection compensation in PACT. These methods enable image reconstruction or significantly improve image quality in numerical simulations, phantom experiments, and most importantly, *in vivo* experiments. It needs to be noted that although all existing PACT systems are imperfect from all perspectives, we need to first focus on those imperfections posing dominant limitations to the image qualities for a certain application, such as those situations indicated by underlined chapter numbers in **Table 1**. Although the homogeneous medium assumption is valid for all the discussed applications, further studies may incorporate acoustic inhomogeneity into the proposed methods for higher-accuracy deep tissue imaging.

Single-shot 3D PACT system using a single transducer element is empowered by a fast forward operator for iterative reconstruction and achieves, so far, the highest 3D frame rate in PACT. This is a representative application of compressed sensing with efforts in both system design (device-level signal compression) and algorithm development (iterative reconstruction with sparsity constraint). This low-cost system breaks the limitation of low frame rate in 3D imaging brought by transducer moving speed and/or laser repetition rate and can facilitate PACT's application in clinical practice.

The radius-dependent and location-dependent spatiotemporal antialiasing methods proposed for the SBH-PACT system and the efficient explicit system matrix expression, fast functional imaging with sparse sampling, and intra-image nonrigid motion correction proposed for the 3D-PACT system are generic. They are not limited to the proposed SBH-PACT or 3D-PACT system but can also be applied to a PACT system of any other detection geometry. Moreover, these methods are directly applicable to CT and radial-sampling MRI due to the high similarities in their system matrices and may be extended to general-sampling MRI through proper mathematical transforms.

BIBLIOGRAPHY

[1] Y. Zhou, J. Yao, and L. V. Wang, "Tutorial on photoacoustic tomography," *J. Biomed. Opt.*, vol. 21, no. 6, p. 061007, 2016.

[2] J. Yao, J. Xia, and L. V. Wang, "Multiscale functional and molecular photoacoustic tomography," *Ultrason. Imaging*, vol. 38, no. 1, pp. 44–62, Jan. 2016.

[3] J. Xia, J. Yao, and L. V. Wang, "Photoacoustic tomography: Principles and advances," *Prog. Electromagn. Res.*, vol. 147, pp. 1–22, 2014.

[4] M. V. Marshall, J. C. Rasmussen, I.-C. Tan, M. B. Aldrich, K. E. Adams, X. Wang, C. E. Fife, E. A. Maus, L. A. Smith, and E. M. Sevick-Muraca, "Near-infrared fluorescence imaging in humans with indocyanine green: A review and update," *Open Surg. Oncol. J.*, vol. 2, no. 2, pp. 12–25, 2010.

[5] L. V. Wang and H. Wu, *Biomedical optics: Principles and imaging*. John Wiley & Sons, 2012.

[6] M. Xu and L. V. Wang, "Universal back-projection algorithm for photoacoustic computed tomography," *Phys. Rev. E*, vol. 71, no. 1, p. 016706, 2005.

[7] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2017," *CA. Cancer J. Clin.*, vol. 67, no. 1, pp. 7–30, 2017.

[8] D. S. Dizon, L. Krilov, E. Cohen, T. Gangadhar, P. A. Ganz, T. A. Hensing, S. Hunger, S. S. Krishnamurthi, A. B. Lassman, M. J. Markham, E. Mayer, M. Neuss, S. K. Pal, L. C. Richardson, R. Schilsky, G. K. Schwartz, D. R. Spriggs, M. A. Villalona-Calero, G. Villani, and G. Masters, "Clinical cancer advances 2016: Annual report on progress against cancer from the American Society of Clinical Oncology," *J. Clin. Oncol.*, vol. 34, no. 9, pp. 987–1011, Mar. 2016.

[9] A. B. Miller, C. Wall, C. J. Baines, P. Sun, T. To, and S. A. Narod, "Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: Randomised screening trial," *BMJ*, vol. 348, p. g366, Feb. 2014.

[10] R. Burton and R. Bell, "The global challenge of reducing breast cancer mortality," *The Oncologist*, vol. 18, no. 11, pp. 1200–1202, Nov. 2013.

[11] R. W. Pinsky and M. A. Helvie, "Mammographic breast density: Effect on imaging and breast cancer risk," *J. Natl. Compr. Canc. Netw.*, vol. 8, no. 10, pp. 1157–1165, Oct. 2010.

[12] P. E. Freer, "Mammographic breast density: Impact on breast cancer risk and implications for screening," *RadioGraphics*, vol. 35, no. 2, pp. 302–315, Mar. 2015.

[13] E. Devolli-Disha, S. Manxhuka-Kërliu, H. Ymeri, and A. Kutllovci, "Comparative accuracy of mammography and ultrasound in women with breast symptoms according to age and breast density," *Bosn. J. Basic Med. Sci.*, vol. 9, no. 2, pp. 131–136, May 2009.

[14] R. J. Hooley, L. M. Scoutt, and L. E. Philpotts, "Breast ultrasonography: State of the art," *Radiology*, vol. 268, no. 3, pp. 642–659, Sep. 2013.

[15] K. J. Murphy, J. A. Brunberg, and R. H. Cohan, "Adverse reactions to gadolinium contrast media: A review of 36 cases," *Am. J. Roentgenol.*, vol. 167, no. 4, pp. 847–849, Oct. 1996.

[16] M. A. Perazella, "Gadolinium-contrast toxicity in patients with kidney disease: Nephrotoxicity and nephrogenic systemic fibrosis," *Curr. Drug Saf.*, vol. 3, no. 1, pp. 67–75, Jan. 2008.
[17] D. Ibrahim, B. Froberg, A. Wolf, and D. E. Rusyniak, "Heavy metal poisoning: Clinical presentations and pathophysiology," *Clin. Lab. Med.*, vol. 26, no. 1, pp. 67–97, Mar. 2006.
[18] R. Choe, A. Corlu, K. Lee, T. Durduran, S. D. Konecky, M. Grosicka-Koptyra, S. R. Arridge, B. J. Czerniecki, D. L. Fraker, A. DeMichele, B. Chance, M. A. Rosen, and A. G. Yodh, "Diffuse optical tomography of breast cancer during neoadjuvant chemotherapy: A case study with comparison to MRI," *Med. Phys.*, vol. 32, no. 4, pp. 1128–1139, 2005.
[19] J. P. Culver, R. Choe, M. J. Holboke, L. Zubkov, T. Durduran, A. Slemp, V. Ntziachristos, B. Chance, and A. G. Yodh, "Three-dimensional diffuse optical tomography in the parallel plane transmission geometry: Evaluation of a hybrid frequency domain/continuous wave clinical system for breast imaging," *Med. Phys.*, vol. 30, no. 2, pp. 235–247, 2003.
[20] 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," *Nat. Commun.*, vol. 9, no. 1, p. 2352, Jun. 2018.
[21] C. D. Lehman and M. D. Schnall, "Imaging in breast cancer: Magnetic resonance imaging," *Breast Cancer Res.*, vol. 7, no. 5, p. 215, Aug. 2005.
[22] Y. Xu, M. Xu, and L. V. Wang, "Exact frequency-domain reconstruction for thermoacoustic tomography. II. Cylindrical geometry," *IEEE Trans. Med. Imaging*, vol. 21, no. 7, pp. 829–833, 2002.
[23] L. Li, L. Zhu, C. Ma, L. Lin, J. Yao, L. Wang, K. Maslov, R. Zhang, W. Chen, J. Shi, and L. V. Wang, "Single-impulse panoramic photoacoustic computed tomography of small-animal whole-body dynamics at high spatiotemporal resolution," *Nat. Biomed. Eng.*, vol. 1, no. 5, pp. 1–11, May 2017.
[24] K. Wang, R. Su, A. A. Oraevsky, and M. A. Anastasio, "Investigation of iterative image reconstruction in three-dimensional optoacoustic tomography," *Phys. Med. Biol.*, vol. 57, no. 17, pp. 5399–5423, Sep. 2012.
[25] 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 Trans. Med. Imaging*, vol. 32, no. 6, pp. 1097–1110, Jun. 2013.
[26] K. Mitsuhashi, J. Poudel, T. P. Matthews, A. Garcia-Uribe, L. V. Wang, and M. A. Anastasio, "A forward-adjoint operator pair based on the elastic wave equation for use in transcranial photoacoustic computed tomography," *SIAM J. Imaging Sci.*, vol. 10, no. 4, pp. 2022–2048, Jan. 2017.
[27] 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," *ArXiv Prepr. ArXiv160500133*, 2016.
[28] M. Pérez-Liva, J. L. Herraiz, J. M. Udías, E. Miller, B. T. Cox, and B. E. Treeby, "Time domain reconstruction of sound speed and attenuation in ultrasound computed

tomography using full wave inversion," *J. Acoust. Soc. Am.*, vol. 141, no. 3, pp. 1595–1604, Mar. 2017.
[29] A. Chambolle, "An algorithm for total variation minimization and applications," *J. Math. Imaging Vis.*, vol. 20, no. 1, pp. 89–97, 2004.
[30] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419–2434, 2009.
[31] N. Davoudi, X. L. Deán-Ben, and D. Razansky, "Deep learning optoacoustic tomography with sparse data," *Nat. Mach. Intell.*, vol. 1, no. 10, pp. 453–460, Oct. 2019.
[32] N. Davoudi, N. Davoudi, B. Lafci, B. Lafci, A. Özbek, A. Özbek, X. L. Deán-Ben, X. L. Deán-Ben, D. Razansky, and D. Razansky, "Deep learning of image- and time-domain data enhances the visibility of structures in optoacoustic tomography," *Opt. Lett.*, vol. 46, no. 13, pp. 3029–3032, Jul. 2021.
[33] V. A. Kelkar and M. Anastasio, "Prior image-constrained reconstruction using style-based generative models," in *Proceedings of the 38th International Conference on Machine Learning*, 2021, pp. 5367–5377.
[34] 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, Mar. 2018.
[35] C. M. Hyun, H. P. Kim, S. M. Lee, S. Lee, and J. K. Seo, "Deep learning for undersampled MRI reconstruction," *Phys. Med. Biol.*, vol. 63, no. 13, p. 135007, Jun. 2018.
[36] T. Würfl, M. Hoffmann, V. Christlein, K. Breininger, Y. Huang, M. Unberath, and A. K. Maier, "Deep learning computed tomography: Learning projection-domain weights from image domain in limited angle problems," *IEEE Trans. Med. Imaging*, vol. 37, no. 6, pp. 1454–1463, Jun. 2018.
[37] P. Hu, L. Li, L. Lin, and L. V. Wang, "Spatiotemporal antialiasing in photoacoustic computed tomography," *IEEE Trans. Med. Imaging*, vol. 39, no. 11, pp. 3535–3547, Nov. 2020.
[38] C. Cai, X. Wang, K. Si, J. Qian, J. Luo, and C. Ma, "Streak artifact suppression in photoacoustic computed tomography using adaptive back projection," *Biomed. Opt. Express*, vol. 10, no. 9, pp. 4803–4814, 2019.
[39] P. Hu, L. Li, and L. V. Wang, "Location-dependent spatiotemporal antialiasing in photoacoustic computed tomography," *IEEE Trans. Med. Imaging*, vol. 42, no. 4, pp. 1210–1224, Apr. 2023.
[40] S. Gottschalk, O. Degtyaruk, B. Mc Larney, J. Rebling, M. A. Hutter, X. L. Deán-Ben, S. Shoham, and D. Razansky, "Rapid volumetric optoacoustic imaging of neural dynamics across the mouse brain," *Nat. Biomed. Eng.*, vol. 3, no. 5, pp. 392–401, May 2019.
[41] X. L. Deán-Ben, T. F. Fehm, S. J. Ford, S. Gottschalk, and D. Razansky, "Spiral volumetric optoacoustic tomography visualizes multi-scale dynamics in mice," *Light Sci. Appl.*, vol. 6, no. 4, p. e16247, 2017.

[42] H.-P. F. Brecht, R. Su, M. P. Fronheiser, S. A. Ermilov, A. Conjusteau, and A. A. Oraevsky, "Whole-body three-dimensional optoacoustic tomography system for small animals," *J. Biomed. Opt.*, vol. 14, no. 6, p. 064007, 2009.
[43] A. P. Jathoul, J. Laufer, O. Ogunlade, B. Treeby, B. Cox, E. Zhang, P. Johnson, A. R. Pizzey, B. Philip, T. Marafioti, M. F. Lythgoe, R. B. Pedley, M. A. Pule, and P. Beard, "Deep in vivo photoacoustic imaging of mammalian tissues using a tyrosinase-based genetic reporter," *Nat. Photonics*, vol. 9, no. 4, pp. 239–246, Apr. 2015.
[44] Y. Matsumoto, Y. Asao, H. Sekiguchi, A. Yoshikawa, T. Ishii, K. Nagae, S. Kobayashi, I. Tsuge, S. Saito, M. Takada, Y. Ishida, M. Kataoka, T. Sakurai, T. Yagi, K. Kabashima, S. Suzuki, K. Togashi, T. Shiina, and M. Toi, "Visualising peripheral arterioles and venules through high-resolution and large-area photoacoustic imaging," *Sci. Rep.*, vol. 8, no. 1, p. 14930, Oct. 2018.
[45] A. Oraevsky, R. Su, H. Nguyen, J. Moore, Y. Lou, S. Bhadra, L. Forte, M. Anastasio, and W. Yang, "Full-view 3D imaging system for functional and anatomical screening of the breast," in *Photons Plus Ultrasound: Imaging and Sensing 2018*, 2018, vol. 10494, pp. 217–226.
[46] S. M. Schoustra, D. Piras, R. Huijink, T. J. P. M. op't Root, L. Alink, W. M. F. Kobold, W. Steenbergen, and S. Manohar, "Twente Photoacoustic Mammoscope 2: System overview and three-dimensional vascular network images in healthy breasts," *J. Biomed. Opt.*, vol. 24, no. 12, p. 121909, Oct. 2019.
[47] R. A. Kruger, C. M. Kuzmiak, R. B. Lam, D. R. Reinecke, S. P. Del Rio, and D. Steed, "Dedicated 3D photoacoustic breast imaging," *Med. Phys.*, vol. 40, no. 11, p. 113301, 2013.
[48] X. L. Deán-Ben and D. Razansky, "Functional optoacoustic human angiography with handheld video rate three dimensional scanner," *Photoacoustics*, vol. 1, no. 3, pp. 68–73, Dec. 2013.
[49] M. Heijblom, W. Steenbergen, and S. Manohar, "Clinical photoacoustic breast imaging: The twente experience," *IEEE Pulse*, vol. 6, no. 3, pp. 42–46, May 2015.
[50] Y. Tan, K. Xia, Q. Ren, and C. Li, "Three-dimensional photoacoustic imaging via scanning a one dimensional linear unfocused ultrasound array," *Opt. Express*, vol. 25, no. 7, pp. 8022–8028, Apr. 2017.
[51] A. M. Winkler, K. Maslov, and L. V. Wang, "Noise-equivalent sensitivity of photoacoustics," *J. Biomed. Opt.*, vol. 18, no. 9, pp. 097003–097003, 2013.
[52] Y. Xu, L. V. Wang, G. Ambartsoumian, and P. Kuchment, "Reconstructions in limited-view thermoacoustic tomography," *Med. Phys.*, vol. 31, no. 4, pp. 724–733, 2004.
[53] X. Pan, Y. Zou, and M. A. Anastasio, "Data redundancy and reduced-scan reconstruction in reflectivity tomography," *IEEE Trans. Image Process.*, vol. 12, no. 7, pp. 784–795, Jul. 2003.
[54] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing MRI," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 72–82, Mar. 2008.
[55] T. Küstner, K. Armanious, J. Yang, B. Yang, F. Schick, and S. Gatidis, "Retrospective correction of motion-affected MR images using deep learning frameworks," *Magn. Reson. Med.*, vol. 82, no. 4, pp. 1527–1540, 2019.

[56] Q. Lyu, H. Shan, Y. Xie, A. C. Kwan, Y. Otaki, K. Kuronuma, D. Li, and G. Wang, "Cine cardiac MRI motion artifact reduction using a recurrent neural network," *IEEE Trans. Med. Imaging*, vol. 40, no. 8, pp. 2170–2181, Aug. 2021.
[57] G.-H. Chen, J. Tang, and S. Leng, "Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets," *Med. Phys.*, vol. 35, no. 2, pp. 660–663, 2008.
[58] H. Yu and G. Wang, "SART-type image reconstruction from a limited number of projections with the sparsity constraint," *J. Biomed. Imaging*, vol. 2010, p. 3:1-3:9, Jan. 2010.
[59] M. Akçakaya, S. Moeller, S. Weingärtner, and K. Uğurbil, "Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: Database-free deep learning for fast imaging," *Magn. Reson. Med.*, vol. 81, no. 1, pp. 439–453, 2019.
[60] Y. Han, L. Sunwoo, and J. C. Ye, "k-space deep learning for accelerated MRI," *IEEE Trans. Med. Imaging*, vol. 39, no. 2, pp. 377–386, Feb. 2020.
[61] S. Winkelmann, T. Schaeffter, T. Koehler, H. Eggers, and O. Doessel, "An optimal radial profile order based on the golden ratio for time-resolved MRI," *IEEE Trans. Med. Imaging*, vol. 26, no. 1, pp. 68–76, Jan. 2007.
[62] H. Lee, J. Lee, H. Kim, B. Cho, and S. Cho, "Deep-neural-network-based sinogram synthesis for sparse-view CT image reconstruction," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 3, no. 2, pp. 109–119, Mar. 2019.
[63] C. Rohkohl, H. Bruder, K. Stierstorfer, and T. Flohr, "Improving best-phase image quality in cardiac CT by motion correction with MAM optimization," *Med. Phys.*, vol. 40, no. 3, p. 031901, 2013.
[64] L. Feng, L. Axel, H. Chandarana, K. T. Block, D. K. Sodickson, and R. Otazo, "XD-GRASP: Golden-angle radial MRI with reconstruction of extra motion-state dimensions using compressed sensing," *Magn. Reson. Med.*, vol. 75, no. 2, pp. 775–788, 2016.
[65] P. Han, R. Zhang, S. Wagner, Y. Xie, E. Cingolani, E. Marban, A. G. Christodoulou, and D. Li, "Electrocardiogram-less, free-breathing myocardial extracellular volume fraction mapping in small animals at high heart rates using motion-resolved cardiovascular magnetic resonance multitasking: A feasibility study in a heart failure with preserved ejection fraction rat model," *J. Cardiovasc. Magn. Reson.*, vol. 23, no. 1, p. 8, Feb. 2021.
[66] W. Wu, D. Hu, C. Niu, H. Yu, V. Vardhanabhuti, and G. Wang, "DRONE: Dual-domain residual-based optimization network for sparse-view CT reconstruction," *IEEE Trans. Med. Imaging*, vol. 40, no. 11, pp. 3002–3014, Nov. 2021.
[67] L. Lin, P. Hu, X. Tong, S. Na, R. Cao, X. Yuan, D. C. Garrett, J. Shi, K. Maslov, and L. V. Wang, "High-speed three-dimensional photoacoustic computed tomography for preclinical research and clinical translation," *Nat. Commun.*, vol. 12, no. 1, pp. 1–10, 2021.
[68] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sci.*, vol. 2, no. 1, pp. 183–202, 2009.

[69] S. Abbas, T. Lee, S. Shin, R. Lee, and S. Cho, “Effects of sparse sampling schemes on image quality in low-dose CT,” *Med. Phys.*, vol. 40, no. 11, p. 111915, 2013.
[70] J. Provost and F. Lesage, “The application of compressed sensing for photo-acoustic tomography,” *IEEE Trans. Med. Imaging*, vol. 28, no. 4, pp. 585–594, Apr. 2009.
[71] G. Yang, S. Yu, H. Dong, G. Slabaugh, P. L. Dragotti, X. Ye, F. Liu, S. Arridge, J. Keegan, Y. Guo, and others, “DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction,” *IEEE Trans. Med. Imaging*, vol. 37, no. 6, pp. 1310–1321, 2017.
[72] G. Wang, M. W. Vannier, and P.-C. Cheng, “Iterative X-ray cone-beam tomography for metal artifact reduction and local region reconstruction,” *Microsc. Microanal.*, vol. 5, no. 1, pp. 58–65, Jan. 1999.
[73] P. Grangeat, “Mathematical framework of cone beam 3D reconstruction via the first derivative of the radon transform,” in *Mathematical Methods in Tomography*, Berlin, Heidelberg, 1991, pp. 66–97.
[74] S. Zhang, D. Zhang, H. Gong, O. Ghasemalizadeh, G. Wang, and G. Cao, “Fast and accurate computation of system matrix for area integral model-based algebraic reconstruction technique,” *Opt. Eng.*, vol. 53, no. 11, p. 113101, Nov. 2014.
[75] K. T. Block, M. Uecker, and J. Frahm, “Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint,” *Magn. Reson. Med.*, vol. 57, no. 6, pp. 1086–1098, 2007.
[76] K. Kalisz, J. Buethe, S. S. Saboo, S. Abbara, S. Halliburton, and P. Rajiah, “Artifacts at cardiac CT: Physics and solutions,” *RadioGraphics*, vol. 36, no. 7, pp. 2064–2083, Nov. 2016.
[77] H. Wang and A. A. Amini, “Cardiac motion and deformation recovery from MRI: A review,” *IEEE Trans. Med. Imaging*, vol. 31, no. 2, pp. 487–503, Feb. 2012.
[78] L. Lin, X. Tong, S. Cavallero, Y. Zhang, S. Na, R. Cao, T. K. Hsiai, and L. V. Wang, “Non-invasive photoacoustic computed tomography of rat heart anatomy and function,” *Light Sci. Appl.*, vol. 12, no. 1, p. 12, Jan. 2023.
[79] S. A. Nehmeh and Y. E. Erdi, “Respiratory motion in positron emission tomography/computed tomography: A review,” *Semin. Nucl. Med.*, vol. 38, no. 3, pp. 167–176, May 2008.
[80] R. Grimm, S. Fürst, M. Souvatzoglou, C. Forman, J. Hutter, I. Dregely, S. I. Ziegler, B. Kiefer, J. Hornegger, K. T. Block, and S. G. Nekolla, “Self-gated MRI motion modeling for respiratory motion compensation in integrated PET/MRI,” *Med. Image Anal.*, vol. 19, no. 1, pp. 110–120, Jan. 2015.
[81] X. Tong, L. Lin, P. Hu, R. Cao, Y. Zhang, J. Olick-Gibson, and L. V. Wang, “Non-invasive 3D photoacoustic tomography of angiographic anatomy and hemodynamics of fatty livers in rats,” *Adv. Sci.*, vol. 10, no. 2, p. 2205759, 2023.
[82] E. D. Brandner, A. Wu, H. Chen, D. Heron, S. Kalnicki, K. Komanduri, K. Gerszten, S. Burton, I. Ahmed, and Z. Shou, “Abdominal organ motion measured using 4D CT,” *Int. J. Radiat. Oncol.*, vol. 65, no. 2, pp. 554–560, Jun. 2006.
[83] G. Cruz, D. Atkinson, C. Buerger, T. Schaeffter, and C. Prieto, “Accelerated motion corrected three-dimensional abdominal MRI using total variation regularized SENSE reconstruction,” *Magn. Reson. Med.*, vol. 75, no. 4, pp. 1484–1498, 2016.

[84] K. Kim, P. A. Habas, F. Rousseau, O. A. Glenn, A. J. Barkovich, and C. Studholme, "Intersection based motion correction of multislice MRI for 3-D in utero fetal brain image formation," *IEEE Trans. Med. Imaging*, vol. 29, no. 1, pp. 146–158, Jan. 2010.
[85] J. Y. Cheng, M. T. Alley, C. H. Cunningham, S. S. Vasanawala, J. M. Pauly, and M. Lustig, "Nonrigid motion correction in 3D using autofocusing with localized linear translations," *Magn. Reson. Med.*, vol. 68, no. 6, pp. 1785–1797, 2012.
[86] A. G. Christodoulou, J. L. Shaw, C. Nguyen, Q. Yang, Y. Xie, N. Wang, and D. Li, "Magnetic resonance multitasking for motion-resolved quantitative cardiovascular imaging," *Nat. Biomed. Eng.*, vol. 2, no. 4, pp. 215–226, Apr. 2018.
[87] J. Y. Cheng, T. Zhang, N. Ruangwattanapaisarn, M. T. Alley, M. Uecker, J. M. Pauly, M. Lustig, and S. S. Vasanawala, "Free-breathing pediatric MRI with nonrigid motion correction and acceleration," *J. Magn. Reson. Imaging*, vol. 42, no. 2, pp. 407–420, 2015.
[88] J. Sanz and Z. A. Fayad, "Imaging of atherosclerotic cardiovascular disease," *Nature*, vol. 451, no. 7181, pp. 953–957, Feb. 2008.
[89] M. Kadem, L. Garber, M. Abdelkhalek, B. K. Al-Khazraji, and Z. Keshavarz-Motamed, "Hemodynamic modeling, medical imaging, and machine learning and their applications to cardiovascular interventions," *IEEE Rev. Biomed. Eng.*, vol. 16, pp. 403–423, 2023.
[90] P. F. Davies, "Hemodynamic shear stress and the endothelium in cardiovascular pathophysiology," *Nat. Clin. Pract. Cardiovasc. Med.*, vol. 6, no. 1, pp. 16–26, Jan. 2009.
[91] M. D. Fox and M. E. Raichle, "Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging," *Nat. Rev. Neurosci.*, vol. 8, no. 9, pp. 700–711, Sep. 2007.
[92] P. Maurovich-Horvat, M. Ferencik, S. Voros, B. Merkely, and U. Hoffmann, "Comprehensive plaque assessment by coronary CT angiography," *Nat. Rev. Cardiol.*, vol. 11, no. 7, pp. 390–402, Jul. 2014.
[93] J. Provost, A. Garofalakis, J. Sourdon, D. Bouda, B. Berthon, T. Viel, M. Perez-Liva, C. Lussey-Lepoutre, J. Favier, M. Correia, M. Pernot, J. Chiche, J. Pouysségur, M. Tanter, and B. Tavitian, "Simultaneous positron emission tomography and ultrafast ultrasound for hybrid molecular, anatomical and functional imaging," *Nat. Biomed. Eng.*, vol. 2, no. 2, pp. 85–94, Feb. 2018.
[94] J. L. Fan, J. A. Rivera, W. Sun, J. Peterson, H. Haeberle, S. Rubin, and N. Ji, "High-speed volumetric two-photon fluorescence imaging of neurovascular dynamics," *Nat. Commun.*, vol. 11, no. 1, p. 6020, Dec. 2020.
[95] X. Shu, L. J. Beckmann, and H. F. Zhang, "Visible-light optical coherence tomography: A review," *J. Biomed. Opt.*, vol. 22, no. 12, p. 121707, Dec. 2017.
[96] C. Wang, B. Qi, M. Lin, Z. Zhang, M. Makihata, B. Liu, S. Zhou, Y. hsi Huang, H. Hu, Y. Gu, Y. Chen, Y. Lei, T. Lee, S. Chien, K. I. Jang, E. B. Kistler, and S. Xu, "Continuous monitoring of deep-tissue haemodynamics with stretchable ultrasonic phased arrays," *Nat. Biomed. Eng.*, vol. 5, no. 7, pp. 749–758, 2021.
[97] H. Hu, H. Huang, M. Li, X. Gao, L. Yin, R. Qi, R. S. Wu, X. Chen, Y. Ma, K. Shi, C. Li, T. M. Maus, B. Huang, C. Lu, M. Lin, S. Zhou, Z. Lou, Y. Gu, Y. Chen, Y. Lei, X.

Wang, R. Wang, W. Yue, X. Yang, Y. Bian, J. Mu, G. Park, S. Xiang, S. Cai, P. W. Corey, J. Wang, and S. Xu, "A wearable cardiac ultrasound imager," *Nature*, vol. 613, no. 7945, pp. 667–675, Jan. 2023.

[98] L. V. Wang and S. Hu, "Photoacoustic tomography: In vivo imaging from organelles to organs," *Science*, vol. 335, no. 6075, pp. 1458–1462, 2012.

[99] L. V. Wang and J. Yao, "A practical guide to photoacoustic tomography in the life sciences," *Nat. Methods*, vol. 13, no. 8, pp. 627–638, Aug. 2016.

[100] J. Weber, P. C. Beard, and S. E. Bohndiek, "Contrast agents for molecular photoacoustic imaging," *Nat. Methods*, vol. 13, no. 8, pp. 639–650, Aug. 2016.

[101] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 83–91, Mar. 2008.

[102] B. Sun, M. P. Edgar, R. Bowman, L. E. Vittert, S. Welsh, A. Bowman, and M. J. Padgett, "3D computational imaging with single-pixel detectors," *Science*, vol. 340, no. 6134, pp. 844–847, May 2013.

[103] M.-J. Sun, M. P. Edgar, G. M. Gibson, B. Sun, N. Radwell, R. Lamb, and M. J. Padgett, "Single-pixel three-dimensional imaging with time-based depth resolution," *Nat. Commun.*, vol. 7, no. 1, p. 12010, Nov. 2016.

[104] D. Stellinga, D. B. Phillips, S. P. Mekhail, A. Selyem, S. Turtaev, T. Čižmár, and M. J. Padgett, "Time-of-flight 3D imaging through multimode optical fibers," *Science*, vol. 374, no. 6573, pp. 1395–1399, Dec. 2021.

[105] P. Kruizinga, Pim van der Meulen, A. Fedjajevs, F. Mastik, G. Springeling, Nico de Jong, J. G. Bosch, and G. Leus, "Compressive 3D ultrasound imaging using a single sensor," *Sci. Adv.*, vol. 3, no. 12, 2017.

[106] X. Luís Dean-Ben and D. Razansky, "Localization optoacoustic tomography," *Light Sci. Appl.*, vol. 7, no. 4, pp. 18004–18004, Apr. 2018.

[107] X. L. Deán-Ben, A. Özbek, H. López-Schier, and D. Razansky, "Acoustic scattering mediated single detector optoacoustic tomography," *Phys. Rev. Lett.*, vol. 123, no. 17, p. 174301, Oct. 2019.

[108] E. Hahamovich, S. Monin, A. Levi, Y. Hazan, and A. Rosenthal, "Single-detector 3D optoacoustic tomography via coded spatial acoustic modulation," *Commun. Eng.*, vol. 1, no. 1, p. 25, Oct. 2022.

[109] G. Montaldo, D. Palacio, M. Tanter, and M. Fink, "Time reversal kaleidoscope: A smart transducer for three-dimensional ultrasonic imaging," *Appl. Phys. Lett.*, vol. 84, no. 19, pp. 3879–3881, May 2004.

[110] G. Montaldo, D. Palacio, M. Tanter, and M. Fink, "Building three-dimensional images using a time-reversal chaotic cavity," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 52, no. 9, pp. 1489–1497, Sep. 2005.

[111] B. T. Cox and P. C. Beard, "Photoacoustic tomography with a single detector in a reverberant cavity," *J. Acoust. Soc. Am.*, vol. 125, no. 3, pp. 1426–1436, Mar. 2009.

[112] M. D. Brown, E. Z. Zhang, B. E. Treeby, P. C. Beard, and B. T. Cox, "Reverberant cavity photoacoustic imaging," *Optica*, vol. 6, no. 6, p. 821, Jun. 2019.

[113] Y. Li, L. Li, L. Zhu, K. Maslov, J. Shi, P. Hu, E. Bo, J. Yao, J. Liang, L. Wang, and L. V. Wang, "Snapshot photoacoustic topography through an ergodic relay for high-

throughput imaging of optical absorption," *Nat. Photonics*, vol. 14, no. 3, pp. 164–170, Mar. 2020.
[114] Y. Li, T. T. W. Wong, J. Shi, H.-C. Hsu, and L. V. Wang, "Multifocal photoacoustic microscopy using a single-element ultrasonic transducer through an ergodic relay," *Light Sci. Appl.*, vol. 9, no. 1, p. 135, Dec. 2020.
[115] L. Li, Y. Li, Y. Zhang, and L. V. Wang, "Snapshot photoacoustic topography through an ergodic relay of optical absorption in vivo," *Nat. Protoc.*, vol. 14, no. 3, pp. 164–170, Apr. 2021.
[116] C. A. Andersen, "Noninvasive assessment of lower extremity hemodynamics in individuals with diabetes mellitus," *J. Vasc. Surg.*, vol. 52, no. 3, pp. 76S-80S, Sep. 2010.
[117] P. Sobieszczyk and J. Beckman, "Carotid artery disease," *Circulation*, vol. 114, no. 7, pp. e244–e247, Aug. 2006.
[118] S. K. Vashist, P. B. Luppa, L. Y. Yeo, A. Ozcan, and J. H. T. Luong, "Emerging technologies for next-generation point-of-care testing," *Trends Biotechnol.*, vol. 33, no. 11, pp. 692–705, Nov. 2015.
[119] P. E. Marik and M. Baram, "Noninvasive hemodynamic monitoring in the intensive care unit," *Crit. Care Clin.*, vol. 23, no. 3, pp. 383–400, Jul. 2007.
[120] C. L. Moore and J. A. Copel, "Point-of-care ultrasonography," *N. Engl. J. Med.*, vol. 364, no. 8, pp. 749–757, Feb. 2011.
[121] A. Trucco, M. Palmese, and S. Repetto, "Devising an affordable sonar system for underwater 3-D vision," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 10, pp. 2348–2354, Oct. 2008.
[122] A. Reigber, R. Scheiber, M. Jager, P. Prats-Iraola, I. Hajnsek, T. Jagdhuber, K. P. Papathanassiou, M. Nannini, E. Aguilera, S. Baumgartner, R. Horn, A. Nottensteiner, and A. Moreira, "Very-high-resolution airborne synthetic aperture radar imaging: Signal processing and applications," *Proc. IEEE*, vol. 101, no. 3, pp. 759–783, Mar. 2013.
[123] Q. Zhou, S. Lau, D. Wu, and K. Kirk Shung, "Piezoelectric films for high frequency ultrasonic transducers in biomedical applications," *Prog. Mater. Sci.*, vol. 56, no. 2, pp. 139–174, Feb. 2011.
[124] Q. Zhou, K. H. Lam, H. Zheng, W. Qiu, and K. K. Shung, "Piezoelectric single crystal ultrasonic transducers for biomedical applications," *Prog. Mater. Sci.*, vol. 66, pp. 87–111, Oct. 2014.
[125] Q. Zhou, X. Xu, E. Gottlieb, L. Sun, J. Cannata, H. Ameri, M. Humayun, P. Han, and K. Shung, "PMN-PT single crystal, high-frequency ultrasonic needle transducers for pulsed-wave Doppler application," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 54, no. 3, pp. 668–675, Mar. 2007.
[126] Y. Zhang, P. Hu, L. Li, R. Cao, A. Khadria, K. Maslov, X. Tong, Y. Zeng, L. Jiang, Q. Zhou, and L. V. Wang, "Single-shot 3D photoacoustic tomography using a single-element detector for ultrafast imaging of hemodynamics." bioRxiv, p. 2023.03.14.532661, 15-Mar-2023.
[127] M. K. Pugsley and R. Tabrizchi, "The vascular system: An overview of structure and function," *J. Pharmacol. Toxicol. Methods*, vol. 44, no. 2, pp. 333–340, Sep. 2000.

[128] M. C. Langham, T. F. Floyd, E. R. Mohler, J. F. Magland, and F. W. Wehrli, "Evaluation of cuff-induced ischemia in the lower extremity by magnetic resonance oximetry," *J. Am. Coll. Cardiol.*, vol. 55, no. 6, pp. 598–606, Feb. 2010.
[129] J. P. Loenneke, R. S. Thiebaud, C. A. Fahs, L. M. Rossow, T. Abe, and M. G. Bemben, "Effect of cuff type on arterial occlusion," *Clin. Physiol. Funct. Imaging*, vol. 33, no. 4, pp. 325–327, Jul. 2013.
[130] Y. Zhou, J. Liang, and L. V. Wang, "Cuffing-based photoacoustic flowmetry in humans in the optical diffusive regime," *J. Biophotonics*, vol. 9, no. 3, pp. 208–212, Mar. 2016.
[131] J. E. Sinex, "Pulse oximetry: Principles and limitations," *Am. J. Emerg. Med.*, vol. 17, no. 1, pp. 59–66, Jan. 1999.
[132] J. M. Ide, "The velocity of sound in rocks and glasses as a function of temperature," *J. Geol.*, vol. 45, no. 7, pp. 689–716, Oct. 1937.
[133] A. Fatima, K. Kratkiewicz, R. Manwar, M. Zafar, R. Zhang, B. Huang, N. Dadashzadeh, J. Xia, and K. (Mohammad) Avanaki, "Review of cost reduction methods in photoacoustic computed tomography," *Photoacoustics*, vol. 15, p. 100137, Sep. 2019.
[134] J. Yao and L. V. Wang, "Photoacoustic microscopy," *Laser Photonics Rev.*, vol. 7, no. 5, pp. 758–778, 2013.
[135] Y. Wang, Z. Li, T. Vu, N. Nyayapathi, K. W. Oh, W. Xu, and J. Xia, "A robust and secure palm vessel biometric sensing system based on photoacoustics," *IEEE Sens. J.*, vol. 18, no. 14, pp. 5993–6000, Jul. 2018.

Appendix A: Mathematical symbols

Table 2 Symbols for the fast forward operator and image reconstruction.

Symbol	Definition	Value
$M\ (m)$	The number of voxels in the image (index)	In functional imaging, $M = 180 \times 240 \times 160$ for numerical simulations and $M = 250 \times 250 \times 200$ for *in vivo* experiments; in motion correction, $M = 224 \times 224 \times 72$ for numerical simulations and $M = 300 \times 300 \times 150$ for *in vivo* experiments
$v_{m,1}$, $v_{m,2}$, and $v_{m,3}$ (m)	Voxel sizes along the x-axis, y-axis, and z-axis, respectively	$(v_{m,1}, v_{m,2}, v_{m,3}) =$ (0.1,0.1,0.1) (mm) for functional imaging $(v_{m,1}, v_{m,2}, v_{m,3}) =$ (0.2,0.2,0.2) (mm) for motion correction
$N_{\mathrm{loc}}\ (n_{\mathrm{loc}})$ and $4N_{\mathrm{loc}}\ (n'_{\mathrm{loc}})$	Rotating location numbers of the four-arc array (index) and single-arc array (index), respectively	In functional imaging, $4N_{\mathrm{loc}} = 396$ for dense sampling and $4N_{\mathrm{loc}} = 76, 40, 28, 20, 16, 12$ for sparse sampling; in motion correction, $4N_{\mathrm{loc}} = 288$ for numerical simulations and $4N_{\mathrm{loc}} = 396$ for *in vivo* experiments
$N_{\mathrm{ele}}\ (n_{\mathrm{ele}})$ and $\frac{N_{\mathrm{ele}}}{4}\ (n'_{\mathrm{ele}})$	The numbers of transducer elements in the four-arc array (index) and single-arc array (index), respectively	For numerical simulations, $\frac{N_{\mathrm{ele}}}{4} = 128$ in functional imaging and $\frac{N_{\mathrm{ele}}}{4} = 72$ in motion correction; for *in vivo* experiments,

		$\frac{N_{\text{ele}}}{4} = 256$
a and b (m)	The length and width of each transducer element	$a = 0.7$ mm and $b = 0.6$ mm
N (n)	The number of virtual elements in the 2D array (index)	$N = N_{\text{loc}}N_{\text{ele}}$
τ (s)	The temporal sampling step size of each element	$\tau = 25$ ns, which corresponds to a 40-MHz sampling frequency
L	The temporal sampling length of each element	$L = 4096$
L'	The effective length of the point source responses for nonzero values	$L' = 151$
A_n (m^2)	The detection area of the n-th virtual element	$A_n = ab = 0.42$ mm^2, $n = 1,2,\dots,N$
c ($\text{m} \cdot \text{s}^{-1}$)	Speed of sound	$c = 1.5$ $\text{mm} \cdot \mu\text{s}^{-1}$ in numerical simulations, tuned for *in vivo* experiments
K	The number of used singular components (Eqs. (72) and (83))	$K = 3$
v_m (m^3)	The volume of the m-th voxel: $v_m = v_{m,1}v_{m,2}v_{m,3}$	
$\mathbf{r}$ (m)	A location for ultrasonic signal detection	
$\mathbf{r}_n$ (m)	The center of the detection surface of the n-th virtual element	
$\mathbf{r}'$ (m)	A location in the image domain	
$\mathbf{r}'_m$ (m)	The center of the m-th voxel in the image	
t (s)	Time	
t_l (s)	The l-th sampling time	

$p_0(\mathbf{r}')$ (Pa)	The initial pressure at $\mathbf{r}'$
$p(\mathbf{r}, t)$ (Pa)	The pressure at location $\mathbf{r}$ and time t
$h_{\mathrm{s},n}(\mathbf{r}', t)$	The SIR of the n-th virtual element (Eq. (59))
$\hat{h}_{\mathrm{s},n}(\mathbf{r}', t)$	The temporally shifted version of $h_{\mathrm{s},n}(\mathbf{r}', t)$ (Eq. (64))
$h_{\mathrm{e},n}(t)$ (s^{-1})	The EIR of the n-th virtual element (Eq. (61))
$h_{\mathrm{e}}(t)$ (s^{-1})	The EIR shared by all virtual elements
$\hat{p}(\mathbf{r}_n, t)$ (Pa)	The pressure detected by the n-th virtual element at time t (Eq. (62))
$h_n(\mathbf{r}', t)$ (m^{-3})	The point source response (from $\mathbf{r}'$) received by the n-th virtual element (Eq. (63))
$\hat{h}_n(\mathbf{r}', t)$ (m^{-3})	The temporally shifted version of $h_n(\mathbf{r}', t)$ (Eq. (65))
$\mathbf{A}_n$ (m)	An orthonormal matrix formed by three vectors representing the local coordinates of the n-th virtual element (Eq. (66))
$\mathbf{r}'_{\mathrm{lcl}}$ (m)	A location in the local coordinate system of a virtual element
$\mathbf{r}'_{\mathrm{lcl},m_{\mathrm{lcl}}}$ (m)	The m_{lcl} -th location of interest in a virtual element's local coordinate system
$\hat{h}_{\mathrm{s,lcl}}(\mathbf{r}'_{\mathrm{lcl}}, t)$	The temporally shifted SIR of a virtual element expressed in its local coordinates (Eq. (68))
$\hat{h}_{\mathrm{lcl}}(\mathbf{r}'_{\mathrm{lcl}}, t)$ (m^{-3})	The temporally shifted point source response of a virtual element expressed in its local coordinates (Eq. (69))
$\hat{h}_{\mathrm{lcl},k}(\mathbf{r}'_{\mathrm{lcl}})$ (m^{-3}), and $\eta_k(t)$	The k-th spatial singular function and the k-th temporal singular function, respectively, in the SVD of $\hat{h}_{\mathrm{lcl}}(\mathbf{r}'_{\mathrm{lcl}}, t)$ (Eq. (72))
$\hat{h}_{\mathrm{lcl},m_{\mathrm{lcl}},l}$ (m^{-3})	The value of $\hat{h}_{\mathrm{lcl}}\left(\mathbf{r}'_{\mathrm{lcl},m_{\mathrm{lcl}}}, t_l\right)$ obtained through Eq. (81)
$\bar{h}_{\mathrm{lcl},k,m_{\mathrm{lcl}}}$ (m^{-3})	The value of $\hat{h}_{\mathrm{lcl},k}\left(\mathbf{r}'_{\mathrm{lcl},m_{\mathrm{lcl}}}\right)$ obtained through Eq. (83)
$\hat{h}_{\mathrm{lcl},k,n,m}$ (m^{-3})	The value of $\hat{h}_{\mathrm{lcl},k}\left((\mathbf{r}'_m - \mathbf{r}_n)\mathbf{A}_n\right)$ obtained through interpolation of $\bar{h}_{\mathrm{lcl},k,m_{\mathrm{lcl}}}$
$\eta_{k,l}$	The value of $\eta_k(t_l)$ obtained through Eq. (83)
$h_{n,m,l}$ (m^{-3})	The value of $h_n(\mathbf{r}'_m, t_l)$

$p_{0,m}$ (Pa)	The value of $p_0(\mathbf{r}'_m)$
$\hat{p}_{n,l}$ (Pa)	The value of $\hat{p}(\mathbf{r}_n, t_l)$
$\hat{p}_{n,l}^{\text{slow}}$ (Pa)	An approximation of $\hat{p}_{n,l}$ using the slow operator defined in Eq. (82)
$\hat{p}_{n,l}^{\text{fast}}$ (Pa)	An approximation of $\hat{p}_{n,l}$ using the fast operator defined in Eq. (86)
$\mathbf{p}_0$ (Pa)	A vector formed by $p_{0,m}, m = 1,2, \dots, M$
$\mathbf{H}$	A system matrix formed by the coefficients in Eq. (82) or Eq. (S30)
$\hat{\mathbf{p}}$ (Pa)	A vector formed by $\hat{p}_{n,l}, n = 1,2, \dots, N, l = 1,2, \dots, L$
$\hat{\mathbf{p}}_0$ (Pa)	A vector formed by the reconstructed voxel values
λ (m)	The regularization parameter in the regularized optimization (Eq. (90))
γ_{c} (Pa^{-1})	The system-specific measurement calibration factor (ignore in practice)
$\lvert\mathbf{p}_0\rvert_{\text{TV}}$ ($\text{Pa} \cdot \text{m}^{-1}$)	The TV norm of $\mathbf{p}_0$ defined in Eq. (91)

The unit of an array represents the unit of its elements.

Table 3 Additional symbols for fast functional imaging with sparse sampling.

Symbol	Definition	Value
L_{f} (l_{f})	The number of sparsely sampled images in functional imaging (index)	$L_{\text{f}} = 36$ for numerical simulations, and $L_{\text{f}} = 60$ for *in vivo* experiments
$\alpha_{\text{f},l_{\text{f}}}$	A functional signal profile	Defined by Eq. (98) for numerical simulations, and obtained from UBP-reconstructed images with $4N_{\text{loc}} = 76$ for *in vivo* experiments
A_{f}	The functional amplitude in Eq. (98)	$A_{\text{f}} = 0.18, 0.06, 0.02$
$\hat{\mathbf{p}}_{NL\times 1}$ (Pa)	Pressures detected by N densely distributed virtual elements at L	

	time points
$\widehat{\mathbf{p}}^{s}_{N'L\times 1}$ (Pa)	Pressures detected by N' sparsely distributed virtual elements at L time points
$\mathbf{p}_{0,M\times 1}$ (Pa)	Values of the M voxels
$\mathbf{H}_{NL\times M}$	A dense-sampling system matrix transferring $\mathbf{p}_{0,M\times 1}$ to $\widehat{\mathbf{p}}_{NL\times 1}$
$\mathbf{H}^{\mathrm{s}}_{N'L\times M}$	A sparse-sampling system matrix transferring $\mathbf{p}_{0,M\times 1}$ to $\widehat{\mathbf{p}}^{s}_{N'L\times 1}$
$\widehat{\mathbf{p}}_{0,M\times 1}$ (Pa)	Reconstructed values of the M voxels (Eq. (92))
$\boldsymbol{\mu}_{0,M'\times 1}$	The column-vector form of a modulation image of size M' ($\ll M$)
$\mathbf{U}_{M\times M'}$	An upsampling operator transferring $\boldsymbol{\mu}_{0,M'\times 1}$ to a smooth modulation image of size M in the column-vector form
$\widehat{\boldsymbol{\mu}}_{0,M'\times 1}$	The modulation image reconstructed for $\widehat{\mathbf{p}}^{s}_{N'L\times 1}$ in Eq. (93)
$\widehat{\mathbf{p}}^{\mathrm{m}}_{0}$ (Pa)	An image reconstructed by UBP from the signals of modulated prior image $\mathbf{H}_{NL\times M}\left(\widehat{\mathbf{p}}_{0,M\times 1}\odot\left(\mathbf{U}_{M\times M'}\widehat{\boldsymbol{\mu}}_{0,M'\times 1}\right)\right)$: the modulated UBP image
$\widehat{\mathbf{p}}^{\mathrm{r}}_{0}$ (Pa)	An image reconstructed by UBP from the residual signals $\widehat{\mathbf{p}}^{s}_{N'L\times 1}-\mathbf{H}^{\mathrm{s}}_{N'L\times M}\left(\widehat{\mathbf{p}}_{0,M\times 1}\odot\left(\mathbf{U}_{M\times M'}\widehat{\boldsymbol{\mu}}_{0,M'\times 1}\right)\right)$: the residual UBP image
$\widehat{\mathbf{p}}^{\mathrm{h}}_{0}$ (Pa)	The hybrid image defined as $\widehat{\mathbf{p}}^{\mathrm{h}}_{0}=\widehat{\mathbf{p}}^{\mathrm{m}}_{0}+\widehat{\mathbf{p}}^{\mathrm{r}}_{0}$ (Eq. (94))
$\widehat{\mathbf{p}}_{0,l_{\mathrm{f}}}$ (Pa)	The l_{f}-th sparsely sampled image
$\mathbf{p}_{0,l_{\mathrm{f}}}$ (Pa)	The l_{f}-th numerical phantom in functional simulation obtained by Eq. (97)
$\mathbf{p}_{0,\mathrm{b}}$ (Pa)	A background numerical phantom obtained in imaging with dense sampling
$\mathbf{p}_{0,\mathrm{f}}$ (Pa)	A functional numerical phantom obtained from smoothing, downsampling, and zero padding $\mathbf{p}_{0,\mathrm{b}}$
$\alpha_{\mathrm{b},l_{\mathrm{f}}}$	The background modulation factor for the l_{f}-th numerical phantom

$\hat{p}_{0,l_\mathrm{f},m}$ (Pa)	The value of the m-th voxel in $\widehat{\mathbf{p}}_{0,l_\mathrm{f}}$
$\underset{l_\mathrm{f}}{\mathrm{PCC}}\left(\hat{p}_{0,l_\mathrm{f},m}, \alpha_{\mathrm{f},l_\mathrm{f}}\right)$	The Pearson correlation coefficient (PCC) between $\hat{p}_{0,l_\mathrm{f},m}$ and $\alpha_{\mathrm{f},l_\mathrm{f}}$ (Eq. (95))
$\underset{l_\mathrm{f}}{\mathrm{PCCR}_{\lambda_\mathrm{f}}}\left(\hat{p}_{0,l_\mathrm{f},m}, \alpha_{\mathrm{f},l_\mathrm{f}}\right)$	The regularized PCC (PCCR) between $\hat{p}_{0,l_\mathrm{f},m}$ and $\alpha_{\mathrm{f},l_\mathrm{f}}$ (Eq. (96))
λ_f	The regularization parameter in PCCR

Table 4 Additional symbols for intra-image nonrigid motion correction.

Symbol	Definition	Value
I_cor (i_cor)	The number of motion corrections (index)	$I_\mathrm{cor} = 16$ for numerical simulations and *in vivo* experiments
M_sub (m_sub)	The number of image subdomains for motion correction (index)	$M_\mathrm{sub} = 3 \times 3 \times 2$ for numerical simulations $M_\mathrm{sub} = 6 \times 6 \times 4$ for *in vivo* experiments
a_x, a_y, and a_z (m)	Translation step sizes of each subdomain along the x-axis, y-axis, and z-axis, respectively	$\left(a_x, a_y, a_z\right) =$ (0.1,0.1,0.1) (mm) for numerical simulations and *in vivo* experiments
K_x, K_y, and K_z	Searching step ranges along the x-axis, y-axis, and z-axis, respectively, in motion correction	$\left(K_x, K_y, K_z\right) = (2,2,2)$ for numerical simulations $\left(K_x, K_y, K_z\right) = (1,1,2)$ for *in vivo* experiments
A_tra (m)	The amplitude of the translation distance in translation-induced motions (Eq. (115))	$A_\mathrm{tra} =$ $0.2, 0.4, 0.6, 0.8, 1.0, 1.2$ (mm) for translation-induced motions
A_def	The amplitude of the	$A_\mathrm{def} =$

	deformation ratio in deformation-induced motions (Eq. (116))	$0.02, 0.04, 0.06, 0.08, 0.10, 0.12$ for deformation-induced motions
N_{per}	The number of motion periods during a 90° rotation of the four-arc array (Eqs. (115) and (116))	$N_{\mathrm{per}} = 0.5, 1, 2, 3$ for translation- and deformation-induced motions
$D_{m_{\mathrm{sub}}}$	The m_{sub}-th subdomain	
$\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}$ (m)	The center of $D_{m_{\mathrm{sub}}}$	
$\mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}}$ (m)	The center of the detection surface of the n_{ele}-th transducer element when the four-arc array is at the n_{loc}-th location	
$\zeta(\mathbf{r}', \mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}})$ (s)	The motion-induced temporal shift of the signals from the source point at $\mathbf{r}'$ that are detected by the n_{ele}-th element when the transducer array is at the n_{loc}-th location	
$\hat{p}_{\zeta}(\mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}}, t)$ (Pa)	Detected signals with motion-induced temporal shifts $\zeta(\mathbf{r}', \mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}})$ (Eq. (99))	
$k_x(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}})$, $k_y(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}})$, and $k_z(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}})$	Translation step numbers of the m_{sub}-th subdomain along the x-axis, y-axis, and z-axis, respectively, when the transducer array is at the n_{loc}-th location	
$\mathbf{k}(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}})$	The translation vector defined as $\left(k_x(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}), k_y(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}), k_z(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}})\right)$	
$\mathbf{k}$	The translation set defined as $\{\mathbf{k}(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}}) \in \mathbb{Z}^3 \mid m_{\mathrm{sub}} = 1,2,\dots,M_{\mathrm{sub}}, n_{\mathrm{loc}} = 1,2,\dots,N_{\mathrm{loc}}\}$	
$\zeta_{\mathbf{k}}(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, \mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}})$ (s)	Temporal shifts caused by the translations in $\mathbf{k}$ (Eq. (101))	
$\zeta_{\mathbf{k}}(\mathbf{r}'_m, \mathbf{r}_n)$ (s)	Temporal shifts obtained from $\tau_{\mathbf{k}}(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, \mathbf{r}_{n_{\mathrm{loc}},n_{\mathrm{ele}}})$ through	

	spatial interpolation
$\hat{p}_{\mathbf{k},n,l}$ (Pa)	Detected signals with motion described by $\mathbf{k}$ (Eq. (102))
$\mathbf{H}_{\mathbf{k}}$	The motion-incorporated system matrix implemented in Eq. (102)
$\hat{k}_x(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}})$, $\hat{k}_y(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}})$, and $\hat{k}_z(\mathbf{r}'_{\mathrm{c},m_{\mathrm{sub}}}, n_{\mathrm{loc}})$	Translation step numbers obtained using Eqs. (111), (112), and (113), respectively
$\hat{\mathbf{k}}$	Reconstructed translation vectors for all subdomains and array locations
$z_{\mathrm{tra},n_{\mathrm{loc}}}$	The translation distance defined in Eq. (115)
$\alpha_{\mathrm{def},n_{\mathrm{loc}}}$	The deformation ratio defined in Eq. (116)