Advancing X-ray Multi-Projection Imaging

Towards Four-Dimensional Reconstruction and Analysis of Rapid Flows

Yao, Zisheng

2026

Document Version:
Publisher's PDF, also known as Version of record

Advancing X-ray Multi-Projection Imaging

Towards Four-Dimensional Reconstruction and Analysis of Rapid Flows

ZISHENG YAO
DEPARTMENT OF PHYSICS | FACULTY OF SCIENCE | LUND UNIVERSITY

Advancing X-ray Multi-Projection Imaging

Advancing X-ray Multi-Projection Imaging

Towards Four-Dimensional Reconstruction and Analysis of Rapid Flows

Zisheng Yao

LUND
UNIVERSITY

DOCTORAL DISSERTATION

To be presented, with the permission of the Faculty of Science, Lund University, Sweden, for public criticism in the Rydberg Lecture Hall at the Department of Physics.

Date and time: 2026-05-22 at 9:15

Faculty opponent
Prof. Dr. Matthieu N. Boone
Ghent University, Ghent, Belgium

Organization
LUND UNIVERSITY

Department of Physics
Box 118
SE–221 00 LUND
Sweden

Document name
DOCTORAL DISSERTATION

Date of disputation
2026-05-22

Sponsoring organization

Author(s)
Zisheng Yao

Title and subtitle
Advancing X-ray Multi-Projection Imaging: Towards Four-Dimensional Reconstruction and Analysis of Rapid Flows

Abstract

Fourth-generation synchrotron light sources deliver flux densities up to two orders of magnitude higher than previous generations, enabling X-ray imaging at unprecedented spatiotemporal resolution. This is attractive for non-destructive 4D (3D + time) studies of fast dynamics under in-situ or operando conditions. State-of-the-art 4D X-ray imaging at large-scale facilities relies on tomography and can achieve 1000 volume reconstructions per second at micrometer resolution. However, the requirement of rapid sample rotation induces large centrifugal forces, limiting studies of rotation-sensitive dynamics. To address this limitation, X-ray Multi-Projection Imaging (XMPI) has emerged as an alternative rotation-free approach. By splitting the primary beam into several angularly separated beams, XMPI enables simultaneous multi-angle acquisitions, enabling ultrafast imaging but making 4D reconstruction and analysis challenging in the sparse-view regime.

In this thesis, we first introduce the basics of X-ray imaging. Second, we present the implementation of the XMPI setup, enabled by the unique capabilities of fourth-generation synchrotron light sources. Third, we present the development of 3D particle tracking velocimetry with XMPI, and demonstrate its capability to quantitatively retrieve flow properties by tracking microscale tracers via triangulation. Lastly, we focus on a more ambitious goal: to develop full-field deep-learning-based 4D reconstruction methods for XMPI, including droplet collisions and flows through porous media, two representative fluid-dynamics examples for validation. This thesis demonstrates that XMPI, together with advanced 4D reconstruction and analysis approaches, becomes an enabling tool for fluid dynamics studies, providing sub-10 micrometer and kHz-to-MHz spatiotemporal resolution.

Key words
X-ray imaging, four-dimensional, deep learning, sparse-view reconstruction, ultrafast dynamics

Classification system and/or index terms (if any)

Supplementary bibliographical information

Language
English

ISSN and key title

ISBN
978-91-8104-771-4 (print)
978-91-8104-772-1 (pdf)

Recipient's notes

Number of pages
248

Price

Security classification

Signature _____ Date 2026-04-13

Advancing X-ray Multi-Projection Imaging

Towards Four-Dimensional Reconstruction and Analysis of Rapid Flows

by Zisheng Yao

LUND
UNIVERSITY

A doctoral thesis at a university in Sweden takes either the form of a single, cohesive research study (monograph) or a summary of research papers (compilation thesis), which the doctoral student has written alone or together with one or several other author(s).

In the latter case the thesis consists of two parts. An introductory text puts the research work into context and summarizes the main points of the papers. Then, the research publications themselves are reproduced, together with a description of the individual contributions of the authors. The research papers may either have been already published or are manuscripts at various stages (in press, submitted, or in draft).

Cover image front: A "table-tennis-style" illustration of X-ray multi-projection imaging, accompanied by a photo showing the first successful implementation of the XMPI setup at the ForMAX beamline, MAX IV.

Cover image back: This is a picture captured by an Axis camera in the experimental hutch at the ForMAX beamline, MAX IV, which vividly shows the high-flux X-ray beam provided by a modern large-scale X-ray facility.

Faculty of Science, Department of Physics

isbn: 978-91-8104-771-4 (print)
isbn: 978-91-8104-772-1 (pdf)

Printed in Sweden by Media-Tryck, Lund University, Lund 2026

"There is only one heroism in the world: to see the world as it is, and to love it."
—— Romain Rolland

Contents

List of Publications iii
Acknowledgements vii
Popular Summary ix
Abbreviations xi

1 Introduction **1**

2 X-ray Imaging: Principles and State-of-the-Art **5**
2.1 X-ray Interaction with Matter 6
2.1.1 Complex Index of Refraction 6
2.1.2 Projection Approximation and Beer–Lambert's Law 7
2.1.3 Imaging Regimes 9
2.2 3D X-ray Imaging - Tomography 11
2.2.1 Reconstruction Methods 11
2.2.2 Crowther Criterion 13
2.2.3 "Missing Wedge" Issue 14
2.3 Extend to 4D - Opportunities and Challenges 15

3 Synchrotron Light Sources and X-ray Multi-Projection Imaging (XMPI) **17**
3.1 Synchrotron Light Source 17
3.1.1 Higher Brilliance and Flux 18
3.1.2 ForMAX Beamline and its Imaging Capabilities 20
3.2 XMPI Setup in a Nutshell 24
3.2.1 Beam-Splitting System 25
3.2.2 X-ray Detector System 27
3.3 XMPI - Opportunities and Challenges 28

4 4D Analysis of Flow Dynamics via Particle Tracking with XMPI **31**
4.1 Particle Tracking Velocimetry 32
4.1.1 Experimental Setup 32
4.1.2 Particle-Tracking Pipeline 33
4.2 Particle Tracking Results 36
4.2.1 Tracking Individual Particles 36
4.2.2 Error Analysis 38

4.2.3 Statistical Behavior of Tracked Particles 41
4.3 Opportunities and Challenges 42

5 AI-Assisted 4D Reconstruction from Sparse Spatiotemporal Data 45
5.1 Overview of Deep Learning 46
5.1.1 Basic Concepts in Deep Learning 46
5.1.2 Neural Network Architectures 47
5.2 "AI + Physics" 4D Reconstruction Framework 52
5.2.1 4D-ONIX 53
5.2.2 4D-PIONIX 54
5.2.3 X-Hexplane 55
5.3 Applications 56
5.3.1 Droplet Collision 57
5.3.2 Flows in Porous Media 61

6 Conclusion and Outlook 67
6.1 Conclusion 67
6.2 Outlook 68

References 71

Scientific Publications 83
Author Contributions 83
Paper I: New opportunities for time-resolved imaging using diffraction-limited storage rings 85
Paper II: Time-resolved 3D imaging opportunities with XMPI at ForMAX 97
Paper III: 4D-ONIX for reconstructing 3D movies from sparse X-ray projections via deep learning 111
Paper IV: Physics-informed 4D x-ray image reconstruction from ultra-sparse spatiotemporal data 137
Paper V: Super Time-Resolved Tomography 149
Paper VI: Synchrotron X-ray Multi-Projection Imaging (XMPI) for High-Resolution 4D Characterization of Multiphase Flows 165
Paper VII: 4D Synchrotron X-Ray Multi Projection Imaging (XMPI) for studying multiphase flow dynamics and flow instabilities in porous network 197

List of Publications

This thesis is based on the following publications, referred to by their Roman numerals:

I **New opportunities for time-resolved imaging using diffraction-limited storage rings**
Z. Yao*, J. Rogalinski*, E. M. Asimakopoulou, Y. Zhang, K. Gordeyeva, Z. Atoufi, H. Dierks, S. McDonald, S. Hall, J. Wallentin, D. Söderberg, K. Nygård, P. Villanueva-Perez
Journal of Synchrotron Radiation, 2024, 31(5), pp. 1299-1307

II **Time-resolved 3D imaging opportunities with XMPI at ForMAX**

J. Rogalinski, **Z. Yao**, Y. Zhang, Z. Hu, K. Gordeyeva, T. Rosén, D. Söderberg, A. Mazzolari, J. Silva, V. Haghighat, S. McDonald, K. Nygård, E. M. Asimakopoulou, P. Villanueva-Perez
Journal of Synchrotron Radiation, 2026, 33(2), pp. 417-428

III **4D-ONIX for reconstructing 3D movies from sparse X-ray projections via deep learning**
Y. Zhang, **Z. Yao**, R. Klöfkorn, T. Ritschel, P. Villanueva-Perez
Communications Engineering, 2025, 4(1), pp. 1-12

IV **Physics-informed 4D x-ray image reconstruction from ultra-sparse spatiotemporal data**
Z. Yao, Y. Zhang, Z. Hu, R. Klöfkorn, T. Ritschel, P. Villanueva-Perez
Measurement Science and Technology, 2025, 36(8), pp. 085403

V **Super Time-Resolved Tomography**

Z. Hu, **Z. Yao**, K. Josefsson, F. García-Moreno, M. Makowska, Y. Zhang, P. Villanueva-Perez
Advanced Science, 2026, 13(3), e11933

VI **Synchrotron X-ray Multi-Projection Imaging (XMPI) for High-Resolution 4D Characterization of Multiphase Flows**

T. Rosén*, **Z. Yao***, J. Tejbo, P. Wegele, J. Rogalinski, F. Nilsson, K. Mom, Z. Hu, S. McDonald, K. Nygård, A. Mazzolari, A. Groetsch, K. Gordeyeva, D. Söderberg, F. Lundell, L. P. Wittberg, E. M. Asimakopoulou, P. Villanueva-Perez*
arXiv preprint, 2026, arXiv:2412.09368v2

VII **4D Synchrotron X-Ray Multi Projection Imaging (XMPI) for studying multiphase flow dynamics and flow instabilities in porous network**
P. Wegele*, **Z. Yao***, J. Tejbo, J. Rogalinski, Z. Hu, Y. Zhang, E. Oliaei, S. Davoodi, T. Rosén, A. Groetsch, K. Nygård, E. M. Asimakopoulou, P. Villanueva-Perez, D. Söderberg*
arXiv preprint, 2026, arXiv:2603.15319

All papers are reproduced with permission from their respective publishers.

Publications not included in this thesis, to which I contributed:

VIII **ONIX: an X-ray deep-learning tool for 3D reconstructions from sparse views**
Y. Zhang, **Z. Yao**, T. Ritschel, P. Villanueva-Perez
Applied Research, 2023, 2(4), e202300016

IX **Megahertz X-ray multi-projection imaging**
P. Villanueva-Perez, V. Bellucci, Y. Zhang, S. Birnsteinova, R. Graceffa, L. Adriano, E. M. Asimakopoulou, I. Petrov, **Z. Yao**, M. Romagnoni, A. Mazzolari, R. Letrun, C. Kim, J. Koliyadu, C. Deiter, R. Bean, G. Giovanetti, L. Gelisio, T. Ritschel, A. Mancuso, H. N. Chapman, A. Meents, T. Sato, P. Vagovic
arXiv preprint, 2023, arXiv:2305.11920

X **Online dynamic flat-field correction for MHz microscopy data at European XFEL**
S. Birnsteinova, D. E. Ferreira de Lima, E. Sobolev, H. J. Kirkwood, V. Bellucci, R. Bean, C. Kim, J. Koliyadu, T. Sato, F. Dall'Antonia, E. M. Asimakopoulou, **Z. Yao**, K. Buakor, Y. Zhang, A. Meents, H. N. Chapman, A. Mancuso, P. Villanueva-Perez, P. Vagovic
Journal of Synchrotron Radiation, 2023, 30(6), pp. 1030-1037

XI **Development towards high-resolution kHz-speed rotation-free volumetric imaging**
E. M. Asimakopoulou, V. Bellucci, S. Birnsteinova, **Z. Yao**, Y. Zhang, I. Petrov, C. Deiter, A. Mazzolari, M. Romagnoni, D. Korytar, Z. Zaprazny, Z. Kuglerova, L. Juha, B. Lukic, A. Rack, L. Samoylova, F. García-Moreno, S. Hall, T. Neu, X. Liang, P. Vagovic, P. Villanueva-Perez
Optics Express, 2024, 32(3), pp. 4413-4426

XII **Development of crystal optics for X-ray multi-projection imaging for synchrotron and XFEL sources**
V. Bellucci, S. Birnsteinova, T. Sato, R. Letrun, J. Koliyadu, C. Kim, G. Giovanetti, C. Deiter, L. Samoylova, I. Petrov, L. L. Morillo, R. Graceffa, L. Adriano, H. Huelsen, H. Kollmann, T. N. T. Calliste, D. Korytar, Z. Zaprazny, A. Mazzolari, M. Romagnoni, E. M. Asimakopoulou, **Z. Yao**, Y. Zhang, J. Ulicny, A. Meents, H. N. Chapman, R. Bean, A. Mancuso, P. Villanueva-Perez, P. Vagovic
Journal of Synchrotron Radiation, 2024, 31(6), pp. 1534-1550

XIII **ForMAX–a beamline for multiscale and multimodal structural characterization of hierarchical materials**
K. Nygård, S. McDonald, J. González, V. Haghighat, C. Appel, E. Larsson, R. Ghanbari, M. Viljanen, J. Silva, S. Malki, Y. Li, V. Silva, C. Weninger, F. Engelmann, T. Jeppsson, G. Felcsuti, T. Rosén, K. Gordeyeva, D. Söderberg, H. Dierks, Y. Zhang, **Z. Yao**, R. Yang, E. M. Asimakopoulou, J. Rogalinski, J. Wallentin, P. Villanueva-Perez, R. Krüger, T. Dreier, M. Bech, M. Liebi, M. Bek, R. Kádár, A. E. Terry, H. Tarawneh, P. Ilinski, J. Malmqvista, Y. Cereniusa
Journal of Synchrotron Radiation, 2024, 31(2), pp. 363-377

XIV **Time-resolved X-ray radiography of through-thickness liquid transport in partly saturated needle-punched nonwovens**
P. Wegele, **Z. Yao**, J. Tejbo, J. Rogalinski, T. Rosén, A. Groetsch, K. Nygård, E. M. Asimakopoulou, P. Villanueva-Perez, D. Söderberg
arXiv preprint, 2026, arXiv:2602.15176

XV **Transverse cracking in glass fibre-reinforced composites monitored with synchrotron X-ray multi-projection imaging**
E. V. Vlierberghe, J. Soete, E. M. Asimakopoulou, **Z. Yao**, J. Rogalinski, Z. Hu, K. Mom, B. Lukić, C. Breite, P. Villanueva-Perez, Y. Swolfs

e-Journal of Nondestructive Testing, 2025, 30(2)

XVI **Constraint-Free Coherent Diffraction Imaging via Physics-Guided Neural Fields**
Z. Hu, **Z. Yao**, Y. Zhang, P. Villanueva-Perez
arXiv preprint, 2025, arXiv:2512.09694

XVII **4D reconstruction of alumina laser melt pools at 25 kHz via operando X-ray multi-projection imaging**
L. Witte, E. Jermann, Z. Hu, **Z. Yao**, E. M. Asimakopoulou, J. Rogalinski, Y. Zhang, K. Nygård, M. G. Makowska, M. Bambach, M. Afrasiabi, P. Villanueva-Perez
arXiv preprint, 2026, arXiv:2603.14391

Acknowledgements

In a table tennis match, players usually call a time-out to sort out their strategy before a decisive point. Analogously, writing this thesis has served as a time-out for me. It provides not only an opportunity to reflect on my PhD journey and the works that lie ahead, but also, more importantly, a chance to acknowledge how fortunate I am to be surrounded by such remarkable teammates, collaborators, friends and family. This thesis work would not have been possible without their support.

First of all, I would express my gratitude to my main supervisor, Pablo Villanueva-Pérez. Thank you for opening the door for me to step into this fascinating and challenging XMPI project. Your guidance and dedication to science has deeply influenced me, far beyond the scope of this thesis work. I would also like to express my sincere gratitude to Tobias Ritschel for all the stimulating brainstorming and discussions, from which I learned how AI can approach scientific problems and how to think more creatively and critically. Many thanks also go to my co-supervisor, Jesper Wallentin, for teaching me modern X-ray physics and for providing helpful suggestions throughout my PhD journey, especially on scientific writing.

Great collaborations are the key to this thesis work. I can hardly imagine what my thesis will look like without your help. I want to thank the team of Robert Klöfkorn for your massive efforts on the droplet collision simulations and all our meetings and after-hours. I also want to thank the teams from KTH, KU Leuven, HZB, ESRF, PSI and ETH Zürich for all the cool experiments. Thank you, Tomas Rosén, Patrick Wegele, Malgorzata Makowska, Eliot Jermann, Valerio Bellucci, Sarlota Birnsteinova, Andrea Mazzolari, Elise Van Vlierberghe, Jeroen Soete, Christian Breite, Yentl Swolfs, Jonas Tejbo, Alexander Groetsch, Korneliya Gordeyeva, Tijana Todorovic, Saeed Davoodi, Lisa Prahl Wittberg, Daniel Söderberg, Francisco García-Moreno, Paul Kamm, Stephen Hall, Robin Krüger and Martin Bech for our exciting beamtimes and collaboration in various applications. Thank you, Kannara Mom and Lars Witte, for our great time working together in Lund. Thank you, Emanuel Larsson, for your wonderful demo of the "KBLT" pipeline and for the support of visualization. Special thanks go to Zdenek Matej for the constant support regarding the computing clusters at MAX IV.

SLJUS has been a wonderful workplace with a full of great colleagues and a welcoming atmosphere. I would like to thank all my current and former colleagues. I would say thank you to Yuhe, Myrto, Julia, Runqing, and Zhe for the routine but valuable weekly discussions on X-ray imaging, and all the fun we had during our coffee breaks. Thank you, my office mates, Stefano and Andrea, for making our office an ideal place to work and for all our daily chatting, especially when we felt

tired. Thank you, Ahmed, Alfred, Annika, Auden, Claire, Eleni, Emma, Gustavo, Hajar, Hanna, Huaiyu, Jeevan, Lingjia, Mehran, Nelia, Nils, Nishant, Oskar, Rohit, Shilpa, Smita, Ulrike, Vidar, Yen-Po, Yi, Zesen, Zhaojun, Ziyun...every SLJUSer for all the great times we have shared together. Thank you, Elizabeth, for your guidance and encouragement during the ISP meetings. Thank you, Patrik and Josefine, for solving millions of issues that I have no idea about.

ForMAX felt like a second home to me, as most of the beamtimes during my PhD journey were conducted there. I would like to express my gratitude and appreciation to Kim, Sam, Vahid, Jackson, Mira, and Myrto (again!) for all the professional support throughout the years. It is the effort of the entire ForMAX team that makes our beamtimes there fruitful. The champagne celebration after a successful beamtime will always be my golden memory, and I fully believe that more exciting works are on the way.

On a more personal level, I would like to thank all the people who have brought joy to my life outside of the PhD work, although it is not possible to name everyone here. Thank you, the "Parternas Gränd Yoga Center" group, including Yue, Huaiyu and Kailun, for all our unforgettable gatherings, the nice food, the funny board games... everything but Yoga. Thank you, Yuchen, Hao, Man, Tianyi, Shuting, and Bo, for helping me in all aspects of my life, especially the cooking skills. Thank you, Ziyun and Jin, for our shared moments immersed in nature, especially during the wonderful "Tour du Mont Blanc". Thank you, the "407" Group from our high school days, for the long-standing companionship over the years, even though we now find ourselves on three different continents most of the time. Thank you, my coaches, teammates and friends from IFK Lund Bordtennis, for helping me gain back the pleasure of playing table tennis and feel more energetic.

Finally, I would like to express my love and gratitude to my family for their unconditional love and support throughout the years. I want to dedicate this work to the memory of my father. Your love and values continue to guide me and shape who I am today. Thank you, my mom, for your constant care, encouragement and trust through every stage of my life. I am deeply proud of being your son. Thank you, my wife, for all the moments we have shared together in our PhD journeys, from walking along the sea to code debugging in front of the screen. I look forward to our journey ahead.

Popular Summary

When talking about X-ray imaging, the first image one can think about is probably the famous "Hand mit Ringen", the first radiograph taken by Wilhelm Röntgen in 1895. This image revealed, for the first time, the invisible interior of the human body without surgery and marked the birth of X-ray imaging. More than a century later, X-rays remain an indispensable tool for exploring the hidden inner structures of objects.

Today, scientists are not only interested in static images, but also in how objects change over time. Interesting dynamics include, for example, how droplets collide with each other, how fluids flow through given materials, and how materials fracture or deform under stress. Capturing and understanding such processes requires four-dimensional (4D) imaging, where time is added to the three-dimensional (3D) space. However, conventional 4D X-ray imaging relies on rapidly rotating the sample to collect many viewing angles. This requirement creates two fundamental problems. First, many systems of interest, such as liquids and ultra-fast dynamic processes, cannot tolerate rapid rotation. Second, the dynamic processes that occur faster than the rotation speed cannot be resolved.

This thesis explores a new solution, called X-ray Multi-Projection Imaging (XMPI), to address the aforementioned challenges. Instead of rapidly rotating the sample, XMPI splits a single X-ray beam into several smaller beams that illuminate the sample simultaneously from different directions. In this way, multiple X-ray images are recorded at exactly the same moment. Using the extremely bright X-rays available at modern synchrotron facilities, this approach enables rotation-free imaging of fast 3D dynamics with micrometer-scale resolution.

Recording only a few images at each moment, however, creates a new challenge. Traditional 3D reconstruction and analysis pipelines no longer work because there is simply not enough information from different angles. This thesis addresses this problem in two complementary ways.

First, this thesis shows that even without reconstructing the entire 3D sample, valuable physical information can be extracted. For example, by observing small tracer particles moving with the flow from as few as two directions, their 3D motion can be determined accurately via triangulation. Second, this thesis develops new artificial intelligence (AI)–based reconstruction methods that combine the power of physics and recent AI developments. These methods learn to reconstruct entire 3D movies directly from very limited X-ray data, while respecting the physical laws governing the observed dynamics. Applications to droplet collisions and fluid flow in porous materials demonstrate that meaningful 4D reconstructions can be achieved far beyond

the limits of conventional techniques.

In summary, this work establishes XMPI as a powerful new approach for studying ultrafast dynamic systems that cannot be imaged using traditional rotation-based methods. By combining the experimental design based on synchrotron light sources with physics-aware AI-assisted reconstruction and analysis, it opens new possibilities for observing and understanding complex dynamic processes hidden inside the studied system.

Abbreviations

2D	two-dimensional
3D	three-dimensional
4D	four-dimensional
XMPI	X-ray multi-projection imaging
FBP	filtered back-projection
SART	simultaneous algebraic reconstruction technique
SIRT	simultaneous iterative reconstruction technique
FRC	Fourier ring correlation
FSC	Fourier shell correlation
tps	tomograms per second
DLSRs	diffraction-limited storage rings
PTV	particle tracking velocimetry
PIV	particle image velocimetry
OF	optical flow
SHGS	silver-coated hollow glass spheres
AI	artificial intelligence
FCNN	fully-connected neural network
CNN	convolutional neural network
GAN	generative adversarial network
PINN	physics-informed neural network
AD	automatic differentiation
NIR	neural implicit representation
PDE	partial differential equation
MSE	mean squared error

DSSIM data structural similarity index measure

ONIX optimized neural implicit X-ray imaging

Chapter 1

Introduction

X-rays, since their discovery in 1895 by Wilhelm Röntgen, have become a powerful probe for imaging in various fields. X-rays are characterized by their short wavelengths and high penetration capabilities. The short wavelengths allow X-rays to interact with matter on sub-micrometer or even atomic length scales, making it possible to resolve structural features that are fundamentally inaccessible to visible-light techniques due to the diffraction limit. Moreover, the high penetration power makes it possible to probe the internal structures of complex systems in a non-destructive manner.

The first reported X-ray image, "Hand mit Ringen" [1], demonstrated such capabilities by revealing the structure of a human hand adorned with a ring, marking the birth of modern X-ray imaging. Since then, continuous development in X-ray generation and detection methods has driven a boost in X-ray imaging, from early laboratory X-ray tubes to large-scale synchrotron radiation facilities. These modern X-ray sources provide dramatically increased photon flux density, laying the foundation for both high-resolution and high-speed imaging.

As X-ray sources evolved, so did the dimensionality of imaging. As a first step, imaging methodologies have expanded from two-dimensional (2D) radiography to three-dimensional (3D) tomographic imaging. X-ray tomography reconstructs volumetric information by combining multiple angular projections, providing access to internal structures that cannot be resolved in a single 2D X-ray image. However, many physical systems of interest are inherently dynamic, and static 3D reconstructions are insufficient to fully describe their behavior. This limitation has motivated the second-step development, i.e., extending to four-dimensional (4D) (3D + time) X-ray imaging to capture the temporal evolution of the 3D structures.

Despite its conceptual simplicity, 4D X-ray imaging encounters significant experi-

mental challenges. The most established approach at large-scale facilities, time-resolved tomography, relies on repeatedly rotating the sample while continuously acquiring 2D projection images. In most of the cases, the achievable temporal resolution is fundamentally limited by the mechanical rotation speed of the sample stage. Moreover, rapid rotation can introduce high centrifugal forces that alter the dynamics or even damage the sample, making this approach unsuitable for many rotation-sensitive systems such as liquids, soft matter, and multiphase flows. In some special cases, e.g., when learning periodic or reproducible dynamics, gated tomography[2] can be employed to avoid fast sample rotation, where projection images are acquired over multiple repeated cycles and then sorted retrospectively. However, this strategy is not applicable to non-reproducible or stochastic dynamics.

To address these limitations, X-ray multi-projection imaging (XMPI) has emerged as a promising rotation-free alternative for 4D imaging. Instead of rotating the sample, XMPI splits a single incident X-ray beam into several angularly separated beamlets that illuminate the sample simultaneously. By recording multiple projections in a single exposure, XMPI enables instantaneous multi-view acquisition, resulting in a temporal resolution determined primarily by the X-ray detector's acquisition rate.

However, the fundamental advantage of XMPI, its ability to acquire multiple projections simultaneously, also introduces a major challenge. In contrast to tomography, which typically utilizes hundreds of 2D projections to extract volumetric information, XMPI is inherently a sparse-view (≤ 3) imaging technique. Recovering volumetric information from sparse projections can be challenging, as directly applying classical approaches suitable for tomography-based setups is usually improper. In other words, even after successfully establishing the XMPI setup, the data interpretation is still a bottleneck in 4D imaging, which, in general, requires strategies that can effectively incorporate additional prior knowledge to compensate for the lack of angular information.

The central objective of this thesis is to address the challenges in 4D reconstruction and analysis, which are crucial to advancing XMPI as a practical and robust tool for high-speed 4D X-ray imaging. The major chapters of this thesis are organized as follows:

Chapter 2 introduces the basics of X-ray imaging, and discusses opportunities and challenges of 4D X-ray imaging relevant to the main content of this thesis.

Chapter 3 introduces the implementation of the XMPI setup enabled by the unique capabilities provided by synchrotron light sources. This chapter is supported by **Papers I and II**.

Chapter 4 introduces triangulation-based particle tracking as a first application exam-

ple, demonstrating the quantitative capabilities of XMPI. This chapter is supported by **Paper VI**.

Chapter 5 focuses on a more ambitious goal: to develop full-field artificial intelligence (AI)-assisted 4D reconstruction methods for XMPI. This chapter also includes representative fluid dynamics examples for validation, supported by **Papers III, IV, V and VII**.

Chapter 2

X-ray Imaging: Principles and State-of-the-Art

The past century has witnessed a boost in X-ray research. Nowadays, X-rays are no longer the "unknown" type of radiation as they were originally called. In essence, as shown in Figure 2.1, X-rays are electromagnetic waves lying on the high-energy side of the electromagnetic spectrum between ultraviolet and gamma rays.

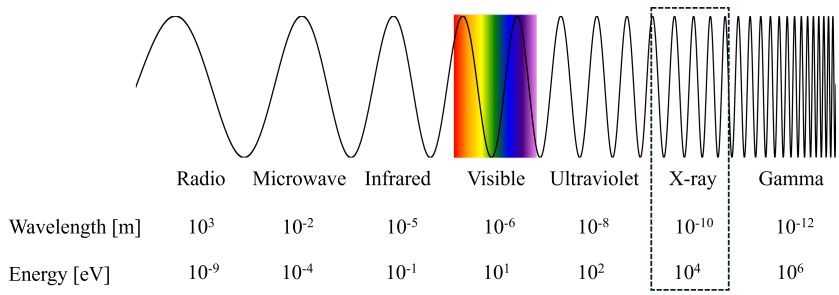


Figure 2.1: The electromagnetic spectrum.

X-rays also follow the basic principles of quantum mechanics, where particle-wave duality holds. From the particle perspective, the energy of the electromagnetic wave is carried by photons. The energy of each photon is given by Planck's energy-frequency relation:

$$E = h\nu = \frac{hc}{\lambda} \approx \frac{12.4}{\lambda\,[\text{Å}]}\,[\text{keV}], \tag{2.1}$$

where h denotes the Planck's constant, ν denotes the frequency, c denotes the speed of light in vacuum and λ denotes the wavelength. Conventionally, in X-ray studies, we describe the photon energy using the unit of keV. It can be easily calculated

using the wavelength in the unit of Å (10^{-10} m), as shown in the right-hand side of Equation 2.1.

The wavelength range for X-rays is typically 0.01 nm to 10 nm. Specifically, X-rays with a wavelength above 1.24 nm are called soft X-rays; below 0.25 nm are called hard X-rays[3]. Such a short wavelength, together with its unique penetration ability, enables non-invasive investigation of microscopic structures, even down to the atomic level. In this thesis work, we will use hard X-rays as a promising probe for X-ray imaging studies.

In this chapter, we will start with some key concepts in X-ray imaging relevant to this thesis work to help readers understand the formalism of 2D X-ray images. Then, we will extend X-ray imaging to higher dimensions and briefly discuss how the state-of-the-art 3D and 4D X-ray imaging method works. We will finish this chapter by discussing the opportunities and challenges of 4D X-ray imaging, which serves as the motivation for the main content of this thesis.

2.1 X-ray Interaction with Matter

Imaging, in a broad sense, refers to the process of creating a visual representation of an object, structure, or phenomenon by mapping its physical properties onto a spatial or spatiotemporal domain. Proper interpretation of such visual representations requires a clear understanding of the imaging probe itself. Hence, for X-ray imaging studies, a fundamental prerequisite is to understand how X-rays interact with matter and how X-ray images are formed. In this section, unlike most of the textbooks on X-ray physics, I will skip the atomistic description of X-ray–matter interactions. Instead, in a more practical way, I will introduce directly what happens when X-rays pass through an object in vacuum, providing insight into the formalism of 2D X-ray imaging. For a more thorough discussion on these topics, readers are kindly referred to relevant textbooks[4,5] on X-ray imaging.

2.1.1 Complex Index of Refraction

Generally, after X-rays pass through an object in vacuum, the X-ray wavefield is modulated by the object. Specifically, one can observe two major differences when comparing the exit wave and the incoming wave, as shown in Figure 2.2. First, the amplitude of the wave is attenuated. Second, the phase of the wave is shifted.

To quantify these two differences, we need to introduce the concept of the index of refraction. It not only describes the property of the material, but also helps describe

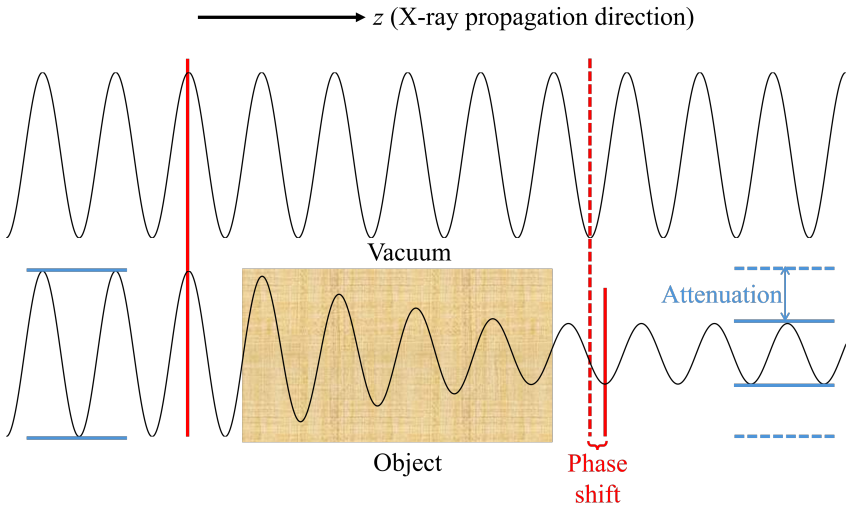


Figure 2.2: Two major differences after X-rays pass through an object in vacuum: attenuation and phase shift.

how an incident X-ray wavefield is modulated by the object. In X-ray physics, the index of refraction is depicted in a complex number, as shown in Equation 2.2:

$$n = 1 - \delta + i\beta, \tag{2.2}$$

where δ and β are positive real numbers and significantly smaller than unity. The real part describes the elastic scattering related to the phase change, while the imaginary part describes the absorption in the material related to the attenuation. As an example, we have $\delta = 2.31 \times 10^{-6}$ and $\beta = 5.01 \times 10^{-9}$ for water (H_2O) at 10 keV.

2.1.2 Projection Approximation and Beer–Lambert's Law

In this subsection, we will quantify how the X-ray wavefield is modulated after passing through an object in vacuum using the example sketched in Figure 2.3. Here, X-rays are depicted as a time-independent plane wave traveling along the z axis before passing through the object:

$$\psi_{\text{in}} = Ae^{ikz}, \tag{2.3}$$

where A denotes the amplitude and k $(= 2\pi/\lambda)$ denotes the wavenumber. Such a scalar form shown in Equation 2.3 is commonly used for imaging studies, as it simplifies the representation of X-rays by ignoring the polarization effects[6].

We further impose the projection approximation, which assumes that the direction of the wave within the sample is not changed compared to the incoming wave. Such

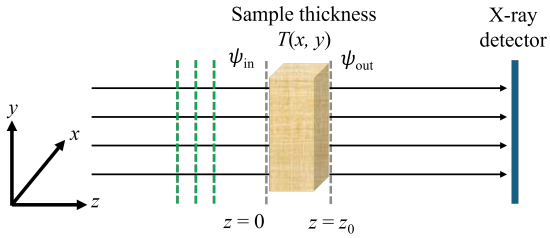


Figure 2.3: X-rays pass through an object in vacuum. Figure adapted from Ref. 7.

an approximation holds under weak interactions when multiple scattering and the diffraction effects within the object can be neglected[8]. Hence, after passing through the object whose index of refraction is a function of the 3D space coordinate (x, y, z), the exit wavefield can be expressed as follows:

$$\begin{aligned}\psi_{\text{out}} &= \psi_{\text{in}} \cdot \exp\left(\int_0^{z_0} ikn(x,y,z)d\tilde{z}\right) \\ &= \psi_{\text{in}} \cdot \exp\left(\int_0^{z_0} ik\left(1-\delta(x,y,z)+i\beta(x,y,z)\right)d\tilde{z}\right) \\ &= \psi_{\text{in}} \cdot \exp\left(-k\int_0^{z_0}\beta(x,y,z)d\tilde{z}\right)\cdot\exp(ikz_0)\cdot\exp\left(-ik\int_0^{z_0}\delta(x,y,z)d\tilde{z}\right),\end{aligned} \tag{2.4}$$

where $\tilde{z}$ denotes the integration variable along the propagation direction.

As shown in Equation 2.4, after passing through the object in vacuum, both amplitude attenuation and phase shift exist in the exit wavefield. In practice, however, the quantity that an X-ray detector can directly provide is the intensity of the wave, which is proportional to the time-averaged squared modulus of the complex wavefield. Phase information cannot be measured directly since no existing X-ray detector can resolve the rapid oscillations of the electromagnetic field, so phase information is only available after using phase-retrieval methods[8].

Given that the measured intensity immediately after the object is directly proportional to the square of its amplitude, we have:

$$I_{\text{out}} = I_{\text{in}} \cdot \exp\left(-2k\int_0^{z_0}\beta(x,y,z)d\tilde{z}\right), \tag{2.5}$$

where $I_{\text{in}} = A^2$ (according to Equation 2.3). Equation 2.5 forms the theoretical basis of absorption-contrast imaging. Regions with larger β values along the propagation direction result in stronger attenuation of the transmitted intensity, whereas regions with smaller β values appear more transparent. These spatial variations in absorption

give rise to intensity contrast in the recorded image. By analyzing such absorption-induced contrast, more information about the imaging object, such as its material composition, can be inferred.

Specifically, when the object is homogeneous (same index of refraction throughout the object) with a thickness T regardless of its position in the (x, y) plane, Equation 2.5 can be simplified as:

$$I_{\text{out}} = I_{\text{in}} \cdot e^{-2k\beta T}. \tag{2.6}$$

With $\mu = 2k\beta$, this becomes the well-known Beer-Lambert's Law, as shown in Equation 2.7:

$$I_{\text{out}} = I_{\text{in}} \cdot e^{-\mu T}, \tag{2.7}$$

where μ denotes the linear absorption coefficient with a unit of m^{-1}. Empirically, the linear absorption coefficient of an element is approximately proportional to the fourth power of the atomic number and inversely proportional to the third power of the X-ray energy.

2.1.3 Imaging Regimes

Up to now, we have discussed the exit wavefield ψ_{out} and the corresponding intensity I_{out} right after X-rays propagate through an object. It is important to note that the recorded X-ray image is determined not only by the exit wavefield itself but also by the detector placement, e.g., the sample-detector distance. In practice, as shown later in Chapter 3, the detector in an X-ray imaging experiment is usually placed at a finite distance away from the studied sample (see Figure 3.3 and 3.7), rather than immediately after the sample. As a consequence, the exit wavefield further evolves through free-space propagation before being recorded by the detector, and it is important to learn how the X-ray wavefield changes at different sample-detector distances.

Figure 2.4 shows a simulation of how the resulting image acquired by the detector changes when increasing the sample-detector distance, using X-rays at an energy of 12.4 keV. The imaging object consists of two disks with a diameter of $280\,\mu\text{m}$. Disk A is an ideal absorbing object with attenuation only ($\delta = 0$ and $\beta \neq 0$). Disk B is an ideal phase object with phase shift only ($\delta \neq 0$ and $\beta = 0$). As shown in Figure 2.4, when the distance is rather small (≤ 10 cm), the phase object B is almost invisible in the detector image, and the recorded intensity closely resembles the exit intensity. Under the assumption of a fully coherent plane-wave illumination (Equation 2.3), however, the phase modulation imprinted by the object leads to interference between different parts of the transmitted wavefront during free-space propagation. As the

propagation distance increases (≥ 1 m in Figure 2.4), this interference gives rise to diffraction-induced intensity modulations, making the phase object B more and more visible, and fringes of both objects (also referred to as edge enhancements) start to appear. Finally, when the distance is sufficiently large (≥ 1 km in Figure 2.4), it is even difficult to recognize the original shapes of the objects from the observed image because diffraction effects dominate the image formation process. In such a scenario, the observed image corresponds to the Fourier transform of the object.

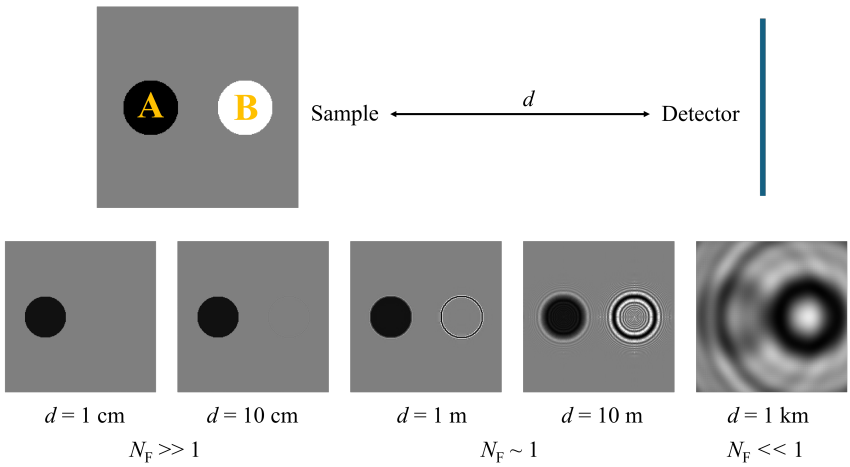


Figure 2.4: Simulation of X-ray images with different sample-detector distances from 1 cm to 1 km. The X-ray energy is 12.4 keV. Each image contains 256×256 pixels with a pixel size of $4\,\mu\text{m}$.

To classify different imaging regimes arising from different propagation distances, it is more appropriate to use the dimensionless Fresnel number rather than the sample-detector distance itself. The Fresnel number is defined as follows:

$$N_{\text{F}} = \frac{a^2}{\lambda d}, \tag{2.8}$$

where λ denotes the wavelength of X-rays, d denotes the sample-detector distance, and a denotes the characteristic size of the imaging feature. In the simulation shown in Figure 2.4, we have $\lambda = 0.1$ nm and $a = 8$ µm (corresponding to two pixels). The dimensionless Fresnel number provides a more physically meaningful classification of imaging regimes, as it allows comparison across different X-ray imaging setups. There are three imaging regimes, as shown below:

(a) $N_{\text{F}} \gg 1$ refers to the contact regime, where diffraction effects can be neglected;

(b) $N_{\text{F}} \sim 1$ refers to the near-field regime, where the recorded intensity contains contributions from diffraction-induced intensity modulations. Such modulations remain locally confined, meaning that each object feature only affects nearby detector pixels;

(c) $N_F \ll 1$ refers to the far-field regime, where diffraction effects dominate, and the recorded intensity becomes non-local.

In this thesis work, all the X-ray imaging experiments lie in the near-field regime, where the diffraction effects are detectable.

2.2 3D X-ray Imaging - Tomography

Up to this point, we have discussed the formalism of the 2D image provided by the X-ray detector. However, when one wants to study the detailed internal structures within the 3D object, a single 2D image is not enough. A fundamental approach to solve this issue is to combine multiple 2D images from different projection angles, which is the basic principle behind X-ray tomography, the state-of-the-art technique for 3D analysis.

The terminology "tomography" originated from two Greek words 'tomos' and 'graphō', where the former means slice or section, and the latter means to write. The idea of X-ray tomography can be dated back to the early 1900s[9], although an actual setup was not yet available. Since the mid-20th century, the increased availability and power of computers started to boost the research on tomography reconstruction. In 1979, the Nobel Prize in Physiology or Medicine was awarded to Allan M. Cormack and Godfrey N. Hounsfield "for the development of computer-assisted tomography", marking a highlight for X-ray tomography. Nowadays, X-ray tomography is widely used when beyond 2D information is needed non-destructively, such as medical diagnosis and industrial inspections.

Practically, there are three major approaches to collect projection images of the object from different angles, as illustrated in Figure 2.5(a)-(c), respectively. The first is to rotate the imaging object without moving the X-ray source and the detector, which is suitable for large-scale X-ray facilities, as moving such big facilities is impossible. The second is to rotate the X-ray source and detector while keeping the imaging object static, as used in e.g., medical CT applications. The third is to stack multiple X-ray beam-detector pairs to simultaneously study the object, which is the main technique to be discussed in the next chapter.

2.2.1 Reconstruction Methods

After collecting 2D images from various angles, one needs to reconstruct the 3D images of an object's internal structures, referred to as tomograms. As will be shown below, there are two main categories of reconstruction methods that are commonly

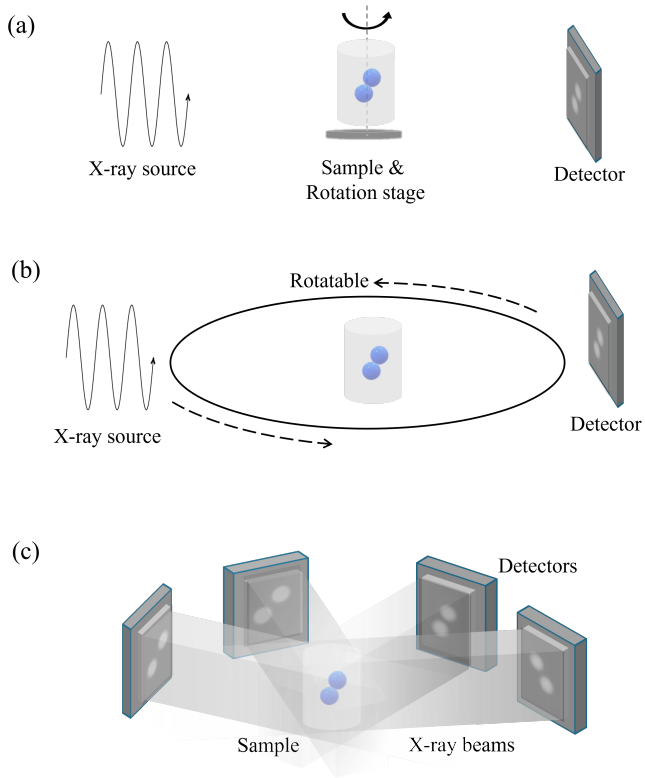


Figure 2.5: Approaches to collect projection images of an object from different angles. (a) To rotate the imaging object without moving the X-ray source and the detector; (b) To rotate the X-ray source and the detector without rotating the object; (c) To stack multiple X-ray beam-detector pairs to study the object simultaneously.

used as the gold standard. The first is Fourier-based reconstruction methods, which in general can provide fast reconstruction due to the high efficiency of computing the fast Fourier transform; the second is iterative methods, which generally require more computational resources but provide more reliable reconstruction by progressively refining the reconstruction.

Fourier-based Method

Derived firstly by Ronald N. Bracewell in 1956[10], the Fourier slice theorem (also referred to as the projection-slice theorem) states that the Fourier transform of an n-dimensional object can be related to the Fourier transforms of its lower-dimensional projections, forming the basis of Fourier-based methods. For example, in a 2D case, if we project a 2D image $f(x, y)$ along a specific angle θ and then perform a 1D Fourier transform, the result is equivalent to taking a 2D Fourier transform of the image and extracting the slice corresponding to the same angle θ.

Using the Fourier slice theorem, one can propose the direct back-projection method, i.e., to reconstruct the 2D slice $f(x, y)$ by assembling the 1D Fourier transforms from many projections at their respective angle and calculating the inverse 2D Fourier transform. However, direct back-projection typically results in a blurring effect in the reconstructed image. The reason is that collecting more projections means sampling more lines passing through the origin in the Fourier domain [4], resulting in oversampling in the low-frequency area and undersampling in the high-frequency area. To address this issue, filtered back-projection (FBP), one of the most famous Fourier-based reconstruction methods, is designed, where an additional filter is applied in the Fourier domain before the inverse Fourier transform. Commonly used filters include the ramp filter, the cosine filter, and the Hamming filter [11].

Iterative Method

Different from the Fourier-based method, the iterative methods treat the reconstruction problem as a linear equation, as shown in Equation 2.9:

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{2.9}$$

where $\boldsymbol{A}$ denotes the projection matrix that convert an image from the image space to the projection space, $\boldsymbol{x}$ denotes the unknown image to be reconstructed, and $\boldsymbol{b}$ denotes the measurements (projections provided by the X-ray detector).

In many cases, solving such a linear equation by calculating the inverse or pseudo-inverse of $\boldsymbol{A}$ is impractical due to high computational cost or ill-posedness of the problem [12]. Instead, iterative approaches can be used to provide a solution by iteratively minimizing the difference between the measured projections and the predicted projections. Examples of widely used iterative methods include simultaneous algebraic reconstruction technique (SART) [13] and simultaneous iterative reconstruction technique (SIRT) [14]. The former updates the reconstructed image after going through each individual projection angle, usually leading to faster convergence; while the latter updates the reconstruction image only after going through all the projection data, usually leading to a smoother and more stable reconstruction. Such methods also provide valuable insights for designing self-supervised AI-assisted reconstruction methods to be presented later in Chapter 5.

2.2.2 Crowther Criterion

In tomography reconstruction, a fundamental question is how many projections are sufficient for a good 3D reconstruction. Based on the principles of the Fourier-based method, the Crowther criterion [15] provides an answer. It states that, for a desired spatial resolution d, the minimum requirement of the number of projections N_θ can

be described in Equation 2.10:

$$N_\theta = \frac{\pi D}{d} \tag{2.10}$$

where D denotes the diameter of the imaging object. Practically, in most imaging scenarios, the required resolution should be several pixels or voxels, meaning that one needs approximately as many projections as the sample diameter measures in the number of pixels or voxels.

Figure 2.6 shows the comparison of a reconstructed slice of a simulated 3D dataset using a different number of projections (N) spanning an angular range of $[0°, 180°]$ with FBP and SART, respectively. The simulated object consists of several randomly placed ellipsoids within a cylinder with a diameter of $D = 200$ pixels. We desire to distinguish the smallest ellipsoid, leading to $d = 10$ pixels and $N_\theta \approx 63$. It can be seen that when the number of projections is far below the value indicated by the Crowther criterion ($N \leq 20$), the reconstruction is either blurry or significantly suffers from artifacts, whereas increasing the number of projections improves reconstruction quality. When comparing the two reconstruction methods, both of them provide reliable reconstruction with sufficient projections ($N = 100$); with insufficient projections ($N \leq 20$), SART outperforms FBP as the former generates fewer artifacts.

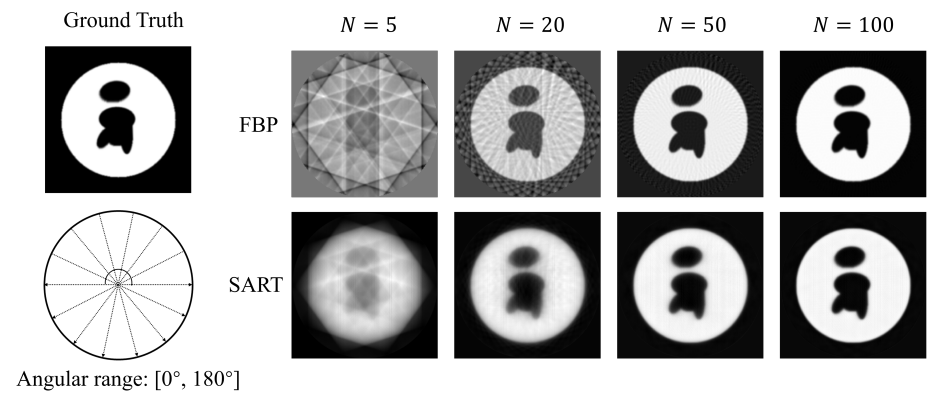


Figure 2.6: Simulation of tomographic reconstruction from a different number of projections equally distributed within $180°$, using FBP and SART, respectively. The simulated object consists of several randomly placed ellipsoids within a cylinder.

2.2.3 "Missing Wedge" Issue

Besides sufficient projection images, the angular range for projections is also crucial for a good tomography reconstruction. With limited projection angles and a gold-standard reconstruction algorithm, even a large number of projections cannot lead to an acceptable reconstructed image. This issue is called "missing wedge"[16]. Figure 2.7

compares the reconstruction using 100 projections equally distributed in different angle ranges. It clearly shows that neither FBP nor SART can provide an acceptable reconstruction, especially when $\Delta\theta \leq 90°$.

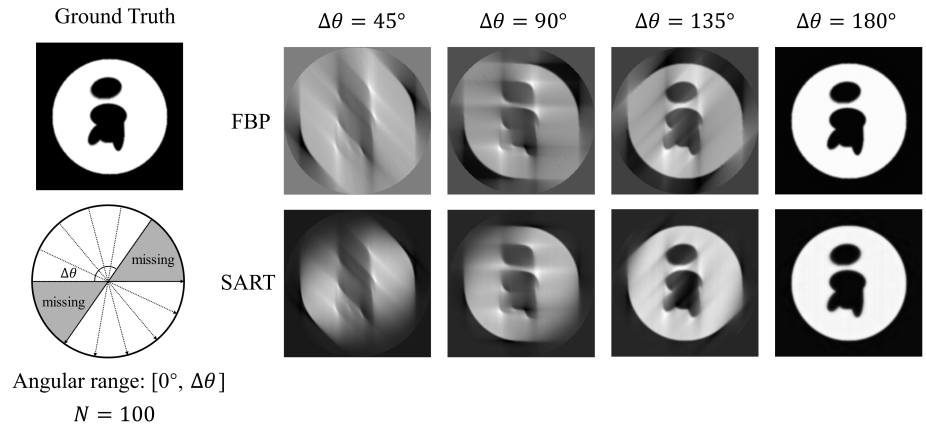


Figure 2.7: Simulation of tomographic reconstruction using 100 projections equally distributed within the angle range $[0°, \Delta\theta]$. Four different $\Delta\theta$ values are included in this figure.

As shown in Figure 2.6 and 2.7, limited projections and angular range can pose significant challenges for image reconstruction. Therefore, to tackle such challenges, it is crucial to include prior knowledge of the sample into the reconstruction workflow rather than directly applying traditional Fourier-based or iterative methods. For example, a recently developed algorithm, optimized neural implicit X-ray imaging (ONIX) [17], combines basic ideas from traditional iterative methods with novel AI approaches. By effectively utilizing prior knowledge of the sample extracted by the AI approaches, this algorithm yields high-quality 3D reconstruction and avoids the "missing wedge" issue with as few as 8 projections over a limited angular range.

2.3 Extend to 4D - Opportunities and Challenges

Using the same principle as X-ray tomography, one can naturally extend X-ray imaging from 3D to 4D by adding the temporal evolution as an extra dimension. This is referred to as time-resolved X-ray tomography, or tomoscopy, resulting in reconstructed 3D movies of the studied dynamics. State-of-the-art tomoscopy experiment is capable of reaching 8 µm resolution at 1000 tomograms per second (tps) [18], thanks to fast image acquisition enabled by high-flux modern X-ray sources, together with fast sample rotation at the speed of 500 revolutions per second. Tomoscopy opens up new opportunities for studying rapid dynamics in 3D within a millisecond timescale, benefiting various fields such as fluid dynamics [19], additive manufacturing [20], and ma-

terial science[21]. Although the prospect of tomoscopy has been widely acknowledged, challenges on both the hardware and software sides need to be carefully addressed.

First, from the sample perspective, withstanding rapid rotations can be challenging due to the induced centrifugal acceleration, which can be calculated as:

$$a_c = r\omega^2, \tag{2.11}$$

where ω denotes the angular velocity of the rotational sample stage and r denotes the distance between the rotational axis and the center of mass of the rotating object. As an example, a distance of 1 mm and a rotational speed of 500 Hz lead to a centrifugal acceleration of $a_c = 9870$ m/s$^2 \approx 1000$ g. Such a high centrifugal acceleration might alter the studied dynamics or even damage the studied samples. Due to this constraint, studying rotation-sensitive samples or further increasing the rotation speed becomes impractical.

Second, from the software perspective, the current reconstruction workflow corresponds every $180°$ or $360°$ sample rotation to a reconstructed time point, and treats each time point independently. The 4D reconstruction is then established by stacking the 3D reconstruction of each time point, respectively. In other words, any dynamic events that occur faster than the rotation speed cannot be precisely reconstructed. Such rapid dynamics can even cause the issue of "motion blur"[22], which poses additional challenges for reconstruction and further analysis.

The key to addressing the aforementioned challenges is to find an alternative method for running 4D experiments without the requirement of fast rotation and to develop the corresponding methods for 4D reconstruction and analysis, which will be presented in Chapters 3-5, respectively.

Chapter 3

Synchrotron Light Sources and X-ray Multi-Projection Imaging (XMPI)

Up to this point, we have discussed the opportunities and challenges of state-of-the-art 4D X-ray imaging (tomoscopy), where fast rotation of the sample is required for monitoring high-speed 3D dynamics. In this chapter, we will bring modern X-ray sources to the table and introduce the possibility of establishing a new rotation-free 4D X-ray imaging setup that addresses the aforementioned limitations.

This chapter begins with a brief introduction to synchrotron light sources and their unique imaging capabilities regarding the spatiotemporal resolution, as all the experiments included in this thesis were conducted using X-rays from synchrotron radiation facilities. Then, we introduce the XMPI setup, which benefits from such unique imaging capability provided by the synchrotron light source. Finally, we conclude this chapter with a brief discussion of the opportunities and challenges brought by XMPI. This chapter is closely related to **Papers I and II**.

3.1 Synchrotron Light Source

Synchrotron radiation refers to the electromagnetic radiation emitted when relativistic charged particles (such as electrons) undergo an acceleration perpendicular to their velocity[23]. For a relativistic electron circulating along a circular trajectory with a

radius R, the total radiated power P can be expressed using the Schwinger's formula[24]:

$$P = \frac{2}{3}\frac{e^2 c}{R^2}\left(\frac{E}{mc^2}\right)^4, \tag{3.1}$$

where E denotes the electron energy, c denotes the speed of light, and mc^2 denotes the electron rest mass energy (0.511 MeV). It was first observed at the General Electric Research Laboratory in Schenectady, New York, in the 1940s[25], although its mechanism was not immediately understood. At an early stage, synchrotron radiation was regarded as an undesirable byproduct in high-energy physics as it led to significant energy losses of charged particles[26] (proportional to E^4/R^2, as indicated in Equation 3.1). However, since the 1950s, its unique properties have gradually been recognized and exploited by the scientific community[27], including its high intensity, broad and continuous spectral range, and narrow angular collimation. After several generations of development, synchrotron light sources have evolved into indispensable facilities for X-ray science, supporting advanced studies in many disciplines. At the time of writing, the most advanced synchrotron light sources are categorized as 4th-generation synchrotron sources, known as diffraction-limited storage rings (DLSRs), which are designed to deliver electron beams with ultralow emittance such that, over a broad range of X-ray energies ($\sim$ 100 eV to a few keV), the emitted photon beam are predominantly limited by diffraction rather than by the phase space of the electron beam. Representative examples of 4th-generation synchrotron light sources worldwide include MAX IV in Sweden[28], SIRIUS in Brazil[29], SLS 2.0 in Switzerland[30], ESRF-EBS in France[31], APS-U in the USA[32], and HEPS in China[33]. Meanwhile, several 3rd-generation synchrotron sources are still in operation.

3.1.1 Higher Brilliance and Flux

To evaluate the quality of an X-ray source, several factors should be considered. The first is the flux, indicating the number of photons emitted per second. The second is the source size, indicating the difficulty of focusing the X-ray to a given size (usually small). The third is the divergence of the source, indicating how the beam diverges during propagation. The fourth is the spectral distribution. A key quantity that summarizes the aforementioned factors of the produced X-ray beam is referred to as brilliance, whose unit is described as follows:

$$\text{Brilliance} = \frac{\text{number of photons}}{(\text{s}) \cdot (\text{mm})^2 \cdot (\text{mrad})^2 \cdot (0.1\%\ \text{bandwidth})}. \tag{3.2}$$

The brilliance is invariant under lossless optical transformations[34], which cannot be improved by passive optical elements after the generation of the X-ray. Such invariance arises from the conservation of phase-space density, as dictated by Liouville's

theorem[6]. The scientific community considers brilliance as a suitable metric for assessing and comparing the quality of different X-ray sources. In practical terms, a higher brilliance corresponds to a photon beam that delivers more photons into a smaller transverse area and a narrower angular spread within a given bandwidth, and thus indicates superior X-ray beam quality. Figure 3.1 shows the evolution of the averaged brilliance of X-ray sources, ranging from lab-based X-ray tubes to different generations of synchrotron light sources since the 1950s[35]. It can be seen that synchrotron light sources have witnessed an increase of more than 10 orders of magnitude regarding the averaged brilliance since the first generation in the 1970s.

In this thesis work, we focus on exploiting the unique imaging capability provided by synchrotron light sources in terms of the spatiotemporal resolution. For this purpose, a key parameter to understand is the flux density, i.e., the number of photons per unit time and area. Intuitively, to achieve a better spatiotemporal resolution, a larger flux density is required. Equation 3.3 gives the relationship among the flux density, spatial resolution, and temporal resolution:

$$\phi = \frac{N}{\tau \cdot \Delta x^2}, \tag{3.3}$$

where ϕ denotes the flux density, N denotes the number of photons required to achieve a certain image quality, τ denotes the temporal resolution, and Δx denotes the spatial resolution. Empirically, to provide high-quality images for X-ray imaging experiments, we can set $N = 10000$, which leads to a signal-to-noise ratio SNR $=$

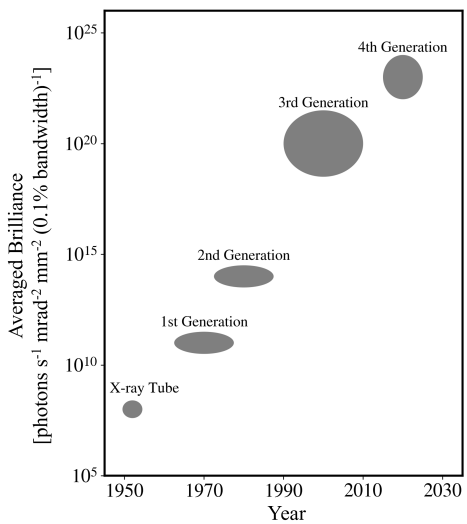


Figure 3.1: Evolution of the averaged brilliance of X-ray sources, ranging from lab-based X-ray tubes to four generations of synchrotron light sources since the 1950s.

$N/\sqrt{N} = 100$ according to Poisson statistics[36].

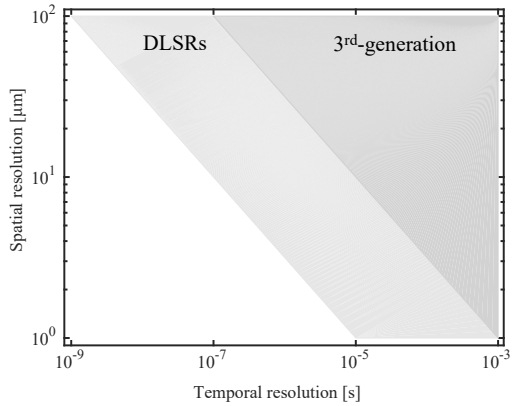


Figure 3.2: Representative spatiotemporal resolution provided by DLSR-based 4th generation synchrotron light sources and 3rd generation synchrotron light sources. The typical values of the flux density at 4th generation and 3rd generation synchrotron light sources we use for this plot are 10^{15} $\mathrm{ph/(s \cdot mm^2)}$ and 10^{13} $\mathrm{ph/(s \cdot mm^2)}$, respectively. Figure from **Paper I**.

Based on Equation 3.3, we can establish Figure 3.2 that illustrates the representative spatiotemporal resolution achievable provided by fourth- and third-generation synchrotron light sources, respectively. It can be seen that when a spatial resolution of several μm is required, the maximum available temporal resolution is at the scale of μs. In other words, MHz imaging at the micrometer scale is possible at modern synchrotron light sources.

3.1.2 ForMAX Beamline and its Imaging Capabilities

In the previous subsection, we discussed the possibility of reaching higher spatiotemporal resolution thanks to the enhanced brilliance and flux density provided by modern synchrotron light sources. In this subsection, we will present the time-resolved imaging experiment to demonstrate this capability.

The experiments were conducted at the ForMAX beamline[37] at MAX IV, Sweden, the first operational 4th-generation synchrotron light source. ForMAX offers multiscale structural characterization, with a particular focus on research involving forest materials. Moreover, ForMAX is optimized for various X-ray scattering and time-resolved X-ray imaging experiments. Using a 3 m long room-temperature in-vacuum undulator as an insertion device, it can provide X-ray beams with energies ranging from 8 keV to 25 keV with a beam size of 1.3 mm $\times$ 1.5 mm. The desired energy can be selected using either a double-crystal monochromator (DCM) or a double-multilayer monochromator (MLM). The latter was used for all the experiments included in this

thesis, as it allows a larger bandpass ($\Delta E/E \approx 0.01$), resulting in a higher flux density crucial for time-resolved imaging experiments. Moreover, such a bandpass is capable of accepting an entire single harmonic of the insertion device, which is crucial for designing the spectral splitting scheme for the XMPI setup (discussed later in Section 3.2.1).

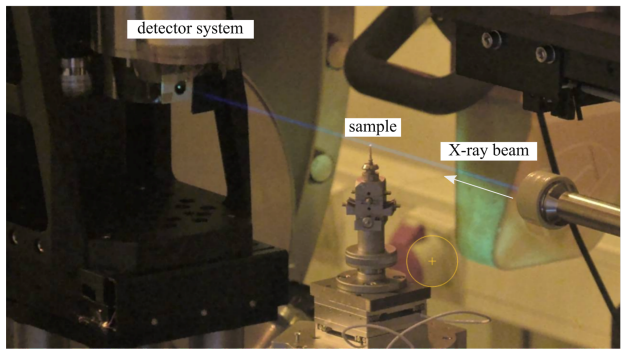


Figure 3.3: A photo captured by an Axis camera in the experimental hutch showing the time-resolved X-ray imaging experiment conducted at the ForMAX beamline. Figure from **Paper I**.

To demonstrate the time-resolved full-field imaging capabilities at ForMAX, a straightforward experimental setup was established, as shown in Figure 3.3. To support both 2D and 3D imaging, the sample was placed on a rotation stage from LAB Motion Systems (RT075s)[38]. The X-ray detector system built upon a Photron Nova S16 camera can support an acquisition rate of up to 1.1 MHz, with an effective pixel size of $4\,\mu m$ and a full dynamic range of 12 bits. More technical details on this setup can be found in Section 3.2.2 and **Paper I**. It is worth noting that the beam path seems "visible" in Figure 3.3. The reason is that high-flux incoming X-rays can ionize air (primarily N_2 molecules), leading to fluorescence in the blue or blue-violet wavelength range.

2D Imaging Capability

We evaluated 2D imaging capability in two steps. First, we used a simple static sample (a metal pin) to evaluate the speed and the corresponding dynamic range at which we can record images with our detector system, as shown in Figure 3.4. The beam energy was set at around 16.3 keV, and the acquisition rate was set as 550 kHz. The maximum and minimum grayscales of the image shown in Figure 3.4(a) were 2504 and 8, respectively, corresponding to a dynamic range of 12 bits. This result matches the estimation shown in Figure 3.2, where MHz imaging at the micrometer scale can be achieved.

Second, we conducted time-resolved 2D imaging of a scientific sample, allowing for the observation of rapid dynamics. Here, we aimed to observe the fracture process

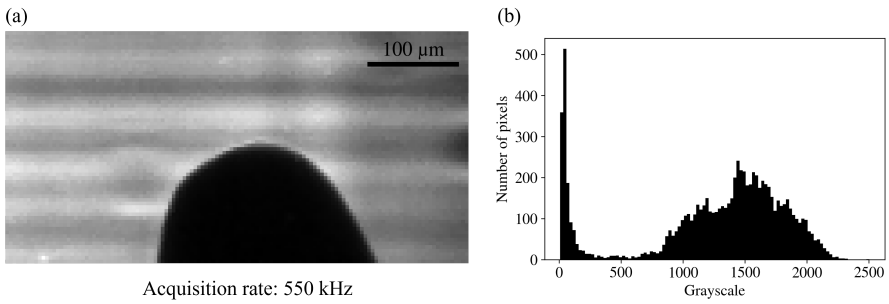


Figure 3.4: Dynamic range evaluation. (a) shows an image of a static metal pin recorded at a frame rate of 550 kHz at around 16.3 keV. (b) shows the histogram of the image in (a). Figure from **Paper I**.

of a multi-filament fiber when an external tensile stress is applied to it. To provide a good sample contrast and capture the desired dynamics, the beam energy and frame rate were set at around 16.3 keV and 36 kHz, respectively. To avoid saturation of the detector and maintain the full dynamic range, we used a total of 1.5 mm Si filters (corresponding to a transmission rate of 0.071) to reduce the flux.

The 2D imaging result is shown in Figure 3.5. Figure 3.5(a) shows three consecutive images during the tensile failure process. The images have been postprocessed using dynamic flat-field correction [39,40] to eliminate fixed-pattern noise. To further analyze the image quality of the acquired images, we used Fourier ring correlation (FRC) [41], a practical method for estimating the spatial resolution. As shown in Figure 3.5(b), the FRC curve and the 1-bit threshold curve intersect at $x = 0.75$. Given that the effective pixel size of the images is $4\,\mu\text{m}$, the resolution is estimated to be $2 \times 4\,\mu\text{m} \div 0.75 = 10.7\,\mu\text{m}$ (the factor of 2 comes from the Nyquist-Shannon theorem [42]), indicating a good image quality.

To sum up, regarding the time-resolved 2D imaging, we demonstrated the potential to achieve high frame rates at the MHz scale, using the full dynamic range of the detector (12 bits) and an effective pixel size of $4\,\mu\text{m}$. The FRC analysis shows an estimated resolution of approximately $10\,\mu\text{m}$ in the scientific case of tensile failure process, indicating a good image quality.

3D Imaging Capability

We further evaluated 3D imaging capability by conducting time-resolved tomography, based on the principles discussed previously in Section 2.2. Here, we aimed to probe the dynamics induced by the radiation damage on the wooden rod at the energy of 9.1 keV. The rotation stage was operated at its maximum speed (3 revolutions per second), resulting in 6 tps. To match the requirement given by the Crowther Criterion (Subsection 2.2.2), the frame rate was set as 750 Hz (to provide 125 projections per

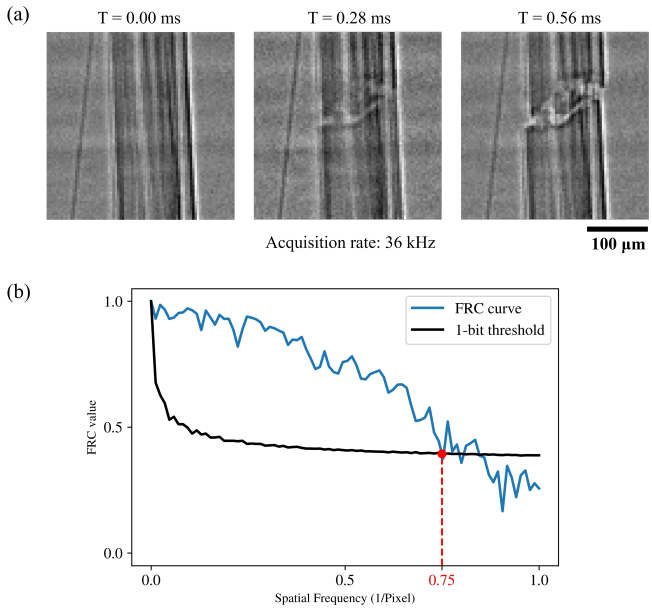


Figure 3.5: Time-resolved 2D imaging studies. (a) shows three consecutive flat-field-corrected frames during the tensile failure process of a multi-filament fiber. (b) shows the FRC analysis of two independent measurements during the tensile failure process, where the red dot denotes the intersection of the FRC curve and the 1-bit threshold curve. Figure adapted from **Paper I**.

tomogram), which is significantly lower than the time-resolved 2D experiment. Since the exposure time was the inverse of the frame rate, a significantly larger X-ray flux went through the sample compared to the time-resolved 2D experiment. To protect the detector, we used a total of 600 µm Si filters with a transmission of 0.0027. The Si filters were placed behind the sample, i.e., between the sample and the detector, to be able to induce radiation damage in the sample while maintaining an acceptable flux on the detector. This provided us with the full dynamic range (12 bits) of the detector at a frame rate of 750 Hz.

Figure 3.6(a) gives flat-field corrected 2D projections and two corresponding reconstructed slices at $t = 0.0$ s and $t = 47.7$ s, respectively, showing the differences before and after the radiation damage. Tomography reconstruction was performed using the Gridrec algorithm[43] in Tomopy[44], a Python-based open-source framework. Similar to FRC in the 2D imaging case, we estimated the resolution of the 3D reconstructions using Fourier shell correlation (FSC), resulting in a resolution of approximately 6 voxels (24 µm).

Although this 3D imaging experiment was performed at 6 tps, significantly slower than the state-of-the-art tomoscopy experiment[45], it is important to note that only less than 0.3% of the flux was utilized. Therefore, the limiting factor here is the

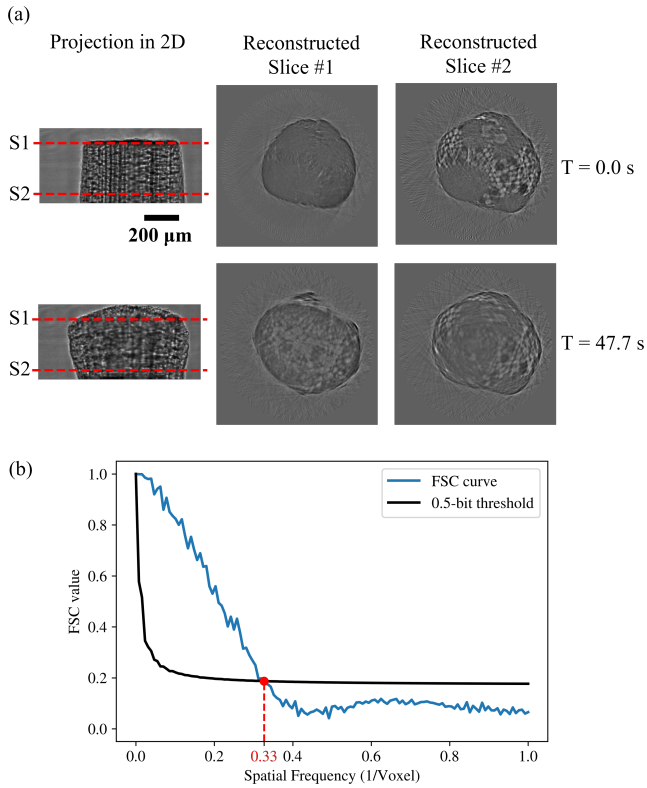


Figure 3.6: Time-resolved 3D imaging test at ForMAX. The flat-field-corrected projections and two corresponding reconstructed slices at two different time points during the radiation damage process of a wooden rod are shown in (a). The FSC analysis of two independent 3D reconstructions of the sample is shown in (b), where the red dot denotes the intersection of the FSC curve and the half-bit threshold curve. Figure adapted from **Paper I**.

rotational speed, rather than the flux. Potentially, if the rotational stage is capable of faster rotations and the full flux is utilized, the acquisition speed can reach 2000 tomograms per second, with the same effective pixel size of 4 µm, full dynamic range (12 bits) and 125 projections per tomogram.

3.2 XMPI Setup in a Nutshell

As is discussed in the previous section, higher flux density provided by synchrotron light sources enables new imaging opportunities in terms of spatiotemporal resolution, and such opportunities have been demonstrated at the ForMAX beamline. However, in time-resolved 3D imaging as shown in Figure 3.6, the requirement of rapid sample rotation is still a limiting factor in the experiment, as it may perturb the intrinsic

dynamics under investigation. To overcome this limitation, an alternative strategy is XMPI, which exploits the available X-ray flux not by illuminating the sample with the direct beam, but by first splitting the primary beam into several beamlets and illuminating the sample simultaneously from multiple directions. This strategy eliminates the need for fast sample rotation to obtain 3D information, although allowing a slow rotation that does not affect the studied dynamics is still beneficial, as will be shown later in Section 5.3.2 and **Paper VII**.

In this section, we will discuss the key modules of XMPI and their implementations at the ForMAX beamline, MAX IV. The schematic is sketched in Figure 3.7(a) together with an annotated photo from one of the XMPI experiments, where four crystals are used to generate three secondary beamlets. The sample is placed on a sample stage (preferably rotatable) at the intersection point of the beamlets. All three cameras are synchronized and able to track the dynamics of the sample from different angles.

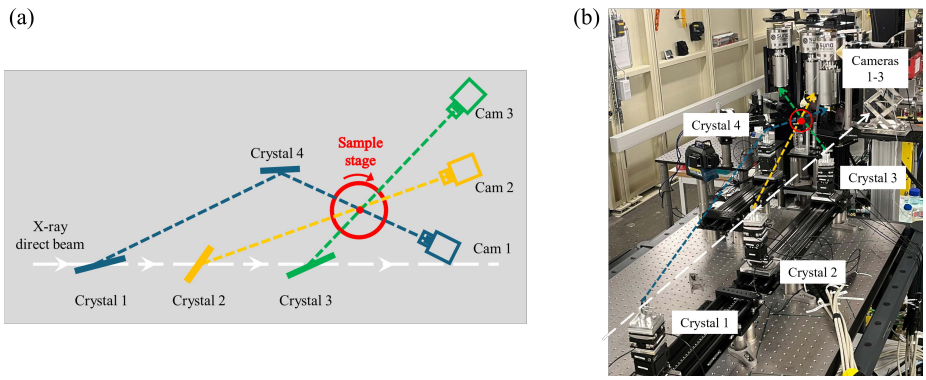


Figure 3.7: Schematic of the XMPI setup. (a) Crystals $1-4$ are used to split the primary beam into three beamlets that intersect within the investigated object. Afterwards, the beamlets are captured by X-ray detector systems Cam $1-3$; (b) A photograph of the experimental setup at the ForMAX beamline with the corresponding annotations. Subfigure (b) adapted from **Paper II**.

As shown in Figure 3.7, the two main modules in the XMPI setup are the beam-splitting system and the X-ray detector system, which will be introduced in the following subsections. For more technical details, we kindly refer readers to **Paper II**.

3.2.1 Beam-Splitting System

There are two main components in the beam-splitting system: the perfect crystals and the nanopositioners. In the context of X-ray optics, a perfect crystal[4] refers to a crystalline material with a highly ordered and periodic lattice structure with negligible impurities, dislocations, and defects. Perfect crystals are employed in the beam-splitting system because they enable precise and controllable manipulation of

the X-ray beam via Bragg or Laue diffraction, while the nanopositioners ensure that all the crystals are placed properly to split the X-ray beam as designed.

Perfect Crystals

The core of the beam-splitting system is the perfect crystals, inspired by concepts developed in interferometers, split-and-delay lines, and specific setups for single-shot imaging[46,47]. Such perfect crystals can split an incident X-ray beam via two different splitting schemes: spectral splitting and amplitude splitting. Spectral splitting refers to the redirection of a portion of the incident beam's spectrum to a new direction due to diffraction by a crystal. This occurs when the crystal is oriented to fulfill the Bragg or Laue condition, i.e., when the incident X-ray beam intersects a family of lattice planes at a specific angle θ_B with respect to the beam direction. The redirected portion of the beam (secondary beamlet) will then travel at an exit angle of $2\theta_B$. Amplitude splitting refers to the positioning of the crystal so that this condition is only fulfilled for a portion of the beam's footprint, i.e., only part of the crystal is in the beam path, allowing the rest of the beam to propagate unhindered[48]. Compared to spectral splitting, amplitude splitting avoids absorption losses at the expense of a reduced field of view.

Upon these splitting schemes, several requirements need to be met when designing the beam-splitting system. First, all the generated beamlets should provide sufficient flux, and the illumination needs to be stable and homogeneous. Second, it is preferred to have a large angular span of the beamlets, as it is beneficial for subsequent 4D analysis of the studied dynamics. Third, practicality needs to be considered, including the heat resistance of the crystal, the available space for placing the crystals and the respective X-ray detectors, the difficulty of crystal fabrication, and associated costs.

Taking into account all these requirements and the available equipment, we arrived at the current solution of beam-splitters at ForMAX when the primary beam energy was set as around 16.5 keV. Regarding the material of the crystals, we selected Ge and Si, as these are among the most accessible materials with high purity. To generate beamlets, all beam splitters employed spectral splitting. At the ForMAX beamline, X-ray beams with high flux density and small band-pass (compared to the white beam) are provided, as discussed in Section 3.1.2. These two factors result in a high spectral flux density (flux density per bandwidth), ensuring sufficient flux density for the spectral-splitted beams. In addition, Crystal 1 also served as an amplitude splitter to reduce absorption losses. Specifically, in Figure 3.7, the first beamlet was established by splitting the direct beam with Crystal 1 (Si-111) and recombining it using Crystal 4 (Ge-400) at $-17.0°$ relative to the direct beam; the second beamlet was established using Crystal 2 (Si-111), resulting in 13.7° with respect to the direct beam; the third beamlet was established using Crystal 3 (Ge-400), resulting in 30.7°

with respect to the direct beam. As a result, all three beamlets were established with a diffraction efficiency of $1\% - 2\%$ and the angular span reached $47.7°$, providing a solid foundation for subsequent 4D analysis.

Nanopositioners

Nanopositioners are also indispensable for successfully performing XMPI experiments for the following two reasons. First, each crystal should be properly placed and oriented to satisfy its Bragg condition. According to the dynamical theory of diffraction, a perfect crystal can only reflect radiation with a near-unity efficiency within a narrow angular range. This angular range is referred to as the Darwin width[49]. For the Ge and Si crystals used in the current XMPI setup at 16.55 keV, the Darwin width is at the scale of a few arcsec, indicating the need to precisely orient all the crystals. Second, all the beamlets should be properly aligned to intersect at a common point in space, which will be used to place the sample for measurement. To precisely position and orient all the crystals, we used a stack of six nanopositioners for each beam splitter, providing six degrees of freedom (three translations, one rotation, and two goniometers). Specifically, the translational stages provide a 1 nm closed-loop positioning resolution with the horizontal (vertical) stage covering a total travel range of 30 mm (8 mm); the rotation stage is capable of continuous rotation with an accuracy of 0.036 arcsec; the goniometers are able to move $\pm 5°$ with an accuracy of 0.0036 arcsec.

3.2.2 X-ray Detector System

Overall, the X-ray detector system is responsible for providing high-quality 2D images at a micrometer resolution at a high acquisition rate (up to MHz), depending on the requirement of the studied sample. The X-ray detector system in the XMPI setup employs an indirect scheme comprising three main components: scintillators, visible-light microscopes, and fast visible-light cameras.

The first component is the scintillators, which convert X-rays into visible light for all three beamlets. Specifically, to achieve a high yield without sacrificing the image quality, the GaGG+ scintillators with a thickness of $250\,\mu\text{m}$ are used.

The second component is visible light microscopes. Each microscope is equipped with an objective holder featuring a motorized focusing unit, compatible with 5X or 10X magnification. Following the objective is a 1X tube lens holder with a motorized camera rotation unit, which covers a range of 90 mrad with an accuracy of 0.02 mrad. Additionally, all the microscopes are mounted on a motorized positioning system for linear motion in two perpendicular directions.

Table 3.1: Image Properties of the XMPI Setup Using Different Cameras. Table adapted from **Paper II**.

Camera	Image Size (W x H)	Effective Pixel Size (W x H)	Full Dynamic Range
Photron Nova S16	1024 pixels × 1024 pixels	5.0 µm × 5.0 µm	12
Andor Zyla 5.5	2560 pixels × 2160 pixels	1.3 µm × 1.3 µm	16

The third component is fast visible-light cameras, which directly affect the acquisition rate and the achievable spatial resolution of the XMPI experiment. It is essential to note that a general trade-off exists between spatial resolution and acquisition rate for most commercial cameras[50]. Higher spatial resolution typically requires smaller pixel sizes, which usually leads to a lower signal-to-noise ratio unless a longer exposure time is used, thereby limiting the achievable acquisition rate. Hence, the selection of cameras is decided by the actual needs of the experiment. In the work presented in this thesis, we used two different commercial cameras: (a) Photron Nova S16 for faster acquisition (above kHz) and slightly bigger pixel size (20 µm); (b) Andor Zyla 5.5 for slower acquisition but smaller pixel size (6.5 µm). Table 3.1 summarizes the image properties provided by the XMPI setup when using these two cameras, respectively, where the effective pixel size is the pixel size of the camera divided by the magnification of the objective.

Beyond the selection of the commercial camera, another requirement in the XMPI experiment is to synchronize all the cameras so that they can either start or finish acquiring images simultaneously. For synchronization, a hardware trigger signal is needed. The ForMAX beamline offers solutions for both cameras: (a) the Zyla cameras were fully integrated in the ForMAX signal control system, ensuring synchronized hardware triggering using a PandABox[51]; (b) the Photron cameras were synchronized in a chain configuration; one camera received the trigger signal and acted as the primary transmitter; the other two cameras acted as receivers triggered by the first camera.

3.3 XMPI - Opportunities and Challenges

Thanks to the high flux density provided by synchrotron light sources, the concept of XMPI has been established, which eliminates the requirement for fast sample rotation to study the dynamics from different angles. At the ForMAX beamline, MAX IV, a stable XMPI setup has been established, where the beam-splitting system can provide three beamlets covering an angular range of 47.7° at an energy of 16.5 keV, and the X-ray detector system can provide high-quality synchronized image sequences with an effective pixel size of several µm at an acquisition rate of above kHz. Potentially, XMPI can create a new frontier for X-ray imaging, offering new scientific opportunities, as

it enables 4D studies of unexplored phenomena due to the limitations of state-of-the-art methods. For example, XMPI can be used to study rotation-sensitive samples, samples with fast (above kHz) dynamics, and even single-shot phenomena, which is not possible using rotation-based X-ray tomoscopy.

It is crucial to note that XMPI is a developing technique and there are still major challenges. Besides ForMAX, XMPI has also been developed or implemented in various modern X-ray sources worldwide, such as SPring-8 in Japan [52,53], ESRF in France [54], and European XFEL in Germany [55], employing similar principles but differing in technical details. In general, XMPI challenges can be categorized into hardware challenges and software challenges.

From the hardware side, there is still space for further improvements of the beam-splitting system. A general principle is to generate as many beamlets as possible covering an angular range as wide as possible (preferably close to 90°), as it is beneficial for retrieving 3D and 4D information of the studied sample. Regarding the number of projections, the setup established in SPring-8, Japan, provides valuable insights, which utilizes bent crystals to generate up to 32 beamlets [52]. However, such a setup relies on a white beam as a primary beam with a much larger bandpass and a much lower spectral flux density. This leads to significantly lower flux density per secondary beamlet, compared to the ForMAX implementation. As a result, the 32-beamlet system compromises its achievable spatiotemporal resolution. At a spatial resolution of 65 µm, only part of the beamlets can provide 1 kHz projection images at a proper image contrast and signal-to-noise ratio [52]. In addition, the limits in terms of available physical space to fit all the equipment and remaining bandwidth available for spectral splitting should always be taken into account. Regarding the angular range, obtaining higher diffraction angles generally requires higher diffraction orders, at the cost of lowering the diffraction efficiency, which results in a decrease in the image quality. Such a trade-off needs to be carefully considered. We can foresee that optimizing the beam-splitting method will be effort-demanding until it reaches the ultimate solution.

From the software side, XMPI also poses challenges regarding the subsequent 4D analysis. It is important to note that the number of projections available in the XMPI setup is far fewer than the requirement stated in the Crowther Criterion (discussed in subsection 2.2.2). It means that the available spatial information of the studied sample is extremely limited compared to cases where state-of-the-art tomoscopy works properly with over 100 projections for reconstruction. In the following two chapters, we will address this challenge by implementing triangulation-based 4D analysis in Chapter 4 and novel AI-assisted 4D reconstruction methods in Chapter 5, respectively, with examples of flow dynamics.

Chapter 4

4D Analysis of Flow Dynamics via Particle Tracking with XMPI

Suppose we only have a single 2D image of a particle; in that case, it is impossible to precisely determine its 3D position via triangulation. However, once an additional image of the same particle is obtained from a second camera, and the relative geometry between the two cameras is known, the 3D position can be uniquely identified, since the location of a point in 3D space corresponds to the intersection of two or more viewing rays. This idea naturally extends to the temporal domain: when an image sequence of a moving particle is recorded from both cameras, its 3D trajectory can be reliably tracked over time.

Tracking moving particles carried by flows using XMPI is conceptually straightforward but scientifically impactful. On one hand, from the perspective of fluid dynamics, analyzing particle trajectories in 3D can help us understand the properties and behavior of the flow, which is crucial for multiphase flow studies where microscopic particles govern the macroscopic behavior. On the other hand, from the perspective of XMPI development, it is the first application example for quantitative analysis enabled by XMPI. Such a quantitative analysis serves as a good indicator of the robustness of the XMPI setup, which will in turn motivate more applications in the near future.

In this chapter, we will present how XMPI is applied to the task of tracking particles in 3D. First, we introduce the basic concepts, experimental setup and the 4D analysis pipeline. Second, we present 3D particle-tracking results, including the analysis of individual and statistical behavior of tracked particles with a quantitative error analysis. We conclude this chapter by summarizing the opportunities and upcoming future works on the study of 4D flow dynamics using XMPI. For more technical details

beyond the content of this chapter, readers are kindly referred to **Paper VI**.

4.1 Particle Tracking Velocimetry

particle tracking velocimetry (PTV) refers to a velocimetry method to measure velocities and trajectories of moving particles. Depending on the number of dimensions, there are two categories of PTV methods: 2D PTV and 3D PTV. The former can be achieved by a single camera capturing a sequence of images, while the latter can only be achieved by a multiple-camera system in a stereoscopic arrangement.

The majority of commonly-used imaging approaches for 3D PTV rely on visible-light setups, including stereoscopic shadowgraphy[56] and optical coherent tomography[57]. Despite their widespread use, these approaches are limited to transparent or semi-transparent systems. X-rays, owing to their short wavelength and high penetration capabilities, serve as a promising tool for high-resolution PTV in flows opaque to visible light. Most of the X-ray-based 3D PTV setups utilize lab sources, such as tomography-based setups[58], which rely on continuous rotation of the sample, and stereographic setups[59], which stack multiple X-ray sources and detectors without the requirement of rotating the sample. However, the achievable spatiotemporal resolution of such setups is ultimately limited by the available flux from the lab-based X-ray sources. Hence, XMPI, established upon synchrotron radiation facilities, can serve as a new imaging modality for 3D PTV, due to its unprecedented spatiotemporal capabilities at the micrometer scale and above kHz. In this section, we will introduce the XMPI experimental setup for 3D PTV and a practical pipeline for particle tracking.

4.1.1 Experimental Setup

Figure 4.1 illustrates the schematic of the XMPI setup for the study of 3D PTV conducted in the ForMAX beamline, MAX IV. Two beamlets were established using crystal beam splitters (corresponding to Camera 1 and 3 in Figure 3.7), covering an angular range of $47.7°$. Two X-ray detectors, positioned several centimeters downstream of the studied sample, were used to simultaneously capture image sequences of particles moving with the flow, providing an effective pixel size of $1.3\,\mu\mathrm{m}$ and an acquisition rate of 40 Hz. As seen in Figure 4.1, the red circles highlight a large particle seen in both detectors. Different edge-enhancement performances for particles are observed due to slightly different sample-detector distances, as discussed in Subsection 2.1.3.

The studied sample consists of a Kapton tube with an inner radius of $R \approx 0.36$ mm

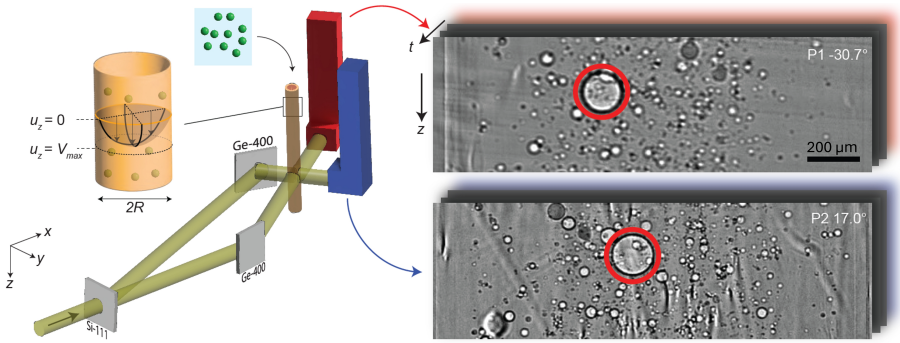


Figure 4.1: Schematic of the XMPI setup for 3D PTV. Figure from **Paper VI**.

and a wall thickness of roughly 25 µm with a flow that can be driven by a syringe pump (New Era Pump Systems NE-4000). The particles used in this work were silver-coated hollow glass spheres (SHGS) with a mean radius of 5 µm and a density of 1400 kg/m^3. These particles are commonly used as tracers in fluid dynamics research and are well-suited for X-ray imaging experiments[60]. Either of the two following media was used to suspend the particles. The first is glycerol with a density of 1260 kg/m^3 and viscosity of 1.4 Pa · s. Glycerol behaves as a homogeneous, Newtonian fluid[61], providing a well-controlled reference flow with a linear relationship between shear stress and shear rate. As a result, the flow field is smooth and predictable, making it suitable for validating the accuracy and precision of the proposed 3D particle-tracking approach. The second is human blood of about 40 percent hematocrit obtained from the Department of Transfusion Medicine at Karolinska University Hospital, Huddinge, Sweden. In contrast to glycerol, blood is a heterogeneous and non-Newtonian fluid composed of suspended cellular constituents, primarily red blood cells. Moreover, at such hematocrit levels, blood behaves as an optically opaque system due to strong absorption and multiple scattering of visible light by cellular components. Consequently, the blood flow not only represents a more challenging test case for particle tracking, but also serves to highlight the advantage of X-ray-based approaches.

4.1.2 Particle-Tracking Pipeline

After conducting the experiment, we need a robust particle-tracking pipeline to convert acquired image sequences into particles' 3D trajectories. In this work, we built a pipeline based on MyPTV[62], which is an open-source toolbox originally designed for a visible light setup, but can be easily adapted for our XMPI experiment.

Figure 4.2 shows the pipeline of the microparticle tracking process, including seven main steps introduced as follows.

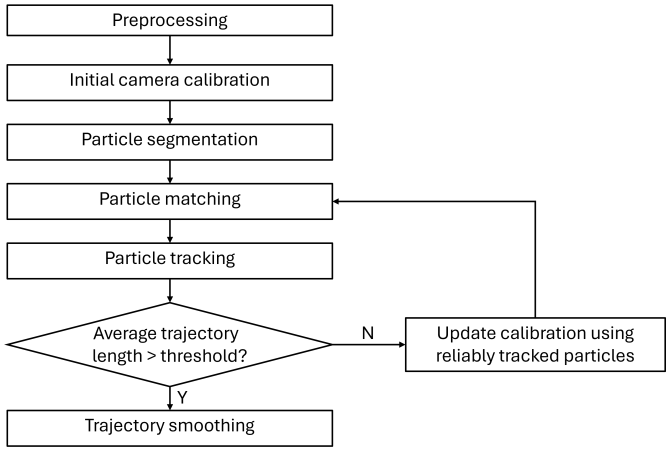


Figure 4.2: The workflow of MyPTV-based microparticle tracking.

Preprocessing

The main purpose of this step is to generate a calibration target for the step "initial camera calibration" and help detect moving particles for the step "particle segmentation".

For the former purpose, first, we manually track a particle in both views and select ten frames with a constant stride when this particle is within the field of view of both cameras. Second, we generate the calibration image for both views by summing up these ten frames so that ten different locations of the particle were clearly seen in both views. Third, we create the calibration target file by recording the particle's vertical positions.

For the latter purpose, we create a binary mask of moving objects to help MyPTV detect the particles more efficiently. The mask is created by thresholding pixels that differ by a certain percentage from the mean image of the entire set of flat-field-corrected frames (5% and 7.5%, respectively, for two detectors). Some problematic regions of the images with fluctuating pixel intensities are manually determined and removed from the binary mask.

Initial Camera Calibration

The main purpose of this step is to get a proper initial guess of the camera models for the two detectors involved in the XMPI setup. For each detector, we need to find a pinhole camera model that maps 3D world coordinates to 2D image pixels. In visible light setups, it is typically done with a standard calibration target, such as a checkerboard. Here, for the XMPI setup, we use the calibration image generated

in the preprocessing step instead, as a suitable calibration target for XMPI was not yet available when conducting the experiment. Although ten particle positions are generally not enough for a precise calibration, we can still get a reasonable initial camera model because the angles of two detectors (47.7°) in the XMPI setup are precisely known due to the diffraction condition of the used splitters. This step is performed once we obtain two camera models that satisfy the angular configuration of the XMPI setup, with an averaged calibration error of less than 1 pixel across all ten particles shown in the calibration images. It is important to note that the initial camera models will be updated later with reliably tracked particles.

Particle Segmentation

In this step, we use the masked images from the preprocessing step as the input. A local mean subtraction filter, a median noise removal filter, and a Gaussian blur filter provided in MyPTV toolbox are applied to boost the performance of particle segmentation. The output of this step is the centroids of each particle with a diameter between 3 pixels and 20 pixels in both cameras.

Particle Matching

The segmentation results from the previous step serve as input to this step. In 3D PTV, matching particles across both cameras is challenging, as the number of such combinations can be large, especially when a large number of particles exist in the field of view, making a brutal search impractical. Therefore, in this step, the "particle marching" algorithm[63] supported in MyPTV toolbox is used to address this challenge in an iterative manner. Matched particles across both cameras are then utilized to triangulate the true 3D positions using epipolar geometry. In practice, the most important parameter in this step is the maximum allowed triangulation error in lab-space coordinates, which was set to 3-5 pixels in this work. It is worth noting that increasing the number of available projections can be beneficial at this stage, as it helps reduce the false particle detections in the 3D space.

Particle Tracking

The 3D trajectories of the particles are formed by linking matched particles over time. At this step, we use the nearest neighbor method[64] provided in MyPTV.

Calibration with Reliably Tracked Particles

Long trajectories from the previous step are used to refine the camera models for each view. Practically, we select trajectories longer than 50 frames for refining the camera model. It is important to note that steps "particle matching", "particle tracking", and this step follow an iterative process. We only proceed to the next step "trajectory smoothing" if the number of trajectories and the average temporal trajectory length

reach a certain level after the "particle tracking" step.

Trajectory Smoothing

Based on the 3D trajectories provided by the "particle tracking" step, linear fitting is implemented to calculate the velocity u_z, as acceleration is negligible in the cases of particles in laminar flows studied in this work. When more complicated flows are studied, polynomial fitting should be used to calculate the velocity and acceleration of each particle.

4.2 Particle Tracking Results

In this section, we present the results obtained by the proposed 3D particle-tracking pipeline. We first track individual particles and analyze their corresponding 3D trajectories to understand their movement. Subsequently, we conduct an error analysis for each tracked particle, focusing on the uncertainties in the detected 3D positions and vertical velocities. Finally, we investigate the statistical behavior by aggregating all tracked particles. This large collection of tracer particles enables the characterization of the flow velocity field and demonstrates consistency with the error analysis.

4.2.1 Tracking Individual Particles

Using the 3D particle-tracking pipeline introduced in Section 4.1.2, we visualized and analyzed the 3D trajectory for individual particles using the following two examples.

First, we studied cylindrical capillary flows with SHGS particles flowing in glycerol at low particle concentration. As shown in Figure 4.3(a)-(b), three particles were matched in two projections, which can be visually verified by the size and the vertical position. To analyze the 3D trajectories in Figure 4.3(c), we plotted the relationship between the vertical position and the time (Figure 4.3(d)), as well as the radial position and the vertical position (Figure 4.3(e)), respectively. It can be observed that the vertical velocities and the radial positions of these three particles are independent of their vertical positions.

Second, we studied the motions of SHGS particles in human blood as an example of a dense suspension. Similarly, the visualization and analysis of 3D trajectories of three matched particles are shown in Figure 4.4. These three particles also keep a constant vertical velocity, as linearity can also be clearly observed in Figure 4.4(d). Figure 4.4(e) shows that particle 1 moves towards the center, while particle 3 moves towards the wall, indicating migration behavior of these particles.

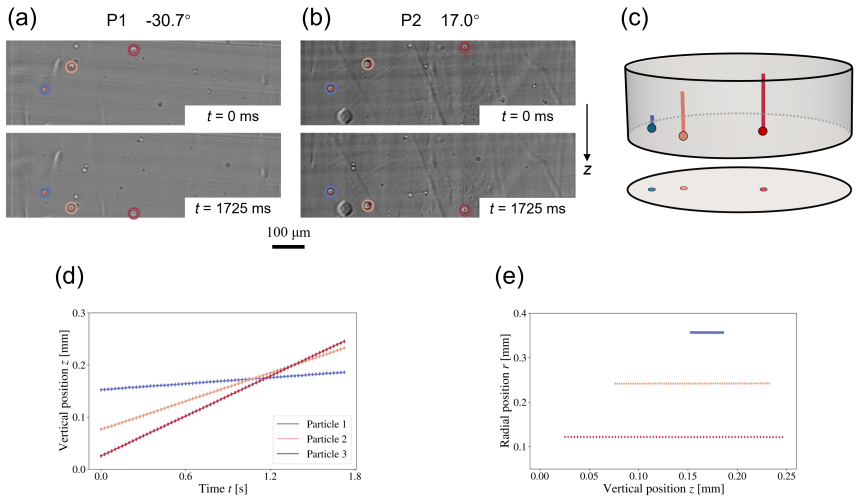


Figure 4.3: 3D individual particle tracking example 1: particles in glycerol. (a)-(b) Three SHGS particles flowing in glycerol were matched in two projections at two time points, respectively, with a red dot indicating the center position of the tracked particle. (c) Visualization of 3D trajectories between two time points shown in (a)-(b). (d) Vertical position vs. time; (e) Radial position vs. vertical position. Figure adapted from **Paper VI**.

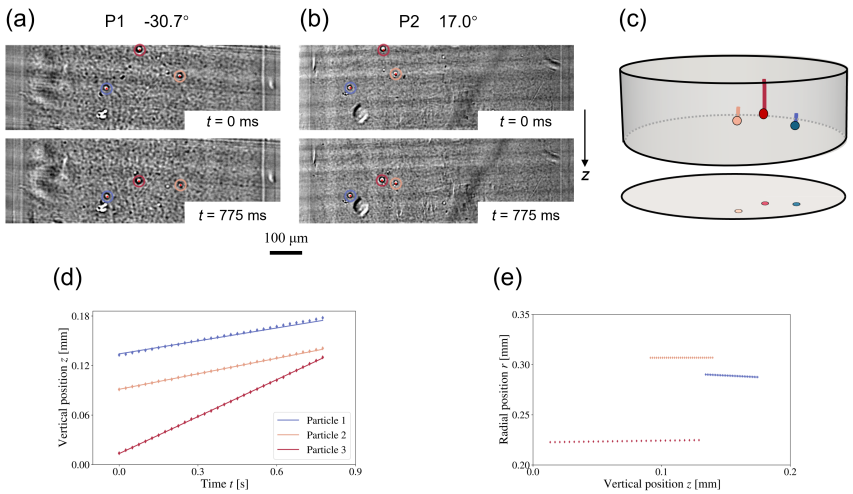


Figure 4.4: 3D individual particle tracking example 2: particles in blood. (a)-(b) Three SHGS particles flowing in blood were matched in two projections at two time points, respectively, with a red dot indicating the center position of the tracked particle. (c) Visualization of 3D trajectories between two time points shown in (a)-(b). (d) Vertical position vs. time; (e) Radial position vs. vertical position. Figure adapted from **Paper VI**.

These two examples demonstrate that XMPI works as a new 3D particle-tracking tool for both dilute and concentrated suspensions. Compared with visible-light-based approaches such as shadowgraphy, XMPI is not constrained by optical transparency and remains applicable in dense suspensions or optically opaque flows. These capabilities highlight the potential of XMPI for quantitative flow characterization in complex multiphase systems. It is important to note that the reliability of such measurements critically depends on an understanding of the associated uncertainties. Therefore, in the next subsection, we will further analyze the sources of errors and quantify their impact on the particle-tracking results.

4.2.2 Error Analysis

To perform error analysis of tracked particles, the coordinate systems, viewed from the direction perpendicular to the optical plane, are defined as shown in Figure 4.5. The red and green arrows denote the directions of two X-ray beamlets generated by our XMPI setup, respectively. Coordinates (X, Y, Z), (X_1, Y_1, Z_1), and (X_2, Y_2, Z_2) denote the lab coordinate, camera 1 coordinate, and camera 2 coordinate, respectively. φ_1 and φ_2 denote the angle between X_1 and X direction, X_2 and X direction, respectively. $\phi = \varphi_2 - \varphi_1$ denotes the angular separation of two projections in the XMPI setup, which is 47.7 degrees in the experiments conducted in this work. Note that the cameras provide only 2D images, so the X_1 and X_2 coordinates cannot be directly read from the acquired images.

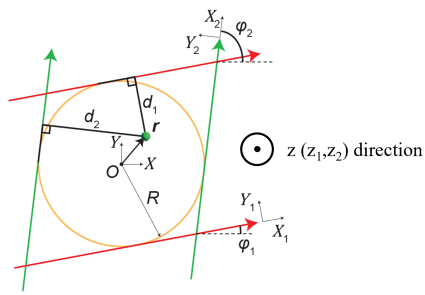


Figure 4.5: Definition of the coordinate system for PTV error analysis. Figure from **Paper VI**.

In this subsection, we first derive the uncertainty of the position of a tracked particle. Then, we can further calculate the uncertainty of particle velocity, since the velocity is determined by linear fitting of the particle trajectory.

Uncertainty of Position

Suppose we have a particle located at $r = (x, y)$ in the XY plane. From two cameras,

y_1 and y_2 can be read to determine the particle position, as it is obvious from Figure 4.5 that $y_1 = R - d_1$ and $y_2 = R - d_2$. Describing the particle position in two camera coordinates (x_1, y_1) and (x_2, y_2), respectively, we have:

$$\begin{cases} x = x_1 \cos\varphi_1 - y_1 \sin\varphi_1 = x_2 \cos\varphi_2 - y_2 \sin\varphi_2 \\ y = x_1 \sin\varphi_1 + y_1 \cos\varphi_1 = x_2 \sin\varphi_2 + y_2 \cos\varphi_2 \end{cases} \tag{4.1}$$

After eliminating x_1 and x_2, which cannot be directly read from the images, we have:

$$\begin{cases} x = \dfrac{y_1 \sin\varphi_1 + y_1 \sin(\varphi_1 - 2\varphi_2) + y_2 \sin\varphi_2 - y_2 \sin(2\varphi_1 - \varphi_2)}{\cos(2\varphi_1 - 2\varphi_2) - 1} \\ y = \dfrac{2\,(y_1 \sin\varphi_2 - y_2 \sin\varphi_1)\sin(\varphi_1 - \varphi_2)}{\cos(2\varphi_1 - 2\varphi_2) - 1} \end{cases} \tag{4.2}$$

In Equation 4.2, the uncertainty of both x and y originates from the uncertainty of y_1 and y_2. Therefore, based on the propagation of uncertainty[65], we have:

$$\begin{cases} \sigma_x = \sqrt{\left(\dfrac{\partial x}{\partial y_1}\,\sigma_{y_1}\right)^2 + \left(\dfrac{\partial x}{\partial y_2}\,\sigma_{y_2}\right)^2} \\ \sigma_y = \sqrt{\left(\dfrac{\partial y}{\partial y_1}\,\sigma_{y_1}\right)^2 + \left(\dfrac{\partial y}{\partial y_2}\,\sigma_{y_2}\right)^2} \end{cases} \tag{4.3}$$

Furthermore, for calculating the uncertainty of the radial position in the XY plane, we have:

$$\begin{cases} r = \sqrt{x^2 + y^2} \\ \sigma_r = \dfrac{\sqrt{\sigma_x^2 x^2 + \sigma_y^2 y^2}}{r} \end{cases} \tag{4.4}$$

In 3D PTV conducted in this experiment, we focus on getting the radial position of each particle, regardless of the orientation of the lab coordinate in the XY plane (defined by φ_1). Therefore, for simplicity in estimating σ_r, we select $\varphi_1 = (90^\circ - \phi)/2 = 21.15^\circ$ so that $\sigma_x = \sigma_y = \sigma_r/\sqrt{2}$.

Given that both X-ray detectors are identical and independent regarding their detection efficiency and spatiotemporal resolution despite their slightly different sample-detector distances, the uncertainties of y_1 and y_2 are the same. Given that the cameras have the same resolution in both axes, we further assume that this uncertainty is equal to σ_z, which describes the uncertainty in vertical position in the lab coordinate. We estimate σ_z by considering three main sources of error. The first is related to the instability of the beamlet, giving $\sigma_1 = 1$ pix. The second is related to the particle

segmentation step in the MyPTV analysis. Considering the difficulty of deciding the center of each particle, especially when several particles contact or interact with each other, a reasonable uncertainty of $\sigma_2 = 2$ pix is used. The third comes from the camera calibration process, resulting in $\sigma_3 = 1$ pix. Therefore, wrapping up all three sources of uncertainty, we have:

$$\sigma_{y_1} = \sigma_{y_2} = \sigma_z = \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2} = 2.45\,\text{pix} \tag{4.5}$$

Using Equations 4.2 - 4.5, the relationship of ϕ and the ratio σ_r/σ_z is shown in Figure 4.6. The uncertainty in radial position decreases when the angular separation ϕ increases towards 90°, validating that increasing the angular coverage of the XMPI can be beneficial, as discussed in Section 3.3.

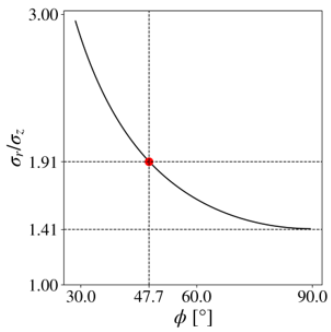


Figure 4.6: The relationship between the XMPI angular separation ϕ and the radial position uncertainty described in the ratio of σ_r/σ_z. Figure from **Paper VI**.

In this XMPI experiment, $\phi = 47.7°$, so the uncertainties of the position of the tracked particle are:

$$\begin{cases} \sigma_z = 2.45\,\text{pix} \\ \sigma_r = \sqrt{2}\,\sigma_x = \sqrt{2}\,\sigma_y = 1.91\,\sigma_z = 4.7\,\text{pix} \end{cases} \tag{4.6}$$

Uncertainty of Velocity

Suppose we have multiple points (z_i, t_i) with $i = 1, 2, \ldots, n$ for a certain trajectory, where n indicates the temporal length of the trajectory (the number of frames per trajectory); z_i is the vertical position given by the tracking results at time $t = t_i$. The vertical velocity u_z is thus the slope of the fitted line with an uncertainty σ_{u_z} given in Equation 4.7, which is dependent on both σ_z and the temporal length of the trajectory, as shown in Figure 4.7.

$$\sigma_{u_z} = \sigma_z \cdot \sqrt{\frac{1}{\sum (t_i - \bar{t})^2}} \tag{4.7}$$

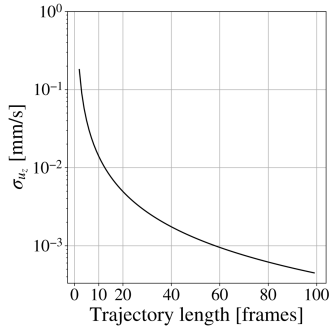


Figure 4.7: The relationship between the temporal length of a specific trajectory and its uncertainty of the vertical velocity. Figure adapted from **Paper VI**.

To sum up, in this work, we can estimate that the vertical position uncertainty is $\sigma_z = 2.45$ pix $= 3.2\,\mu$m; the radial position uncertainty is $\sigma_r = 4.7$ pix $= 6.1\,\mu$m. The uncertainty of vertical velocity is proportional to σ_z and depends on the temporal length of the trajectory. A longer trajectory results in lower velocity uncertainty.

4.2.3 Statistical Behavior of Tracked Particles

Up to now, we have demonstrated the ability to track individual micrometer-sized particles in 4D and discussed the sources of error when using XMPI. We can further extend the study from individual behavior to statistical behavior, which is crucial for researchers to understand the properties of the studied flows when small particles primarily follow the fluid motion. For instance, by tracking sufficient particles at different locations and performing statistical analysis, we can estimate the velocity field of the flow. In this subsection, we study the statistical behavior of thousands of tracked SHGS particles in glycerol through the Poiseuille flow[66] in the cylindrical capillary. In theory, the Poiseuille flow follows a parabolic axial velocity profile[67], as stated in Equation 4.8:

$$u_z = \frac{2Q}{\pi R^2}\left(1 - \frac{r^2}{R^2}\right), \tag{4.8}$$

where R denotes the radius of the cylindrical capillary, r ($r \leq R$) denotes the radial position from the center of the cylindrical capillary, u_z denotes the vertical velocity along the capillary, and Q denotes the flow rate.

Figure 4.8(a)-(b) demonstrates the 3D tracking results of 2471 SHGS particles (minimum trajectory length: 20 frames) in glycerol flowing through the cylindrical capillary at a flow rate of 0.1 mL/h and a 0.1 wt.$\%$ particle concentration. Among all 2471 trajectories, the average Pearson correlation coefficient between vertical coordinate z and time t is 0.9992 (very close to 1), indicating that the acceleration of particles

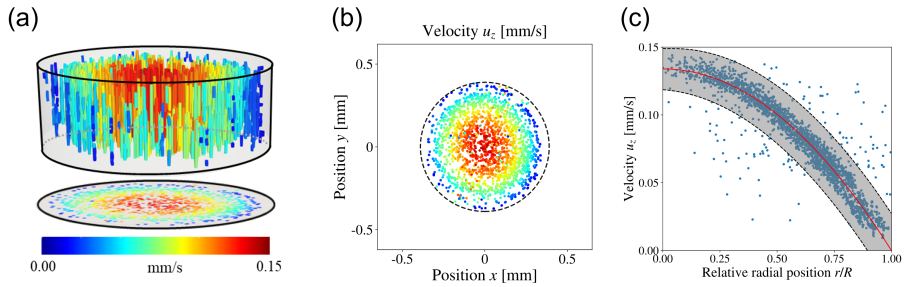


Figure 4.8: Statistical behavior of tracked particles. (a) Stacking all 3D trajectories of tracked particles; (b) Horizontal distribution of particles and their vertical velocities; (c) Vertical velocity as a function of radial position compared with the theoretical parabolic velocity profile (red curve).

in the vertical direction is negligible. Figure 4.8(c) gives the vertical velocity as a function of radial position, with a comparison of the theoretical parabolic Poiseuille flow profile (the red curve). Based on Subsection 4.2.2, a 3σ error band is created by considering both uncertainties of radial position (6.1 µm) and velocity (0.005 mm/s). As a result, 95.8% of all tracked particles are inside the error band, which validates the ability to retrieve statistical properties when conducting 3D PTV using XMPI. More experimental results with different flow rates and different particle concentrations using the same workflow as described in Section 4.1.2 can be found in the appendix of **Paper VI**. All of these experiments result in over 90% of tracked particles lying inside the 3σ error band, validating the robustness of extracting statistical behavior in XMPI-based 3D PTV study.

To sum up, the statistical analysis of a large number of tracked particles demonstrates that XMPI-based 3D PTV is capable of reliably retrieving macroscopic flow properties from microscopic particle trajectories. Using the glycerol flow as a reference case, we reached an agreement between the experimentally obtained velocity distribution and the theoretical Poiseuille flow profile, which confirms the validity of the proposed 3D particle-tracking pipeline.

4.3 Opportunities and Challenges

In this chapter, we have discussed the first 3D PTV study using the XMPI experimental setup and the MyPTV-based particle-tracking pipeline. Using the current XMPI setup at the ForMAX beamline, we validate the capability of tracking individual particles with an uncertainty of 3.2 µm regarding the vertical position and an uncertainty of 6.1 µm regarding the radial position, which further enables determining the statistical

flow properties, such as the velocity profile. Such a study shows the potential of using XMPI as a new imaging modality for quantitative 4D analysis of flow dynamics. For example, similar studies can be used to investigate various flows in more complex geometries, such as porous media or other custom-made 3D-printed geometries at the micro-meter scale. Moreover, XMPI provides unique opportunities to bridge the gap between experimental studies and particle-resolved computational fluid dynamics (CFD) simulations.

There are also several challenges regarding the 4D analysis to be addressed. First, there is still room to improve the MyPTV-based particle-tracking workflow. For camera calibration, it is beneficial to have a grid-like calibration target suitable for an X-ray imaging experiment. This will not only result in a more precise initial camera model to reduce the iterations among particle matching, particle tracking, and recalibration, but also help with the correction of commonly existing lens distortion in high-resolution X-ray detector optics used in synchrotron light sources[68]. Second, the current particle-tracking workflow primarily works for relatively low particle concentrations. At higher particle concentrations, segmentation of each particle becomes impractical. In such a scenario, alternative methods for analysis need to be considered, such as optical flow (OF)[69] and particle image velocimetry (PIV)[70]. OF can assign a 2D velocity vector to every pixel by enforcing the conservation of intensity between consecutive images to estimate the velocity, while PIV calculates the 2D velocity by tracking the average displacement of tracer particles within a small window between two frames. **Paper VI** provides both OF and PIV using 2D image sequences provided by each detector in the XMPI setup; however, combining such 2D velocity fields to a 3D velocity fields with respect to time remains challenging and requires further investigation in the near future.

Chapter 5

AI-Assisted 4D Reconstruction from Sparse Spatiotemporal Data

Up to this point, we have discussed the motivation and implementation of XMPI and introduced 3D particle tracking as an example of 4D analysis enabled by XMPI. In this chapter, we move to a more ambitious goal: full-field 4D reconstruction using image sequences acquired by only two to three detectors with a limited angular range in XMPI.

While classical iterative algorithms shown in Figures 2.6 and 2.7 are unable to properly address this severely ill-posed 4D reconstruction problem, they offer valuable insights for developing a progressively refined, self-consistent reconstruction strategy. When combined with the power of AI, or more specifically, deep learning, which provides unprecedented capability to capture complex prior knowledge for 4D reconstruction, reaching an end-to-end solution from acquired 2D image sequences directly to a 4D reconstruction becomes feasible.

This chapter is organized as follows. First, we start with an overview of deep learning and the neural network architectures that help build the 4D reconstruction solution. Second, we introduce our proposed "AI + Physics" solution to address the 4D reconstruction challenge posed by XMPI. Third, we present how the proposed solutions work for several applications in flow dynamics. This chapter is closely related to **Papers III, IV, V, and VII**.

5.1 Overview of Deep Learning

Over the recent decade, deep learning has emerged as a major research focus in AI research, owing to the revolutionary development of hardware enabling high-performance computing, the availability of massive high-quality datasets, and the development of relevant algorithms[71]. Deep learning refers to a type of machine learning that utilizes a deep neural network, inspired by the structure of the human brain, to learn from data automatically with minimum human intervention. In deep learning research, a fascinating theory is called the universal approximation theorem[72], which states that neural networks with a certain structure can approximate any continuous function to any desired degree of accuracy, providing a mathematical foundation for using neural networks in real-world applications.

Such a capability of deep learning provides invaluable insights for tackling the 4D reconstruction challenge, where we seek to reconstruct in higher dimensions (4D) from limited observations in lower dimensions (2D image sequences). The extracted spatial and temporal patterns from the given data in imaging applications, such as the shape and the motion of the studied object, are also referred to as 4D priors[73]. Such prior information can potentially provide effective constraints in the solution space, helping to offer a reliable end-to-end 4D reconstruction solution that addresses the limitation of classical reconstruction methods, where a 4D reconstruction problem is decomposed into many unrelated 3D reconstruction problems, as discussed in Chapter 2.

In this section, we will go through some basic concepts and neural network architectures in deep learning as preparation for proposing AI–based solutions for 4D reconstruction discussed later in Section 5.2.

5.1.1 Basic Concepts in Deep Learning

Loss Function

A deep learning model can be regarded as a complex (usually nonlinear) parametric system M_θ mapping the inputs and the outputs, by combining linear and nonlinear transformations. The parameter set θ contains all learnable components of the model, determining how the deep learning model processes the information. To search for the optimum θ in the deep learning model, we need to define a loss or cost function with the help of the available data. A common way to express the loss function is shown as follows:

$$\mathcal{L}(\theta) = \mathbb{E}_{(\boldsymbol{x},\boldsymbol{y})\sim p_{\text{data}}} L(M_\theta(\boldsymbol{x}), \boldsymbol{y}) \tag{5.1}$$

where $M_\theta(\boldsymbol{x})$ denotes the model output using the input $\boldsymbol{x}$; $\boldsymbol{y}$ denotes the target output; $L(\cdot)$ denotes a pointwise discrepancy measure (e.g., the mean squared error (MSE)) that quantifies the error between the model output and the target output; $\mathbb{E}_{(\boldsymbol{x},\boldsymbol{y})\sim p_{\text{data}}}(\cdot)$ denotes the expectation over the data distribution, which is usually approximated by averaging finite samples within the available dataset.

In fact, the choice of the loss function is task-oriented, which can potentially determine the performance of the neural network. In imaging, particularly in 4D reconstruction, the loss function may include additional terms, e.g., a penalty term to enhance spatial or temporal sharpness or to enforce physics-informed consistency of the reconstructed dynamics. As presented later in Section 5.2, designing the loss functions plays a crucial role in our proposed 4D reconstruction workflow.

Model Training

Mathematically, the deep learning process can be formulated as an optimization problem of finding optimal model parameters by iteratively minimizing the loss function. This minimization procedure is also referred to as model training. During model training, the most important step is to calculate the gradient of the loss function with respect to the model parameters $\nabla_\theta \mathcal{L}(\cdot)$, using backpropagation[74], a mechanism that applies the chain rule throughout the deep learning model. Since the gradient determines how small changes of each parameter influence the loss function, it provides guidance for parameter updates. Gradient-based optimizers[75], such as stochastic gradient descent (SGD) or the adaptive moment estimation (ADAM) optimizer, can then be implemented to refine the model. By repeatedly computing gradients and adjusting the parameters accordingly until the loss function falls below a set threshold, we complete the model training.

In practice, the model training requires a huge amount of computational resources and is often performed on graphics processing units (GPUs). For all AI-assisted works presented in this thesis, we utilized NVIDIA V100 GPUs with 32 GB of RAM installed in the high-performance cluster at MAX IV[76].

5.1.2 Neural Network Architectures

In this section, we introduce some basic neural network architectures that are commonly used in imaging tasks, which serve as basic elements of building the 4D reconstruction solution for X-ray imaging presented later in Section 5.2.

Neuron - the Fundamental Unit

To start with, we introduce the neuron (also known as the perceptron), which is the

fundamental unit of any deep neural network. The structure of a neuron is illustrated in Figure 5.1, including an input vector $\boldsymbol{x}$, a weight vector $\boldsymbol{w}$, a bias value b, an activation function f, and an output value y. The formalism can be expressed as Equation 5.2:

$$y = f(\boldsymbol{w} \cdot \boldsymbol{x} + b). \tag{5.2}$$

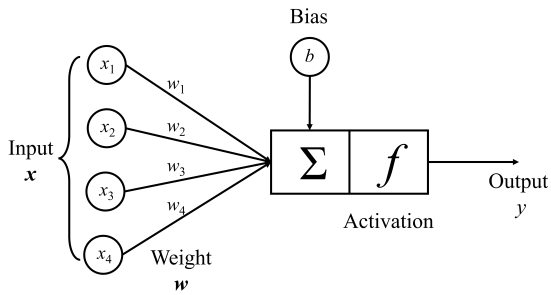


Figure 5.1: Illustration of a neuron - the fundamental unit of a neural network.

It is important to note that the activation function f needs to be non-linear; otherwise, no matter how many neurons are connected to each other, it would collapse into an equivalent single linear transformation, limiting the representation ability. Commonly-used activation functions include the sigmoid function $f(x) = 1/(1 + e^{-x})$ and the ReLU function $f(x) = \max(0, x)$.

FCNN and CNN

Neurons are organized into layers, which are interconnected to create the neural networks. Two of the most basic architectures of neural networks are fully-connected neural network (FCNN)[72] and convolutional neural network (CNN)[77].

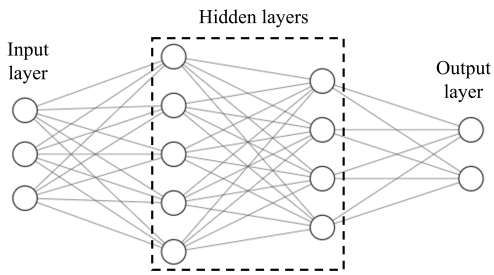


Figure 5.2: Example of an FCNN consisting of one input layer, two hidden layers, and one output layer.

An FCNN is characterized by its dense interconnectivity, which means that every neuron in one of the layers (except the output layer) interconnects with every neuron

in the subsequent layer. Figure 5.2 shows an example of FCNN, consisting of an input layer (containing three neurons), an output layer (containing two neurons), and two hidden layers in between. The hidden layers are responsible for learning complex representations from the input data. Usually, the early hidden layers are responsible for capturing simple patterns or structures, while the latter hidden layers utilize the simple patterns to learn more complex representations. Such a hierarchical feature extraction process enables the FCNN to reliably generate the output, which is helpful for implementing the concepts of neural implicit representation (NIR) and physics-informed neural network (PINN) to be discussed later in this section.

FCNN also has its limitations, especially when images serve as the input. Let us consider a scenario where a 2D image is flattened to feed into the hidden layers. Fully-connected hidden layers not only induce a huge number of trainable parameters posed by the interconnectivity, but also ignore the spatial structure of images, meaning that we might lose valuable information from the available data.

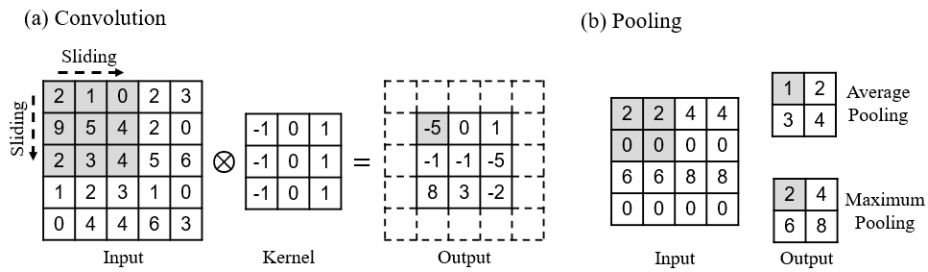


Figure 5.3: Example of a convolution layer (a) and a pooling layer (b) used in CNN.

Such limitations motivate the design of CNN, where we replace some early fully-connected layers with sequences of convolutional layers and pooling layers based on locally-connected operations, as shown in Figure 5.3(a) and (b), respectively. The convolutional layers apply a set of learnable kernels that slide across the image, producing feature maps that respond to local structures of the image, such as edges and contours[78]. Beyond that, introducing convolutional kernels (usually with a size of 3×3 or 5×5) also significantly reduces the number of parameters compared to fully-connected layers, as weights are shared across spatial locations rather than learned independently for every pixel. The pooling layers perform spatial downsampling by aggregating local regions, usually by taking the maximum or the average. This operation reduces the size of intermediate feature maps while retaining the important features extracted, making the workflow even more computationally efficient.

Generative Adversarial Network (GAN)

Introduced by Ian Goodfellow et al. in 2014, generative adversarial network (GAN)[79] quickly became one of the highlights in deep learning research, due to its eye-catching

structural design and its outstanding generative capability. GAN follows a two-network architecture, including a generator network and a discriminator network. Both FCNN and CNN can be used to implement these components, although in imaging applications, it is more common to employ CNN for the discriminator due to the ability to capture spatial structures and local features.

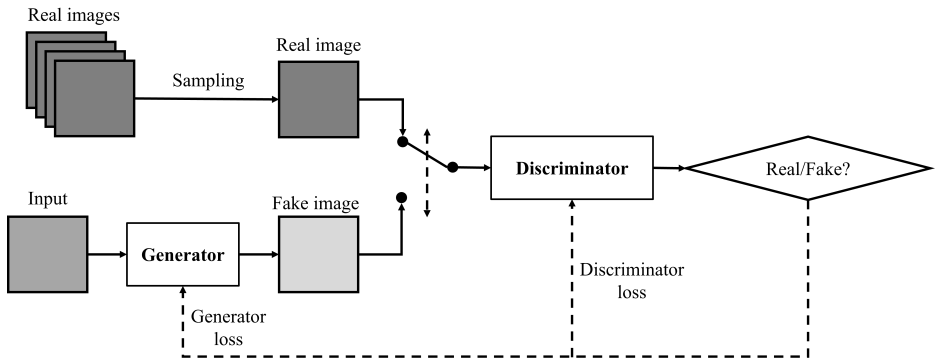


Figure 5.4: The general workflow of GAN. Figure adapted from Ref. 80.

The workflow of GAN is depicted in Figure 5.4. In essence, two competing neural networks are engaged in a zero-sum game: the generator tries to minimize the chance that the discriminator detects fake images, while the discriminator tries to maximize its detection accuracy. Both networks are trained simultaneously based on their respective loss functions until an equilibrium is reached, at which point the discriminator can no longer correctly identify the generated images, and we obtain high-quality images from the well-trained generator.

GAN serves as a crucial component when designing 4D reconstruction methods using sparse views, as presented in **Papers III and IV**.

Neural Implicit Representation (NIR)

In imaging applications, the information of an object is typically stored explicitly using a discretized pixel (2D) or voxel (3D) representation. However, such grid-based representations become inefficient as the required resolution or dimensionality increases. In high-dimensional applications (e.g., 3D or 4D), the number of required grid points grows exponentially This phenomenon leads to huge memory consumption and computational cost, making grid-based representation impractical in many applications.

To address these limitations, NIR provides a more elegant and flexible alternative. Instead of storing data explicitly on discrete grids, an implicit deep learning model learns a continuous function of the coordinates (e.g., a function of space coordinate

(x, y, z) for 3D applications or a function of space-time coordinate (x, y, z, t) for 4D applications). Specifically, in the 4D reconstruction task presented in this chapter, we need to learn the mapping from the 4D coordinate to the index of refraction of the studied sample:

$$n = N_\theta(x, y, z, t), \tag{5.3}$$

where $N_\theta(\cdot)$ denotes the deep learning model with a parameter set θ, and n represents the index of refraction characterized by δ and β, as discussed in Equation 2.2. Compared to grid-based representations, this coordinate-based representation $N_\theta(\cdot)$ is more memory-efficient and resolution-independent. Namely, once such a continuous function is successfully learned, we can sample it according to our desired (spatial or temporal) resolution.

The ability NIR to represent complex objects as continuous functions is supported by the universal approximation theorem. This theoretical foundation explains why NIR can capture fine-scale structures and complex geometries without explicit discretization, making it particularly suitable for high-dimensional imaging problems where grid-based representations suffer severely from the curse of dimensionality.

Neural radiance field (NeRF)[81] is one of the influential works using the concept of NIR. It represents a 3D scene for a visible-light setup by using an FCNN that maps spatial coordinates and viewing directions to volumetric density and radiance. During model training, NeRF optimizes this implicit function to match multi-view observations, enabling reliable 3D reconstruction from sparse 2D observations. The concept of NIR serves as a key aspect for the 4D reconstruction framework for X-ray imaging, as discussed later in Section 5.2.

Physics-Informed Neural Network (PINN)

While NIR provides a flexible way to represent an object or a physical field as a continuous implicit function of coordinates, the training of such a function is usually challenging due to a large solution space. As a result, the learned function $N_\theta(x, y, z, t)$ alone does not naturally guarantee physical plausibilities. To address this issue, the concept of PINN[82] is proposed, which incorporates known physical laws directly into the learning process to constrain the solution space. The physical laws are typically expressed as a partial differential equation (PDE):

$$\mathcal{F}\big(N_\theta(x, y, z, t)\big) = 0, \tag{5.4}$$

where $\mathcal{F}$ denotes the differential operator.

A key enabler of PINN is automatic differentiation (AD)[83], which allows the derivatives required by $\mathcal{F}$ to be computed exactly with respect to either spatial or temporal coordinates. To integrate the PDE constraint into the model learning process, the

PINN loss can be designed as:

$$\mathcal{L}_{\mathrm{PDE}} = \mathbb{E}_{\boldsymbol{x}\sim\Omega}\Big[\mathcal{F}\big(N_\theta(x,y,z,t)\big)^2\Big] \tag{5.5}$$

where Ω denotes the continuous domain over which the physical law is enforced. In practice, this means that a set of points $(x, y, z, t) \in \Omega$ is sampled within the domain, and automatic differentiation is used to evaluate the differential operator $\mathcal{F}$ at those points, enabling the neural network to satisfy the PDE constraint throughout the entire continuous domain.

The power of utilizing the concept of PINN is demonstrated later in Section 5.2.2 and in **Paper IV**.

5.2 "AI + Physics" 4D Reconstruction Framework

Up to this point, we have discussed the fundamentals of X-ray imaging physics (Chapter 2) and deep learning (Section 5.1). In this section, we focus on combining the power of AI and physics to address the 4D reconstruction challenges posed by the sparse-view measurements from the XMPI setup. As discussed previously, in such 4D reconstruction problems, only a very limited number ($\leq$ 3) of projections are available at each time point, making the reconstruction problem severely ill-posed if each time frame is treated independently. In fact, the objective of 4D reconstruction extends beyond recovering the spatial structures of the studied sample; it also aims to capture the temporal evolution of the studied system. Simply stacking independently reconstructed objects at successive time points may lead to spatiotemporally inconsistent or physically implausible 3D movies. Therefore, these considerations motivate the development of reconstruction approaches that operate directly in the 4D domain, rather than relying on a volume-by-volume reconstruction scheme.

Here, we will present three solutions: 4D-ONIX (4D Optimized Neural Implicit X-ray Imaging), 4D-PIONIX (4D Physics-Informed Optimized Neural Implicit X-ray Imaging), and X-Hexplane. The core of all three methods is to utilize NIR to progressively learn the mapping from the 4D coordinate (x, y, z, t) to the index of refraction $n(\delta, \beta)$ of the studied dynamic object. In the X-ray regime, the electron density is proportional to the refractive index decrement terms δ and β[84], which in turn scales with the macroscopic mass density of the material. Hence, reconstructing the index of refraction is equivalent to reconstructing the material composition within the studied object. All three methods are capable of reconstructing dynamics in 4D in an "end-to-end" manner directly from 2D image sequences acquired by XMPI setup, but they lie in different "comfort zones".

5.2.1 4D-ONIX

In this subsection, we present the workflow of 4D-ONIX. As shown in Figure 5.5(a), 4D-ONIX consists of two main neural networks: a 4D generator and a discriminator. The 4D generator $N_\theta(x, y, z, t)$ is formed by an FCNN, generating the mapping from the 4D coordinate to the index of refraction as shown in Equation 5.3. As discussed in Chapter 2, the index of refraction dictates how X-rays propagate through the matter. In other words, it performs as a "bridge" between the 3D object and the 2D projection image. Once we sample sufficient query points along a ray within the object in a certain direction at a given time, as shown in Figure 5.5(b), we can compute the line integral based on the projection approximation to form a pixel in the projection plane. By assembling all the rays along one direction, the predicted projection image is formed. In principle, such a projection image can be either absorption-contrast or phase contrast, but we only showcase the absorption-contrast in the context of this work. When a generated projection image is obtained from an angle at which we have collected a real image, we compare them directly to enforce self-consistency; when a real image is unavailable, we enforce random-view consistency with the help of the discriminator.

The discriminator is formed by a CNN. The discriminator aims to distinguish image patches[85] extracted from real (measured) projection images from those generated by the generator. Using feedback from the discriminator, the generator can be trained to provide a higher-quality 4D representation, leading to more indistinguishable projection images for the discriminator. It is important to note that introducing adversarial learning also helps achieve multiple-experiment consistency in the 4D reconstruction. For example, when several similar samples or dynamics are studied using XMPI, all the collected 2D projections can potentially contribute to the competition between the generator and the discriminator, which results in better performance of the 4D generator. We will discuss this point in more detail in Section 5.3.1.

In 4D-ONIX, two different loss functions are involved to constrain the 4D reconstruction. The first is the self-consistency loss (Equation 5.6), aiming to compare the pixel-wise difference between the real projection image and the generated projection image:

$$\mathcal{L}_{\mathrm{MSE}} = \sum_{\nu \in A} \|\mathbf{c}_v - \hat{\mathbf{c}}_\nu\|_2^2 \,, \tag{5.6}$$

where A denotes the set containing available angles from which the projections are recorded by the XMPI setup; $\mathbf{c}_v$ and $\hat{\mathbf{c}}_\nu$ denote the 2D image of the real and generated projections, respectively. Specifically, each pixel value $\hat{c}_\nu(x_\mathrm{c}, y_\mathrm{c}; t)$ in image $\hat{\mathbf{c}}_\nu$ at a given time point t can be expressed as follows:

$$\hat{c}_\nu(x_\mathrm{c}, y_\mathrm{c}; t) = \mathcal{P}_\nu(N_\theta(x, y, z, t)) \tag{5.7}$$

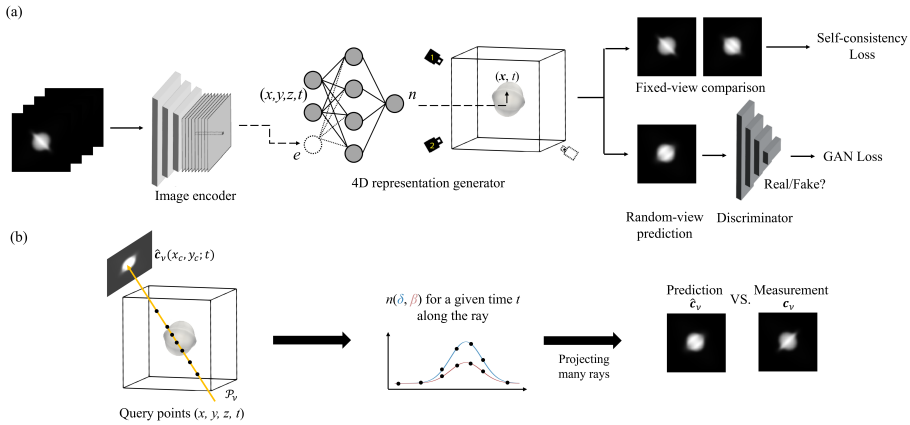


Figure 5.5: The workflow of 4D-ONIX. (a) The 4D representation generates the mapping from a space-time coordinate with optional encoded information to the index of refraction of the sample. The representation is constrained by i) self-consistency between the generated and recorded 2D image at given angles, and ii) feedback from the discriminator based on the generated 2D image at random angles. (b) The process of generating a predicted projection image. Figure adapted from **Papers III and IV**.

where (x_c, y_c) denotes the 2D camera coordinate; and $\mathcal{P}_\nu$ denotes the projection operator at a given angle ν, as illustrated in Figure 5.5(b).

The second is the GAN-loss (Equation 5.8), aiming to evaluate whether a generated projection image is feasible or not:

$$\mathcal{L}_{\text{GAN}} = \mathbb{E}_{\mathbf{c}_\nu \sim p_D} \log(\mathbf{D}(\mathbf{c}_\nu)) + \mathbb{E}_{\hat{\mathbf{c}}_\nu \sim p_\nu} \log(1 - \mathbf{D}(\hat{\mathbf{c}}_\nu)), \tag{5.8}$$

where $\mathbb{E}$ denotes the expectation; p_D and p_ν denote the data distribution over the real (measured) projection images and projection images generated by 4D-ONIX, respectively; $\mathbf{D}$ denotes the discriminator.

Additionally, 4D-ONIX provides the option to include latent features of the sample[17] extracted from the input projection images using a CNN as a third neural network, as shown in the left-hand side of Figure 5.5(a). Such an option enhances the representation capability, which is particularly useful when dealing with imaging objects with complex structures.

5.2.2 4D-PIONIX

In this subsection, we present the workflow of 4D-PIONIX, the successor of 4D-ONIX. As mentioned previously, achieving a physically plausible 3D movie is one of the main considerations in 4D reconstruction. In a dynamic system, the governing

full-physical model is typically expressed as a series of PDE, which can work as a crucial sample prior for the 4D reconstruction task.

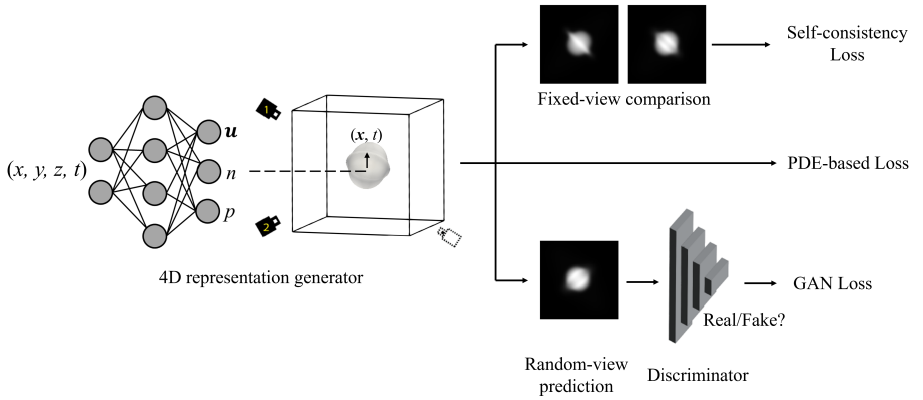


Figure 5.6: The workflow of 4D-PIONIX. Compared with 4D-ONIX, the 4D representation is further constrained by the PDE-based loss from the full physical model of the studied dynamic system. Figure from **Paper IV**.

4D-PIONIX, as shown in Figure 5.6, has an additional constraint to the 4D reconstruction using the full physical model of the dynamics. To include this additional constraint, we modify the output layer of the 4D generator by adding auxiliary variables involved in the full physical model. In fluid dynamics applications shown later in Section 5.3.1, the velocity field $\boldsymbol{u}$ and the pressure field p are included. By exploiting automatic differentiation, we can establish $\mathcal{L}_{\mathrm{PDE}}$ as a third term in the loss function according to the PDE that describes the dynamics. For example, $\mathcal{L}_{\mathrm{PDE}}$ can be built upon the Navier-Stokes equation in fluid dynamic applications. In Section 5.3.1, we will present how $\mathcal{L}_{\mathrm{PDE}}$ is established in the droplet-collision example and validate the power of implementing it.

5.2.3 X-Hexplane

Although both 4D-ONIX and 4D-PIONIX can provide an end-to-end solution for the 4D reconstruction challenge posed by XMPI, they are computationally heavy as they learn the representation of the index of refraction directly in the 4D space. Inspired by TensoRF[86] and Hexplane[87], two pioneering works designed for visible-light setups that demonstrate the power of factorizing the 4D scene tensor into multiple compact, low-rank tensor components, we incorporate essential adaptations for XMPI and establish X-Hexplane.

X-Hexplane modifies the 4D generator of 4D-ONIX and 4D-PIONIX by decomposing a 4D spacetime grid into six feature planes (XY, ZT, XZ, YT, XT, YZ) spanning

each pair of coordinate axes, as shown in Figure 5.7. Each sampled point (x, y, z, t) is projected onto all six feature planes. Bilinear interpolation is applied to extract six corresponding feature vectors, which are then aggregated to form a fused feature vector, followed by an FCNN to generate the index of refraction. Such an architecture jointly encodes spatial structures and temporal evolution and allows the network to share information across different time steps, making it suitable for 4D reconstruction under XMPI configurations where spatial information at each time point is limited. To optimize X-Hexplane, we keep using the self-consistency loss as done in 4D-ONIX and 4D-PIONIX (Figure 5.5(b) and Equation 5.6). In practice, such a modification of the 4D generator usually leads to a significantly faster training process (approx. 20 minutes) compared to 4D-ONIX or 4D-PIONIX (approx. 1 day), which potentially opens up new possibilities of implementing online 4D reconstruction during XMPI experiment.

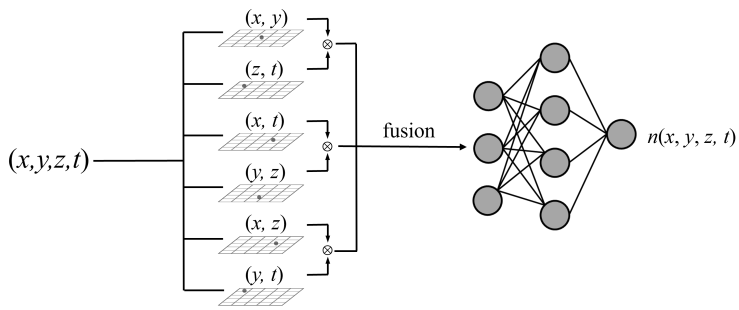


Figure 5.7: 4D generator used in X-Hexplane. Each sampled point (x, y, z, t) is projected onto six feature planes and fused to form a feature vector before feeding into the FCNN. Figure adapted from **Paper V**.

5.3 Applications

In this section, we focus on two representative applications of fluid dynamics that benefit from the development of XMPI and discuss the 4D reconstruction results provided by our proposed "AI + Physics" solutions. These applications also help us understand the "comfort zones" of each reconstruction method more clearly and provide insights for future XMPI experiment and reconstruction strategies.

The first example is the 4D reconstruction of the droplet collision process, which is important for various natural and industrial processes, such as raindrop formation [88] and drug delivery [89]. Most studies focusing on droplet collisions are conducted in 2D with visible-light setups [55], such as shadowgraphy, which generally suffer from a lack of volumetric information about the collision dynamics and from opacity to visible light. Such limitations can be addressed using XMPI thanks to its rotation-free property

and its ability to penetrate. For this example, we simulate the XMPI experiment to observe the droplet collision process, and then validate the 4D reconstruction using 4D-ONIX and 4D-PIONIX described in Section 5.2.1 and 5.2.2.

The second example is to reconstruct and analyze the process of flows penetrating through porous media in 4D. Studying liquid movement in porous media is important across various fields, such as fuel cells[90] and carbon storage[91]. In such studies, one key mechanism during the penetration through porous media is called the "Haines Jump"[92], describing instantaneous filling of pores once a threshold pressure[93] is overcome. XMPI provides unique capabilities to investigate the "Haines Jump" in 3D in the subsecond timescale. For this example, we present the XMPI experiment, the subsequent 4D reconstruction using X-Hexplane (described in Section 5.2.3), and the analysis of the pore filling process.

5.3.1 Droplet Collision

Regarding the droplet collision procedure as a first validation of the proposed 4D reconstruction workflows, we consider a controlled and reproducible test case that provides access to both 4D ground truth and corresponding 2D projection images compatible with an XMPI acquisition geometry. Numerical simulations naturally fulfill these requirements by offering full control over the underlying dynamics while providing reference solutions. In this subsection, we investigate the 4D reconstruction of the simulated droplet collision procedure, which serves as a physically relevant benchmark for ultra-fast dynamics. We first describe the simulation procedure used to generate the 4D ground truth and the corresponding 2D projection data. Subsequently, we introduce the quantitative metrics for evaluation and compare the performance of 4D-ONIX and 4D-PIONIX.

Simulation

The simulation was conducted in two steps. First, we simulated the droplet collision process in the 4D domain, where two identical droplets at a diameter of $80\,\mu\mathrm{m}$ collide head-on at a constant speed. This process is governed by the following non-dimensionalized Navier-Stokes equation for the incompressible fluid with potential surface tension $\eta\nabla\psi$ (η denotes the chemical potential as defined in Cahn-Hilliard equation[94]):

$$\rho(\psi)\big(\partial_t \boldsymbol{u} + \boldsymbol{u}\cdot\nabla\boldsymbol{u}\big) - \frac{\mu(\psi)}{\mathrm{Re}}\nabla\cdot\nabla\boldsymbol{u} + \nabla p + \frac{\eta\nabla\psi}{\mathrm{We}} = \mathbf{0} \tag{5.9}$$

$$\nabla\cdot\boldsymbol{u} = 0 \tag{5.10}$$

In Equations 5.9 and 5.10, $\psi \in [-1, 1]$ is the phase variable, with $\psi = 1$ representing

pure water, $\psi = -1$ representing pure air, and $\psi \in (-1, 1)$ representing a combination phase of water and air; Re and We denote non-dimensional Reynolds number and Weber number, respectively; $\boldsymbol{u}$ and p denote the vectorial velocity and scalar pressure fields, respectively; the densities (ρ) and the viscosities (μ) are expressed as a function of ψ:

$$\rho(\psi) = \frac{1}{2}\big((1+\psi)\rho_1 + (1-\psi)\rho_2\big) \quad \text{and} \quad \mu(\psi) = \frac{1}{2}\big((1+\psi)\mu_1 + (1-\psi)\mu_2\big) \tag{5.11}$$

To numerically solve the phase variable ψ and the field variables $\boldsymbol{u}$ and p, we used the open-source framework DUNE[95,96]. The 4D simulation contains 75 time points or frames in total, and the time difference between two adjacent frames is 0.075 µs. At each time point, the simulated 3D object contains $128 \times 128 \times 128$ voxels with a voxel size of 4 µm, corresponding to a size of 0.5 mm^3. Such settings match the targeted field of view (~ 1 mm^2), acquisition rate (beyond kHz), and spatial resolution (sub-10 µm) of state-of-the-art XMPI experiments. The 3D objects at different time points illustrating different stages of the droplet collision process are shown in the first row of Figure 5.8.

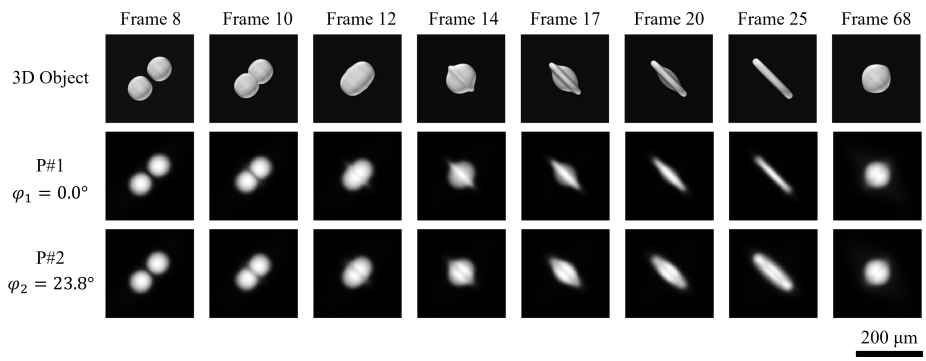


Figure 5.8: Examples of simulated 3D objects and projection images at eight different time points. The first row shows the simulated 3D object. The second and the third rows show the simulated 2D projection images at $\varphi_1 = 0^\circ$ and $\varphi_2 = 23.8^\circ$, respectively. Figure from **Paper IV**.

Second, we simulated the XMPI acquisition. Based on the 4D simulation, projection images with a pixel size of 4 µm were generated using the projection approximation. We mimicked the challenging conditions of existing MHz XMPI experiments[55] with only two projections (23.8° apart) per time point, as shown in the last two rows of Figure 5.8. To analyze the capabilities of both 4D-ONIX and 4D-PIONIX, we established three datasets with different sparsity in terms of time points based on the same XMPI experiment. Dataset 1 (75-frame dataset) contains projection images for all 75 time points, corresponding to a frame rate of 13.3 MHz; Dataset 2 (15-frame

dataset) is a subset of the first dataset where we took the first dataset at a stride of 5 time points, corresponding to a frame rate of 2.7 MHz; Dataset 3 (16-exp dataset) contains the largest number of projections by reproducing 16 similar XMPI experiments with unknown relative angles among these experiments.

Quantitative Metrics for 4D Reconstruction

To evaluate the quality of the 4D reconstruction, we used three quantitative metrics comparing the reconstruction and the ground truth provided by the 4D simulation for all 75 time points: MSE, data structural similarity index measure (DSSIM)[97], and the estimated resolution based on FSC with the half-bit threshold criterion[41]. The MSE and DSSIM indicate the accuracy of the 4D reconstruction; while the FSC indicates the 3D resolution of the reconstruction at each time point. For all three quantitative metrics, lower values indicate better reconstruction quality.

Comparison of 4D Reconstruction Results

Here, we present the 4D reconstruction results provided by 4D-ONIX and 4D-PIONIX, respectively, to get a better understanding of their "comfort zones". Specifically, in this droplet collision case, to implement 4D-PIONIX, the PDE-loss in Figure 5.6 is built on Equations. 5.9 and 5.10:

$$\mathcal{L}_{\mathrm{PDE}} = \left\| \rho(\psi)\left(\partial_t \boldsymbol{u} + \boldsymbol{u}\cdot\nabla\boldsymbol{u}\right) - \frac{\mu(\psi)}{\mathrm{Re}}\nabla\cdot\nabla\boldsymbol{u} + \nabla p + \frac{\eta\nabla\psi}{\mathrm{We}} \right\|_2^2 + \|\nabla\cdot\boldsymbol{u}\|_2^2 \tag{5.12}$$

First, we compare the performance of 4D-ONIX on Datasets 1 (containing 1 experiment and 75 frames) and 3 (containing 16 experiment and 1200 frames), respectively. The reconstructed objects at several time points are shown in the rows (4) and (5) of Figure 5.9(a), respectively, which clearly shows that using Datasets 3 with multiple similar XMPI experiments leads to a significantly better 4D reconstruction. The reason lies in the adversarial learning that is included in the 4D-ONIX workflow (Figure 5.5(a)). By combining similar experiments into the dataset, more projection images are available to the discriminator, which in turn helps the 4D generator learn from different experiments. This result also indicates a potential path to addressing the challenge posed by sparse projections in XMPI by aggregating similar experiments to enable the AI-based reconstruction algorithm to learn the pattern across them.

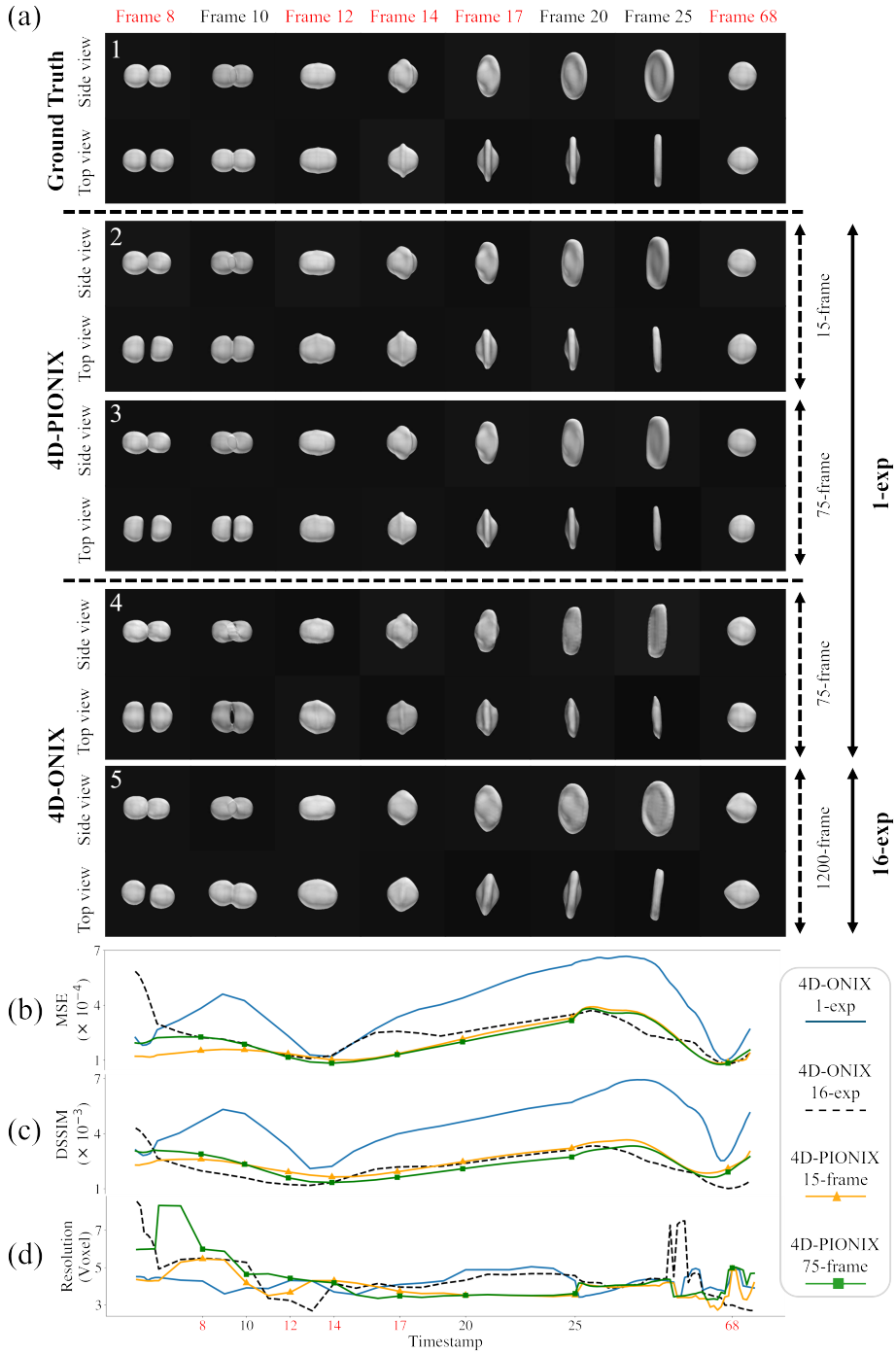


Figure 5.9: Reconstruction results. (a) Ground truths (1) and reconstructions using 4D-PIONIX, 15-frame dataset (2); 4D-PIONIX, 75-frame dataset (3); 4D-ONIX, 1-experiment dataset (4); 4D-ONIX, 16-experiment dataset (5) at eight time points. At the time points marked in red, projection images are unavailable in the 15-frame dataset, while projection images are available for all four datasets at the time points marked in black. (b)-(d) Comparison of the distribution as a function of time of 3D MSE (b), 3D DSSIM (c), and 3D resolution estimated by FSC analysis (d) under all four reconstruction settings. Figure from **Paper IV**.

Second, we compare the performance of 4D-PIONIX on Datasets 1 (containing 1 experiment and 75 frames) and 2 (containing 1 experiment and 15 frames), respectively. The reconstructed 3D objects using Datasets 2 and 1 are shown in rows (2) and (3) of Figure 5.9(a), respectively, where two datasets lead to similar reconstruction quality comparable to the 4D-ONIX reconstruction using Dataset 3, the dataset containing much more information for reconstruction. It is also important to note that projection images at time points marked in red in Figure 5.9 are not available in Dataset 2, but 4D-PIONIX is still able to provide reliable reconstruction at these time points. Furthermore, in order to compare the distributions of the 3D metrics (3D-MSE, 3D-DSSIM, and 3D resolution) with time, Figures 5.9(b)-(d) are given. It shows that, except for the case of using 4D-ONIX on single-experiment data, all three reconstruction conditions yield high-quality 4D reconstructions across space and time. In other words, 4D-PIONIX provides high-quality 4D reconstruction comparable to 4D-ONIX, even when 80 times (1200 vs. 15) more projection images are available for the latter, showing the power of including the physical model of the studied dynamics into the reconstruction workflow via the PDE-based loss.

In summary, both 4D-ONIX and 4D-PIONIX can provide reliable 4D reconstruction despite the challenges posed by limited views in XMPI. From the application perspective, 4D-PIONIX is preferable when the full physical model of the studied dynamics is available, or at least partially available. In particular, compared to 4D-ONIX, 4D-PIONIX enables reconstructing non-reproducible 4D dynamics, where repeated experiments are not feasible. Conversely, when the governing physics of the dynamics is unavailable but the studied dynamics are fully or partially reproducible, 4D-ONIX can provide a robust 4D reconstruction solution due to its ability to learn across multiple experiments. It is worth noting that both 4D-PIONIX and 4D-ONIX may experience relatively long training times. For 4D-PIONIX, the calculation of the PDE loss, which usually requires automatic differentiation of higher-order (≥ 2) derivatives, can be computationally demanding. For 4D-ONIX, the training time generally scales with the size of available datasets. Such limitations motivate the development of X-Hexplane, as will be presented in the following application.

5.3.2 Flows in Porous Media

In this subsection, we extend the proposed 4D reconstruction framework to the study of flows through porous media and use this application to validate X-Hexplane, the most computationally efficient workflow among the presented 4D reconstruction methods. The validation is performed using data acquired in a real XMPI experiment conducted at the ForMAX beamline of MAX IV, thereby demonstrating the applicability of the proposed approach under realistic experimental conditions. There are

two main highlights to be noted. First, from the experimental perspective, inspired by rotation-based tomoscopy, we applied a slow and continuous rotation to the sample stage, which does not perturb the intrinsic pore-filling dynamics of interest. Although we claim that XMPI is a rotation-free technique, this adaptation demonstrates that richer spatial information can be acquired to boost the performance of the 4D reconstruction from a single experiment. Second, beyond showing the reconstructed 3D movie, we also demonstrate a subsequent quantitative analysis of the pore-filling dynamics enabled by the 4D reconstruction, showing the potential of achieving scientific insights through an integrated workflow of XMPI experiment and its tailored 4D reconstruction and analysis.

XMPI Experiment

Figure 5.10 illustrates the schematic of the experimental setup for the study of flows penetrating through porous media. Two beamlets were established, covering an angular range of 47.7°. Two X-ray detectors worked at an effective pixel size of 1.3 µm and an acquisition rate of 50 Hz. The sample is the additive-manufactured porous network, as shown on the right-hand side of Figure 5.10. A polished PMMA cylinder is used as an X-ray transparent sample holder. The flow was driven by a syringe pump to create a constant flow of deionized water. To enrich the spatial information acquired by the XMPI setup without altering the studied dynamics, we placed the sample on a rotational stage whose rotational axis was aligned with the intersection point of two beamlets and allowed the sample to rotate slowly and continuously during the image acquisition.

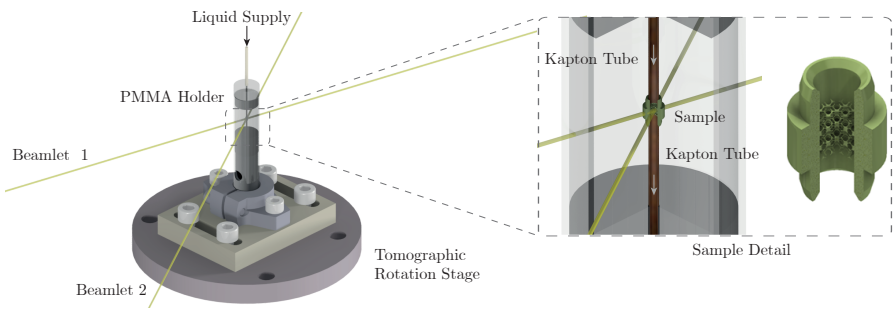


Figure 5.10: Schematic of the XMPI experiment for studying flows in porous media. Figure adapted from **Paper VII**.

Using the experimental setup depicted in Figure 5.10, we acquired movies from both cameras simultaneously, with the sample stage rotating at a constant speed of 12°/s. The entire movie can be categorized into two stages. During stage 1, no flow enters, and only rotation of the sample can be observed. During stage 2, the flow enters so that both the investigated dynamic behavior and the rotation can be clearly observed.

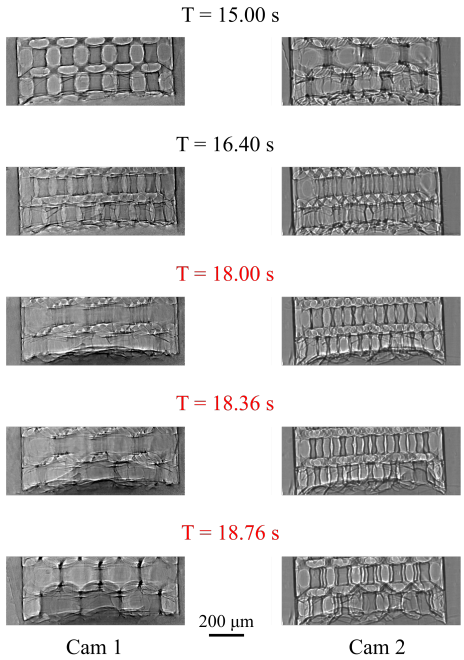


Figure 5.11: Flat-filed corrected images at several time points acquired by the XMPI setup. The first two time points (marked in black) lie in stage 1 before the flow arrives, while the last three time points (marked in red) lie in stage 2 after the flow arrives. The original time of the experiment is used to show stages 1 and 2. Figure from **Paper VII**.

Although the main scope of this experiment is to study the dynamic behavior during stage 2, images at stage 1 still play a crucial role, as they help ensure the self-consistency of the images collected from different cameras in the pre-processing step as discussed in more details in **Paper VII**, and can significantly boost the performance of the 4D reconstruction. Examples of flat-field corrected[39] images from both cameras at some typical time points are shown in Figure 5.11, where the time points marked in black are within stage 1, while the time points marked in red are within stage 2.

4D Reconstruction and Analysis

We use the X-Hexplane workflow described in Section 5.2.3 to reconstruct the pore filling dynamics for the first layer of the porous network in 4D. It is important to note that our reconstruction yields a 3D movie at a frame rate of 50 Hz, matching the acquisition rate of the X-ray detector but independent of the rotational speed of the sample stage. The reconstructed volumes at several time points are shown in Figure 5.12(a). The 4D reconstruction reveals that the imbibition process of the first layer starts with pores 6, 7, and 5, which are completely filled at $t = 0.84$ s, and ends with pore 1 at $t = 1.72$ s. The filling order of all 9 pores in the layer is: $6 \rightarrow 7 \rightarrow 5 \rightarrow 8 \rightarrow 4 \rightarrow 9 \rightarrow 2 \rightarrow 3 \rightarrow 1$, as shown in Figure 5.12(b).

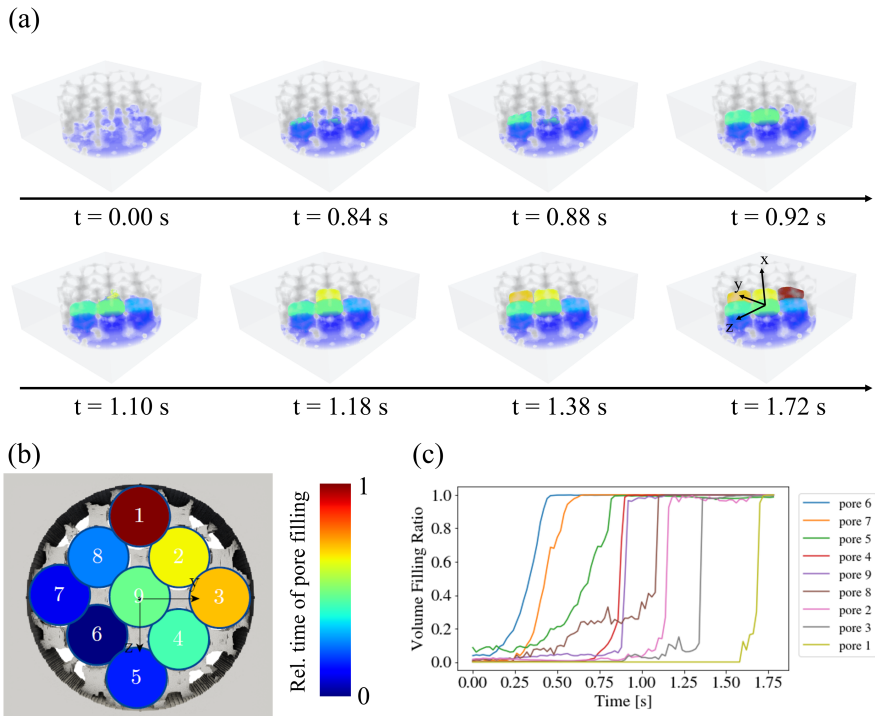


Figure 5.12: 4D reconstruction and analysis of the pore-filling process for the first layer of the porous network. (a) 4D reconstruction of eight time points; (b) Normalized pore filling order for all 9 pores in the layer. Colors indicate the relative start time normalized to total layer filling; (c) The volume filling ratio of all 9 pores with respect to time. Figure adapted from **Paper VII**.

Based on the 4D reconstruction, we can further analyze the flow instabilities during imbibition of porous networks by examining the volume filling ratio of each individual pore as a function of time, as shown in Figure 5.12(c). We can clearly see the instant filling behavior in pores $1, 2, 3, 4$, and 9, which occurs within only $1 - 2$ frames at a 50 Hz acquisition rate. This suggests that these events may approach the intrinsic Haines jump timescales in the millisecond range.

To sum up, we demonstrate the XMPI experiment (with a slow rotation) for studying liquid penetrating through a porous network and subsequent 4D reconstruction and analysis enabled by X-Hexplane. From the experimental perspective, the adaptation of applying a slow rotation enriches the spatial information accessible to the setup. This strategy provides a practical extension to XMPI and can be applied to similar dynamic systems where the slow rotation does not interfere with the underlying physical processes. In the long term, XMPI offers a unique platform for studying fluid dynamics in 4D, helping bridge the gap between experiments and numerical simulations, where

simulations are often more accessible but require experimental validation. From the software perspective, as the most computation-friendly 4D reconstruction workflow presented in this chapter, X-Hexplane only required approximately 20 minutes of training. Such an efficiency opens up new possibilities of getting timely feedback by implementing online 4D reconstruction right after the XMPI experiment.

Chapter 6

Conclusion and Outlook

6.1 Conclusion

In this thesis, we have presented new pathways toward 4D X-ray imaging by advancing XMPI and developing complementary 4D analysis and reconstruction methods. Motivated by the fundamental limitations of time-resolved tomography approaches that require fast rotation, XMPI is introduced as a viable alternative for acquiring time-resolved volumetric information. Taking advantage of the high-brilliance X-rays provided in synchrotron radiation facilities, a practical XMPI implementation was realized at the ForMAX beamline at MAX IV, where multiple X-ray beamlets can illuminate the sample simultaneously from different angles, providing micrometer spatial resolution at an acquisition rate of over 10 kHz.

This thesis addressed the central challenge posed by the intrinsic sparsity of angular information in XMPI using two main strategies. First, a triangulation-based 4D analysis framework was established for particle-resolved flow measurements, without requiring full-field 4D reconstruction. By implementing 3D PTV enabled by XMPI, we performed quantitative 4D tracking of microscale tracer particles in opaque flows. The resulting spatial and velocity uncertainties were analyzed and experimentally validated, demonstrating that XMPI can provide statistically reliable flow characterization despite its limited number of projections. This approach demonstrates that XMPI can work as a powerful tool for flow analysis in optically opaque systems.

Second, this thesis presented a family of AI-assisted, physics-aware reconstruction frameworks to tackle the more ambitious goal of full-field 4D reconstruction from limited spatiotemporal data. By integrating NIR with forward X-ray image formation models, the proposed methods, 4D-ONIX, 4D-PIONIX, and X-Hexplane, enable

end-to-end reconstruction of 4D dynamics directly from multi-view 2D image sequences. Through representative fluid dynamics examples (binary droplet collisions and flows in porous media), these methods were validated to recover both spatial structures and rapid temporal evolution under conditions far beyond the reach of conventional reconstruction techniques (the Crowther Criterion stated in Section 2.2.2).

6.2 Outlook

XMPI is still a developing technique, and its full potential has yet to be explored. In the near future, several promising directions for further development can be identified, spanning XMPI experimental setups, data analysis and reconstruction methodologies, and scientific applications.

From the perspective of the XMPI setup, as already stated in Section 3.3, further optimization of crystal-based beam-splitting strategies may allow an increased angular coverage or a larger number of projections without compromising image quality. Moreover, the development of the synchrotron light sources themselves should not be neglected. Major upgrades of existing DLSRs, such as MAX 4^{U} [98], are expected to deliver further enhanced coherent flux, pushing the spatiotemporal resolution of XMPI to the micrometer scale at above 100 kHz or even MHz in the near future. Beyond the imaging regime presented in this thesis, the concept of XMPI might be extended to far-field imaging [99,100], enabling the spatial resolution to be switched from the micrometer scale to the nanometer scale.

From the perspective of 4D reconstruction and analysis, several exciting directions can be explored. The ultimate goal of full-field 4D reconstruction is not merely to recover spatiotemporal structures, but also to help us understand the underlying physics, and potentially to reveal new mechanisms behind the dynamics. Inspired by 4D-PIONIX, a promising direction is to extend the use of PINN to extract unknown physical parameters of a system [101] from XMPI measurements. In parallel, incorporating recent advances in AI may further improve the performance of 4D reconstruction in the field of X-ray imaging. Promising examples include using 4D Gaussian splatting [102] as an alternative to NIR and using more advanced generation models, such as diffusion models [103], as alternatives to GAN-based methods, as presented in Section 5.2. These developments may offer increased expressive power and improved robustness in 4D reconstruction under ultra-sparse scenarios with XMPI. Moreover, from an engineering perspective, another important direction concerns online or near-real-time 4D analysis. Inspired by X-Hexplane discussed in Section 5.2.3, with the increasing efficiency of low-rank and factorized representations, as well as advances in computational hardware, it becomes possible to provide feedback via rapid preliminary 4D

reconstruction during XMPI experiments, enabling adaptive experiment control and early assessment of data quality, which can result in much more efficient use of limited beamtime at large-scale facilities [104].

From the perspective of applications, although this thesis primarily showcases examples on fluid dynamics, the applicability of XMPI extends far beyond this domain. In essence, any system characterized by rapid 3D dynamics and limited tolerance to mechanical rotation may benefit from XMPI. Potential applications include operando studies of functional materials [21], mechanical fracture of fiber materials [105,106], as well as transient phenomena in additive manufacturing [20,107]. By providing rotation-free access to volumetric dynamics at high temporal resolution, XMPI has the potential to bridge the gaps between experiments, theory, and simulation across a wide range of scientific and engineering disciplines.

To conclude, this thesis demonstrates that XMPI, together with its tailored 4D analysis pipelines and AI-assisted reconstruction frameworks, constitutes an enabling tool at an unprecedented spatiotemporal resolution (sub-10 µm, over 10 kHz) for the fluid dynamics community [108]. Rather than replacing traditional time-resolved tomography methods at large-scale X-ray facilities, XMPI complements them by extending 4D imaging into regimes where fast rotation based approaches fail, thereby substantially broadening the range of dynamic phenomena that can be quantitatively investigated using X-rays.

References

[1] Alexi Assmus. Early history of X rays. *Beam Line*, 25(2):10–24, 1995.

[2] Rajmund Mokso, Daniel A Schwyn, Simon M Walker, Michael Doube, Martina Wicklein, Tonya Müller, Marco Stampanoni, Graham K Taylor, and Holger G Krapp. Four-dimensional in vivo X-ray microscopy with projection-guided gating. *Scientific Reports*, 5(1):8727, 2015.

[3] Paul Northrup, Alessandra Leri, and Ryan Tappero. Applications of "tender" energy (1-5 keV) X-ray absorption spectroscopy in life sciences. *Protein & Peptide Letters*, 23(3):300–308, 2016.

[4] Jens Als-Nielsen and Des McMorrow. *Elements of modern X-ray physics*. John Wiley & Sons, 2011.

[5] Chris Jacobsen. *X-ray Microscopy*. Cambridge University Press, 2019.

[6] Max Born and Emil Wolf. *Principles of optics: Electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.

[7] Hanna Dierks. *X-ray Phase Contrast Tomography*. PhD thesis, Lund University, Lund, Sweden, March 2023.

[8] David Paganin. *Coherent X-ray optics*. Oxford University Press, 2006.

[9] Steve Webb. *From the watching of shadows: The origins of radiological tomography*. CRC Press, 2024.

[10] Ronald N Bracewell. Strip integration in radio astronomy. *Australian Journal of Physics*, 9(2):198–217, 1956.

[11] Bernard Mulgrew, Peter Grant, and John Thompson. *Digital signal processing: Concepts and applications*. Bloomsbury Publishing, 2002.

[12] Thomas Hobiger, Tetsuro Kondo, and Yasuhiro Koyama. Constrained simultaneous algebraic reconstruction technique (C-SART) — a new and simple algorithm applied to ionospheric tomography. *Earth, Planets and Space*, 60(7):727–735, 2008.

[13] Anders H Andersen and Avinash C Kak. Simultaneous algebraic reconstruction technique (SART): a superior implementation of the art algorithm. *Ultrasonic Imaging*, 6(1):81–94, 1984.

[14] Peter Gilbert. Iterative methods for the three-dimensional reconstruction of an object from projections. *Journal of Theoretical Biology*, 36(1):105–117, 1972.

[15] Richard Anthony Crowther, DJ DeRosier, and Aaron Klug. The reconstruction of a three-dimensional structure from projections and its application to electron microscopy. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 317(1530):319–340, 1970.

[16] Andreas Kupsch, Axel Lange, Manfred P Hentschel, Sebastian Lück, Volker Schmidt, Roman Grothausmann, André Hilger, and Ingo Manke. Missing wedge computed tomography by iterative algorithm DIRECTT. *Journal of Microscopy*, 261(1):36–45, 2016.

[17] Yuhe Zhang, Zisheng Yao, Tobias Ritschel, and Pablo Villanueva-Perez. ONIX: an X-ray deep-learning tool for 3D reconstructions from sparse views. *Applied Research*, 2(4):e202300016, 2023.

[18] Francisco García-Moreno, Paul Hans Kamm, Tillmann Robert Neu, Felix Bülk, Mike Andreas Noack, Mareike Wegener, Nadine von der Eltz, Christian Matthias Schlepütz, Marco Stampanoni, and John Banhart. Tomoscopy: Time-resolved tomography for dynamic processes in materials. *Advanced Materials*, 33(45):2104659, 2021.

[19] Theodore (Ted) J Heindel. X-ray flow visualization: Techniques and applications. *Journal of Fluids Engineering*, 146(1):010801, 2024.

[20] Malgorzata G Makowska, Fabrizio Verga, Stefan Pfeiffer, Federica Marone, Cynthia ST Chang, Kevin Florio, Christian M Schlepütz, Konrad Wegener, Thomas Graule, and Steven Van Petegem. Operando tomographic microscopy during laser-based powder bed fusion of alumina. *Communications Materials*, 4(1):73, 2023.

[21] Nikhilesh Chawla and Eshan Ganju. Four-dimensional materials science: Time-resolved X-ray microcomputed tomography. *MRS Bulletin*, pages 1–18, 2025.

[22] Walter Huda and R Brad Abrahams. X-ray-based medical imaging and resolution. *American Journal of Roentgenology*, 204(4):W393–W397, 2015.

[23] Antonella Balerna and Settimio Mobilio. Introduction to synchrotron radiation. In *Synchrotron Radiation: Basics, Methods and Applications*, pages 3–28. Springer, 2014.

[24] Julian Schwinger. On the classical radiation of accelerated electrons. *Physical Review*, 75(12):1912, 1949.

[25] Philip Willmott. *An introduction to synchrotron radiation: Techniques and applications*. John Wiley & Sons, 2019.

[26] S Mobilio and A Balerna. Introduction to the main properties of synchrotron radiation. In *Conference Proceedings-Italian Physical Society*, volume 82, pages 1–24. Editrice Compositori; 1999, 2003.

[27] Do H Tomboulian and PL Hartman. Spectral and angular distribution of ultraviolet radiation from the 300-MeV cornell synchrotron. *Physical Review*, 102(6):1423, 1956.

[28] Pedro F Tavares, Simon C Leemann, Magnus Sjöström, and Åke Andersson. The MAX IV storage ring project. *Journal of Synchrotron Radiation*, 21(5):862–877, 2014.

[29] Lin Liu, Natalia Milas, Afonso HC Mukai, Ximenes R Resende, and Fernando H de Sá. The Sirius project. *Journal of Synchrotron Radiation*, 21 (5):904–911, 2014.

[30] Andreas Streun, Terence Garvey, Lenny Rivkin, Volker Schlott, Thomas Schmidt, Philip Willmott, and Albin Wrulich. SLS-2 - the upgrade of the Swiss Light Source. *Journal of Synchrotron Radiation*, 25(3):631–641, 2018.

[31] Pantaleo Raimondi. ESRF-EBS: The extremely brilliant source project. *Synchrotron Radiation News*, 29(6):8–15, 2016.

[32] Robert Hettel et al. Status of the APS-U project. In *Proceedings of the 12th International Particle Accelerator Conference. Campinas, Brazil*, pages 7–12, 2021.

[33] Yi Jiao, Gang Xu, X-H Cui, Zhe Duan, Y-Y Guo, Ping He, D-H Ji, J-Y Li, X-Y Li, Cai Meng, et al. The HEPS project. *Journal of Synchrotron Radiation*, 25(6): 1611–1618, 2018.

[34] Christopher J Gilmore, James A Kaduk, and Henk Schenk. *International tables for crystallography, volume H: Powder diffraction*. John Wiley & Sons, 2019.

[35] David Norman. Prospects for X-ray absorption with the super-bright light sources of the future. *Journal of Synchrotron Radiation*, 8(2):72–75, 2001.

[36] Thilo Michel, Gisela Anton, M Böhnel, J Durst, M Firsching, A Korn, B Kreisler, A Loehr, F Nachtrab, D Niederlöhner, et al. A fundamental method to determine the signal-to-noise ratio (SNR) and detective quantum efficiency (DQE) for a photon counting pixel detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 568(2):799–802, 2006.

[37] K Nygård, SA McDonald, JB González, Vahid Haghighat, Christian Appel, Emanuel Larsson, Reza Ghanbari, Mira Viljanen, José Silva, Suleyman Malki, et al. Formax–a beamline for multiscale and multimodal structural characterization of hierarchical materials. *Journal of Synchrotron Radiation*, 31 (2):363–377, 2024.

[38] Hanna Dierks, Philip Stjärneblad, and Jesper Wallentin. A versatile laboratory setup for high resolution X-ray phase contrast tomography and scintillator characterization. *Journal of X-ray Science and Technology*, 31(1):1–12, 2023.

[39] Vincent Van Nieuwenhove, Jan De Beenhouwer, Francesco De Carlo, Lucia Mancini, Federica Marone, and Jan Sijbers. Dynamic intensity normalization using eigen flat fields in X-ray imaging. *Optics Express*, 23(21):27975–27989, 2015.

[40] Sarlota Birnsteinova, Danilo E Ferreira de Lima, Egor Sobolev, Henry J Kirkwood, Valerio Bellucci, Richard J Bean, Chan Kim, Jayanath CP Koliyadu, Tokushi Sato, Fabio Dall'Antonia, et al. Online dynamic flat-field correction for MHz microscopy data at European XFEL. *Journal of Synchrotron Radiation*, 30(6):1030–1037, 2023.

[41] Marin Van Heel and Michael Schatz. Fourier shell correlation threshold criteria. *Journal of Structural Biology*, 151(3):250–262, 2005.

[42] Damber Thapa, Kaamran Raahemifar, and Vasudevan Lakshminarayanan. Less is more: Compressive sensing in optics and image science. *Journal of Modern Optics*, 62(6):415–429, 2015.

[43] F Marone and M Stampanoni. Regridding reconstruction algorithm for real-time tomographic imaging. *Journal of Synchrotron Radiation*, 19(6):1029–1037, 2012.

[44] Doga Gürsoy, Francesco De Carlo, Xianghui Xiao, and Chris Jacobsen. TomoPy: A framework for the analysis of synchrotron tomographic data. *Journal of Synchrotron Radiation*, 21(5):1188–1193, 2014.

[45] Francisco García-Moreno, Tillmann Robert Neu, Paul Hans Kamm, and John Banhart. X-ray tomography and tomoscopy on metals: A review. *Advanced Engineering Materials*, 25(8):2201355, 2023.

[46] P Oberta and Rajmund Mokso. A Laue–Bragg monolithic beam splitter for efficient X-ray 2-beam imaging. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 703:59–63, 2013.

[47] Rajmund Mokso and Peter Oberta. Simultaneous dual-energy X-ray stereo imaging. *Journal of Synchrotron Radiation*, 22(4):1078–1082, 2015.

[48] Sebastian Roling, Liubov Samoylova, Björn Siemer, Harald Sinn, Frank Siewert, Frank Wahlert, Michael Wöstmann, and Helmut Zacharias. Design of an X-ray split-and delay-unit for the European XFEL. In *X-Ray Free-Electron Lasers: Beam Diagnostics, Beamline Instrumentation, and Applications*, volume 8504, pages 34–43. SPIE, 2012.

[49] André Authier. Dynamical theory of X-ray diffraction. In *International Tables for Crystallography, Volume B: Reciprocal Space*, pages 534–551. Springer, 2006.

[50] Margie P Olbinado, Xavier Just, Jean-Louis Gelet, Pierre Lhuissier, Mario Scheel, Patrik Vagovic, Tokushi Sato, Rita Graceffa, Joachim Schulz, Adrian Mancuso, et al. MHz frame rate hard X-ray phase-contrast imaging using synchrotron radiation. *Optics Express*, 25(12):13857–13871, 2017.

[51] Vincent Hardion, Paul Bell, Thomas Eriksson, Mirjam Lindberg, Peter Sjöblom, Darren Spruce, et al. MAX IV Laboratory's control system evolution and future strategies. In *19th International Conference on Accelerator and Large Experimental Physics Control Systems (ICALEPCS'23), Cape Town, South Africa, 09-13 October 2023*, pages 395–400. JACOW Publishing, Geneva, Switzerland, 2024.

[52] Wolfgang Voegeli, Kentaro Kajiwara, Hiroyuki Kudo, Tetsuroh Shirasawa, Xiaoyu Liang, and Wataru Yashiro. Multibeam X-ray optical system for high-speed tomography. *Optica*, 7(5):514–517, 2020.

[53] Hiroki Sumiishi, Wolfgang Voegeli, Kentaro Kajiwara, Yoshimasa Urushihara, Xiaoyu Liang, Hiroyuki Kudo, and Wataru Yashiro. Multi-projection imaging of a woodlouse with an improved multibeam X-ray optical system. In *2024 Conference on Lasers and Electro-Optics Pacific Rim (CLEO-PR)*, pages 1–4. IEEE, 2024.

[54] Eleni Myrto Asimakopoulou, Valerio Bellucci, Sarlota Birnsteinova, Zisheng Yao, Yuhe Zhang, Ilia Petrov, Carsten Deiter, Andrea Mazzolari, Marco Romagnoni, Dusan Korytar, et al. Development towards high-resolution kHz-speed rotation-free volumetric imaging. *Optics Express*, 32(3):4413–4426, 2024.

[55] Pablo Villanueva-Perez, Valerio Bellucci, Yuhe Zhang, Sarlota Birnsteinova, Rita Graceffa, Luigi Adriano, Eleni Myrto Asimakopoulou, Ilia Petrov, Zisheng Yao, Marco Romagnoni, et al. Megahertz X-ray multi-projection imaging. *arXiv preprint arXiv:2305.11920*, 2023.

[56] Y Wu, CY Zhao, and Q Wang. Volumetric particle tracking velocimetry with improved algorithms using a two-view shadowgraph system. *Measurement Science and Technology*, 33(8):085301, 2022.

[57] Kevin C Zhou, Brendan K Huang, Ute A Gamm, Vineet Bhandari, Mustafa K Khokha, and Michael A Choma. Particle streak velocimetry-optical coherence tomography: A novel method for multidimensional imaging of microscale fluid flows. *Biomedical Optics Express*, 7(4):1590–1603, 2016.

[58] Simo A Mäkiharju, Jan Dewanckele, Marijn Boone, Christian Wagner, and Andreas Griesser. Tomographic X-ray particle tracking velocimetry: Proof-of-concept in a creeping flow. *Experiments in Fluids*, 63(1):16, 2022.

[59] Xi Chen, Wenqi Zhong, and Theodore J Heindel. Orientation of cylindrical particles in a fluidized bed based on stereo X-ray particle tracking velocimetry (XPTV). *Chemical Engineering Science*, 203:104–112, 2019.

[60] Jason T Parker and Simo A Mäkiharju. Experimentally validated X-ray image simulations of 50 μm X-ray PIV tracer particles. *Measurement Science and Technology*, 33(5):055301, 2022.

[61] Ronald L Panton. *Incompressible flow*. John Wiley & Sons, 2024.

[62] Ron Shnapp. MyPTV: A Python package for 3D particle tracking. *Journal of Open Source Software*, 7(75):4398, 2022.

[63] Mickaël Bourgoin and Sander G Huisman. Using ray-traversal for 3D particle matching in the context of particle tracking velocimetry in fluid mechanics. *Review of Scientific Instruments*, 91(8), 2020.

[64] Nicholas T Ouellette, Haitao Xu, and Eberhard Bodenschatz. A quantitative study of three-dimensional Lagrangian particle tracking algorithms. *Experiments in Fluids*, 40(2):301–313, 2006.

[65] HH Kv. Notes on the use of propagation of error formulas. *NBS Special Publication*, 70(300-301):331, 1969.

[66] Richard J Leite. An experimental investigation of the stability of Poiseuille flow. *Journal of Fluid Mechanics*, 5(1):81–96, 1959.

[67] SM Richardson. *Fluid mechanics*. Hemisphere Publishing Corp, 1989.

[68] Nghia T Vo. Discorpy: Algorithms and software for camera calibration and correction. *Journal of Synchrotron Radiation*, 32(3), 2025.

[69] Zachary Teed and Jia Deng. RAFT: Recurrent all-pairs field transforms for optical flow. In *European conference on computer vision*, pages 402–419. Springer, 2020.

[70] William Thielicke and René Sonntag. Particle image velocimetry for MATLAB: Accuracy and enhanced algorithms in PIVlab. *Journal of Open Research Software*, 9(1), 2021.

[71] Jeffrey Dean. A golden decade of deep learning: Computing systems & applications. *Daedalus*, 151(2):58–74, 2022.

[72] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5): 359–366, 1989.

[73] Siwei Zhang, Yan Zhang, Federica Bogo, Marc Pollefeys, and Siyu Tang. Learning motion priors for 4D human body capture in 3D scenes. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 11343–11353, 2021.

[74] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *Nature*, 323(6088):533–536, 1986.

[75] Sebastian Ruder. An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*, 2016.

[76] Zdeněk Matěj, Rajmund Mokso, Krister Larsson, Vincent Hardion, and Darren Spruce. The MAX IV imaging concept. *Advanced Structural and Chemical Imaging*, 2(1):16, 2016.

[77] Jiuxiang Gu, Zhenhua Wang, Jason Kuen, Lianyang Ma, Amir Shahroudy, Bing Shuai, Ting Liu, Xingxing Wang, Gang Wang, Jianfei Cai, et al. Recent advances in convolutional neural networks. *Pattern Recognition*, 77:354–377, 2018.

[78] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86 (11):2278–2324, 2002.

[79] Ian J Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in Neural Information Processing Systems*, 27, 2014.

[80] Yuhe Zhang. *Advancing X-ray imaging with deep learning: Physics-inspired reconstruction approaches*. PhD thesis, Lund University, Lund, Sweden, May 2024.

[81] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. NeRF: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1):99–106, 2021.

[82] Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.

[83] Atilim Gunes Baydin, Barak A Pearlmutter, Alexey Andreyevich Radul, and Jeffrey Mark Siskind. Automatic differentiation in machine learning: A survey. *Journal of Machine Learning Research*, 18(153):1–43, 2018.

[84] Malcolm R Howells, Tobias Beetz, Henry N Chapman, C Cui, JM Holton, CJ Jacobsen, J Kirz, Enju Lima, Stefano Marchesini, Huijie Miao, et al. An assessment of the resolution limitation due to radiation-damage in X-ray diffraction microscopy. *Journal of Electron Spectroscopy and Related Phenomena*, 170(1-3):4–12, 2009.

[85] Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, volume 2017, pages 4681–4690, 2017.

[86] Anpei Chen, Zexiang Xu, Andreas Geiger, Jingyi Yu, and Hao Su. TensoRF: Tensorial radiance fields. In *European Conference on Computer Vision*, pages 333–350. Springer, 2022.

[87] Ang Cao and Justin Johnson. Hexplane: A fast representation for dynamic scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 130–141, 2023.

[88] Zailiang Hu and RC Srivastava. Evolution of raindrop size distribution by coalescence, breakup, and evaporation: Theory... *Journal of the Atmospheric Sciences*, 52(10), 1995.

[89] Carole Planchette, Elise Lorenceau, and Günter Brenn. Liquid encapsulation by binary collisions of immiscible liquid drops. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 365(1-3):89–94, 2010.

[90] Puneet K Sinha and Chao-Yang Wang. Pore-network modeling of liquid water transport in gas diffusion layer of a polymer electrolyte fuel cell. *Electrochimica Acta*, 52(28):7936–7945, 2007.

[91] Martin J Blunt, Branko Bijeljic, Hu Dong, Oussama Gharbi, Stefan Iglauer, Peyman Mostaghimi, Adriana Paluszny, and Christopher Pentland. Pore-scale imaging and modelling. *Advances in Water Resources*, 51:197–216, 2013.

[92] William B Haines. Studies in the physical properties of soils. V. The hysteresis effect in capillary properties, and the modes of moisture distribution associated therewith. *The Journal of Agricultural Science*, 20(1):97–116, 1930.

[93] David Wilkinson and Jorge F Willemsen. Invasion percolation: A new form of percolation theory. *Journal of Physics A: Mathematical and General*, 16(14):3365, 1983.

[94] Aleksander Lovrić, Wulf G Dettmer, and Djordje Perić. Low order finite element methods for the Navier-Stokes-Cahn-Hilliard equations. *arXiv preprint arXiv:1911.06718*, 2019.

[95] Peter Bastian, Markus Blatt, Andreas Dedner, Nils-Arne Dreier, Christian Engwer, René Fritze, Carsten Gräser, Christoph Grüninger, Dominic Kempf, Robert Klöfkorn, et al. The Dune framework: Basic concepts and recent developments. *Computers & Mathematics with Applications*, 81:75–112, 2021.

[96] Andreas Dedner and Robert Klöfkorn. Extendible and efficient Python framework for solving evolution equations with stabilized discontinuous galerkin methods. *Communications on Applied Mathematics and Computation*, pages 1–40, 2021.

[97] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

[98] E Al-Dmour, A Mitrovic, A Vorozhtsov, A Martinez Carboneres, A Rosborg, A Robert, F Cullinan, H Tarawneh, H Duarte, J Schnadt, et al. MAX 4U:

An upgrade of the MAX IV 3 GeV ring. In *IPAC'25: International Particle Accelerator Conference, Taipeh, 1st-6th June 2025*, page 752, 2025.

[99] Jianwei Miao, Richard L Sandberg, and Changyong Song. Coherent X-ray diffraction imaging. *IEEE Journal of Selected Topics in Quantum Electronics*, 18 (1):399–410, 2011.

[100] Zhe Hu, Zisheng Yao, Yuhe Zhang, and Pablo Villanueva-Perez. Constraint-free coherent diffraction imaging via physics-guided neural fields. *arXiv preprint arXiv:2512.09694*, 2025.

[101] Jeremy Yu, Lu Lu, Xuhui Meng, and George Em Karniadakis. Gradient-enhanced physics-informed neural networks for forward and inverse PDE problems. *Computer Methods in Applied Mechanics and Engineering*, 393:114823, 2022.

[102] Guanjun Wu, Taoran Yi, Jiemin Fang, Lingxi Xie, Xiaopeng Zhang, Wei Wei, Wenyu Liu, Qi Tian, and Xinggang Wang. 4D Gaussian splatting for real-time dynamic scene rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20310–20320, 2024.

[103] Rundi Wu, Ruiqi Gao, Ben Poole, Alex Trevithick, Changxi Zheng, Jonathan T Barron, and Aleksander Holynski. CAT4D: Create anything in 4D with multi-view video diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 26057–26068, 2025.

[104] Aymeric Robert, Yngve Cerenius, Pedro Fernandes Tavares, Anna Hultin Stigenberg, Olof Karis, Ann-Christine Lloyd Whelan, Caroline Runéus, and Marjolein Thunnissen. MAX IV Laboratory. *The European Physical Journal Plus*, 138(6):495, 2023.

[105] Jinling Gao, Nesredin Kedir, Cody D Kirk, Julio A Hernandez, Junyu Wang, Shane Paulson, Xuedong Zhai, Todd Horn, Garam Kim, Kamel Fezzaa, et al. High-speed synchrotron X-ray phase-contrast imaging for evaluating microscale damage mechanisms and tracking cracking behaviors inside cross-ply GFRCs. *Composites Science and Technology*, 210:108814, 2021.

[106] Elise Van Vlierberghe, Jeroen Soete, Eleni Myrto Asimakopoulou, Zisheng Yao, Julia Rogalinski, Zhe Hu, Kannara Mom, Bratislav Lukić, Christian Breite, Pablo Villanueva Perez, et al. Transverse cracking in glass fibre-reinforced composites monitored with synchrotron X-ray multi-projection imaging. *e-Journal of Nondestructive Testing*, 30(2), 2025.

[107] Naying An, Sansan Shuai, Tao Hu, Chaoyue Chen, Jiang Wang, and Zhongming Ren. Application of synchrotron X-ray imaging and diffraction in additive manufacturing: A review. *Acta Metallurgica Sinica (English Letters)*, 35(1):25–48, 2022.

[108] Outi Supponen. Visualizations of ultrafast bubble dynamics. *Physical Review Fluids*, 11(2):023601, 2026.

Scientific Publications

Author Contributions

Paper I: New opportunities for time-resolved imaging using diffraction-limited storage rings

Z. Yao*, J. Rogalinski*, E. M. Asimakopoulou, Y. Zhang, K. Gordeyeva, Z. Atoufi, H. Dierks, S. McDonald, S. Hall, J. Wallentin, D. Söderberg, K. Nygård, P. Villanueva-Perez
Journal of Synchrotron Radiation, 2024, 31(5), pp. 1299-1307

I contributed to the experiments, data processing, and manuscript writing.

Paper II: Time-resolved 3D imaging opportunities with XMPI at ForMAX

J. Rogalinski, **Z. Yao**, Y. Zhang, Z. Hu, K. Gordeyeva, T. Rosén, D. Söderberg, A. Mazzolari, J. Silva, V. Haghighat, S. McDonald, K. Nygård, E. M. Asimakopoulou, P. Villanueva-Perez
Journal of Synchrotron Radiation, 2026, 33(2), pp. 417-428

I contributed to the experiments, data processing, and manuscript revision.

Paper III: 4D-ONIX for reconstructing 3D movies from sparse X-ray projections via deep learning

Y. Zhang, **Z. Yao**, R. Klöfkorn, T. Ritschel, P. Villanueva-Perez
Communications Engineering, 2025, 4(1), pp. 1-12

I contributed to the generation, processing, and analysis of the data.

Paper IV: Physics-informed 4D x-ray image reconstruction from ultra-sparse spatiotemporal data

Z. Yao, Y. Zhang, Z. Hu, R. Klöfkorn, T. Ritschel, P. Villanueva-Perez
Measurement Science and Technology, 2025, 36(8), pp. 085403

I developed the algorithm and contributed to data generation, data analysis, and manuscript writing.

Paper V: Super Time-Resolved Tomography

Z. Hu, **Z. Yao**, K. Josefsson, F. García-Moreno, M. Makowska, Y. Zhang, P. Villanueva-Perez
Advanced Science, 2026, 13(3), e11933

I contributed to code development, data generation, processing, and analysis.

Paper VI: Synchrotron X-ray Multi-Projection Imaging (XMPI) for High-Resolution 4D Characterization of Multiphase Flows

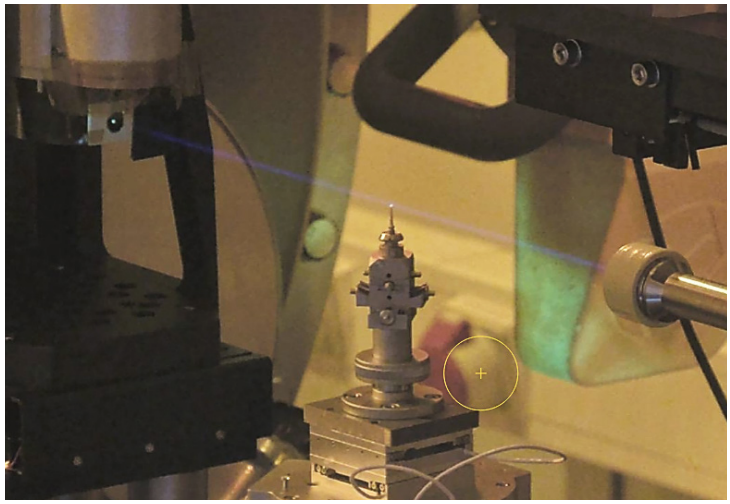
T. Rosén*, **Z. Yao***, J. Tejbo, P. Wegele, J. Rogalinski, F. Nilsson, K. Mom, Z. Hu, S. McDonald, K. Nygård, A. Mazzolari, A. Groetsch, K. Gordeyeva, D. Söderberg, F. Lundell, L. P. Wittberg, E. M. Asimakopoulou, P. Villanueva-Perez*
arXiv preprint, 2026, arXiv:2412.09368v2

I contributed to algorithm development, data processing and analysis, error analysis, and manuscript writing.

Paper VII: 4D Synchrotron X-Ray Multi Projection Imaging (XMPI) for studying multiphase flow dynamics and flow instabilities in porous network

P. Wegele*, **Z. Yao***, J. Tejbo, J. Rogalinski, Z. Hu, Y. Zhang, E. Oliaei, S. Davoodi, T. Rosén, A. Groetsch, K. Nygård, E. M. Asimakopoulou, P. Villanueva-Perez, D. Söderberg*
arXiv preprint, 2026, arXiv:2603.15319

I contributed to algorithm development, data processing and analysis, and manuscript writing.



Printed by Media-Tryck, Lund 2026 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

Synchrotron Radiation Research
Department of Physics
Faculty of Science

ISBN 978-91-8104-771-4