MHz - Tomoscopy

D2.3

Machine learning framework

Project information

Project full title	MHz rate mulTiple prOjection X-ray MicrOSCOPY
Project acronym	MHz-TOMOSCOPY
Grant agreement no.	101046448
Instrument	EIC Pathfinder Open
Duration	42 months
Website	https://tomoscopy.eu/

Deliverable information

Deliverable no.	D2.3
Deliverable title	Machine learning framework
Deliverable responsible	LU
Related Work-Package/Task	WP2/Task 2.3
Type (e.g. report; other)	Other
Author(s)	Pablo Villanueva-Perez, Yuhe Zhang, and Zhe Hu
Dissemination level	Sensitive
Document Version	1.0

Funded by
the European Union

Date	31/01/2025
Download page	https://github.com/yuhez/4D-ONIX

Document information

Version no.	Date	Author(s)	Comment
1.0	2025-01-31	P. Villanueva-Perez, Y. Zhang, and Z. Hu	Initial version

Abstract

Executive Summary

This report summarizes the development of machine-learning frameworks to reconstruct 3D movies (4D) from the data provided by the MHz-XMPI instrument developed by WP1. First, we present our 4D approach, named 4D-ONIX[1], which was used to analyze and retrieve the first MHz 4D movie with XMPI of a binary droplet collision using the WP1 prototype[2]. This approach, which has been accepted for publication in Communications Engineering, has been validated using simulated data and experimental data provided by conventional time-resolved tomography at synchrotron radiation facilities. We also report the limitations of 4D-ONIX and our current developments in adapting state-of-the-art 4D machine-learning approaches to overcome them. Our recent 4D framework, which is based on a machine-learning architecture known as Hexplane, will be used to address the data coming from WP3.

In the first section, we describe the 4D-ONIX approach and architecture, the code of which will be publicly available upon publication. The second section describes the application to MHz-XMPI data coming from WP1. The third section analyzes the impact and limitations of 4D-ONIX. Section 4 describes our novel developments using state-of-the-art 4D machine-learning approaches to overcome key limitations of 4D-ONIX.

Table of Contents

Funded by
the European Union

1. 4D-ONIX

4D-ONIX is a self-supervised deep learning model designed to reconstruct 4D (3D + time) data from sparse X-ray projections acquired using the MHz-XMPI instrument developed by WP1. As shown in Figure 1a, unlike traditional time-resolved tomography, which relies on sample rotation to collect hundreds of projections, XMPI enables volumetric imaging without rotation by simultaneously capturing multiple projections using split beamlets[2,3]. However, XMPI was limited to recording only two per timestamp initially, as initially implemented by WP1[2]. This limited number of projections makes conventional reconstruction methods ineffective.

4D-ONIX has been developed to address this challenge by integrating several key factors: (i) using state-of-the-art machine learning concept (neural implicit representation[4]) to model the refractive index as a continuous function of space and time, (ii) using self-supervised learning to learn from only recorded projections, eliminating the need for ground truth 3D/4D data, (iii) incorporating the physics of X-ray interaction with matter into the model, (iv) learning latent sample features by generalizing across all timestamps and experiments, and (v) applying adversarial learning to ensure spatial and temporal consistency.

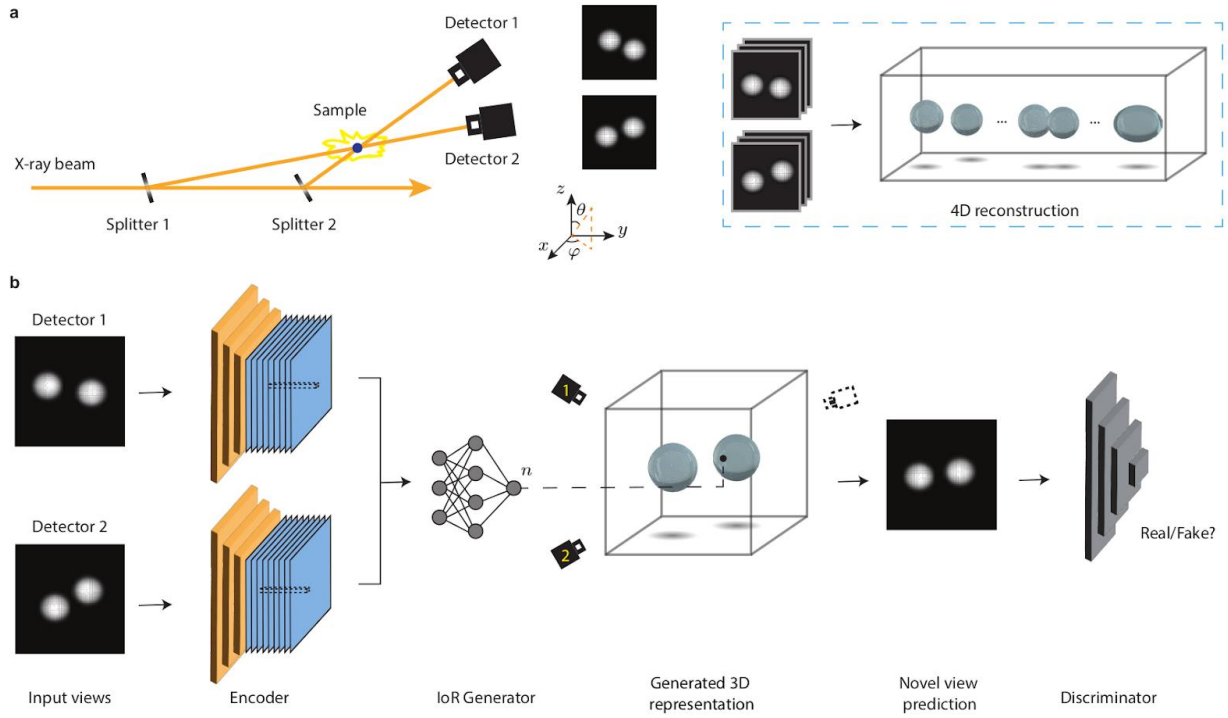

Figure 1: (a) Conceptual illustration of the prototype for MHz-XMPI setup used to collect the data[2]. The inset shows the goal of 4D-ONIX to reconstruct direct 4D from 2D projections. (b) Overview of the 4D-ONIX approach with the key components to reconstruct 4D from two projections.

As shown in Figure 1b, the architecture of 4D-ONIX consists of three main components: an encoder, an Index of Refraction (IoR) generator, and a discriminator. The recorded projections first pass through a convolutional neural network encoder, which transforms the 2D images into stacks of downscaled feature maps. These extracted features are then processed by the IoR generator, which reconstructs the 3D representation at any spatial-temporal point. A physics-based

Funded by the European Union

forward propagation model is applied to predict projections from different angles within the plane of the incoming X-ray beam. Finally, the discriminator helps to optimize the reconstruction over novel/predicted views by minimizing the differences between real and predicted projections.

2. 4D-ONIX on MHz-XMPI

The performance of 4D-ONIX was evaluated using both simulated and experimental MHz-XMPI data of water droplet collisions, demonstrating its ability to reconstruct high-resolution 4D information from highly sparse projections.

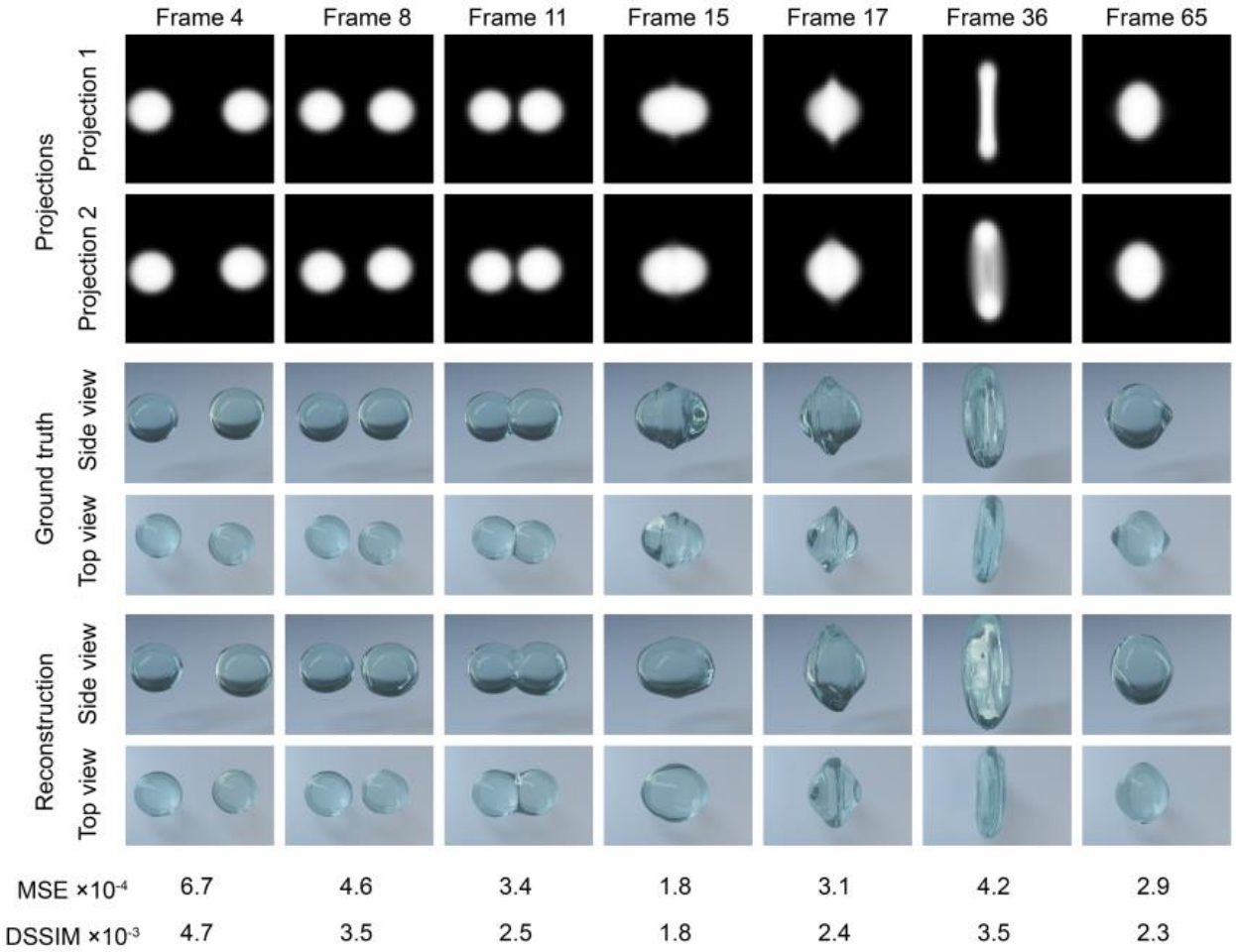

Figure 2: 4D-ONIX results on simulated data of water droplet collisions.

Simulated Data

As aforementioned, Figure 1a illustrates an experiment conducted at EuXFEL to study the collision dynamics of water droplets, capturing two X-ray projections spaced 23.8° apart. For validation, 4D data of water droplet collisions were simulated using the incompressible Navier-Stokes equations, and two projections were generated to match the geometry of the EuXFEL experiments.

4D-ONIX was applied to reconstruct the full 4D dynamics of the simulated water droplet collisions, with the results presented in Figure 2. The reconstruction results were evaluated using Mean Squared Error (MSE) and Dissimilarity Structure Similarity Index Metric (DSSIM). As can be seen, 4D-ONIX successfully reconstructed the core stages of water droplet collision, providing invaluable insights for studying the 3D dynamics of the collision process. This reconstruction is based on 16 simulated experiments of water droplet collisions, each containing 75 timestamps and two projections 23.8° apart.

Experimental Data

Following the validation using simulated data, 4D-ONIX was applied to experimental MHz-XMPI data collected at EuXFEL, where water droplet collisions were imaged at a frame rate of 1.128 MHz using single X-ray pulses. The preprocessed projections from the two detectors and the corresponding reconstructions generated by 4D-ONIX are shown in Figure 3.

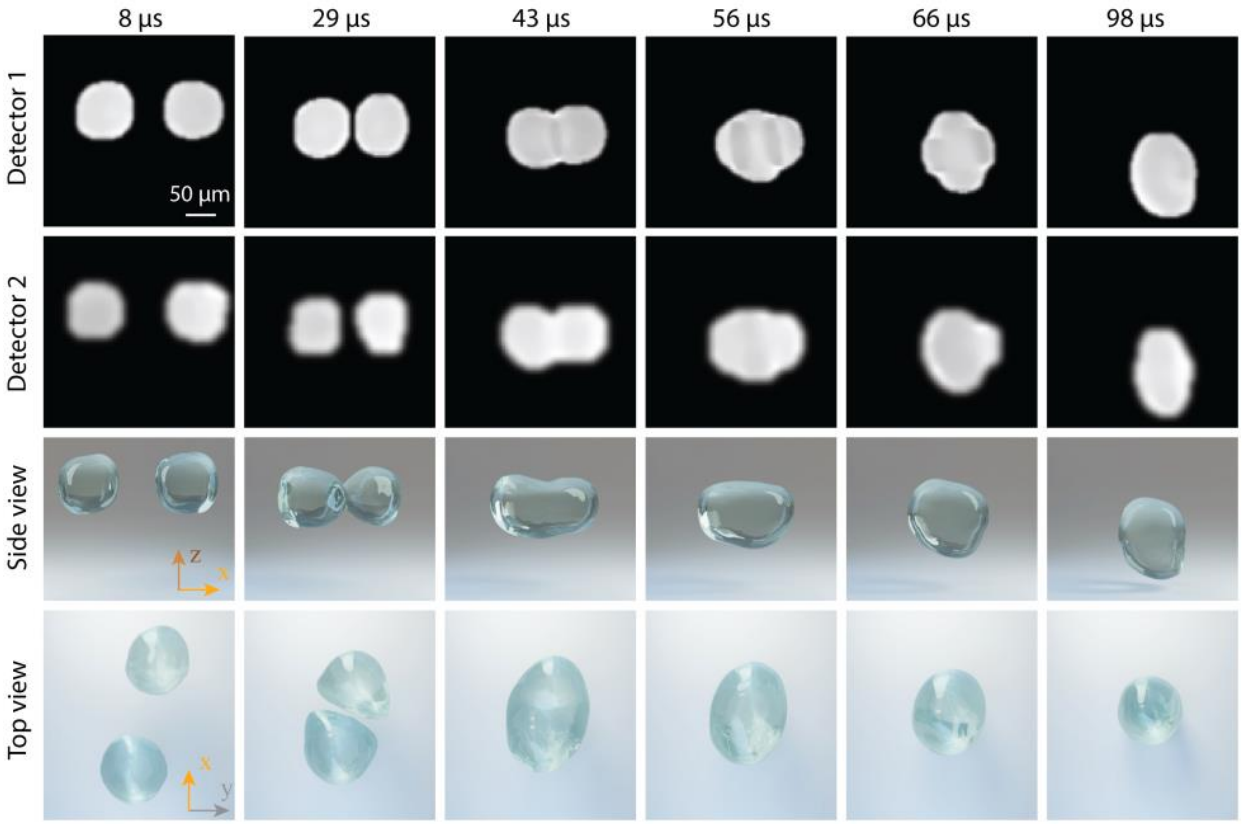

Figure 3: 4D-ONIX results on experimental MHz-XMPI data of water droplet collisions collected at EuXFEL.

Despite the extreme sparsity of the recorded projections, 4D-ONIX successfully reconstructed a full 3D movie of the collision, capturing the evolution of the droplets at microsecond timescales together with micrometer resolution. The reconstructed movie achieved a temporal resolution of 0.89 µs, which is three orders of magnitude faster than conventional time-resolved tomography[5] demonstrating the potential of 4D-ONIX for MHz-rate 4D imaging.

However, the retrieved reconstructions are subject to certain limitations. First, due to experimental constraints, only two sequences of the droplet collision were captured, with minimal variation in sample orientation, restricting the diversity of training data. Second, the recorded projections

contained noise and imaging artifacts, particularly from the second detector, which introduces imperfections in the reconstructed 4D movies. This can be mitigated in future work by incorporating additional projections and improving data quality using the new setup coming from WP1.

3. 4D-ONIX impact and limitations

The development of 4D-ONIX represents a step forward in 4D X-ray imaging, enabling high-resolution MHz-rate reconstructions from extremely sparse data. It opens up new opportunities for studying ultrafast processes in a variety of fields, including fluid dynamics and material science. The methodology can also be extended to other time-resolved X-ray imaging techniques, such as phase contrast imaging and coherent diffraction imaging, by modifying the underlying physical propagation model.

Despite its advantages, 4D-ONIX has several limitations that hinder its applicability to the MHz-TOMOSCOPY project. One of the primary challenges is its computational complexity. The integration of multiple convolutional neural networks increases the computational cost, making real-time or online reconstructions impossible. For example, the reconstruction presented in the previous section was retrieved after several days of training on an NVIDIA A100 GPU with 80 GB memory. Another limitation is the requirement for multiple experiments of similar dynamic processes. Since 4D-ONIX learns from generalizing across multiple timestamps and experiments, it requires acquiring similar dynamics over several experiments. However, obtaining such data can be extremely difficult because of i) the highly stochastic nature of some physical processes, which makes them challenging to reproduce, and ii) the limited access to facilities like EuXFEL. Besides, the accuracy of 4D-ONIX reconstructions is inherently constrained by the sparsity of the input data. For more complex dynamics, additional projections and experiments would be necessary to enhance reconstruction fidelity. Although the model has been successfully applied to water droplet collisions, its applicability to more complex materials and dynamics remains to be explored.

To address these challenges in the context of MHz-TOMOSCOPY, we have developed an improved 4D machine-learning framework based on Hexplane, as described in the next section. This new framework aims to enhance reconstruction quality while reducing computational demands. Additionally, future improvements will explore the integration of Physics-Informed Neural Networks[6] to impose constraints derived from known physical laws. This will not only improve reconstruction accuracy but also enable interpolation between timestamps, enhancing the temporal resolution of reconstructed 4D movies beyond the acquisition rate of the setup.

4. Beyond 4D-ONIX: Hexplane for 4D X-ray imaging

In this section, we present X-Hexplane, a state-of-the-art 4D reconstruction algorithm tailored to fulfill the needs of the MHz-TOMOSCOPY project and overcome critical limitations of 4D-ONIX.

X-Hexplane, an extension of Hexplane[7], adopts a tensorial approach and processes 4D dynamic sample data efficiently. Representing dynamic samples as an explicit voxel grid of features significantly reduces memory consumption through tensorial factorization. Specifically, X-

Hexplane decomposes a 4D space-time grid into six feature planes spanning each pair of coordinate axes (e.g., XY, ZT). This factorization approach reduces the space-time complexity from $O(n^3T)$ to $O(n^2F)$, where n, T, and F denote the spatial resolution, temporal resolution, and feature size, respectively, leading to a significant reduction in memory footprint. Figure 4 provides an overview of the X-Hexplane approach. First, a 4D point in space-time is projected onto each of these feature planes, generating six feature vectors that are then aggregated into a fused feature vector. Second, a tiny multilayer perceptron is used to regress the IoR at that point from the fused feature vector. After computing the value for all points in space-time, 2D projections can be rendered via a line integral along an X-ray propagation direction, assuming projection approximation. For that, we have included the X-ray-mater interactions in our approach.

To address the sparse-view challenge inherent to XMPI, X-Hexplane ensures information consistency across space and time by sharing low-rank tensors over the space-time grid. This overcomes the problem of modeling each timestep independently as done, which fails to provide sufficient information for high-quality and consistent 4D reconstruction. By sharing features across timesteps, X-Hexplane achieves temporal coherence. The unique capabilities of X-Hexplane are achieved by: i) incorporating the physics of X-ray propagation into the model, ii) using a tensorial representation of the dynamic process to reduce memory usage and accelerate the training and reconstruction process, and iii) sharing features across time to overcome the sparse-view limitations. When applied to XMPI, X-Hexplane enables efficient and high-quality 4D reconstructions in a timely manner (around 100 times faster than 4D-ONIX). To validate the capabilities of X-Hexplane for MHz-TOMOSCOPY and compare it to 4D-ONIX, we used the simulation dataset presented in Section 2.

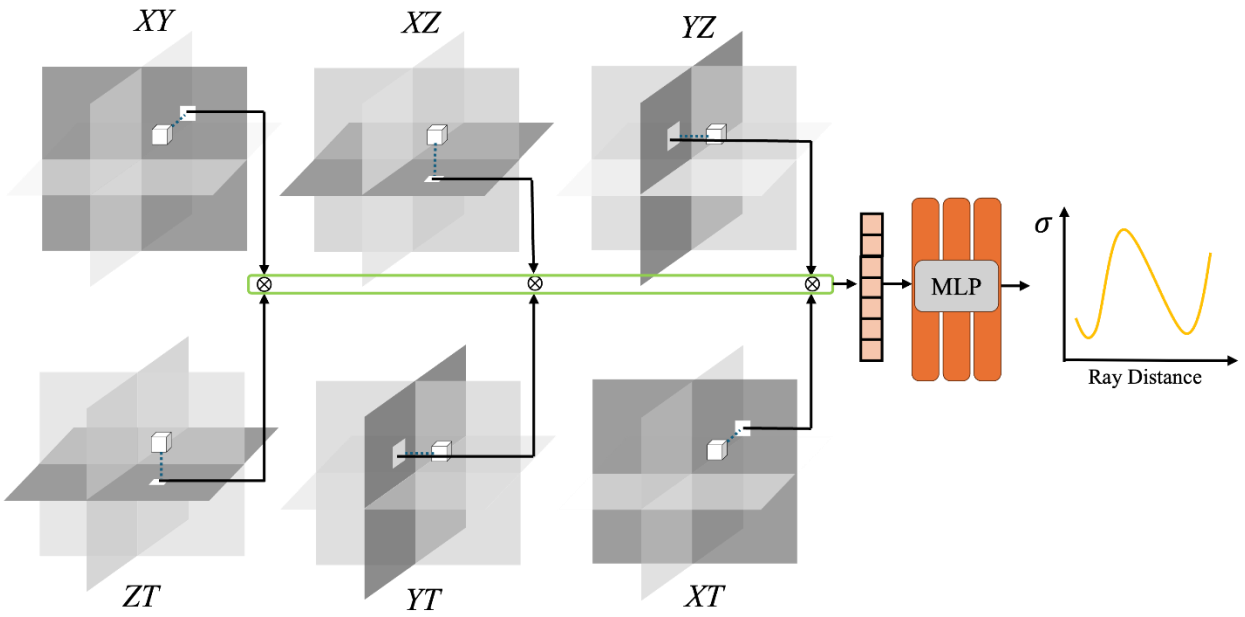

Figure 4: Overview of X-Hexplane. X-Hexplane contains six feature planes spanning each pair of coordinate axes (e.g. XY, ZT). Rays are marched from each pixel of the captured images. Points are sampled along with the rays. For any point in space-time, it is projected to six planes. Features extracted from paired planes are multiplied and then concatenated (green box). IoR at certain points are regressed from fused features using a tiny MLP.

Our new approach can train and retrieve results comparable to 4D-ONIX while requiring only 10 minutes to train on the same dataset instead of several days as required by 4D-ONIX.

Funded by
the European Union

An optional Generative Adversarial Network (GAN) architecture is also introduced to improve the precision of the reconstruction results. The GAN framework consists of two competing networks: X-Hexplane, which works as a generator, and a convolutional-neural-network discriminator, which learns from the ground truth data and attempts to differentiate between the input projections and the generated ones by X-Hexplane. The inclusion of the GAN will increase the training time to several hours, as the discriminator requires more computational resources, while X-Hexplane without GAN will only take 10 minutes. The discriminator was trained to learn from the ground truth information available at specific angles (0 and 23.8 degrees) and tasked with assessing the novel views generated by the X-Hexplane. This adversarial process significantly improved the output of X-Hexplane at the cost of longer training time.

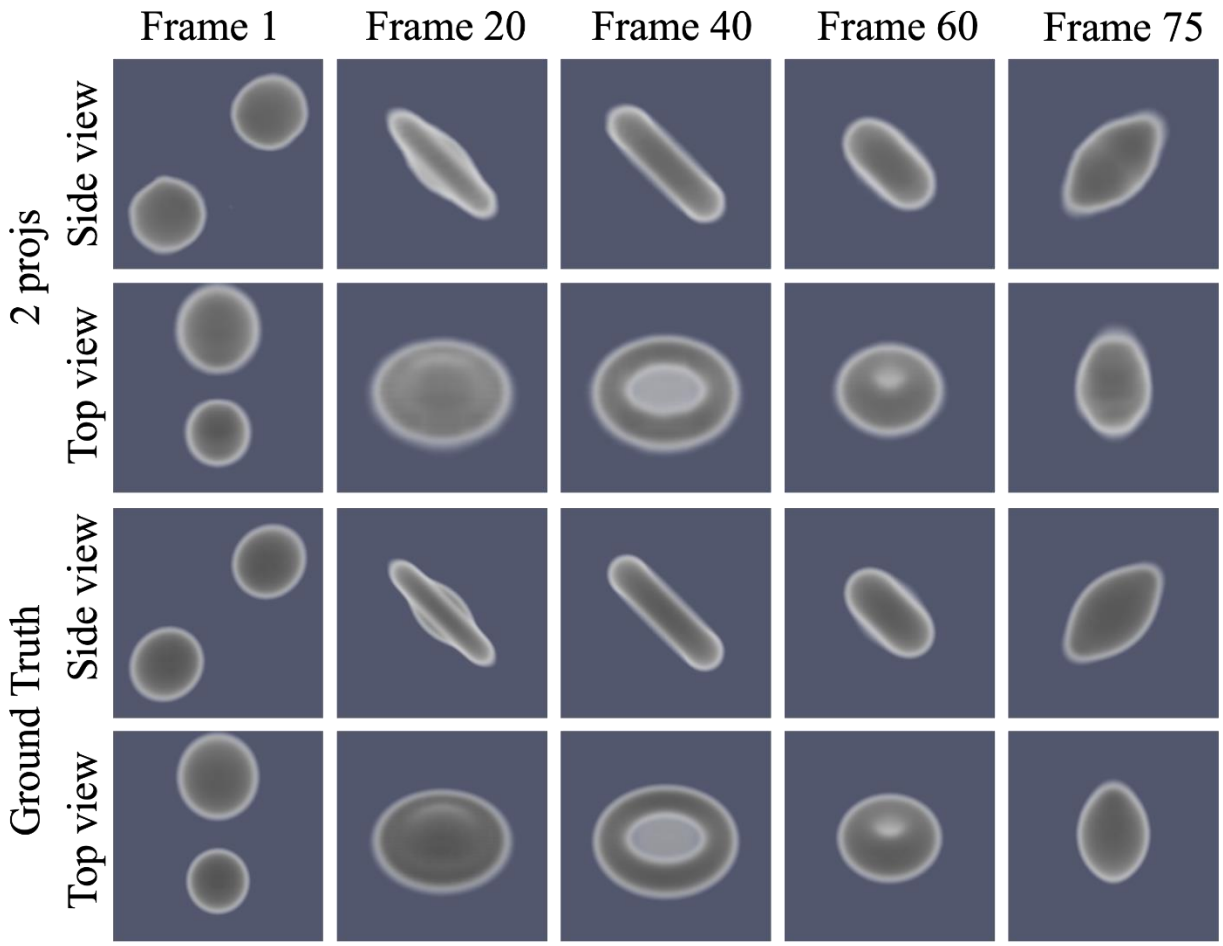

Figure 5: Reconstruction of X-Hexplane with GAN compared with ground truth for different time frames.

The adversarial process forced the generator to improve its outputs and produce more self-consistent and realistic reconstructions. The results in Fig. 5 demonstrate that the X-Hexplane with GAN successfully reconstructed the droplet collision process with high precision. It preserved the correct shape and produced consistent reconstructions across all viewpoints. Furthermore, the overall dynamics of the droplet, including the collision outcomes, closely matched the expected results based on physical simulations based on physical simulations.

Funded by
the European Union

In conclusion, the X-Hexplane offered a fast solution for sparse view 4D reconstruction tasks. The incorporation of a GAN significantly improved the reconstruction quality, yielding results that were more accurate and physically consistent with the expected droplet behavior, albeit at the cost of increased training time. These results underscore the potential of XMPI for rapid 4D imaging, especially when paired with advanced machine-learning techniques. The improvements in accuracy and shape preservation achieved with the GAN-based method indicate that XMPI, combined with adversarial training methods, holds great promise for real-time 4D imaging of dynamic processes like fluid dynamics. The 4D nature of X-Hexplane, combined with its incorporation of X-ray physics, presents exciting opportunities for further development. The temporal information derived from reconstructions can be leveraged to enhance model accuracy by introducing additional constraints. For instance, if the sample dynamics follow a partial differential equation (PDE), integrating a PDE-based loss term within the loss function using Physics-Informed Neural Networks (PINNs)[6,8] can not only improve alignment with physical laws but also facilitate interpolation between dynamic states. This approach enables the generation of a continuous 3D time-lapse reconstruction with a temporal resolution exceeding the original tomography experiment's recording rate. Finally, X-Hexplane offers adaptability and flexibility, making it applicable to a broad range of time-resolved imaging experiments and different spectral domains. It can be extended to techniques such as coherent diffraction imaging[9], phase-contrast imaging[10], MRI, and ultrasound imaging[11], where the propagation model is explicitly known.

We envision for the rest of the MHz-TOMOSCOPY project within WP2-Task 2.3 to:

1. Apply X-Hexplane to the data acquired from WP3 with the new prototype provided by WP1. The increased number of projections of the new setup and large angular diversity will enable the use of X-Hexplane without GAN based on our simulations and provide fast reconstructions within a few minutes from a single MHz movie.
2. We will develop a validation protocol to evaluate 4D reconstruction quality (resolution and accuracy) without ground truth based on ablation studies by extending current ideas used in cryo-electron microscopy for 3D evaluation.
3. As suggested in the project review report and evaluation sessions, we will assess the application of X-Hexplane to, e.g., standard tomography and its application to the biomedical sector (medical CT) and security (airport scanners).
4. We will aim to establish collaborations to extend the use of X-Hexplane to other probes like MRI or ultrasound.

.

6. Bibliography

1. Zhang, Y., Yao, Z., Klöfkorn, R., Ritschel, T. & Villanueva-Perez, P. 4D-ONIX: A deep learning approach for reconstructing 3D movies from sparse X-ray projections. Preprint at https://doi.org/10.48550/arXiv.2401.09508 (2024).
2. Villanueva-Perez, P. *et al.* Megahertz X-ray Multi-projection imaging. Preprint at https://doi.org/10.48550/arXiv.2305.11920 (2023).
3. Villanueva-Perez, P. *et al.* Hard x-ray multi-projection imaging for single-shot approaches. *Optica, OPTICA* **5**, 1521–1524 (2018).
4. Mildenhall, B. *et al.* NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* vol. 12346 LNCS 405–421 (Springer Science and Business Media Deutschland GmbH, 2020).
5. García-Moreno, F. *et al.* Tomoscopy: Time-Resolved Tomography for Dynamic Processes in Materials. *Advanced Materials* **33**, 2104659 (2021).
6. Raissi, M., Perdikaris, P. & Karniadakis, G. E. Physics Informed Deep Learning (Part I): Data-driven Solutions of Nonlinear Partial Differential Equations. Preprint at https://doi.org/10.48550/arXiv.1711.10561 (2017).
7. Cao, A. & Johnson, J. HexPlane: A Fast Representation for Dynamic Scenes. Preprint at https://doi.org/10.48550/arXiv.2301.09632 (2023).
8. Raissi, M., Perdikaris, P. & Karniadakis, G. E. 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).
9. Miao, J. Computational microscopy with coherent diffractive imaging and ptychography. *Nature* **637**, 281–295 (2025).
10. Bravin, A., Coan, P. & Suortti, P. X-ray phase-contrast imaging: from pre-clinical applications towards clinics. *Physics in medicine and biology* **58**, R1-35 (2013).
11. Wysocki, M. *et al.* Ultra-NeRF: Neural Radiance Fields for Ultrasound Imaging. Preprint at https://doi.org/10.48550/arXiv.2301.10520 (2023).

Funded by the European Union