

TomoBank: a tomographic data repository for computational x-ray science

Francesco De Carlo¹, Doğa Gürsoy¹, Daniel J Ching²,
K Joost Batenburg³, Wolfgang Ludwig⁴, Lucia Mancini⁵,
Federica Marone⁶, Rajmund Mokso⁷, Daniël M Pelt⁸, Jan Sijbers⁹
and Mark Rivers¹⁰

¹ Advanced Photon Source, Argonne National Laboratory, Lemont, IL, United States of America

² Oregon State University, Corvallis, OR, United States of America

³ Centrum Wiskunde & Informatica, Science Park 123, 1098 XG Amsterdam, Netherlands

⁴ European Synchrotron Radiation Facility, Grenoble, France

⁵ Elettra—Sincrotrone Trieste, Basovizza, Italy

⁶ Swiss Light Source, Paul Scherrer Institut, Villigen, Switzerland

⁷ MAX IV Laboratory, Lund, Sweden

⁸ Lawrence Berkeley National Laboratory, Berkeley, CA, United States of America

⁹ imec-Vision Lab, University of Antwerp, Antwerp, Belgium

¹⁰ University of Chicago, Chicago, IL, United States of America

E-mail: decarlo@aps.anl.gov

Received 27 June 2017, revised 17 September 2017

Accepted for publication 21 November 2017

Published 8 February 2018



Abstract

There is a widening gap between the fast advancement of computational methods for tomographic reconstruction and their successful implementation in production software at various synchrotron facilities. This is due in part to the lack of readily available instrument datasets and phantoms representative of real materials for validation and comparison of new numerical methods.

Recent advancements in detector technology have made sub-second and multi-energy tomographic data collection possible (Gibbs *et al* 2015 *Sci. Rep.* **5** 11824), but have also increased the demand to develop new reconstruction methods able to handle *in situ* (Pelt and Batenburg 2013 *IEEE Trans. Image Process.* **22** 5238–51) and dynamic systems (Mohan *et al* 2015 *IEEE Trans. Comput. Imaging* **1** 96–111) that can be quickly incorporated in beamline production software (Gürsoy *et al* 2014 *J. Synchrotron Radiat.* **21** 1188–93).

The x-ray tomography data bank, tomoBank, provides a repository of experimental and simulated datasets with the aim to foster collaboration among computational scientists, beamline scientists, and experimentalists and to accelerate the development and implementation of tomographic reconstruction methods for synchrotron facility production software by providing easy access to challenging datasets and their descriptors.

Keywords: tomography, tomographic data archive, tomographic data analysis, x-ray tomography

(Some figures may appear in colour only in the online journal)

1. Introduction

Current tomographic experiments include a broad range of advanced setups, combining new imaging modes and specialized sample environments, yielding datasets that are far larger and more complex than what has been designed for standard

software and algorithms. As examples, we mention complex environment control cells for flow and pressure experiments producing high resolution 3D imaging of dynamic systems and *in vivo* experiments for which dose reduction is essential.

These new types of tomographic data lead to a kind of paradox: on one hand, the dataset size is dramatically increasing,

while, on the other hand, these datasets contain *limited data*, in the sense that for each time-frame, the available data is noisy or contains just a limited set of images.

Besides the experimental part, the quality of the resulting images depends crucially on the algorithms that are employed when processing and reconstructing the data. As experiments increase in complexity, more complex and sophisticated algorithms [3] are needed to obtain the highest possible quality reconstructions leading to better quality data and, ultimately, to faster scientific discoveries.

The often complex parameter configuration required to use sophisticated reconstruction algorithms and the fact that their validity is often tested on a limited set of measurements or phantoms creates a disconnect: while sophisticated reconstruction algorithms are constantly being developed by applied mathematicians and computational scientists, experimental x-ray imaging facilities often lag behind, and, in most cases, continue to use classical reconstruction methods like filtered back projection [5] or gridrec [6].

A series of Lorentz-workshops [7–9] brought together active researchers from x-ray experimental facilities and from applied mathematical and computational sciences. The topics of discussion included x-ray tomography experiment, modeling and algorithms design, artefacts correction, and quantitative tomography including dimensional metrology and quantification of dynamic processes. The existence of the gap between sophisticated algorithms and experimental x-ray imaging became clear during these workshops and motivated a group of synchrotron facility scientists to informally collect and share challenging tomographic datasets.

These datasets were grouped under various ‘challenge’ categories, including round-Robin appendix A.1.1, *in situ* studies appendix A.1.2, dynamic systems appendix A.1.3, small field of view appendix A.1.4, phase contrast appendix A.1.5, etc to include the most challenging experimental scenarios like limited view, noisy projection, sample motion, instrument instability, beam hardening effects, strong ring artefacts etc.

2. TomoBank

In this paper, we formalize that effort by introducing tomoBank, a tomographic data bank aiming to provide applied mathematicians and computational scientists with an extensive catalog of challenging x-ray tomography raw datasets from different x-ray instruments, as well as software-generated phantoms that more closely represent material and live science samples compared to the standard phantoms used in the mathematical community.

The distribution of raw data will enable computational scientists to access all information available as collected during the measurement. This is critical as new algorithms sometimes need more/different information than existing standard algorithms require.

2.1. Comparison with other data-banks

Many of the existing data banks focus on the science surrounding the sample by providing data of biological systems [10, 11], materials [12–14] and simulations [15–18] but are rarely focused on the science surrounding the experimental techniques [19–21] and/or targeted to computational experts in that technique. TomoBank provides tomography technique specific datasets collected at different facilities and is directly aimed at computational scientists by including a ‘reference’ catalog containing datasets and phantoms to be used for algorithm validation and comparison and a ‘challenge’ catalog containing datasets requiring, in most cases, the development of new algorithms to obtain good quality reconstructions. While datasets transition from ‘challenge’ to ‘reference’ catalogs we also aim to create a network through tomoBank in which computational scientists can not only use existing data but also request a new ad hoc measurement to validate a new method or vice-versa to optimize the experiment design. For some advanced tomographic algorithms, such new measurements are crucial for experimental validation: optimizing the experiment for the chosen algorithm is either required, or can lead to significantly more accurate results [2, 3].

2.2. Data format

Datasets are collected and distributed in their original file format e.g. multiple TIFFs [22], dxfile [23], EDF [24], XRM [25], etc. However, for standard equally spaced angular scans, they are also converted and distributed as dxfile (a single HDF5 file containing raw, dark, and flat field arrays with all the required meta-data).

All parallel beam tomography datasets and phantoms are also distributed with a Python script to read and perform a basic tomographic reconstruction using tomoPy [4].

2.3. Simulated phantoms

Although having access to a variety of experimental data sets will be helpful, knowing the ground truth may be even more valuable when evaluating the performance of novel reconstruction algorithms or other post-processing techniques. This is why tomoBank also includes a selection of simulated phantoms constructed with XDesign [26] appendix A.2. XDesign is a Python module which can simulate the interaction of an x-ray probe with simulated phantoms constructed from simple geometric shapes. Developers testing new or on-the-fly reconstruction algorithms that may rely on unconventional data acquisition schemes will find these more useful.

2.4. Browsing tomoBank

Access to tomoBank can be found at [27]. All datasets and phantoms are listed with a unique ID number and can be sorted

Table A1. Datasets currently available in tomoBank.

Category	Facility	Name/Description	tomo_ID
Round Robin	APS 2-BM	B1 hornby	00001
— — — — —	APS 2-BM	N1 blakely	00002
— — — — —	SLS Tomcat	B1 hornby	00003
— — — — —	SLS Tomcat	N1 blakely	00004
— — — — —	ALS 8.3.2	B1 hornby	00005
— — — — —	ALS 8.3.2	N1 blakely	00006
Limited view	APS 32-ID	High pressure at 0.3 GPa	00007
— — — — —	APS 32-ID	High pressure at 0.57 GPa	00008
— — — — —	APS 32-ID	High pressure at 2 GPa	00009
— — — — —	APS 32-ID	High pressure at 8.59 GPa	00010
— — — — —	APS 32-ID	High pressure at 13.37 GPa	00011
— — — — —	APS 32-ID	High pressure at 17.44 GPa	00012
— — — — —	APS 32-ID	High pressure at 19 GPa	00013
— — — — —	APS 32-ID	High pressure at 21.39 GPa	00014
— — — — —	APS 32-ID	High pressure at 26.17 GPa	00015
— — — — —	APS 32-ID	High pressure at 29.5 GPa	00016
— — — — —	APS 32-ID	High pressure at 33.07 GPa	00017
— — — — —	APS 32-ID	High pressure at 41.88 GPa	00018
— — — — —	APS 32-ID	High pressure at 47.89 GPa	00019
— — — — —	APS 32-ID	High pressure at 54.73 GPa	00020
— — — — —	APS 32-ID	High pressure at 59 GPa	00021
Small field of view	Elettra symmep	Double bamboo tin wire	00022
— — — — —	Elettra symmep	Bamboo wires 2	00023
— — — — —	Elettra symmep	Polystyrene	00024
— — — — —	Elettra symmep	Rock no oil	00025
— — — — —	Elettra symmep	Rock oil saturated	00026
— — — — —	Elettra symmep	bone local	00027
— — — — —	Elettra tomolab	bone local stone	00028
— — — — —	Elettra tomolab	candie local	00029
— — — — —	Elettra tomolab	smiling sample	00030
Fast tomography	APS 13-BM	dorthe_F_002	00031

Table A2. Phantoms currently available in tomoBank.

phantom_ID	Facility	Name/Description
00001	XDesign	SlantedSquares(count = 16, angle = $5/360 \cdot 2 \cdot \pi$, gap = 0.01)
00002	XDesign	UnitCircle(radius = 0.4, mass_atten = 1)
00003	XDesign	HyperbolicConcentric()
00004	XDesign	DynamicRange(steps = 16, jitter = True, shape = 'square')
00005	XDesign	DogaCircles(n_sizes = 8, size_ratio = 0.5, n_shuffles = 0)
00006	XDesign	SiemensStar(32)
00007	XDesign	WetCircles()
00008	XDesign	Foam(size_range = [0.05, 0.01], gap = 0, porosity = 1)
00009	XDesign	Foam(size_range = [0.07, 0.01], gap = 0, porosity = 0.75)
00010	XDesign	Foam(size_range = [0.1, 0.01], gap = 0, porosity = 0.5)
00011	XDesign	Foam(size_range = [0.1, 0.01], gap = 0.015, porosity = 0.5)

by facility, instrument, or ‘challenge’ category. Once a dataset or phantom is selected to be part of tomoBank, a digital object identifier [28] will be assigned to point to a permanent dataset web page containing its meta data (authors, supported article, data citation, description and the data download link, etc). To efficiently download tomographic datasets, tomoBank relies on Globus [29].

2.5. Contributing data to tomoBank

TomoBank provides links to published tomographic datasets valuable for the advancement of computational methods. To submit a dataset follow the instruction at [27] by providing author list, dataset title, description, DOI number, license, link to the raw data and, if the dataset was used to support a scientific publication, the reference to the supported publication.

Table A3. Limited view high-pressure experiment conditions.

Instrument	APS 32-ID TXM
Energy	8000 eV
Monochromator	Double crystal Si (1,1,1) 180°
Scan range	
Number of projections	359
White fields	20 before
Dark fields	8 before
Exposure time	15 s
PixelSize	13.8 nm
Comment	10x 60 nm ZP

For datasets hosted on Argonne National Laboratory servers, the DOI number and link to a permanent download web page can be provided by the Argonne Research Library (an example is at [30]).

3. Conclusions

We introduced tomoBank, a repository of tomographic datasets and phantoms designed to provide computational scientists with a reference data bank to verify, improve, and design new advanced computational methods for reconstruction. We also plan to extend tomoBank with an extended library of reconstructed datasets that we hope will be useful for the development of new segmentation and visualization toolkits.

Acknowledgments

This research used resources of the Advanced Photon Source, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Portions of this work were performed at GeoSoilEnviroCARS (The University of Chicago, Sector 13), Advanced Photon Source (APS), Argonne National Laboratory. GeoSoilEnviroCARS is supported by the National Science Foundation—Earth Sciences (EAR-1128799) and Department of Energy- GeoSciences (DE-FG02-94ER14466). Peter Cloetens is acknowledged for the interesting discussion at the Lorentz workshop that lead to this paper. Networking support was provided by the EXTREMA COST Action MP1207.

Appendix. Datasets and phantoms currently available in tomoBank

Below is a brief description of the datasets (table A1) and phantoms (table A2) already available in tomoBank. We expect this list to continue to grow, so we recommend following this project on [31].

Table A4. Elettra Syrmep experimental conditions.

tomo_ID	00022	00023	00024	00025	00026
X-ray energy	— White beam mode —				
Filter	— 1.5 mm Si + 0.025 mm Mo —				
Exposure time	— 1 s —				
Detector	— sCMOS 16-bit —				
Sample-to-detector distance	— 150 mm —				
Pixel size	2.4 μm	2.6 μm	2.4 μm	2.4 μm	2.4 μm
Projections	1800	1800	1800	400	400
Total scanned angle	— 180° —				

A.1. Datasets

A.1.1. Round-Robin. As in a round-Robin tournament where each player faces every other player, these are series datasets of specimens imaged at multiple facilities.

The *Hornby-Blakely* datasets (tomo_00001 to tomo_00006) [30] consist of two shale samples obtained from the North Sea and the Upper Barnett Formation in Texas [32] (see figure A1). These specimens were imaged under similar conditions at the advanced photon source (APS) of Argonne National Laboratory, USA, at the swiss light source (SLS) of the Paul Scherrer Institut, Switzerland and at the advanced light source (ALS) of Lawrence Berkeley National Laboratory, USA. The goals of this round-Robin project were to (i) characterize microstructures and porosity on the micrometer scale, (ii) compare results measured at three synchrotron facilities, and (iii) identify optimal experimental conditions of high-resolution tomography (microCT) for fine-grained materials. Shale was chosen for the round-Robin because it is a challenging material due to its multiphase composition, small grain size, low but significant amount of porosity, as well as strong shape- and lattice-preferred orientation. All instruments used a 10 \times objective lens with an effective pixel size of $\sim 0.7 \mu\text{m}$.

A.1.2. In situ studies. Environment control cells, which can control physical parameters like pressure, temperature chemical environment etc around a sample, allow for 3D characterizations of materials and biological systems in real conditions; however, these conditions or the chamber itself can create difficult data acquisition conditions.

The *Limited View* datasets (tomo_00007 to tomo_00021) contain nano tomography (nanoCT) data of a small particle of Ce_6Al_4 undergoing a pressure increase (see figure A2). Data is collected at 15 pressure levels from 0.3 GPa to 59 GPa. Because in 86 of the 359 projections, the particle partially hidden from the field of view by the high-pressure cell's load frame, the challenge of this dataset is to determine accurately the particle volume as a function of pressure. The *Limited View* datasets are described in tables A1 and A3 and are available for download at [27] under tomo_00007 to tomo_00021.

Table A5. TomoLab @Elettra experimental conditions.

tomo_ID	00027	00028	00029	00030
Voltage	110 kV	60 kV	70 kV	60 kV
Current	145 mA	250 mA	114 mA	300 mA
Filter	0.1 mm Cu	1.0 mm Al	1.5 mm Al	1.5 mm Al
Focal spot	16 μm	16 μm	8 μm	20 μm
Exposure time/proj	4 s	4 s	6 s	1 s
Detector	— CCD 12-bit —			
Source-to-sample distance	100 mm	100 mm	100 mm	250 mm
Source-to-detector distance	250 mm	250 mm	250 mm	350 mm
Pixel size	10 μm	10 μm	10 μm	58 μm
Projections	1800	1800	1800	720
Total scanned angle	— 360° —			
Horizontal image size	2004 pixels	2004 pixels	2004 pixels	668 pixels
Vertical image size	1335 pixels	1335 pixels	1335 pixels	445 pixels
Detector pixel size	25 μm	25 μm	25 μm	75 μm

Table A6. The Dorthe_F_002 experimental conditions.

tomo_ID	00031
Data file name	Dorthe_F_002.nc
Sample name	Glass beads from Dorthe Wildenschild, Oregon State University
Instrument	APS 13-BM-D
Energy	33.269 keV
Monochromator	Double crystal Si (1,1,1)
Scan range	180°
Number of projections	900
White fields	20
Dark fields	None
Exposure time	0.006 s
PixelSize	3.18 μm /pixel
Comment	80 frames/s, 11.25 s total collection time

The *rock with and without oil* (tomo_00025 to tomo_00026) datasets with experimental conditions reported in table A4 represent a microCT geo-science example (figure A6) where the background fluid is replaced.

A.1.3. Dynamic systems. Datasets in this category aim to capture fast evolving samples in 3D by adjusting several parameters including x-ray energy, exposure time, frame rate, rotary stage speed etc.

The *Dorthe_F_002* dataset (tomo_00031) has an exposure time that is much shorter than what is optimal for data acquisition (figure A7). The sample description and the experimental conditions for *Dorthe_F_002* are reported in table A6.

A.1.4. Small region of interest. These datasets have specimens which are larger than the field of view.

Double bamboo tin wire (figure A3), *bamboo wire 2* (figure A4), *polystyrene* (figure A5) (tomo_00022 to tomo_00024) are microCT datasets acquired at the SYRMEP beamline of Elettra-Sincrotrone Trieste (Elettra) laboratory (Italy) in nearly-parallel beam geometry. The related sample description and the experimental conditions are reported in table A4.

Bone local, bone local stone, candie local, and smiling sample (tomo_00027 to tomo_00030) were acquired by microfocus CT at the TomoLab station of Elettra using a cone-beam geometry. The sample description and the experimental conditions are reported in table A5.

A.1.5. Phase contrast. When performing in-line (propagation-based) phase-contrast x-ray tomographic measurements, images are obtained with both absorption and refraction contribution. This allows one to obtain an enhanced contrast (Fresnel diffraction) at the edges of sample features producing local phase variations. By this technique, 3D variations of the electron density in the sample can be resolved and the experiments aim at a quantitative 3D reconstruction of the complex refractive index. The recorded sample radiographs are determined by the projection of the real and imaginary part of the refractive index within the sample volume. The Fresnel fringes visible at the edges of the phase objects can be considered as an artefact in the reconstructed slices and should be reduced or removed before any quantitative post-processing and analysis of the images.

Radiographs can be recorded at a single or multiple distances. If a single distance is employed, a phase retrieval

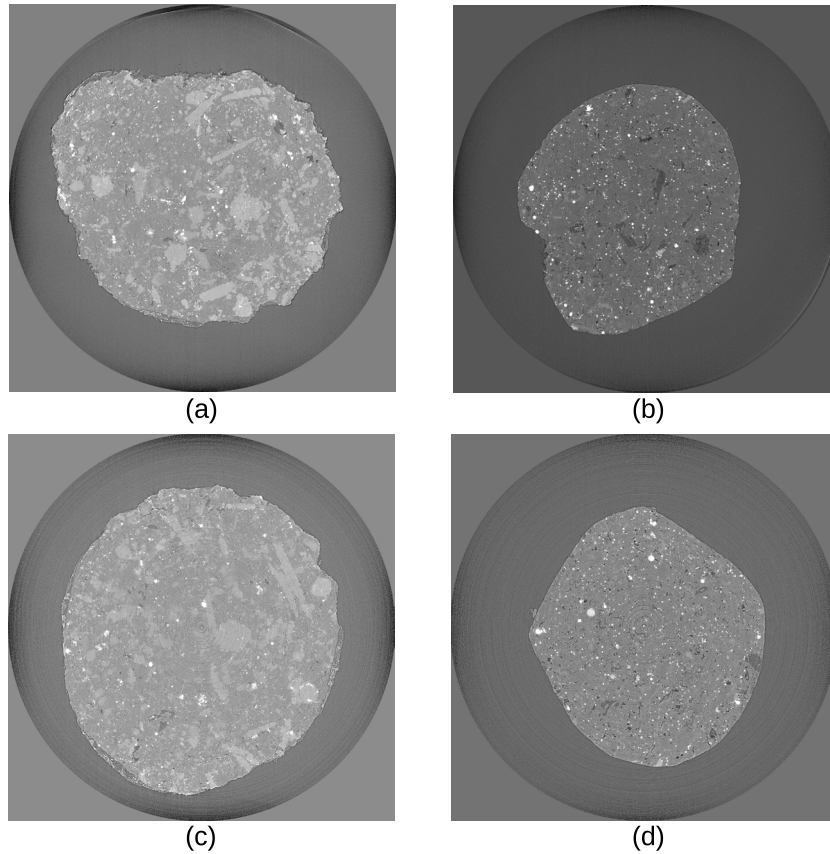


Figure A1. The round-Robin dataset [30]. (a) and (b) are slices of rock samples measured by TOMCAT at the swiss light source. (c) and (d) are slices of the same rock sample measured at advanced photon source beamline 2-BM. (a) and (c) are from sample Upper Barnett Formation in Texas and (b) and (d) are from the North Sea.

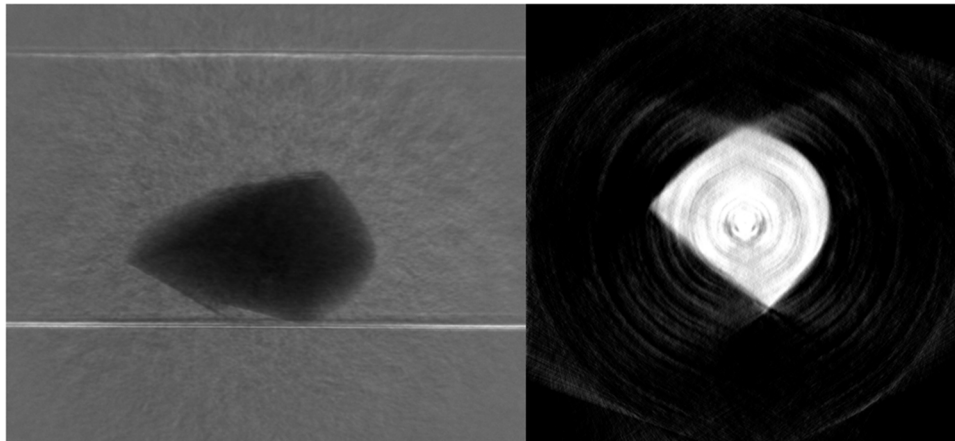


Figure A2. Projection and reconstructions of a single slice of a sample from tomo_00007 in a high-pressure diamond anvil cell which blocks 86 of the 359 projections over 180° , with 2560 detector pixels per projection [36]. Reconstructions are computed with SIRT with a non-negativity constraint using the ASTRA toolbox [37].

algorithm is applied most often based on the homogeneity assumption and the transport of intensity equation proposed by [33]. In the case of multiple distances [34], the problem is afforded by using contrast transfer functions. Because the phase-retrieval application usually produces a blurring of the reconstructed slices, it is crucial to optimize the algorithm

parameters in order to obtain a good compromise between spatial resolution and discrimination of the phases of interest in the final images.

The common reconstruction approach foresees a two-steps procedure: (1) a phase-retrieval algorithm is individually applied to the acquired projections; (2) a full

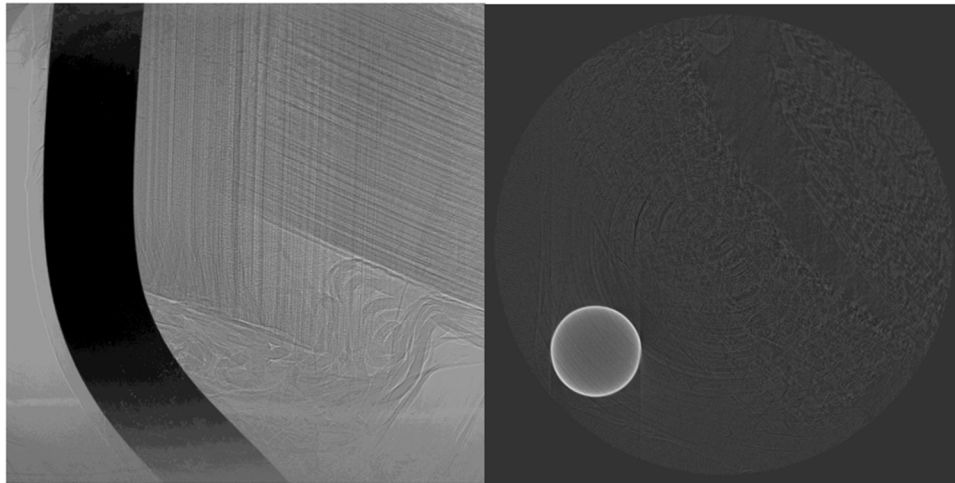


Figure A3. Projection and reconstructed axial slice of a sample constituted of two pieces of bamboo with different fibers orientation plus a tin wire. The sample is named Double Bamboo Tin Wire (tomo_0022) and experimental conditions are reported at table A4.

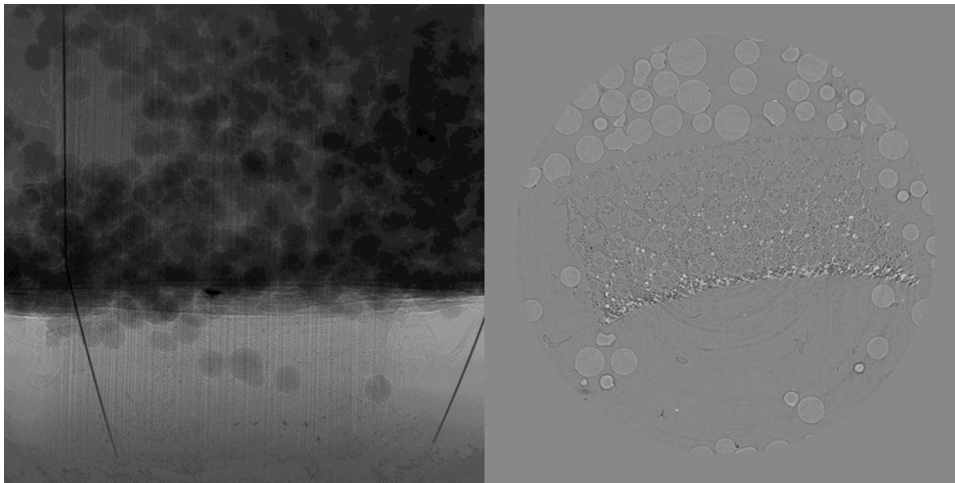


Figure A4. Projection and reconstructed axial slice of a sample constituted of a piece of bamboo plus two thin metallic wires inserted in a box partially filled with Si spheres. The sample is named Bamboo Wires 2 (tomo_0023) with experimental condition at table A4.

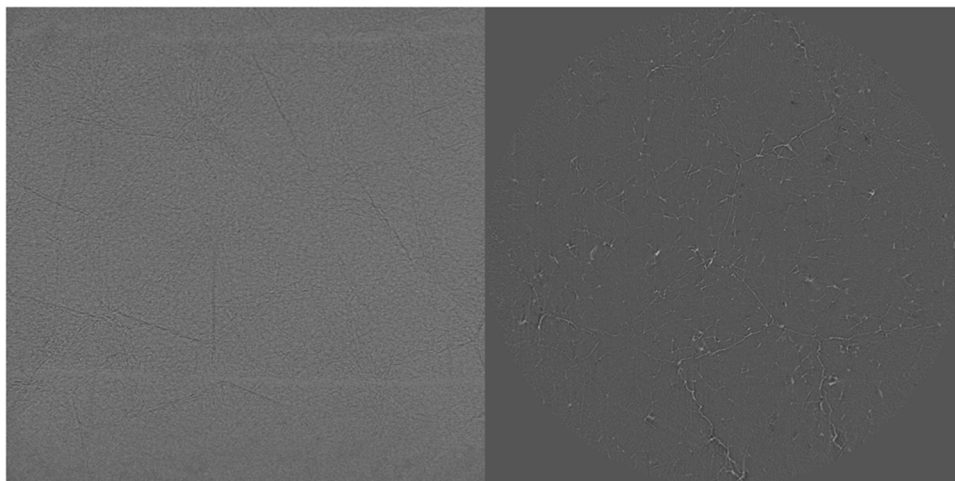


Figure A5. Polystyrene foam (tomo_0024) with experimental condition at table A4.

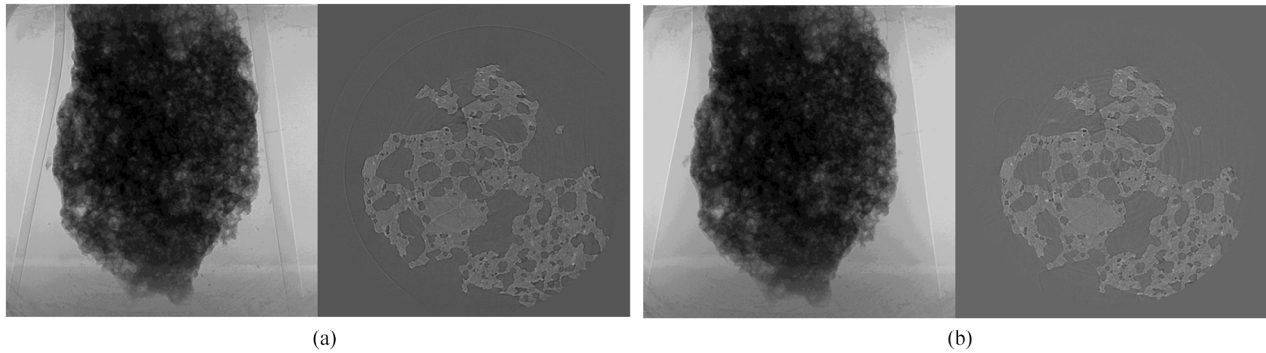


Figure A6. A dry rock sample in a plastic tube (tomo_0025) and the same rock sample with the plastic tube filled with vegetable oil (tomo_0026), with experimental condition at table A4 (a) tomo_0025. (b) tomo_0026.

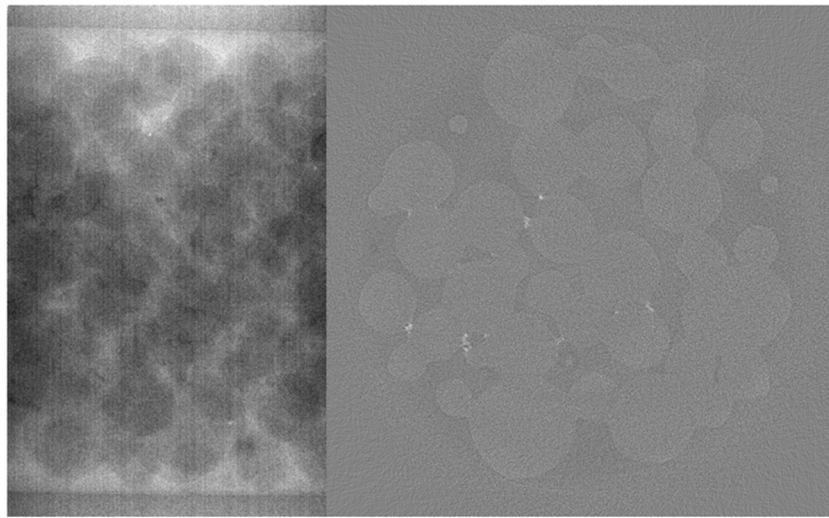


Figure A7. Dorthe_F_002 dataset (tomo_0031) with experimental condition at table A6.

tomographic reconstruction is then performed using as input data the phase-retrieved projections. [27] includes several phase-contrast datasets (see tomo_00064 to tomo_00068).

A.1.6. Beam hardening and ring artefacts. When using polychromatic x-ray sources for tomographic experiments, the reconstructed slices in many cases display the so-called beam hardening artefacts [35]. These artefacts are due to the differential absorption of the x-ray spectrum by the sample and lead to a misleading recovery of the linear absorption coefficients, mainly appearing as brighter sample borders in the slices. For regular objects the effect is easily detected, but for irregular objects it is commonly difficult to differentiate between beam hardening artefacts and actual material variations because of the additional appearance of dark stripes in the images. Often it is also possible to observe the presence of concentric rings which are typical artefacts arising from inhomogeneities in

the individual pixel response of detector elements. These artefacts, named *ring artefacts*, seriously complicate the visibility of fine details in the sample and the post-processing of the images.

Several datasets containing beam hardening and ring artefacts are available in [27], including data obtained by using the TomoLab station at Elettra.

A.2. The phantoms

All phantoms available in tomoBank are generated using XDesign [26]. Each phantom is discretized to obtain the ground truth reconstructed slice and is projected on a 256 pixel detector in 512 equally spaced angles. All phantoms are saved as dxfile [23] and distributed together with the XDesign Python code used to generate them together with a basic reconstruction script using [4]. Some of the phantoms included in tomoBank are listed in table A2 and shown in figure A8.

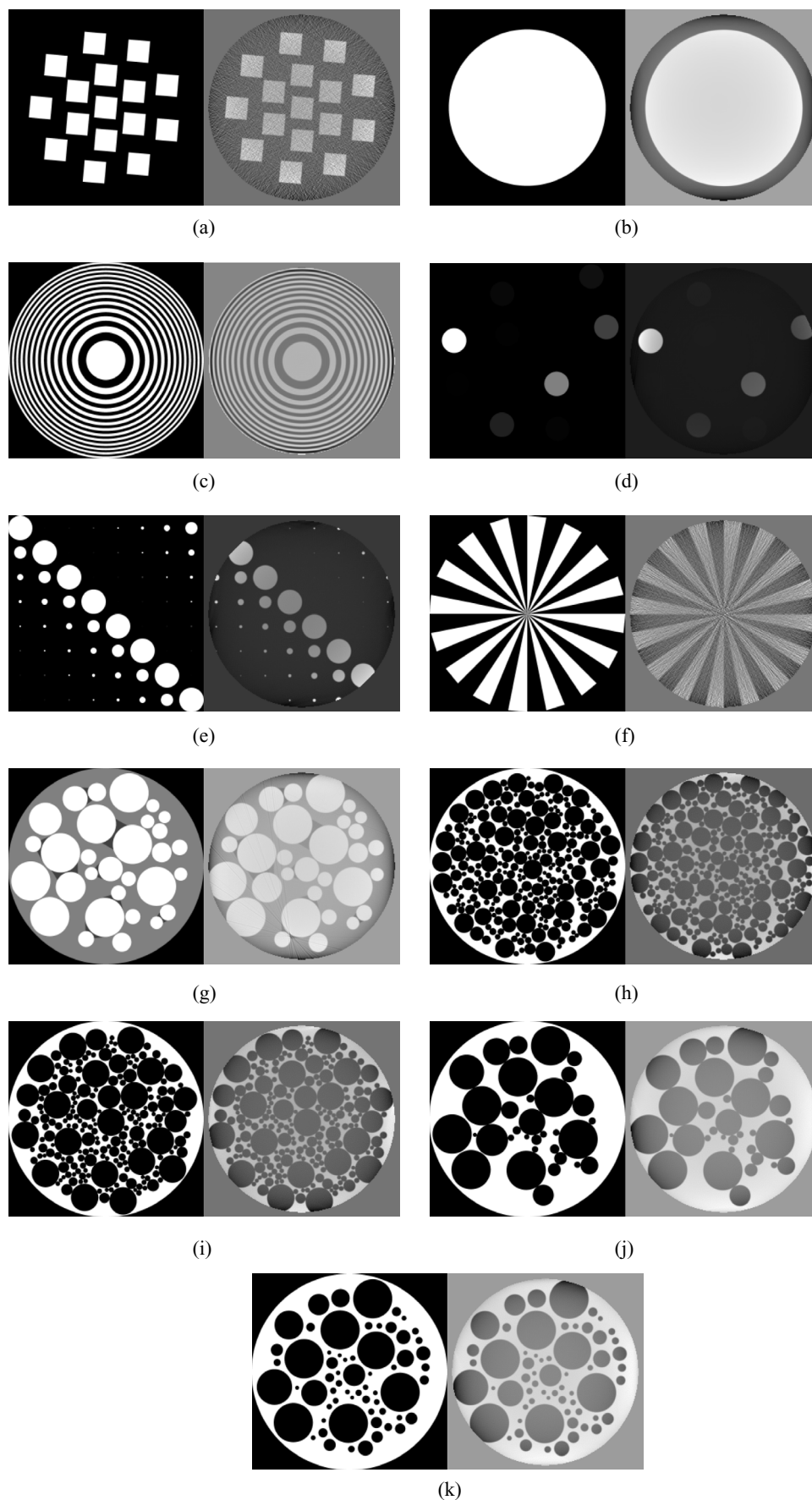


Figure A8. XDesign phantoms (left) and projections reconstructed with tomoPy (right). (a) 00001. (b) 00002. (c) 00003. (d) 00004. (e) 00005. (f) 00006. (g) 00007. (h) 00008. (i) 00009. (j) 00010. (k) 00011.

ORCID iDs

Francesco De Carlo  <https://orcid.org/0000-0003-1068-7785>
 Daniel J Ching  <https://orcid.org/0000-0003-3992-190X>
 Lucia Mancini  <https://orcid.org/0000-0003-2416-3464>
 Daniël M Pelt  <https://orcid.org/0000-0002-8253-0851>

References

- [1] Gibbs J W, Mohan K A, Gulsoy E B, Shahani A J, Xiao X, Bouman C A, De Graef M and Voorhees P W 2015 The three-dimensional morphology of growing dendrites *Sci. Rep.* **5** 11824
- [2] Pelt D M and Batenburg K J 2013 Fast tomographic reconstruction from limited data using artificial neural networks *IEEE Trans. Image Process.* **22** 5238–51
- [3] Mohan K A, Venkatakrishnan S V, Gibbs J W, Gulsoy E B, Xiao X, Graef M D, Voorhees P W and Bouman C A 2015 Timbir: a method for time-space reconstruction from interlaced views *IEEE Trans. Comput. Imaging* **1** 96–111
- [4] Gürsoy D, De Carlo F, Xiao X and Jacobsen C 2014 Tomopy: a framework for the analysis of synchrotron tomographic data *J. Synchrotron Radiat.* **21** 1188–93
- [5] Shepp L A and Logan B F 1974 The fourier reconstruction of a head section *IEEE Trans. Nucl. Sci.* **21** 21–43
- [6] Dowd B A, Campbell G H, Marr R B, Nagarkar V V, Tipnis S V, Axe L and Siddons D P 1999 Developments in synchrotron x-ray computed microtomography at the national synchrotron light source *Developments in X-Ray Tomography II* vol 3772 pp 224–36
- [7] Lorentz 2014 Advanced x-ray tomography: experiment, modeling and algorithms www.lorentzcenter.nl/lc/web/2014/614/info.php3?wsid=614
- [8] Lorentz 2015 Artefacts in x-ray tomography www.lorentzcenter.nl/lc/web/2015/689/info.php3?wsid=689
- [9] Lorentz 2016 Quantitative 3D x-ray imaging: from tomographic images to metrics www.lorentzcenter.nl/lc/web/2016/781/info.php3?wsid=781
- [10] Howe D G et al 2013 ZFIN, the Zebrafish model organism database: increased support for mutants and transgenics *Nucl. Acids Res.* **41** D854–60
- [11] Bryson-Richardson R, Salgado D, Currie P and Marcelle C 2016 The ZebraFish Anatomy Portal <http://zebrafish.anatomyportal.org/>
- [12] Materials 2016 The Materials Project <https://materialsproject.org/>
- [13] Prisms 2016 The Materials Commons Projects www.prisms-center.org/
- [14] Nist 2016 NIST databases www.calphad.org/
- [15] Kaufman L 2016 Computer Coupling of Phase Diagrams and Thermochemistry www.calphad.org/
- [16] Morgan D and Curtarolo S 2016 High-throughput Calculations for Materials Properties <http://materials.duke.edu/afLOW.html>
- [17] Kirklin S, Saal J E, Meredig B, Thompson A, Doak J W, Aykol M, Rühl S and Wolverton C 2015 The open quantum materials database (oqmd): assessing the accuracy of dft formation energies *NPJ Comput. Mater.* **1** 15010
- [18] iastate 2016 A Repository for 3D Experiments and Simulations on a Variety of Material Systems <https://cosmicweb.mse.iastate.edu/wiki/display/home/Materials+Atlas+Home>
- [19] Maia F R N C 2012 The coherent x-ray imaging data bank *Nat. Methods* **9** 854–5
- [20] Berman H M, Westbrook J, Feng Z, Gilliland G, Bhat T N, Weissig H, Shindyalov I N and Bourne P E 2000 The protein data bank *Nucleic Acids Res.* **28** 235–42
- [21] Lawson C L et al 2011 EMDDataBank.org: unified data resource for CryoEM *Nucleic Acids Res.* **39** D456–64
- [22] tiff 2016 The tiff file format <https://en.wikipedia.org/wiki/TIFF>
- [23] De Carlo F et al 2014 Scientific data exchange: a schema for HDF5-based storage of raw and analyzed data *J. Synchrotron Radiat.* **21** 1224–30
- [24] ESRF 2016 The EDF file format www.esrf.eu/Instrumentation/software/data-analysis/OurSoftware/SAXS/SaxsHeader
- [25] Liu Y, Meirer F, Williams P A, Wang J, Andrews J C and Pianetta P 2012 TXM-Wizard: a program for advanced data collection and evaluation in full-field transmission x-ray microscopy *J. Synchrotron Radiat.* **19** 281–7
- [26] Ching D J and Gürsoy D 2017 XDesign: an open-source software package for designing x-ray imaging phantoms and experiments *J. Synchrotron Radiat.* **24** 537–44
- [27] De Carlo F 2016 The x-ray Tomography Data Bank www.aps.anl.gov/tomobank
- [28] DOI 2016 The Digital Object Identifier <http://doi.org/>
- [29] Globus 2016 www.globus.org/
- [30] Kanitpanyacharoen W, Parkinson D Y, De Carlo F, Marone F, Stapanoni M, Mokso R, MacDowell A and Wenk H R 2017 The round-Robin data set <https://doi.org/10.17038/XSD/1344306>
- [31] github 2017 The TomoBank Repository <https://github.com/tomography/tomobank/watchers>
- [32] Kanitpanyacharoen W, Parkinson D Y, De Carlo F, Marone F, Stapanoni M, Mokso R, MacDowell A and Wenk H R 2013 A comparative study of x-ray tomographic microscopy on shales at different synchrotron facilities: ALS, APS and SLS *J. Synchrotron Radiat.* **20** 172–80
- [33] Paganin D, Mayo S C, Gureyev T E, Miller P R and Wilkins S W 2002 Simultaneous phase and amplitude extraction from a single defocused image of a homogeneous object *J. Microsc.* **206** 33–40
- [34] Cloetens P, Ludwig W, Baruchel J, Van Dyck D, Van Landuyt J, Guigay J P and Schlenker M 1999 Holotomography: quantitative phase tomography with micrometer resolution using hard synchrotron radiation x rays *Appl. Phys. Lett.* **75** 2912–4
- [35] Baruchel J, Buffiere J Y and Maire E 2000 *X-Ray Tomography in Material Science* (France: Hermes science publications)
- [36] Pelt D, Gürsoy D, Palenstijn W, Sijbers J, Carlo F D and Batenburg K 2016 Integration of tomopy and the astra toolbox for advanced processing and reconstruction of tomographic synchrotron data *J. Synchrotron Radiat.* **23** 842–9
- [37] van Aarle W, Palenstijn W J, Cant J, Janssens E, Bleichrodt F, Dabavolski A, De Beenhouwer J, Batenburg K J and Sijbers J 2016 Fast and flexible x-ray tomography using the ASTRA toolbox *Opt. Express* **24** 25129–47