Using Generative Adversarial Networks to match experimental and simulated inelastic neutron scattering data

Andy S. Anker[1*], Keith T. Butler[2], Duc Le[3], Jola Mirecka[2], Tom Burnley[2], Antonis Vamvakeros[4], Mohsen Danaie[5], Jeyan Thiyagalingam[2]

1: Nano-Science Center and Department of Chemistry, University of Copenhagen
2: Scientific Computing Department, Rutherford Appleton Laboratory
3: ISIS Neutron and Muon Source, Rutherford Appleton Laboratory
4: Finden Limited, Abingdon OX14 5EG, United Kingdom
5: Electron Physical Science Imaging Centre, Diamond Light Source Ltd & Department of Materials, University of Oxford
*andy@chem.ku.dk

Science and Technology Facilities Council

UNIVERSITY OF COPENHAGEN

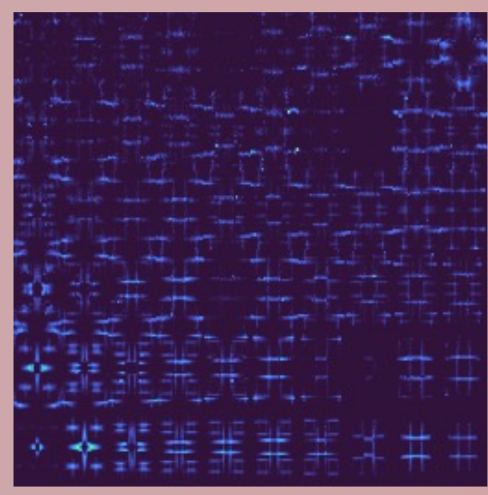
Scientific Computing
ISIS Neutron and Muon Source

Introduction

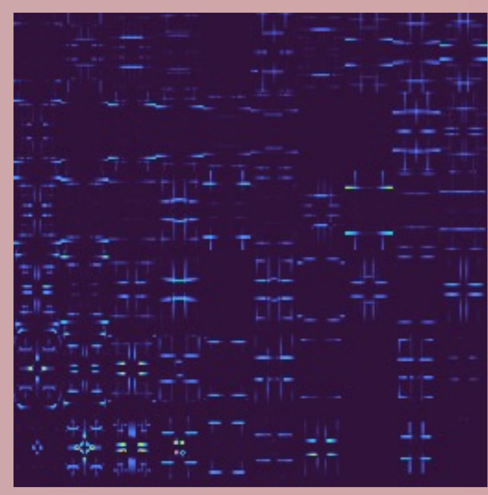
During the past decades, research in materials science has been accelerated by the rapid development of synchrotron and neutron sources.[1] Conventional data analysis approaches using minimization techniques, such as least-squares fitting algorithms, cannot keep up with the increasing size of measured datasets. Consequently, data analysis is becoming a bottleneck for research in materials science.[2-3] Therefore, it is of great importance to improve the current state-of-the-art for data analysis for materials science, particularly utilizing recent developments in artificial intelligence and machine learning (ML).[2-4] One of the unsolved problems in this context is to match the simulated datasets that the ML algorithms are trained on to the experimental datasets. This has particularly been a problem for the analysis of inelastic neutron scattering (INS), where it is computationally expensive to ensure that simulated data correctly mimics the experimental signal and background.[5]

[1] Wang, C. et al., *Synchrotron Big Data Science,* Small **2018**. [2] Agrawal, A. et al., *Perspective: Materials informatics and big data: Realization of the "fourth paradigm" of science in materials science,* APL Materials **2016**. [3] Butler, K. T. et al., *Machine learning for molecular and materials science,* Nature **2018**
[4] Hey, T. et al, *Machine learning and big scientific data*, Phil. Trans. R. Soc. Lond. A **2020**. [5] Butler, K. T. et al., *Interpretable, calibrated neural networks for analysis and understanding of inelastic neutron scattering data* , J. Phys.: Condens. Matter **2021.** [6] https://spinw.org/

Simulated data including the resolution function



Simulated data excluding the resolution function

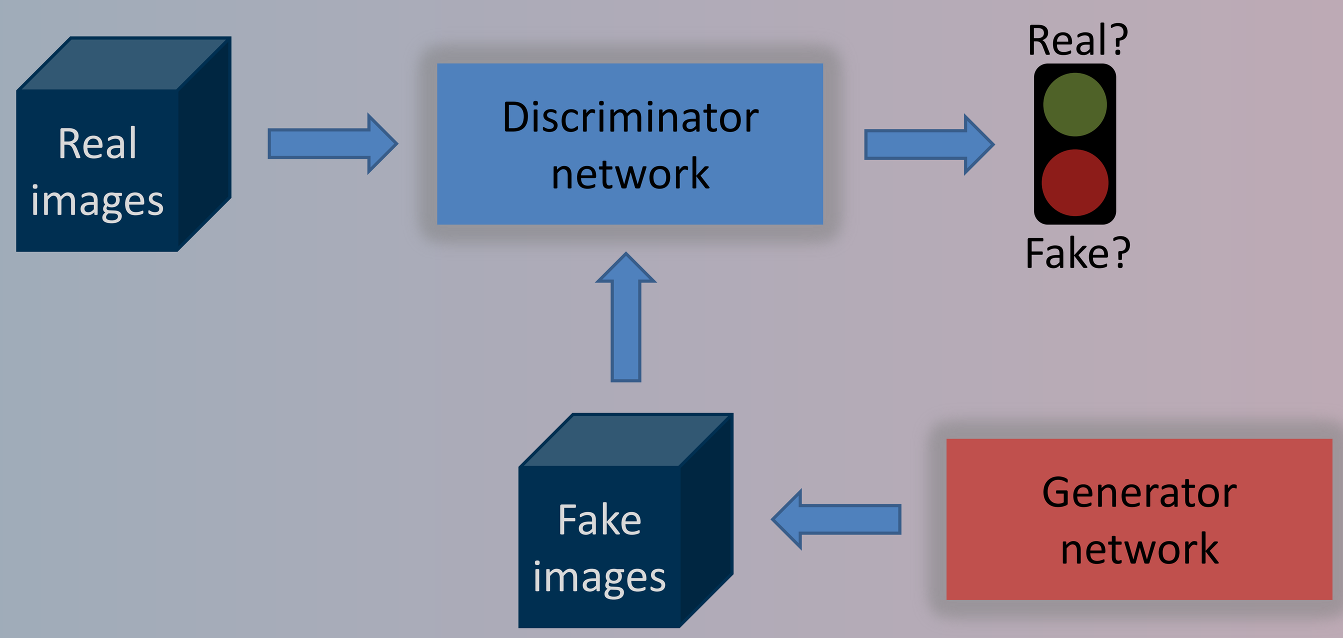


The concept of a generative adversarial network (GAN)

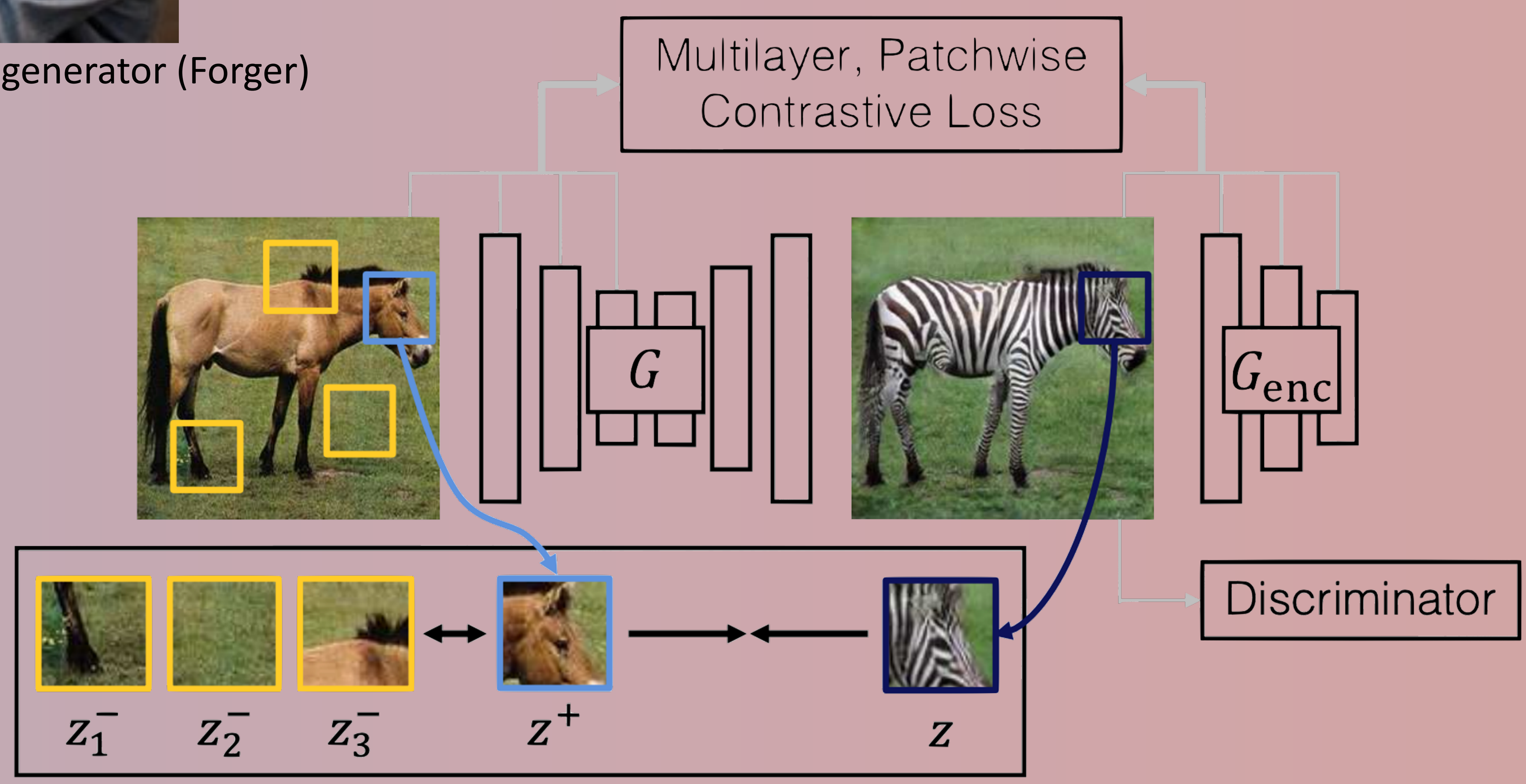
Real data (Paintings)

Fake data (painting) by generator (Forger)

General architecture of a GAN



Learning from patches



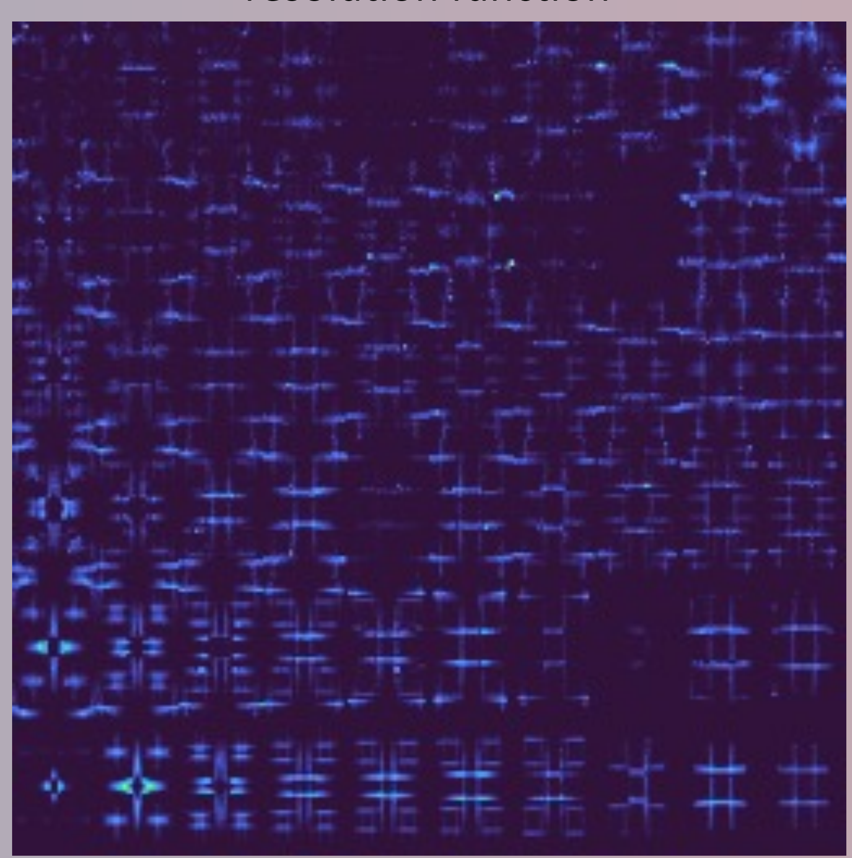
Patchwise Contrastive Learning

Park, T., et al., *Contrastive Learning for Unpaired Image-to-Image Translation, ArXiv* ***2020***

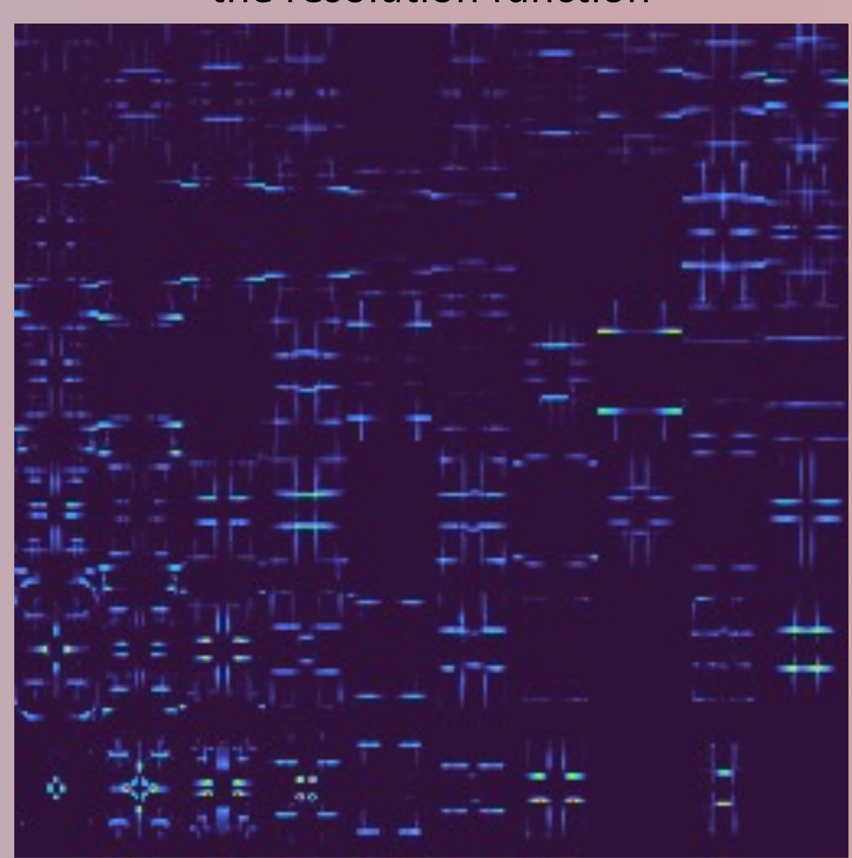
Matching experimental and simulated inelastic neutron scattering (INS) data

We are developing generative adversarial networks (GANs) that can learn to make simulated INS data that matches experimental INS dataset under a second. This GAN-based approach, once trained, will be deployed in a range of scenarios for analysing and understanding INS dataset. The GAN can be used to help classify materials structures from the INS datasets and to work with other ML and non-ML (e.g. Spin-W[6]) algorithms which can estimate magnetic Hamiltonian parameters from INS data.

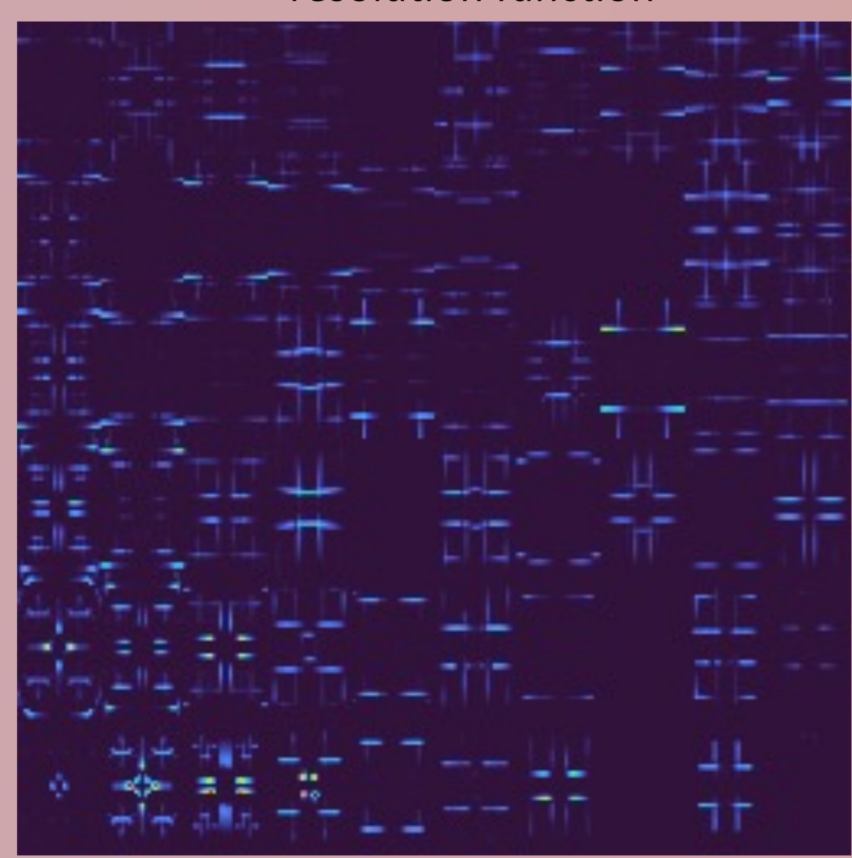
Simulated: data including the resolution function



GAN generated: data excluding the resolution function



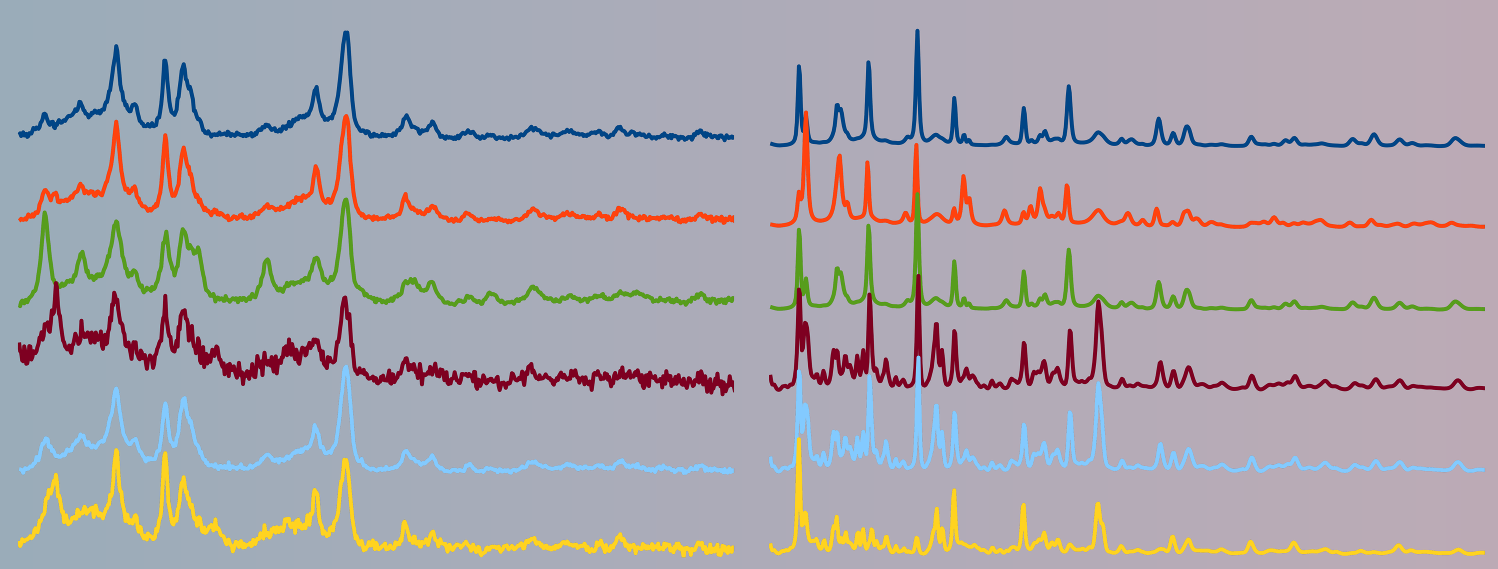
Simulated: data excluding the resolution function



Work in progress: Applying the framework on data from scattering (1D), spectroscopy (2D) and imaging (3D)

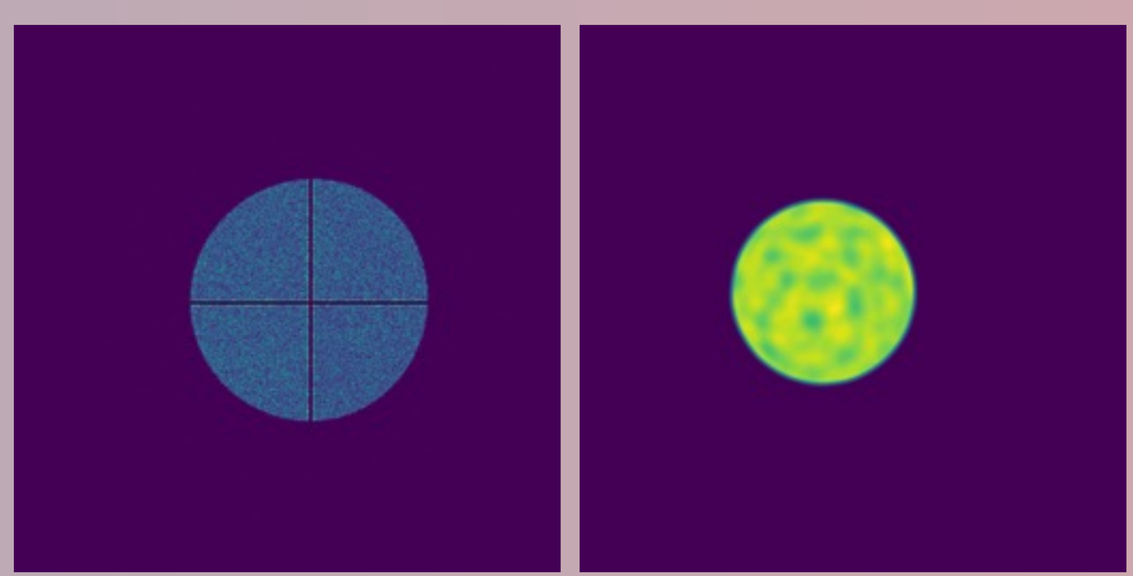
X-ray tomography (1D noise removal)

Noisy — Clean



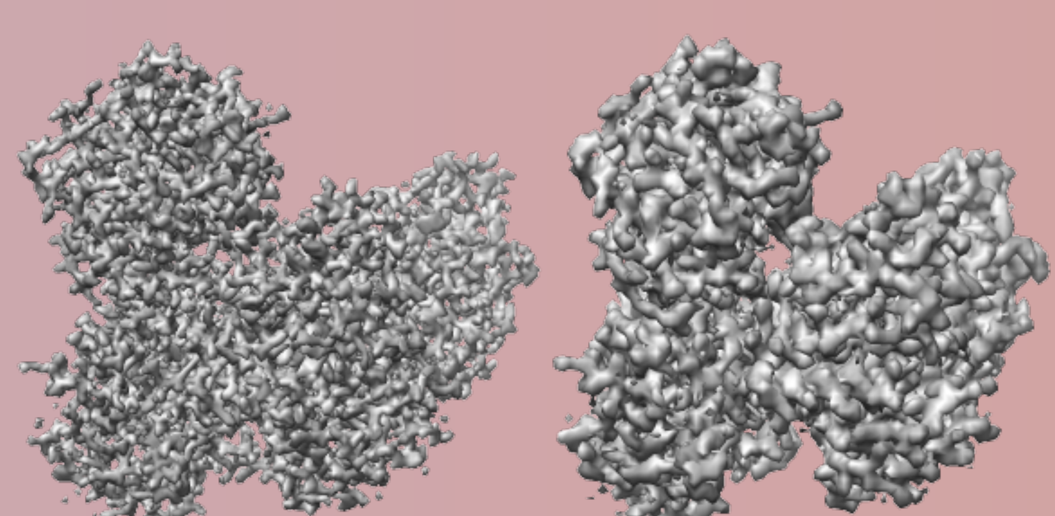
Ptychography (2D simulated ⟷ experimental)

Experimental — Simulated



Cryo EM (3D super resolution)

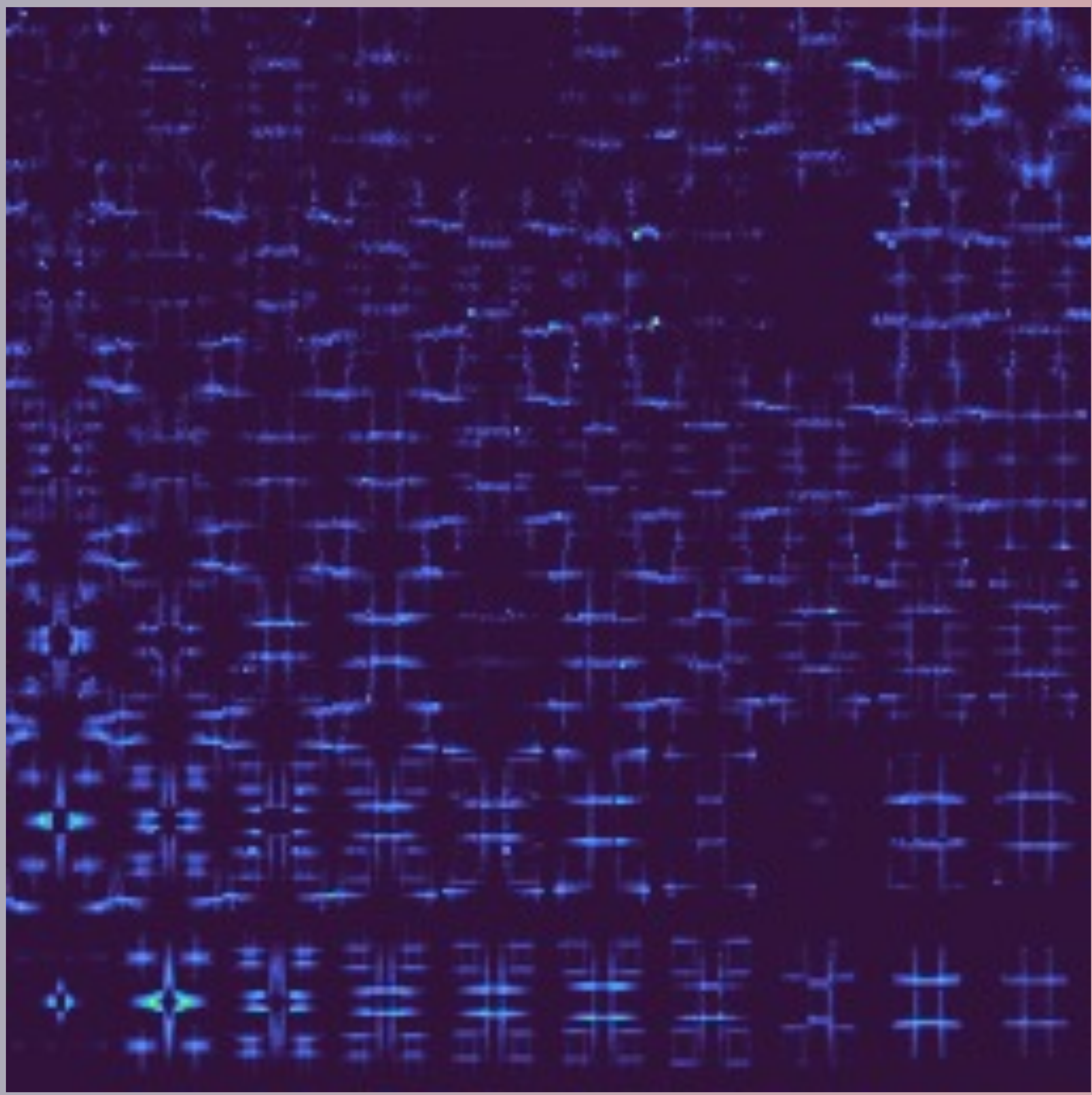
3 Å resolution — 4 Å resolution



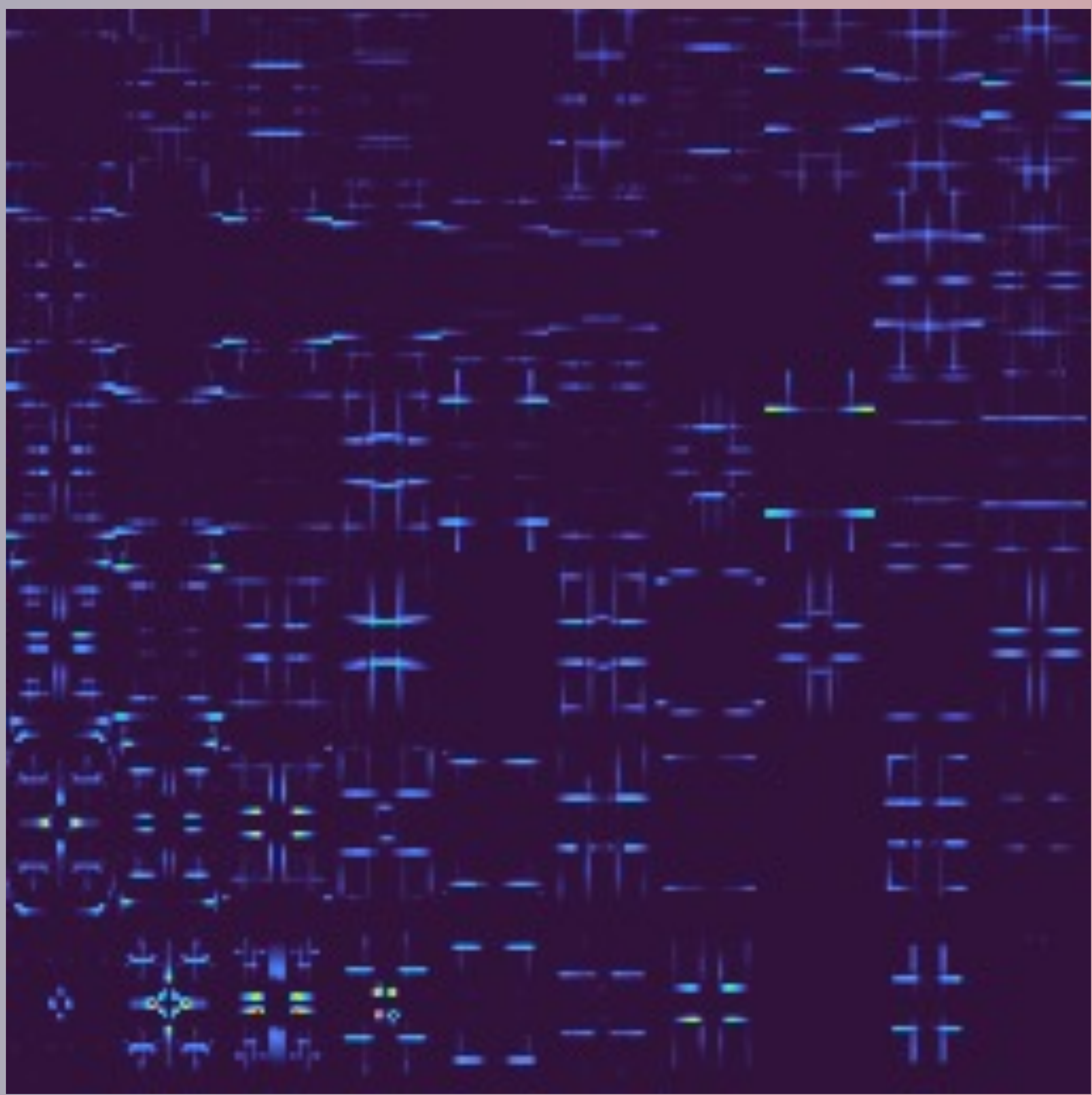
Acknowledgments

A. S. A. would like to thank the Augustinus Foundation, the Fabrikant Vilhelm Pedersen og hustrus Foundation, the Haynmann Foundation, the Henry og Mary Skovs Foundation, the Knud Højgaard Foundation, the Thomas B. Thriges Foundation and the Viet Jacobsen Foundation for financial support to this research project. This work was partially supported by Wave 1 of The UKRI Strategic Priorities Fund under the EPSRC Grant EP/T001569/1, particularly the "AI for Science" theme within that grant and The Alan Turing Institute. The ML models were trained using computing resources provided by STFC Scientific Computing Department's SCARF cluster and the PEARL cluster.

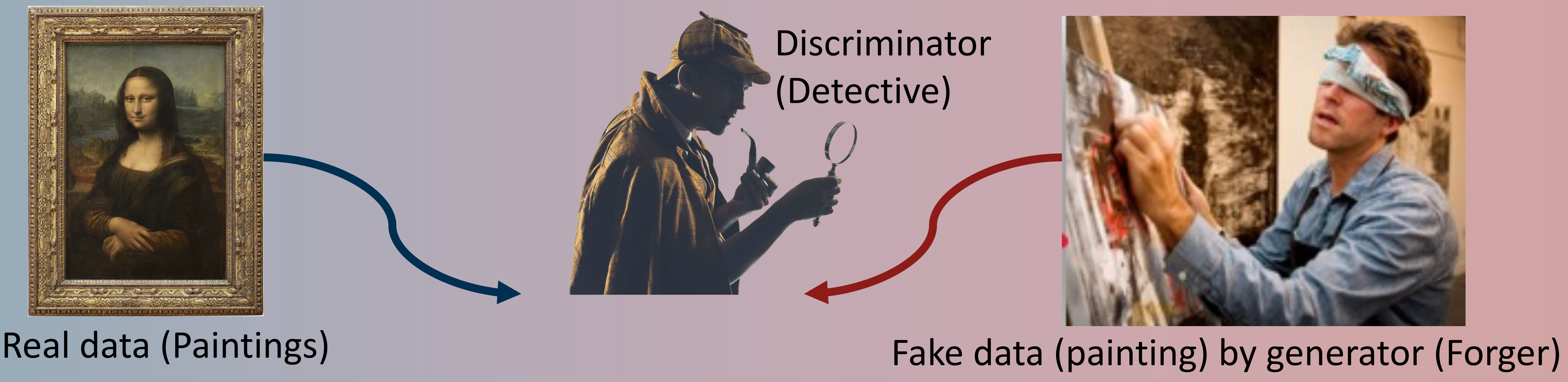
Simulated data including the resolution function



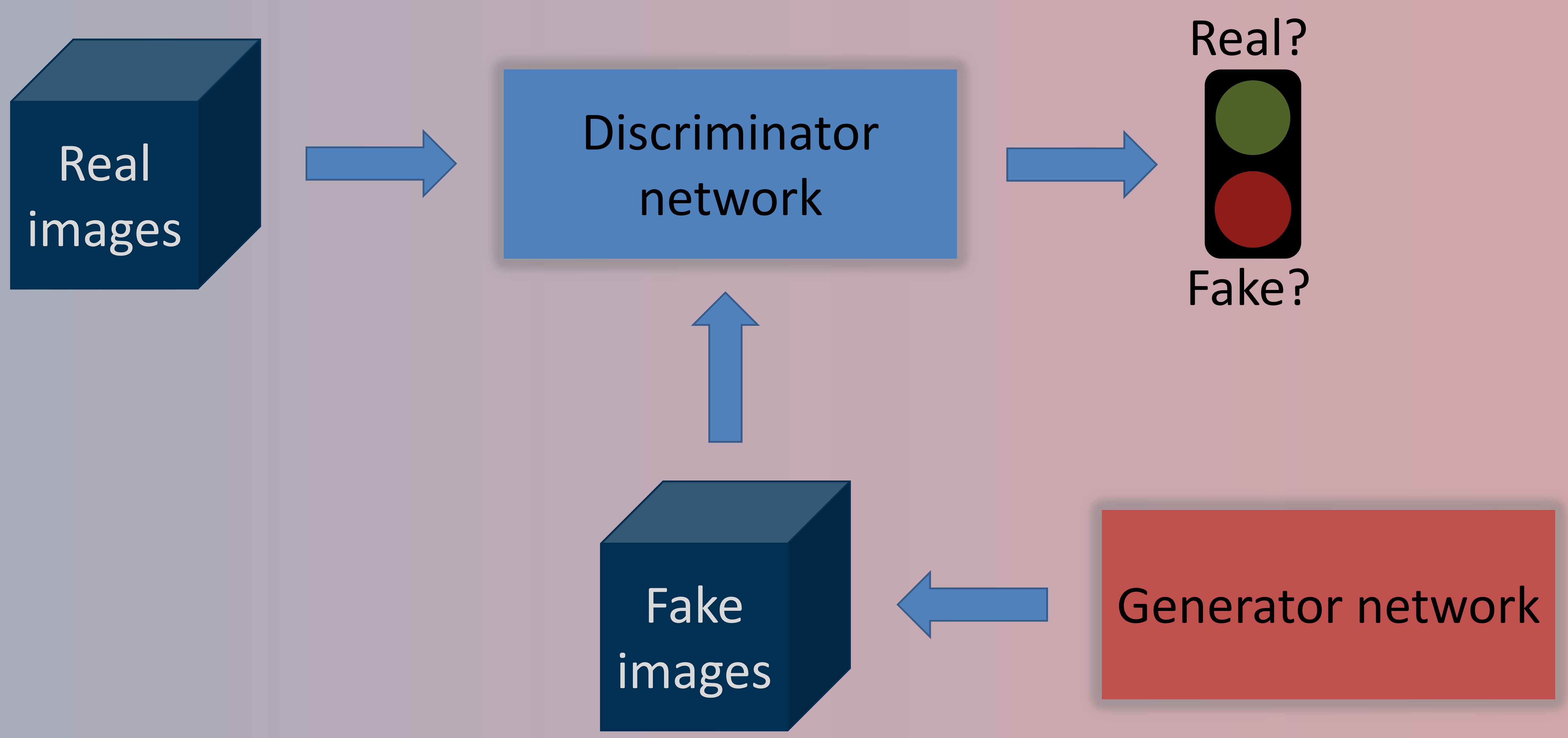
Simulated data excluding the resolution function



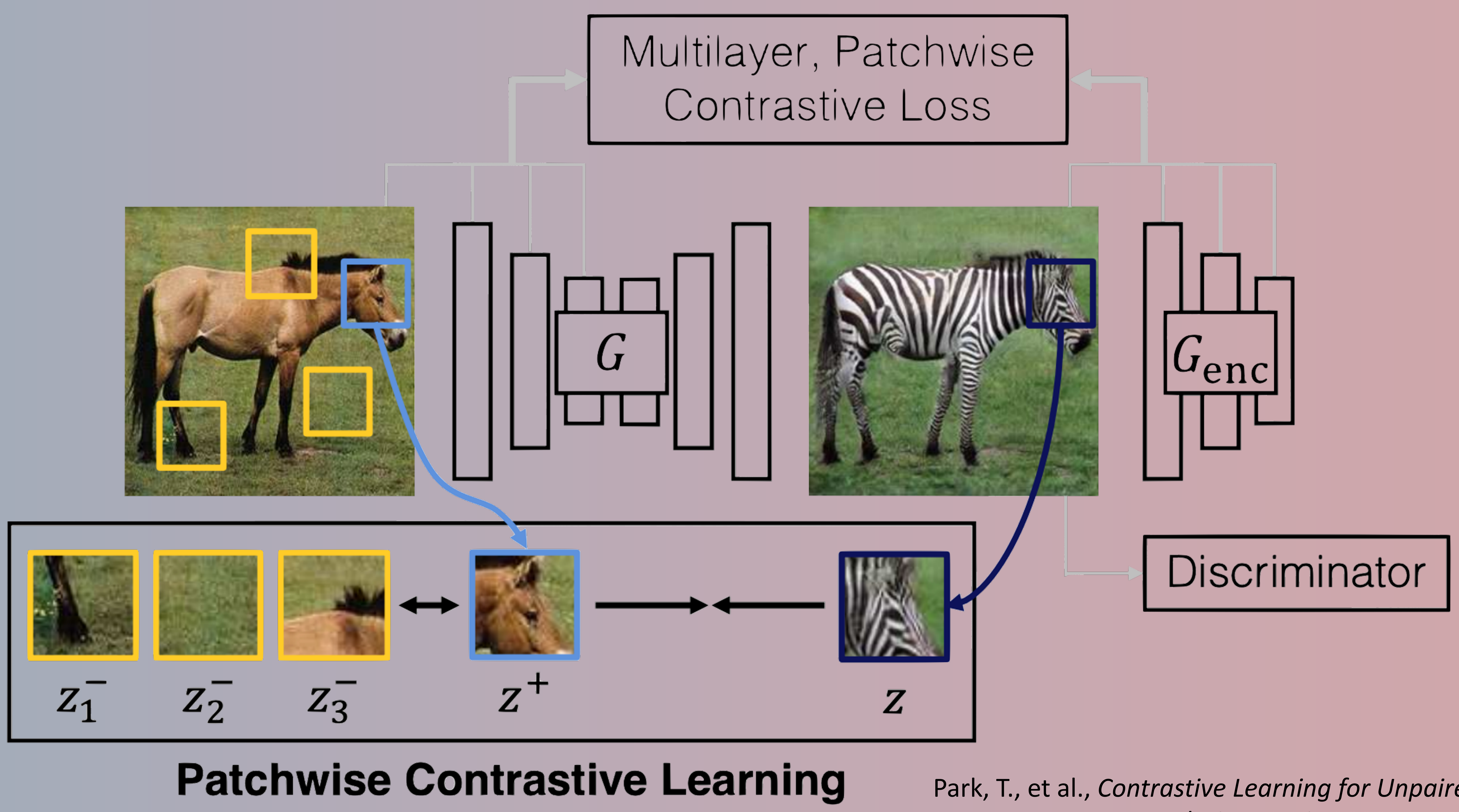
The concept of a generative adversarial network (GAN)



General architecture of a GAN



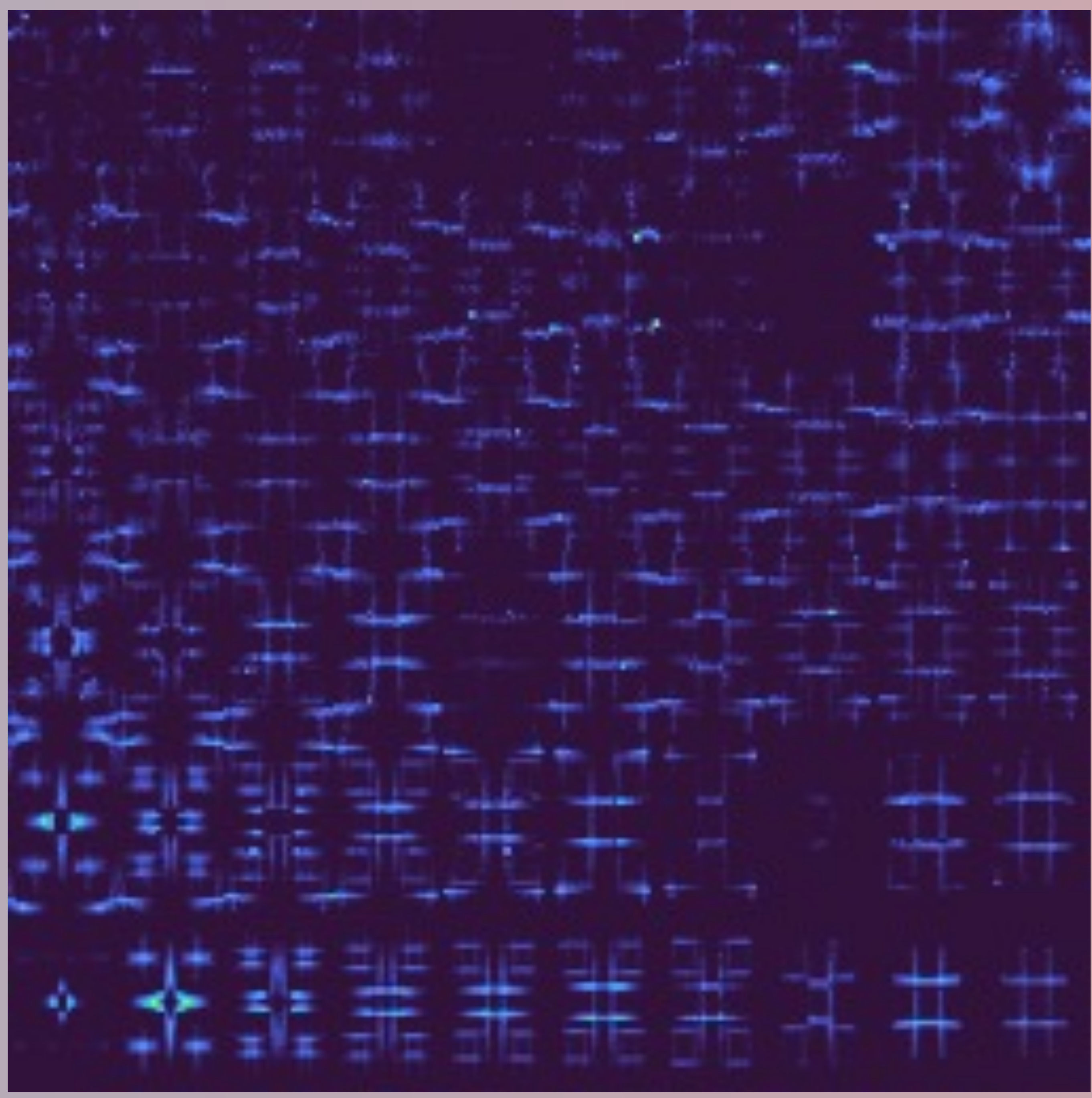
Learning from patches



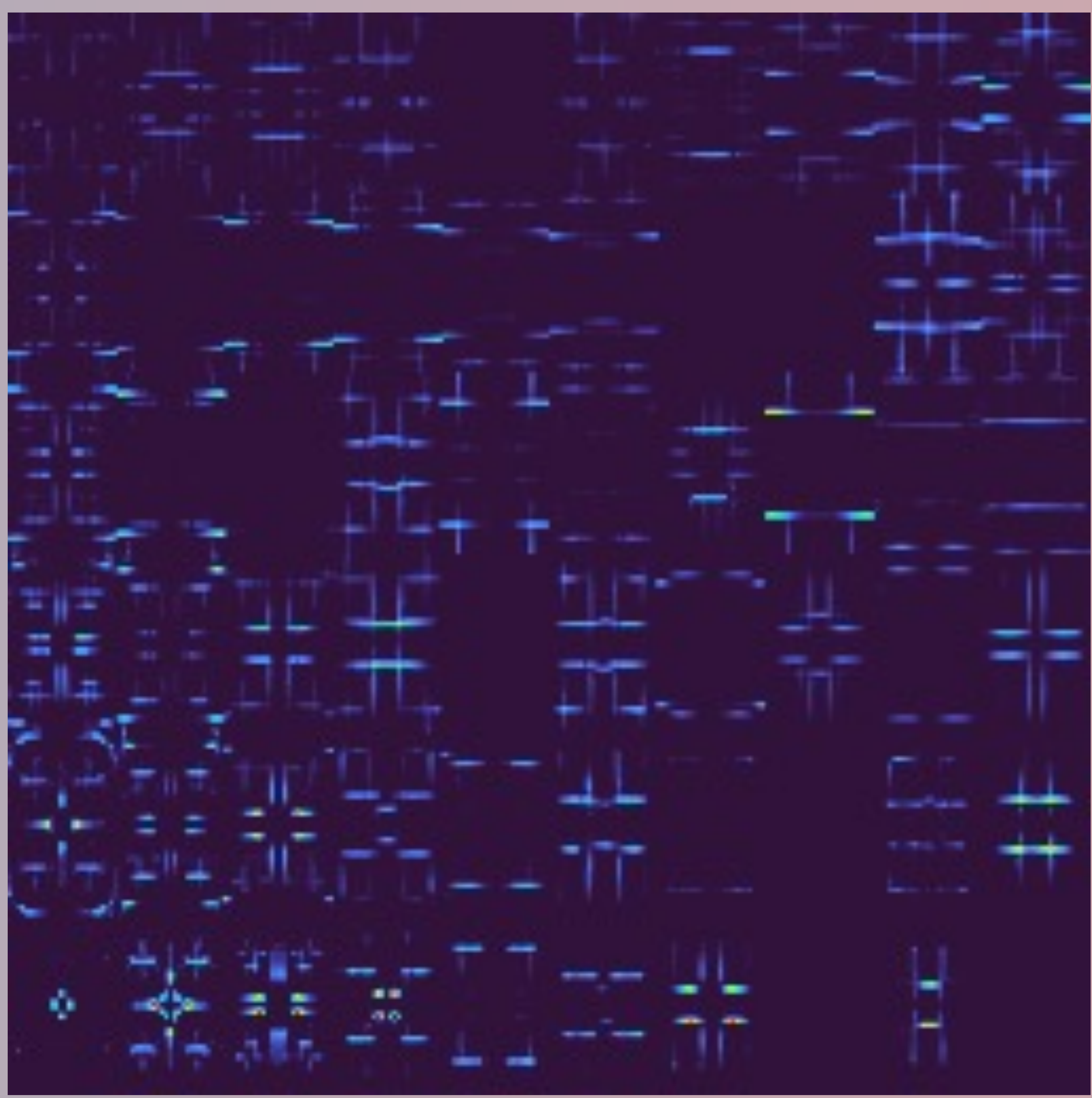
Park, T., et al., *Contrastive Learning for Unpaired Image-to-Image Translation, ArXiv **2020***

Matching experimental and simulated inelastic neutron scattering (INS) data

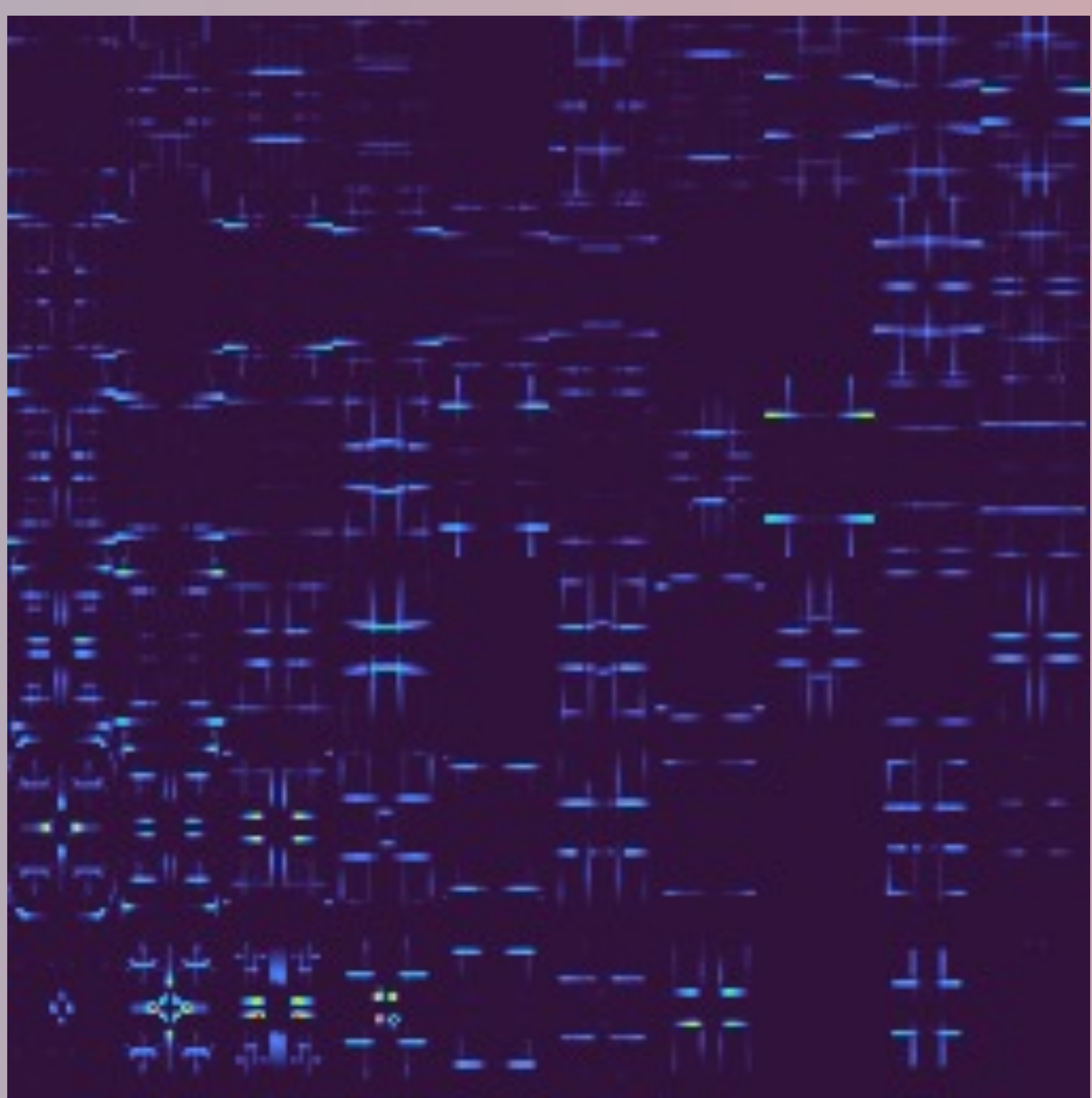
Simulated: data including the resolution function



GAN generated: data excluding the resolution function



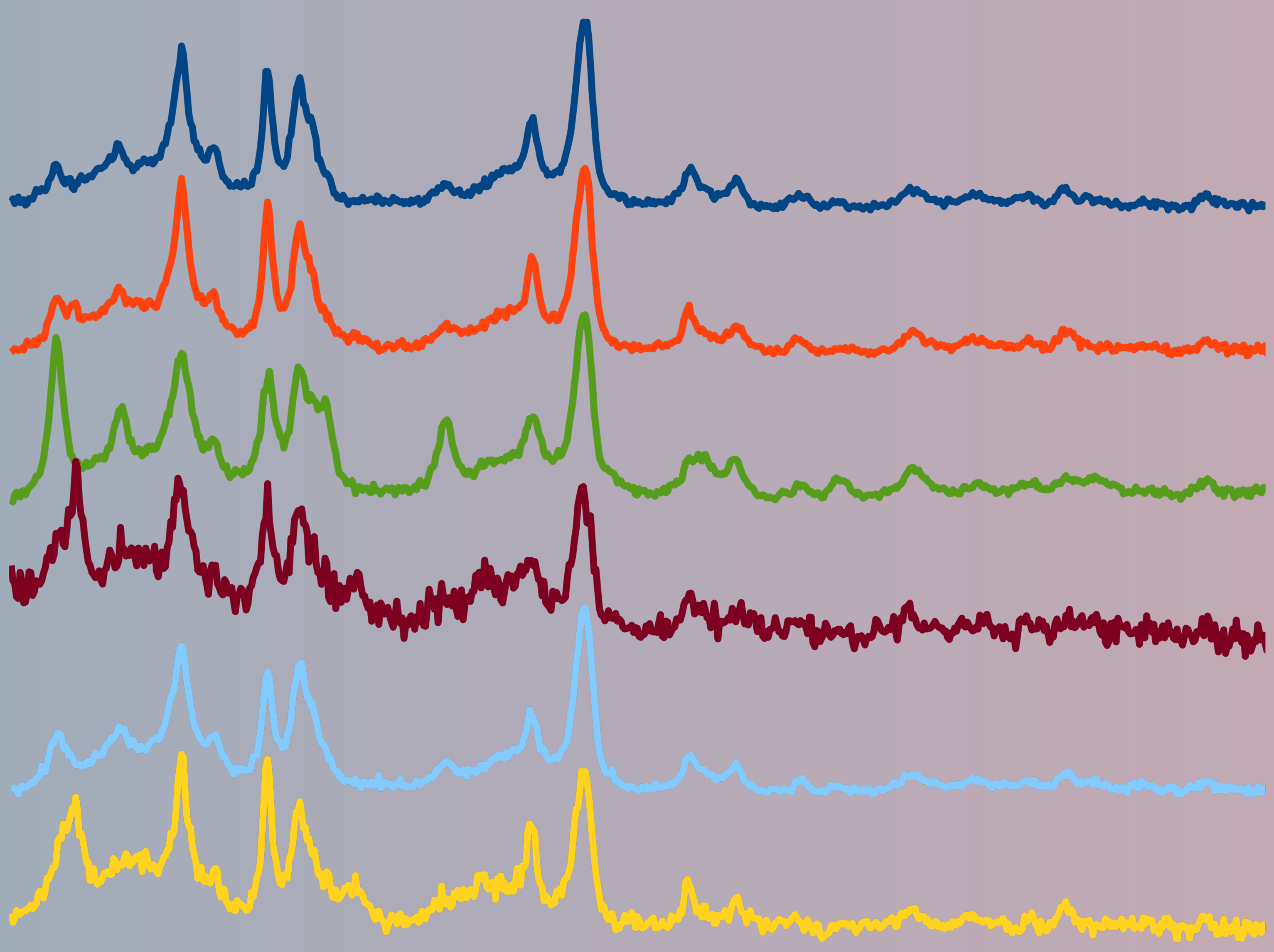
Simulated: data excluding the resolution function



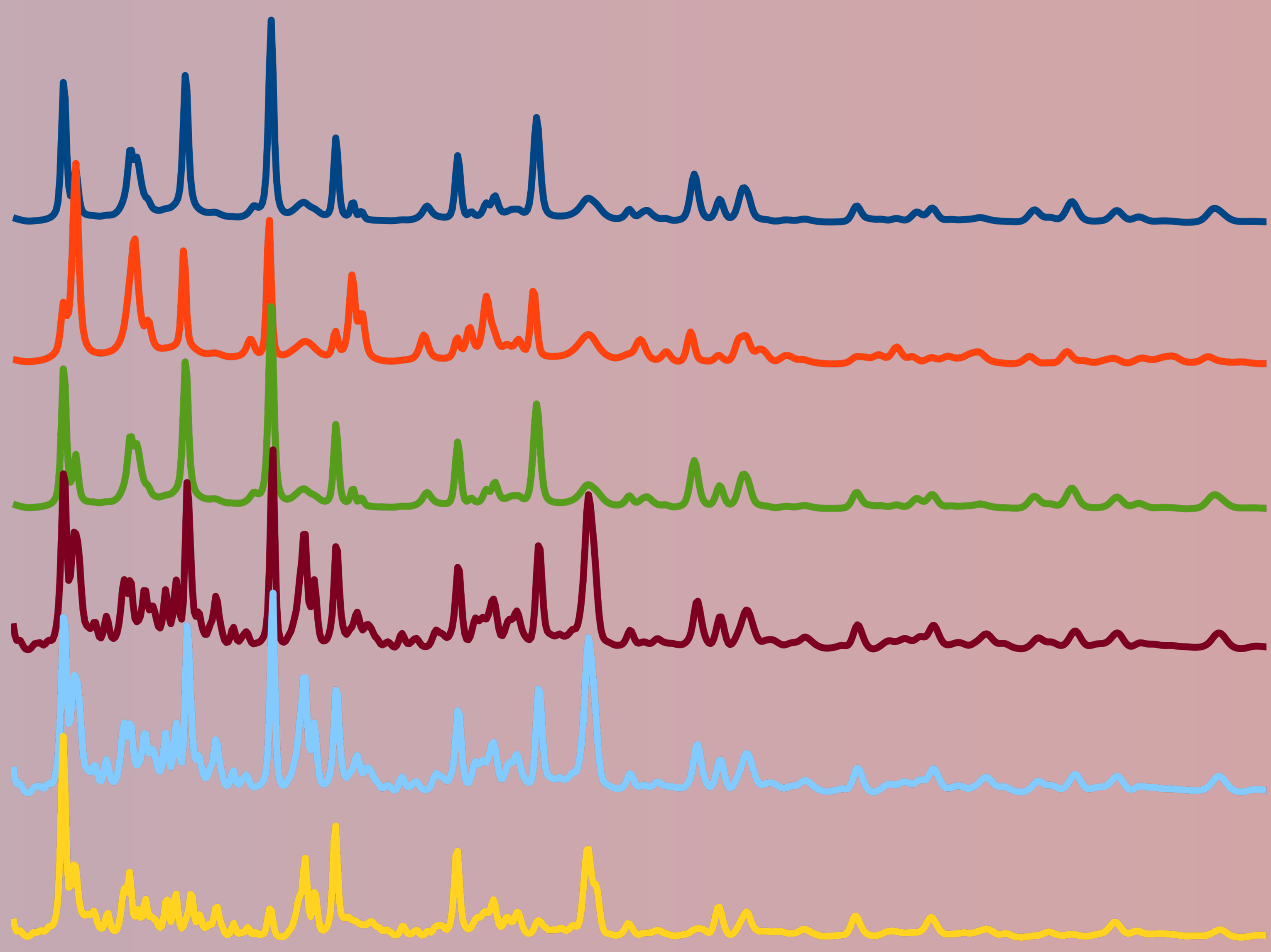
Work in progress: Applying the framework on data from scattering (1D), spectroscopy (2D) and imaging (3D)

X-ray tomography (1D noise removal)

Noisy

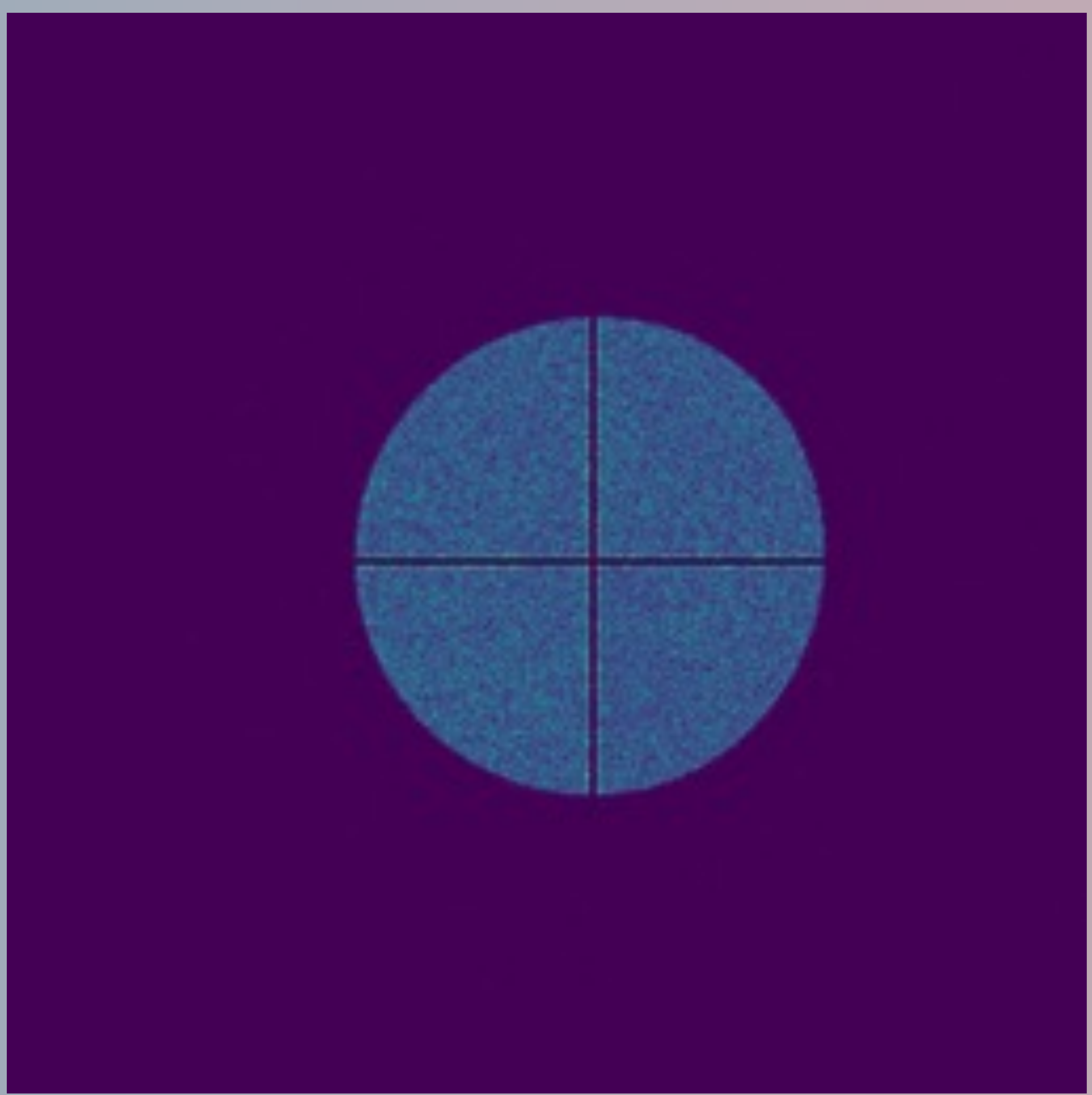


Clean

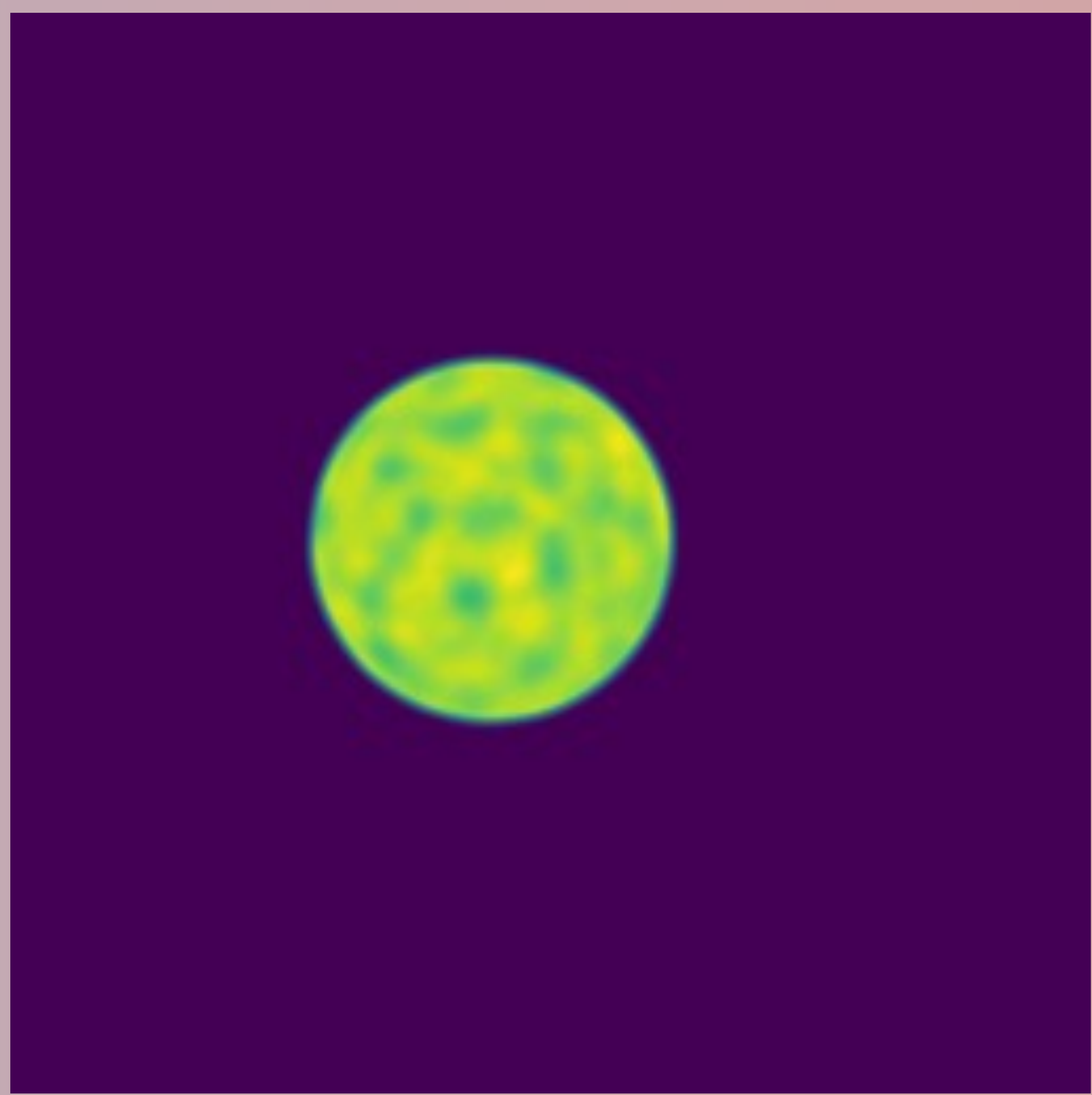


Ptychography (2D simulated ⟷ experimental)

Experimental

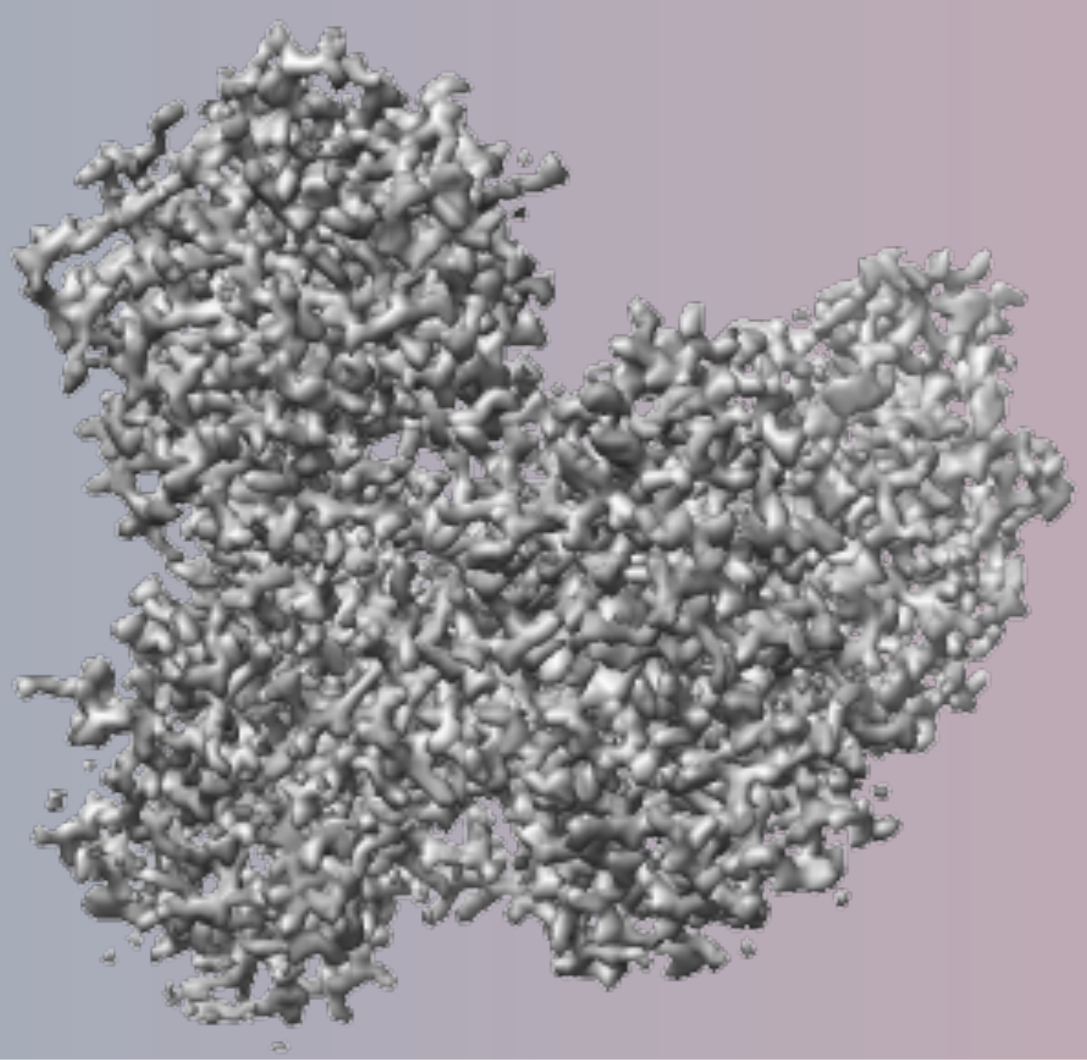


Simulated



Cryo EM (3D super resolution)

3 Å resolution



4 Å resolution