Radio Image Reconstruction for Multi-Messenger Astronomy

PhD Topic

Laboratory: IRFU/DAP/CosmoStat, CEA Paris-Saclay

Supervisors: Samuel Farrens and Jean-Luc Starck

Contact:

- jstarck@cea.fr ☏ : 01 69 08 57 64
- samuel.farrens@cea.fr ☏ : 01 69 08 83 77

Keywords: *machine learning - big data - radio astronomy - cosmology*

Context

Cosmology in the 21st century aims to better our understanding of the Universe by seeking to answer open questions concerning the nature of dark matter and dark energy, and the precise expansion rate of the Universe. In order to tackle these questions, it is essential to take advantage of all the data made available in the current era of multi-messenger astronomy, capitalising on the latest advances in signal processing, machine learning and the handling of big data.

Current and upcoming optical surveys, such as KiDS-450 [1], the Dark Energy Survey Year 1 [2], the Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST) [3] and Euclid [4], are probing wider and deeper patches of the late-time Universe to improve constraints on cosmological parameters by measuring the shapes and distributions of galaxies. These parameters can be compared to the latest analyses of the cosmic microwave background radiation (*i.e.* the early Universe) by surveys like Planck [5]. Recent studies [6] highlight some discrepancy between early and late-time analyses indicating some systematic uncertainty or an incomplete model of cosmology.

Additionally, In recent years gravitational wave interferometers, such as LIGO and Virgo, have pushed astronomical observations beyond the electromagnetic spectrum. This has made it possible to detect the interaction of distant neutron stars and black holes.

Radio wavelengths provide a complementary and independent probe of the late-time Universe. Radio astronomy provides the advantage of probing higher redshifts, having a deterministic point spread function (PSF) and being less sensitive to PSF anisotropies [7]. Cross-correlations between radio and optical surveys can additionally alleviate systematics effects such as intrinsic alignments improving cosmological constraints [8-9]. Upcoming radio surveys, such as the Square Kilometre Array (SKA), are designed to reach an order of magnitude greater sensitivity and survey speed than existing instruments. SKA has the potential to add significant additional constraints on cosmological parameters given the vast sky area it will cover (~75%). This, however, comes at the cost of having to manage extremely large scales of data and complicated image reconstruction. SKA is expected to produce ~1 TB of data every second. With typical observations taking ~6h and a total lifespan of 15 years, **SKA will produce data in the Exabyte (10^{18} bytes) scale** [10], **making it one of the biggest data management problems in modern science.**

The CosmoStat team has been pioneering the use of signal processing and machine learning techniques for solving inverse problems in astronomical image reconstruction. Applying these methods to radio-interferometric data, however, brings a host of new challenges, in part due to the additional complexity of the inverse problems to solve, but also due to the extremely large-scale dimensions of the problem. Conventional deep learning approaches will not be able to scale to the typical size of an SKA field, and developing efficient model-parallelism approaches will be necessary.

Project

The CosmoStat lab of CEA Paris-Saclay proposes a PhD project under the supervision of Dr. Samuel Farrens and Dr. Jean-Luc Starck to work on developing radio image reconstruction tools as part of the European project ARGOS. The objective being to build fast and robust solutions for analysing SKA-scale radio data and providing this to the community in the form of user-friendly and efficient software.

The student will be able to take advantage of cutting-edge image reconstruction methods already developed at CosmoStat [11-14] using both sparsity and machine learning. Additionally, the student can investigate other machine learning approaches, such as deep unrolled networks or automatic differentiation, to extract information directly from visibilities (*i.e.* interferometric observations). Both approaches can be compared to the current state of the art in radio image reconstruction to assess the performance.

To handle the scale of the inverse problems posed by radio data, the student will additionally explore distribution strategies for model parallelism on GPU clusters, and in particular on the IDRIS Jean-Zay computer. The student will be able to take advantage of the existing expertise in CosmoStat on distributed TensorFlow (through Mesh TensorFlow, and Horovod libraries). The model-parallelism methods developed for this project will be highly interdisciplinary and may also find natural applications for high-resolution MRI developed at CosmoStat and NeuroSpin [15]. Additional time-saving methods, such as transfer learning, can be explored to further reduce the processing times required for handling data on the scales expected.

The prospective student will benefit from the abundance of experience in scientific software development, signal processing, cosmology, radio astronomy, and machine learning available in CosmoStat and the ARGOS partner institutions in Greece and Germany. Students aiming for an academic career track will be well situated to take on leading roles in the field, while students aiming for the private sector will learn valuable techniques in data science, software development and handling big data.

Environment

CEA Paris-Saclay is located 20 km south of Paris, France, in the vicinity of various universities and other research centres. The CosmoStat group is a diverse and multi-disciplinary team of researchers working on various topics in cosmology and data science. Our group is committed to diversity and equality, and encourages applications from women and underrepresented minorities. We support a flexible and family-friendly work environment. Benefits for this position include retirement, health care, parental leave, vacation and sick days, subsidised meals, discount for public transport, sport and culture, and French language classes.

References

[1] Hildebrandt et al., KiDS-450: Cosmological parameter constraints from tomographic weak gravitational lensing, 2017

[2] Troxel et al., Dark Energy Survey Year 1 Results: Cosmological Constraints from Cosmic Shear, 2018

[3] Ivezić et al., LSST: From Science Drivers to Reference Design and Anticipated Data Products, Astrophysical Journal 873, 2019

[4] Laureijs et al., Euclid Definition Study Report, 2011

[5] Planck Collaboration, Planck 2018 results. VI. Cosmological parameters, A&A 641, 2020

[6] Raveri and Hu, Concordance and discordance in cosmology, Physical Review 99, 2019

[7] Brown et al., Weak gravitational lensing with the Square Kilometre Array, AASKA14, 2015

[8] Casas et al., Linear and non-linear Modified Gravity forecasts with future surveys, Physics of the Dark Universe 18, 2017

[9] Harrison et al., SKA Weak Lensing I: Cosmological Forecasts and the Power of Radio-Optical Cross-Correlations, MNRAS 463, 2016

[10] Scaife, Big telescope, big data: towards exascale with the Square Kilometre Array, Philos. Trans. R. Soc A, 2020

[11] Sureau et al., Deep Learning for space-variant deconvolution in galaxy surveys, A&A 641, 2020

[12] Gardsen et al., LOFAR Sparse Image Reconstruction, A&A 575, 2015

[13] Farrens et al., Space variant deconvolution of galaxy survey images, A&A 601, 2017

[14] Nammour et al., Galaxy Image Restoration with Shape Constraint, Journal of Fourier Analysis and Applications, 2021

[15] Ramzi et al., Density Compensated Unrolled Networks for Non-Cartesian MRI Reconstruction, 2021