AKPIK 2: Parallel Talks

Time: Monday 16:00–18:00 Location: KS 00.003

AKPIK 2.1 Mon 16:00 KS 00.003

Metadata Analysis for the Square Kilometre Array Max Planck Institute Telescope — •Felix Wersig and Luca Davide Di Bella — TU Dortmund University

The Square Kilometre Array Max Planck Institute (SKAMPI) telescope has been operating since 2019 at the Karoo Radio Astronomy Reserve, collecting metadata from the receiver, the antenna control unit and the MeerKAT weather sensors for many observations under diverse conditions. Searching for unusual patterns in the metadata could be a key method for assessing data quality and identifying maintenance needs.

The SKAMPI telescope is designed as a 15 m prototype dish for MeerKAT+ and the SKA-Mid component of the Square Kilometre Array Observatory. The SKA-Mid component is going to consist of 197 parabolic dishes incorporating the MeerKAT+ radio telescope. Its layout will consist of a ∼1 m central core and antennas radiating out from it up to a diameter of 150 km.

Ensuring that a large fraction of the antennas remain operational at all times is a key challenge given the large number of antennas and extensive distances present in the planned layout. Leveraging the large dataset collected by the SKAMPI telescope, a preliminary tool for quick analysis of the collected metadata from a dish is presented.

AKPIK 2.2 Mon 16:15 KS 00.003

Calibration in radio interferometer simulations using pyvisgen — •Sadiah Azeem[1], Kevin Schmitz[1,2], Anno Knierim[1,2], and Christian Arauner[1,2] — [1]TU Dortmund University, Dortmund, Germany — [2]Lamarr Institute for Machine Learning and Artificial Intelligence, Dortmund, Germany

Modern physics experiments greatly benefit from the ability to reproduce entire measurement processes through realistic simulations. In radio interferometry, such simulations enable testing of calibration and data analysis pipelines and allow estimation of the discovery potential for faint astrophysical sources. `pyvisgen` is a Python framework designed to generate radio interferometric visibilities based on the radio interferometer measurement equation (RIME) and is a part of the `radionets-project`, which uses deep learning based methods for radio interferometric imaging. Realistic simulations must include instrumental and atmospheric effects that corrupt the measured visibilities such as antenna gains and tropospheric delays. They also provide the ground truth needed to test and evaluate calibration methods. This talk will outline how to generate RIME-consistent corrupted visibilities and apply them to the development of calibration methods.

AKPIK 2.3 Mon 16:30 KS 00.003

Simulating Protoplanetary Disks as Training Data for Neural Network based Reconstruction of Radio Interferometer Measurements — •Tom Gross[1,2], Kevin Schmitz[1,2], Christian Arauner[1,2], and Anno Knierim[1,2] — [1]TU Dortmund University — [2]Lamarr Institute for Machine Learning and Artificial Intelligence

Understanding the evolution of planetary systems is an important research field in modern astronomy. Planets form through accretion processes of dust and gas in protoplanetary disks around young stars.

The thermal radiation of the disks is measurable in the radio spectrum using radio interferometers. These enable the observation of small structures but only measure samples of an incomplete Fourier space. This intrinsic limitation results in noisy images, unsuitable for physical analyses.

The `radionets-project` uses a modern deep-learning framework to reconstruct the missing information. To enhance the scientific interpretation of the reconstructions, the training data needs to represent realistic source distributions. The software `FARGO3D` simulates the distribution of dust and gas in a protoplanetary disk using a hydrodynamical model. From these distributions, the thermal radiation can be simulated using Monte Carlo methods. This presentation shows the results of the simulations and how they are integrated into the data pipeline for the deep-learning based reconstruction.

AKPIK 2.4 Mon 16:45 KS 00.003

ParaO: In situ reproducible workflows through user-centric design — Martin Erdmann, •Benjamin Fischer, and Felix Zinn — III. Physikalisches Institut A, RWTH Aachen University

In High Energy Physics (HEP) and other fields, data analyses require many different steps to produce a final result. As such, it is sensible to leverage workflow tools to codify these complex structures of tasks and their dependencies. Of course, such intricate analyses are not created ad hoc; rather, their development is a lengthy and iterative process.

Here, ParaO's design empowers users to rapidly grow and readily adapt their work while avoiding verbose boilerplate. For example, integrating a new parameter or dependency in the middle of a ten-step workflow needs just a single line of code. There, ParaO automatically propagates the resulting knock-on effects, such as changed implicit parameterization of later tasks—unlike Luigi, which requires a more manual treatment.

Thus, ParaO enables in situ usage of workflows and their benefits, such as bookkeeping and reproducibility, even in early and chaotic stages of development. We demonstrate ParaO's principles and features in the context of a comprehensive top-Higgs analysis.

AKPIK 2.5 Mon 17:00 KS 00.003

Cleaning MOJAVE observations with neural networks — •Christian Arauner[1,2], Kevin Schmitz[1,2], Anno Knierim[1,2], and Tom Gross[1,2] — [1]TU Dortmund University — [2]Lamarr Institute for Machine Learning and Artificial Intelligence

Radio interferometry enables observations of astronomical objects with high angular resolution. However, the inherent incompleteness of the (u, v)-plane sampling leads to significant noise and artifacts in the reconstructed images. Existing state-of-the-art cleaning algorithms effectively mitigate these effects but are computationally demanding, not easily reproducible and not readily scalable to the data volumes anticipated from next-generation radio telescopes. As a promising alternative, neural network-based approaches offer the potential to automate and accelerate the image reconstruction process.

The Python package `radionets` implements a deep-learning framework for the reconstruction of calibrated data from radio observations. In this approach, ResNets are used to reconstruct the missing values directly in the (u, v)-plane. The validity of the neural network-based approach has been shown with simulated data, now it is possible to apply it to real measurments from the MOJAVE program. The observations of the MOJAVE archive are particular suitable for the development and testing of these neural networks as it comprises a large data set of high-quality data, measured over a long period of time.

In this talk, I will present first observations of the MOJAVE program, reconstructed with `radionets`.

AKPIK 2.6 Mon 17:15 KS 00.003

Methodological Preparation for Measurements with the DSA-2000: Simulation, Analysis, and Reconstruction Using Neural Networks — •Jonas Grossmann[1], Kevin Schmitz[1,2], Anno Knierim[1,2], Christian Arauner[1,2], and Tom Gross[1,2] — [1]TU Dortmund University, Dortmund, Germany — [2]Lamarr Institute for Machine Learning and Artificial Intelligence, Dortmund, Germany

A new generation of radio telescopes is set to begin operations later this decade. Among these is the Deep Synoptic Array 2000 (DSA-2000), a radio survey interferometer. It comprises 2000 small antennas, each 5 m in diameter, distributed over an area of 160 square kilometres in the Nevada Desert, USA. This configuration is optimised for rapid, high-resolution surveys of the entire northern sky. The DSA-2000 will achieve at least an order of magnitude improvement in angular resolution, enabling the discovery of billions of previously unknown radio sources. The resulting data volumes are unprecedented, surpassing the capabilities of conventional analysis methods and necessitating new approaches for translating interferometric measurements into cleaned images. The `radionets-project` group at TU Dortmund University aims to develop neural networks that, could automatically reconstruct dirty radio images. Training these networks requires large, realistic datasets. Given that all 2000 antennas cannot operate flawlessly at all times, the goal is to train the networks to deliver stable and reliable reconstructions even when subsets of antennas are inactive. This talk outlines the framework in `radionets-project` required to create the datasets necessary for training these neural networks.

AKPIK 2.7 Mon 17:30 KS 00.003

Resource Awareness in Deep Learning-based Imaging in Radio Interferometry — •Anno Knierim[1,2], Kevin

Schmitz[1,2], Christian Arauner[1,2], Raphael Fischer[1,2], Dominik Elsässer[1,2], and Wolfgang Rhode[1,2] — [1]TU Dortmund University, Dortmund, Germany — [2]Lamarr Institute for Machine Learning and Artificial Intelligence, Dortmund, Germany

Recent approaches in radio interferometry aim to improve image cleaning of measurements using machine learning techniques. Reconstructing sources using these novel techniques has the advantage of being agnostic to the initial parameters used in traditional cleaning algorithms.

The `radionets-project` is a multi-software environment developed at TU Dortmund University. The main deep-learning framework, `radionets`, reconstructs calibrated data from radio observations using convolutional neural networks (CNNs). The framework aims to achieve a high dynamic range and produce high-resolution source images.

Due to their high complexity, applying deep learning models sustainably requires balancing predictive performance with resource consumption. The environmental impacts extend well beyond training, as model complexity also affects the efficiency during data simulation and inference. This talk presents ongoing efforts to track and estimate the environmental impact of our models using `CodeCarbon`, `PyTorch Lightning`, and `MLFlow`.

AKPIK 2.8 Mon 17:45 KS 00.003

FAIR-Driven Retrospective Metadata Enrichment for KCDC Digital Resources — •Victoria Tokareva — Karlsruhe Institute of Technology, Institute for Astroparticle Physics, 76021 Karlsruhe, Germany

The KASCADE Cosmic-Ray Data Centre (KCDC) is an established open data archive in astroparticle physics with more than a decade of experience in long-term data preservation. It aims to apply modern technologies to maximise the accessibility, usability, and scientific value of research data for a broad audience, from professional researchers to students, citizen scientists, and science enthusiasts, while advancing the use of FAIR (Findable, Accessible, Interoperable, Reusable) practices. Operating within a data-intensive high-energy physics environment, we address community-wide challenges such as the large data volumes produced in astroparticle physics and the complexity of formats like ROOT and HDF5, which require domain-specific software and specialised approaches for machine-actionable metadata extraction. To support interdisciplinary discovery and integration into infrastructures such as PUNCH4NFDI, we extend established metadata with additional discovery-oriented elements and references, including domain knowledge historically dispersed in scientific publications. In this contribution, we present an enhanced metadata model for KCDC digital resources featuring expanded use of persistent identifiers and increased semantic connectivity. We also demonstrate how natural language processing and large language models support retrospective metadata enrichment for archival KASCADE datasets.