Posters

01	Markus Blank-Burian	**WWU Cloud - Open Source based Clou Services at the University of Münster** Poster not available due to illness
02	Harry Enke	**FAIR in astronomy context**
03	Anastasia Galkin	**Daiquiri - Python based framework for the publication of scientific databases**
04	Zohreh Ghaffari + Catalina Sobrino Figaredo	**A new multi-band optical image pipeline for the Magellan 6.5 m telescope**
05	Niraj Kandpal	**Molecular Outflow Detection in G327: A 3D Approach**
06	Lars Künkel	**Searching Pulsars Using Neural Networks**
07	Jörn Künsemöller	**Metadata and User-Provided Data in the LOFAR Long Term Archive**
08	Man I Lam	**PyParadise: A simultaneous pipeline of stellar and gas kinematics**
09	Abhishek Malik	**Exoplanet detection using Machine Learning**
10	Jan Mayer	**Entering NeuLAND: Analysis workflow preservation for a fair FAIR**
11	Martin Müller	**Data Infrastructure at the University of Cologne's Institute for Nuclear Physics**
12	Annika Oetjens	**Past planetary engulfment as a possible explanation for observed high stellar rotation on metal poor main sequence stars**
13	Oleksandra Razim	**Towards reliable photometric redshifts with machine learning methods**

Posters

14	Benedikt Riedel	**Cloud burst for Multi-Messenger Astrophysics**
15	Bernd Schleicher	**Machine Learning in Cherenkov Astronomy**
16	Kilian Schwarz	**The PAHN-PaN consortium at the NFDI**
17	Kseniia Sysoliatina	**Astronomy meets big data: Improving the Milky Way model with the billion-star surveyor Gaia**
18	Alexander Trettin	**From 2D to 3D to Graph Networks: Representing Detector Geometries in Neural Networks**
19	Meetu Verma	**Classification of high-resolution solar Hα spectra using t-SNE**
20	Christoph Welling	**NuRadioMC and NuRadioReco – A Software Framework for the Radio Detector Community**
21	Tobias Winchen	**Prototypes for the Next Generation of Computing Backends in Radio-Astronomy**
22	Martin Wolf	**BoostNumpy: Big Data Processing in C++ with Python convenience**

Out of competition: additional poster

WWU Cloud - Open Source based Cloud Services at the University of Münster

- IaaS self-service platform based on OpenStack/Ceph/Kubernetes
- Free of charge for faculty members
- Virtual machines (mainly for services) and data storage
- JupyterHub running in cloud for interactive data analysis, ideal for shared resource usage
- Hardware provided by DH-NRW via RDI-NRW project (5 cloud locations)
    - 1600 Cores, 25TB RAM, 10PB HDD, 2PB NVMe, 2x NVidia Tesla M10
    - Scientific consortia may use cloud hardware at multiple locations

- Poster: Overview, Architecture, JupyterHub, Outlook

FAIR Data

Findable

F1. (Meta)data are assigned a globally unique and persistent identifier
F2. Data are described with rich metadata
F3. Metadata clearly and explicitly include the identifier of the data they describe
F4. (Meta)data are registered or indexed in a searchable resource available

Interoperable

I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (Meta)data use vocabularies that follow FAIR principles
I3. (Meta)data include qualified references to other (meta)data

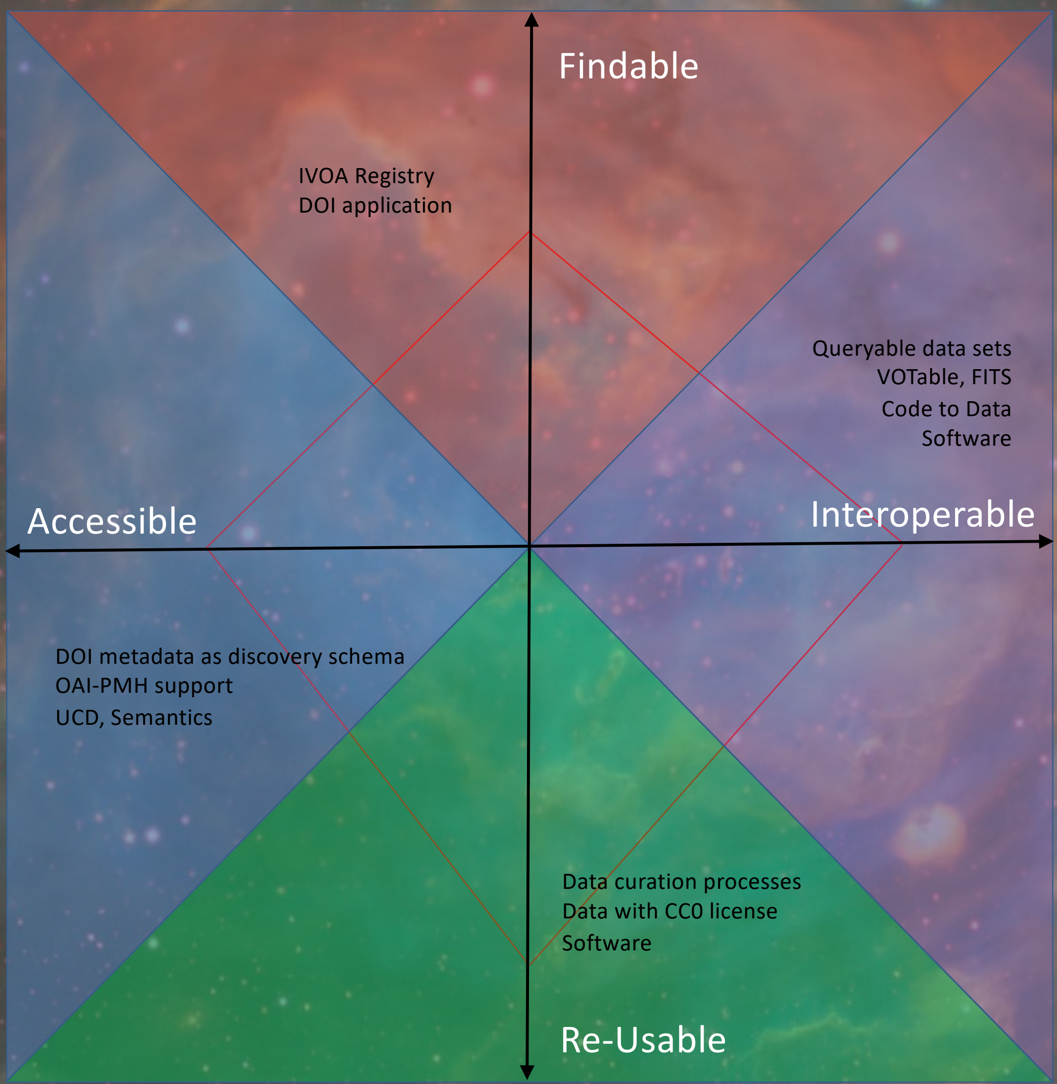
Accessible

A1. (Meta)data are retrievable by their identifier using a standardised communications
A1.1 The protocol is open, free, and universally implementable
A1.2 The protocol allows for an authentication and authorization procedure, where necessary
A2. Metadata are accessible, even when the data are no longer available

Re-Usable

R1. (Meta)data are richly described with a pluralityof accurate and relevant attributes
R1.1. (Meta)data are released with a clear and accessible data usage license
R1.2. (Meta)data are associated with detailed provenance
R1.3. (Meta)data meet domain-relevant community standard

https://www.astro-nfdi.de/how-fair-is-your-data



Use the QR-Code

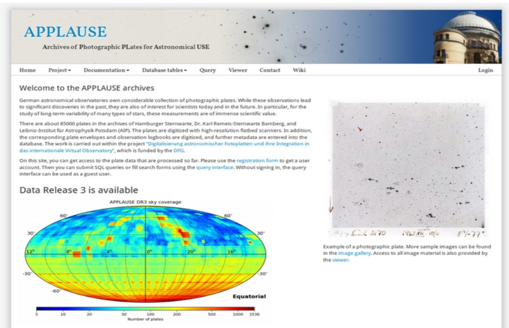
- access the website
- select your favorite wavelength
- provide your estimate of the FAIRness of the data in the field

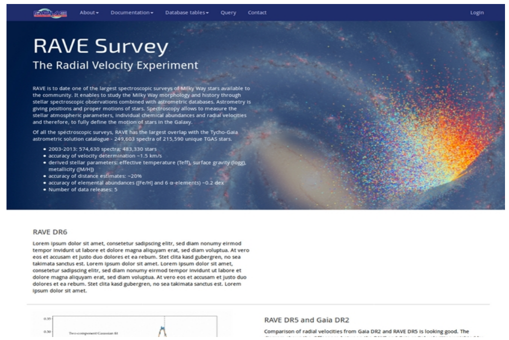
Example:

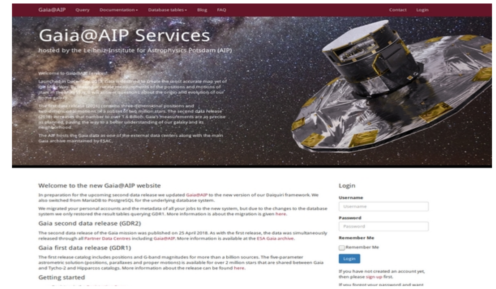
- data in the field of visible wavelength
- some pointers to arguments

Daiquiri - Python based framework for the publication of scientific databases

A. Galkin, J. Klar, K. Riebe, G. Matijevic, H. Enke – Leibniz Institute for Astrophysics Potsdam (AIP)



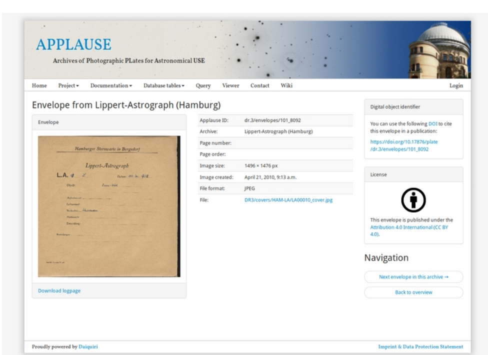


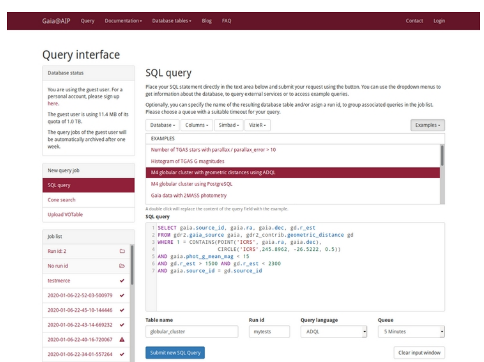


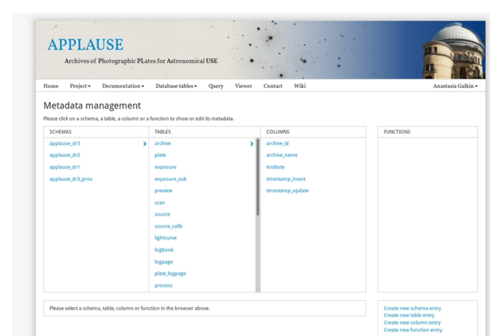
Scientific archives hosted @AIP: APPLAUSE https://www.plate-archive.org, RAVE https://www.rave-survey.org, Gaia@AIP https://gaia.aip.de, MuseWIDE https://musewide.aip.de, CosmoSim https://www.cosmosim.org

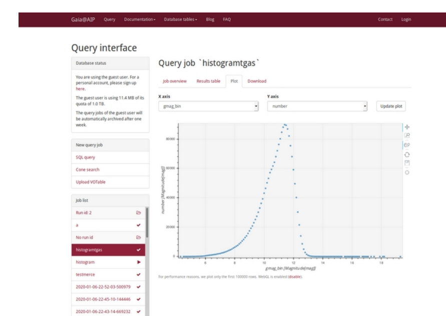
Daiquiri features: Queryparser package to translate ADQL to backend SQL, an interactive query interface, asynchronous database queries, visualization tools, an IVOA compliant cone search API and a registry of registries implementation, metadata, access rights and user management, UCDs support, a cut-out API for data cubes, a contact form, DOI integration, an OAI PMI interface for harvesters, a meeting module and an integrated Wordpress for the documentation.

Technologies: Python 3, Django Framework, webserver apache, NGINX, gunicorn, Astropy, RabbitMQ, Celery, Redis, PostgresQL, MariaDB, Wordpress, Docker.









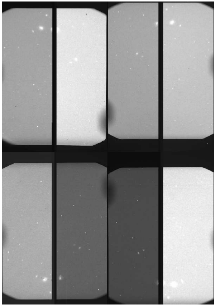
Open Source and available on GitHub https://github.com/django-daiquiri

Also as docker compose setup https://github.com/django-daiquiri/daiquiri-docker-compose

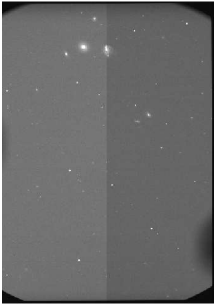
A new multi-band optical image pipeline for the Magellan 6.5 m telescope.

Zohreh Ghaffari[1], *Catalina Sobrino Figaredo*[1], *Martin Haas*[1], *Rolf Chini*[1,2], *Steven Willner*[3],

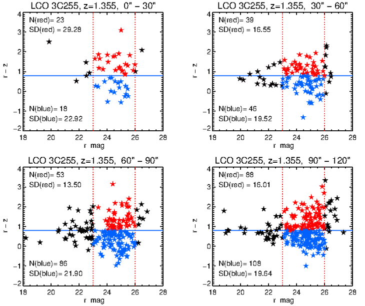
[1] Ruhr-University Bochum, Germany,[2] Universidad Catolica del Norte, Antofagasta, Chile,[3] Harvard-Smithsonian Center for Astrophysics, Cambridge, USA.



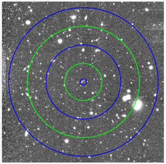
1. PISCO 5'x8' raw image in filters griz



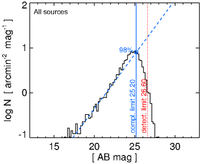
2. PISCO r image after bias, dark, flat, geometric distortion correction



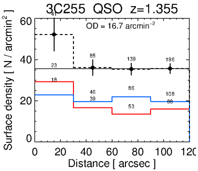
5. Color-magnitude diagrams r-z / r for 4 radial bins



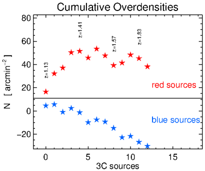
3. Final r image 4'x4' combined from 21 dithered frames of 120 s



4. Detection limits $r \sim 26$ **mag**



6. Radial Surface Density



7. Clustering evolution of 3C sources ($1 < z < 2$**)**

Outflow Detection in G327 : A 3D Approach

N. Kandpal[1], A. Sanchez Monge[1], Peter Schilke[1]

1. First Institute Physics, Cologne, Germany

2 Dimensional Analysis

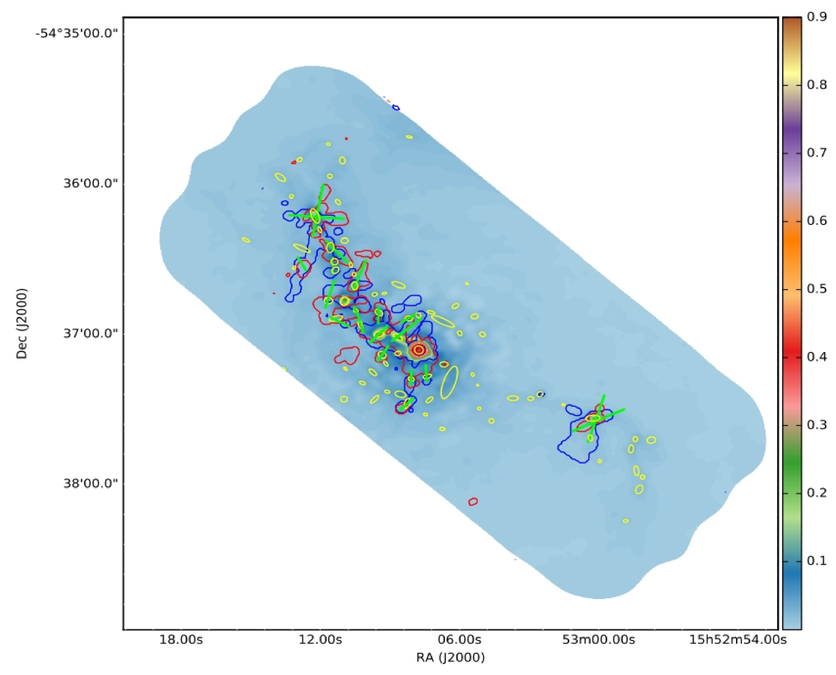


Figure 1: Red and blue shifted contour in SiO overlayed over ALMA+ATLASGAL continuum. Green lines show direction of the outflow and in yellow are the continuum sources using Sextractor.

- We used 2D contour analysis to detect the outflows in G327.
- In Fig1 we can see blue and red shifted outflow distribution.
- In all around 20 outflows were found which we can see in green in Fig 2 showing their direction.
- Method was not efficient specially in most crowded regions to assign and detect outflows.
- We need alternative methods to detect outflows.

3 Dimensional Analysis

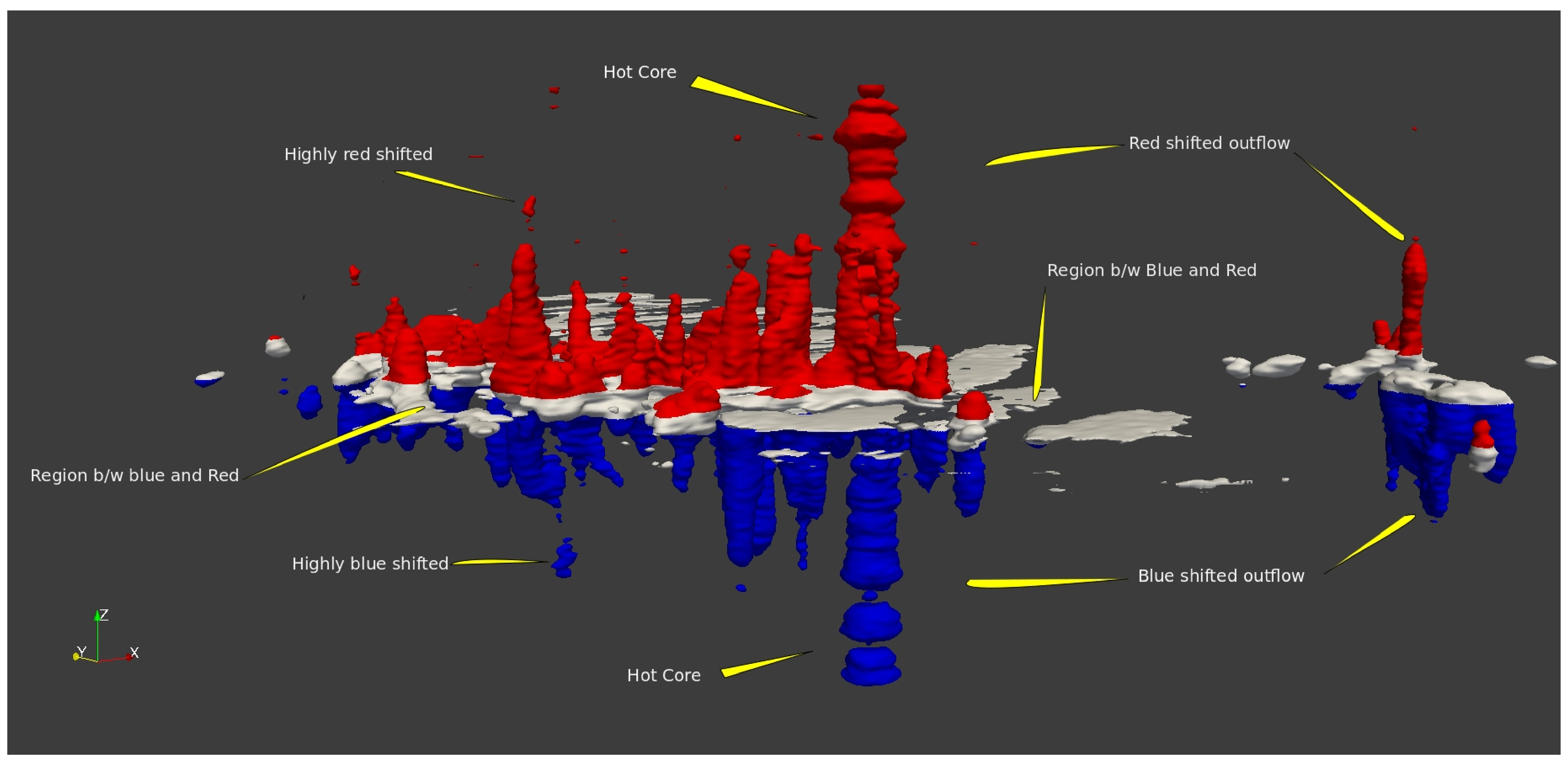


Figure 2: 3 Dimensional representation of SiO outflow. Red and blue colors represent red and blue shifted outflow. Pancake like structure in the middle is hot core and the region between red and blue shifted outflow is around systemic velocity.

- We needed to disentangle the outflows in order to get a complete picture of all the outflows.
- A 3D conversion of ALMA data was one of the solutions we found.
- We could detect around 43 outflows which is twice the number which we got using 2D analysis.
- We divided the G327 full region into subregions for better analysis.
- The outflow direction and relation can be concluded using skeleton structure of the subregions.

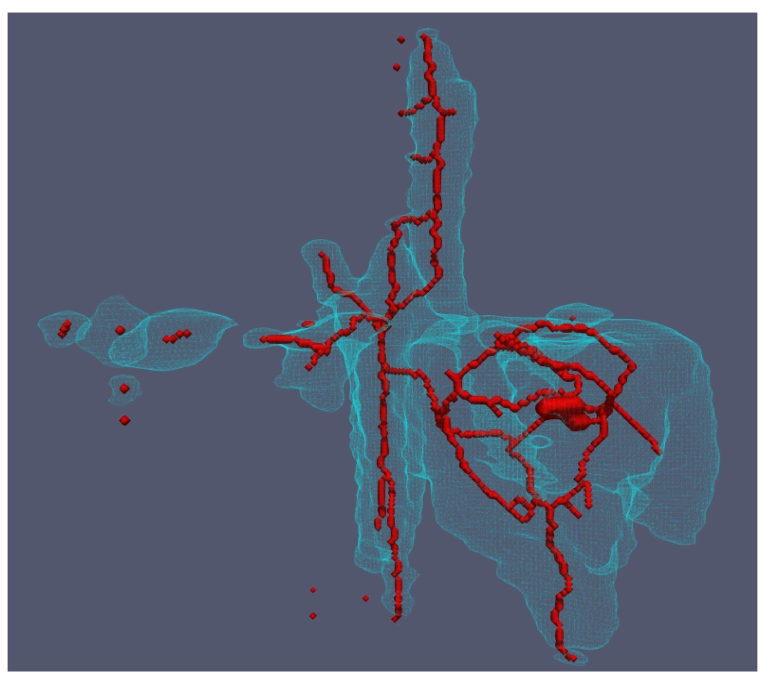


Figure 3: Subregion of G327 on the right showing skeletal structure of outflow.

Outflow Detection in 3D : Method

- We divided the G327 into subregions. We can see on such a subregion in Fig4.
- We found that individual outflows in 3D had complex geometry with branched outflows and multiple outflows.
- In order to calculate the outflow parameters we developed method using rotating ellipsoids.
- The ellipsoid was rotated using 2 angles in the direction of outflow and ellipsoid parameters were adjusted to fit the outflow lobe.
- The flux was calculated by using ellipsoid as mask which gave us outflow parameters

Outflow Detection in 3D : Method

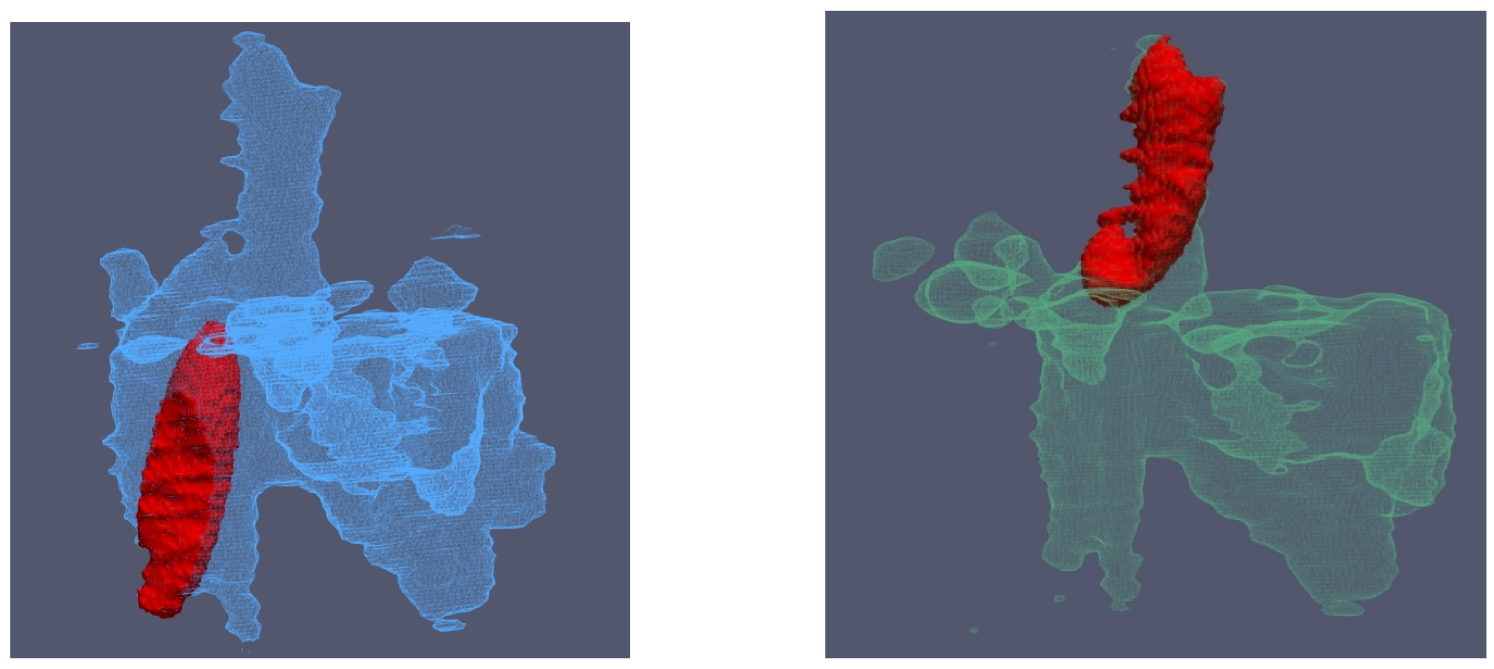


Figure 4: Section of G327 with ellipsoid in red being rotated to the position of the outflow in both the figures.

Results : Using 3D method

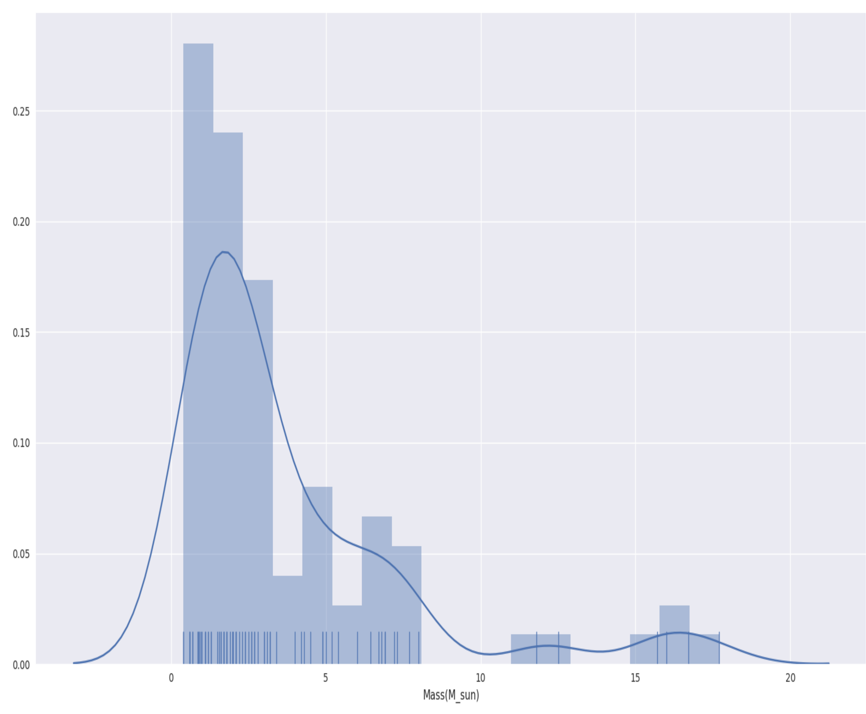


Figure 5: Kernel Density Estimation(KDE) for all the individual outflow lobes. Rugs in red are individual outflow lobes.

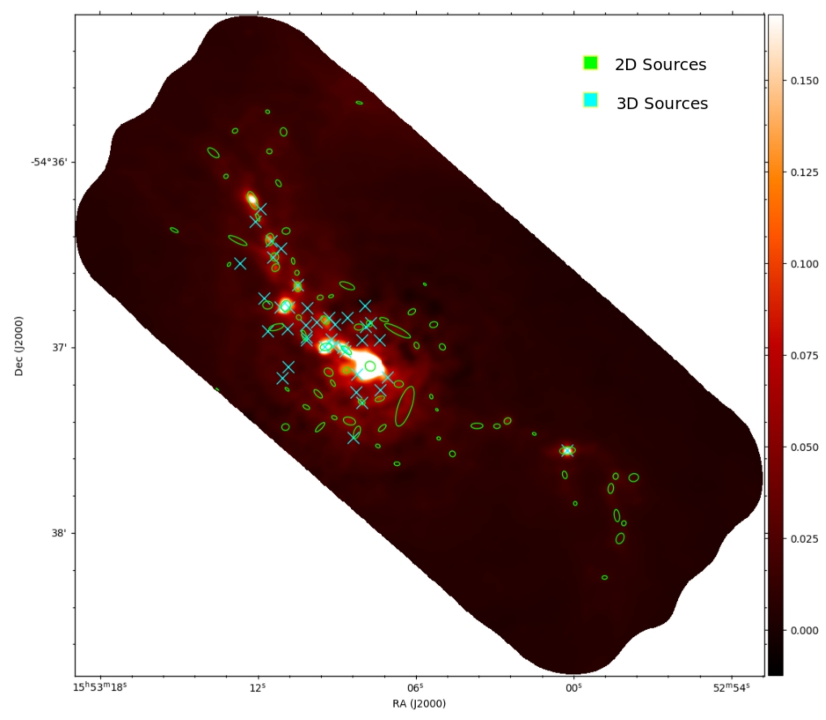


Figure 6: G327 ALMA+ATLASGAL continuum image with 2D sources(SExtractor) and possible 3D sources.

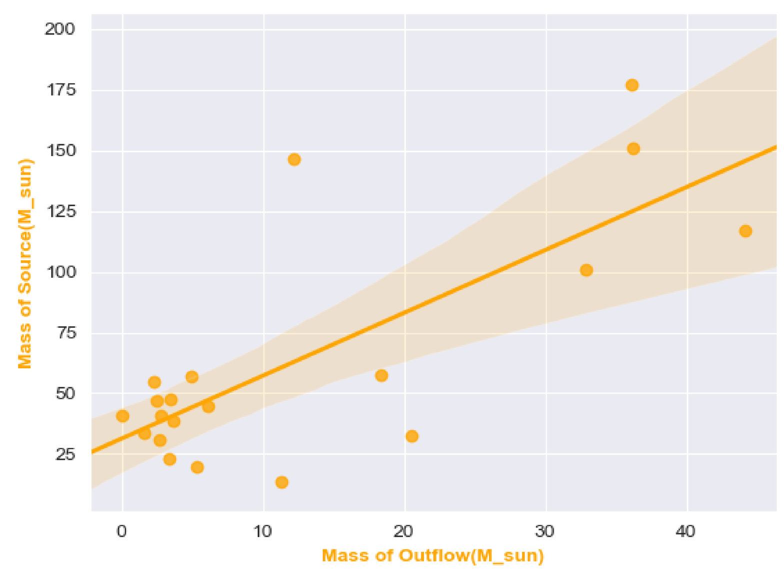


Figure 7: Plot of source mass from Sextractor vs Outflow mass form 3D method with linear fit.

- Kernel Density Estimation in Fig 5 shows masses of individual outflows in the range ‘0.4-17.7’ solar mass.
- It was found in Fig6 that 2D Sextractor sources do not always overlap with possible 3D sources implies there could be more sources driving the outflow.
- A plot of source mass to the outflow mass(Fig 7) shows outflow mass not very high which could imply that our sources are in early stages of stellar evolution.

Position Angle : Monte Carlo Simulation

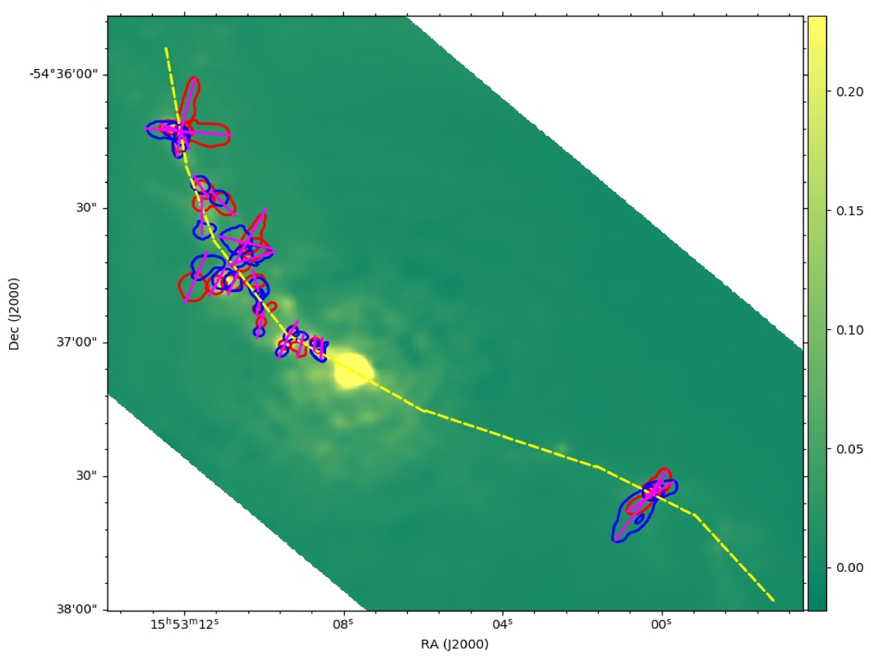


Figure 8: Outflow direction represented by pink line on top of blue and red shifted lobes close to filament in yellow.

- The 3D lobes selected from rotated ellipsoid are projected back in 2D.
- Using Monte Carlo simulation form (Kong et. el. 2019) we found that outflows in G327 are preferentially oriented orthogonal to filament.
- This can suggest that angular momentum of protostellar accretion disk is correlated with host filament.

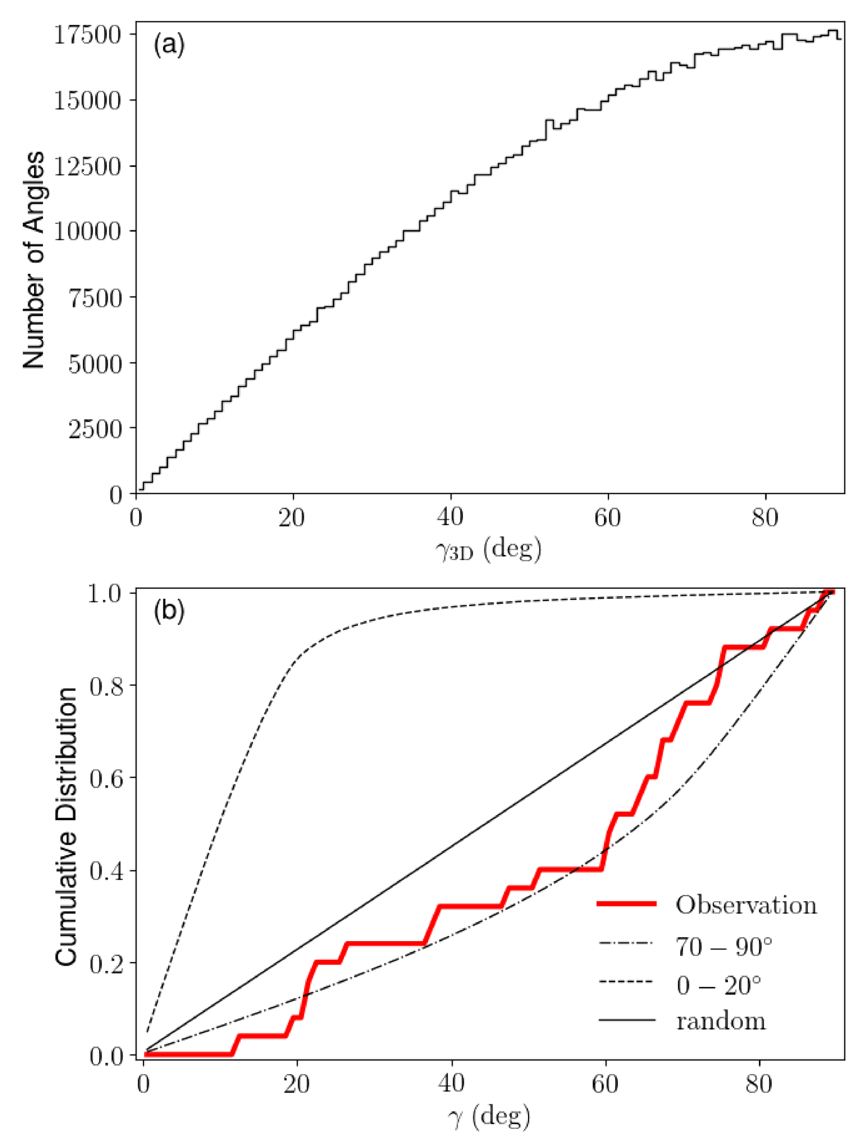


Figure 9: (a) CDF of the simulated gamma3D. **(b)** CDF of gamma2D based on randomly generated gamma3D.

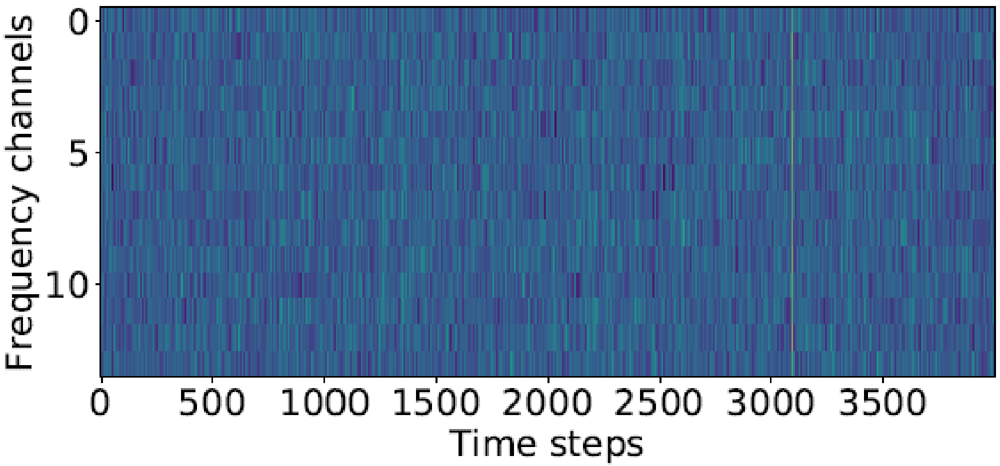
■ Link for 3D model and Virtual Reality : https://sketchfab.com/3d-models/2g327-sio-colored2-b510a2f11b1447fa8908afb01f79ef19

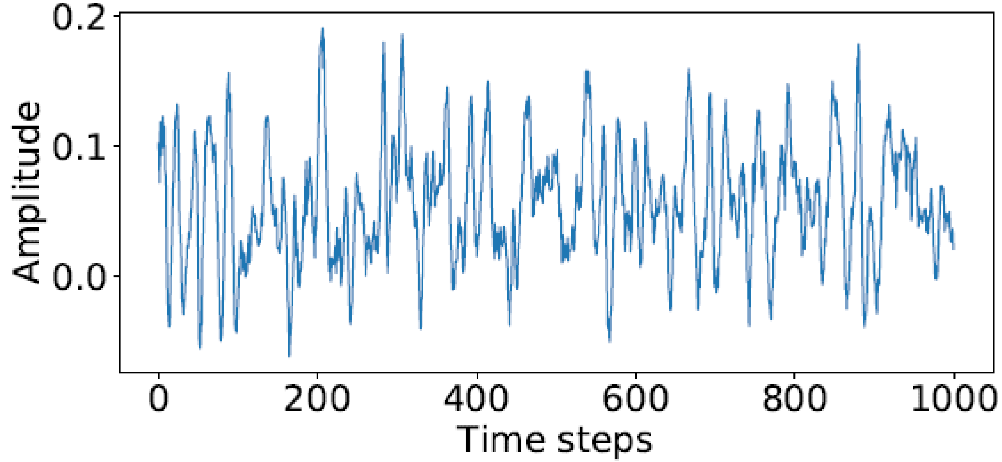
Searching Pulsars Using Neural Networks

Lars Künkel[1], Joris P. W. Verbiest[1,2] & Rajat M. Thomas[3]

[1]Universität Bielefeld, [2]MPIfR Bonn,
[3]Department of Psychiatry, Amsterdam UMC, University of Amsterdam

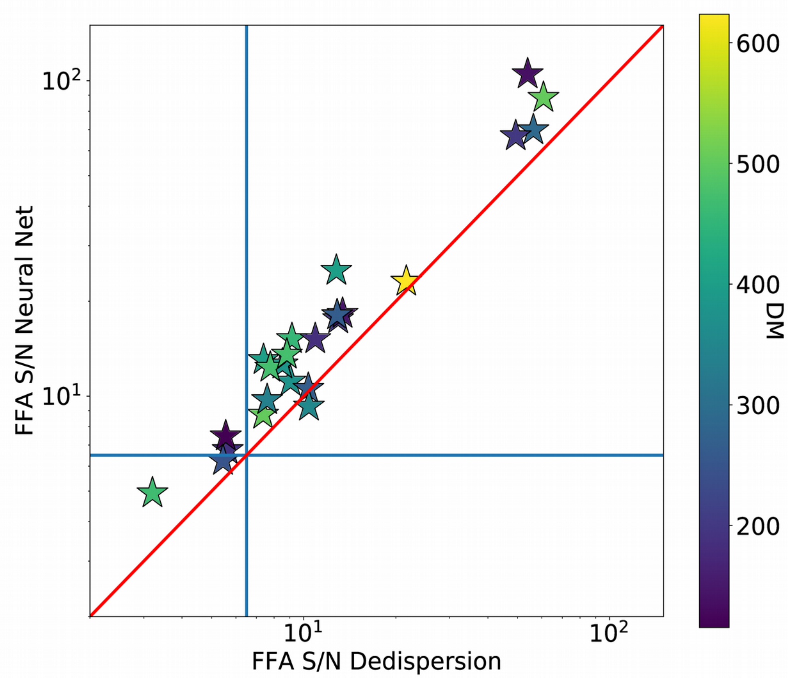
lars.kuenkel@uni-bielefeld.de





Filterbank file

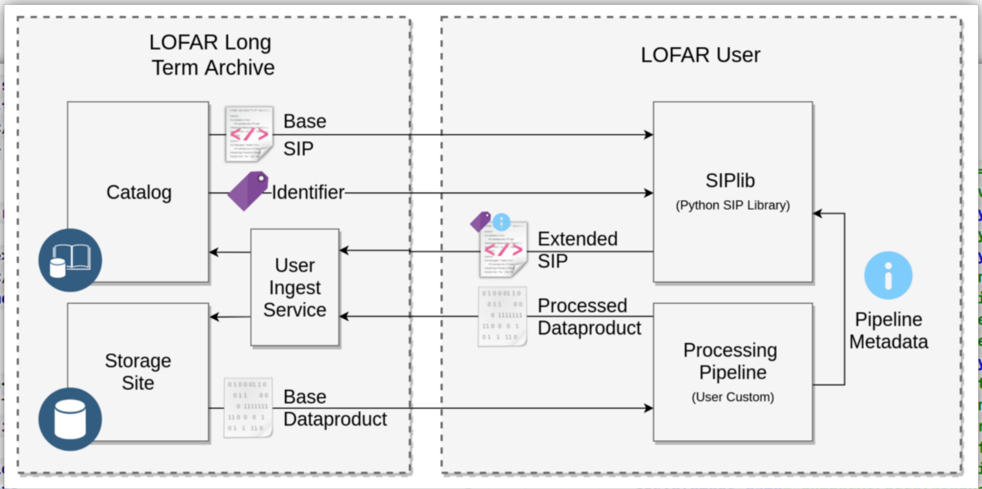
With neural net dedispersed time series



Metadata and User-Provided Data in the LOFAR Long Term Archive

Jörn Künsemöller[1], H.A. Holties[2] and G.A. Renting[2]

[1]Bielefeld University, Bielefeld, Germany; [2]ASTRON, Dwingeloo, Netherlands




```
<strategyName>BeamObservation</s
<strategyDescription>default</s
<startTime>2016-07-30T00:09:00<
<duration>PT3H51M30S</duration>
<relations>
    <relation>
        <relationType>GroupID</
        <identifier>
            <source>MoM</source
            <identifier>697898<
            <name>W63+CygA</nam
        </identifier>
    </relation>
</relations>
<observingMode>Beam Observation
<observationDescription>CygA/1/
<instrumentFilter>110-190 MHz</
<clock units="MHz">200</clock>
<stationSelection>Custom</stati
<antennaSet>HBA Dual</antennaSet>
<timeSystem>UTC</timeSystem>
<numberOfStations>38</numberOfStations>
<stations>
    <station>
        <name>RS210</name>
        <stationType>Remote</stationType>
```

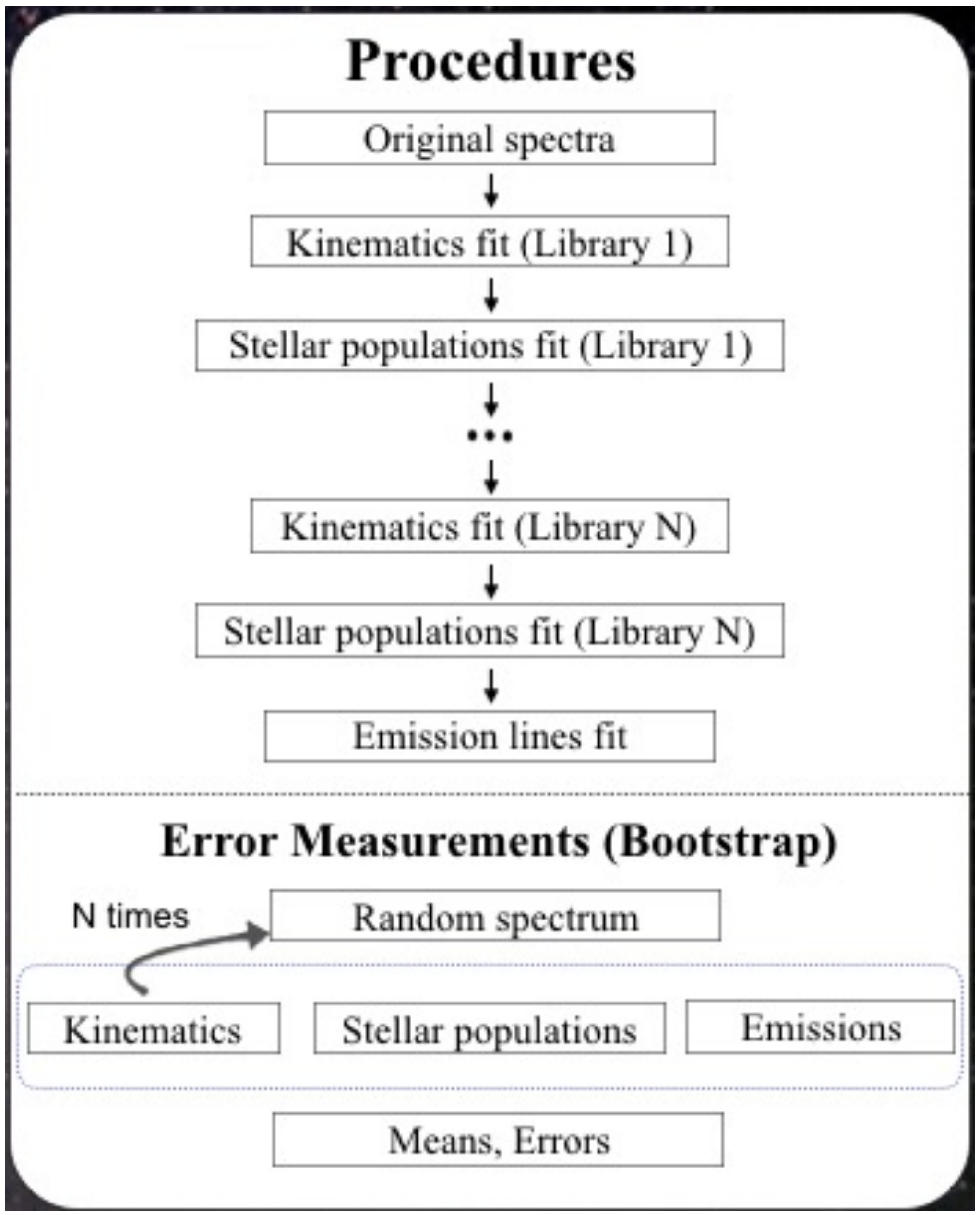
```
="ObservingModeType"/>
vationDescription" type="xs:string"/>
ype="FilterSelectionType"/>
ype"/>
ype="StationSelectionType"/>
ntennaSetType"/>
imeSystemType"/>
elWidth" type="Frequency"/><!--BlueGen
elsPerSubband" type="xs:unsignedShort"
ype="xs:unsignedByte"/>
tions"/>
ntings" type="xs:unsignedShort"/>
rayPointings" type="SubArrayPointings"
fferBoardEvents" type="xs:unsignedShor
ientBufferBoardEvents" type="Transient
        <xs:element name="numberOfBeamFormedDataProducts" type="xs:unsignedShort"/
        <xs:element name="numberOfBitsPerSample" type="xs:unsignedShort"/>
      </xs:sequence>
    </xs:extension>
  </xs:complexContent>
</xs:complexType>
```

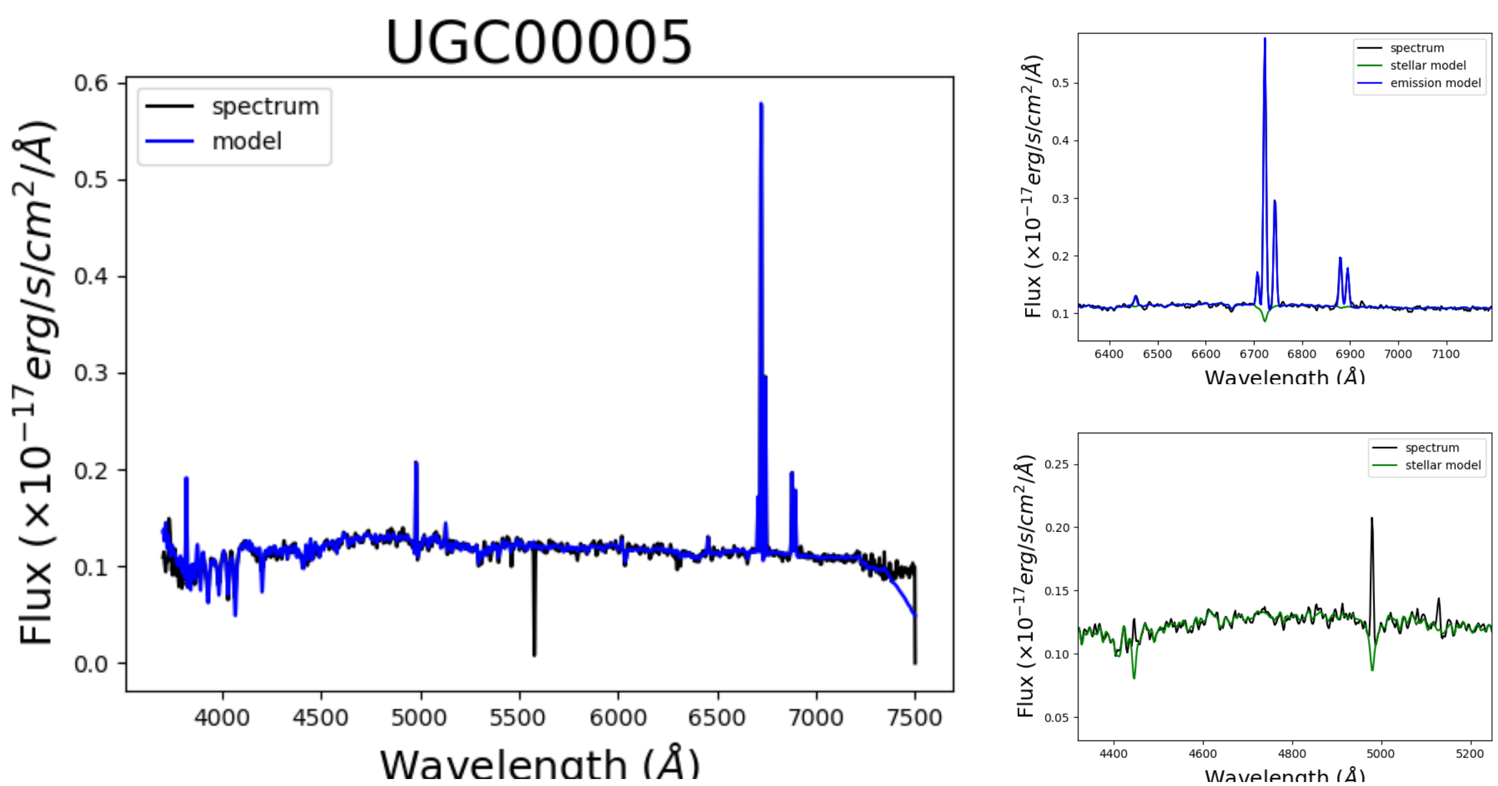
PyParadise: A simultaneous pipeline of stellar and gas kinematics

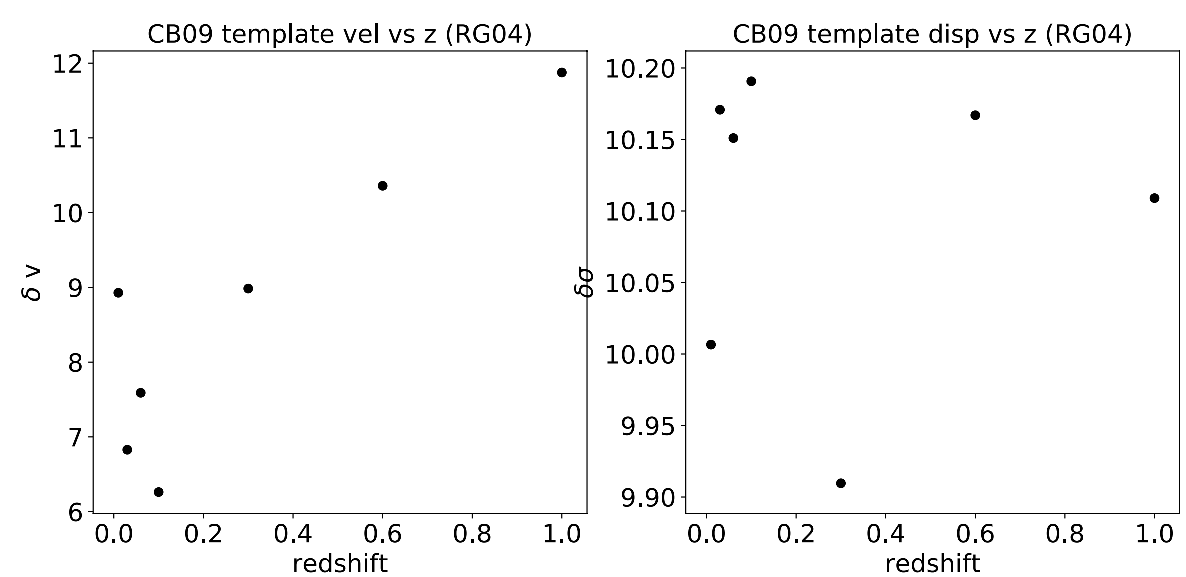
Man I Lam[1], **Bernd Husemann**[2], **Omar S Choudhury**[1], **Anika Beer**[1] and **C. Jakob Walcher**[1]

[1]Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, 14482 Potsdam, Germany

[2] European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching b. München, Germany







Redshift accuracy

Exoplanet detection using Machine Learning

A. Malik[1]

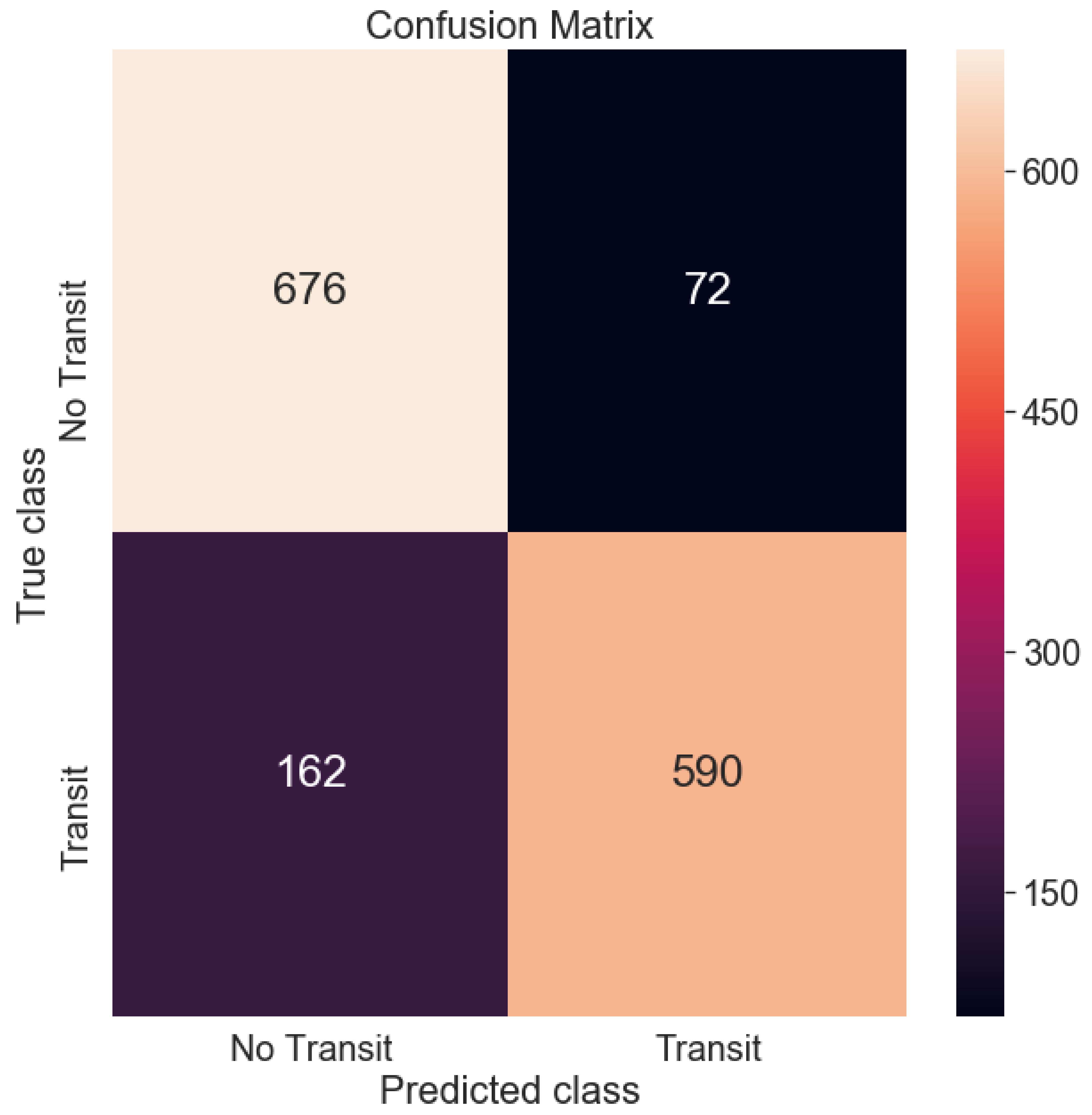
[1]Universitäts-Sternwarte, Ludwig-Maximilians-Universität, München, Germany

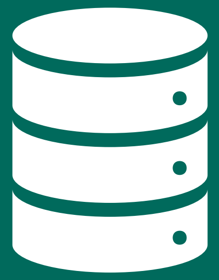
DATA PREPARATION

- Used data from the K2 mission.
  - Removed all known sources from it
  - Flattened the lightcurves and removed 3σ outliers
  - Randomly injected transits
- Used the library 'TSFresh' to extract time series features from each lightcurve.
- After pre-processing, those features served as input to the model.
- Training set: 4500, Validation set: 1500

ANALYSIS

- Trained an XGBoost model, it was able to predict with an accuracy of 85%.
- Detecting a planet with a precision of 0.78
- Large number of false negatives because transits were randomly injected:
  - where signal-noise ratio is very low as planet injection was random.
  - where inclination angle is very low in some cases
  - Next step: Higher quality data
- These methods are efficient as well as robust, can be easily extended to classify single vs multi-planet transit signals without removing any detected transit signals





"I think you should be more explicit here in step two."

THEN A MIRACLE OCCURS

Entering NeuLAND: Analysis work-flow preservation for a fair FAIR

Jan Mayer, Andreas Zilges
University of Cologne, Institute for Nuclear Physics
jan.mayer@ikp.uni-koeln.de

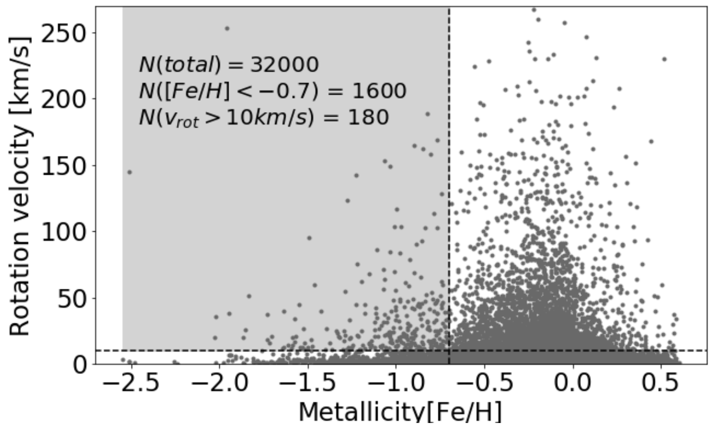
Data Infrastructure at the University of Cologne's Institute for Nuclear Physics

Martin Müller, Jan Mayer and Andreas Zilges
University of Cologne, Institute for Nuclear Physics

A Case Study

Ghost from the past

Planetary engulfment as a possible explanation for observed high stellar rotation in metal poor main sequence stars

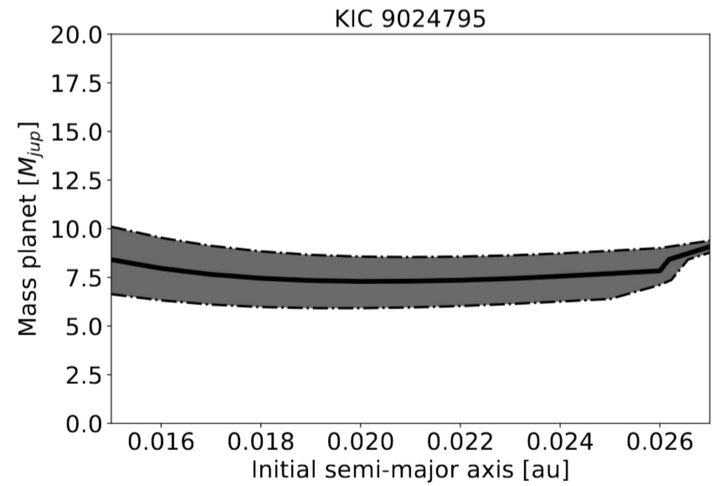


High-rotating main sequence stars from the Kepler mission.

Solution:
Spin-up due to tidal interaction with a high-mass, close-in planet.

Rotation velocity: 26 km/s
Gyrochronology age: ~4 Gyr
Age-Metallicity relationship: > 8 Gyr

Our model presents a solution to this paradox



References:
Angus et al. (2019)
Bensby et al. (2014)
Bouvier et al. (1997)
Carone et al. (2007)
Privitera et al. (2016)
McQuillan et al. (2014)
Hubert et al. (2014)

A.Oetjens and M. Bergemann and L. Carone
Max-Planck-Institute for Astronomy, Heidelberg
Ruprecht-Karls University Heidelberg

Towards reliable photo-z with SOM

O. Razim[1] **and G. Longo**[1]
[1]*Department of Physics, Strada Vicinale Cupa Cintia, 21, 80126, University Federico II, Napoli, Italy*

This project has received financial support from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 721463 to the SUNDIAL ITN network.

Goal: ensure reliability of ML photo-z for the catalog with non-homogeneous spectr-z sample

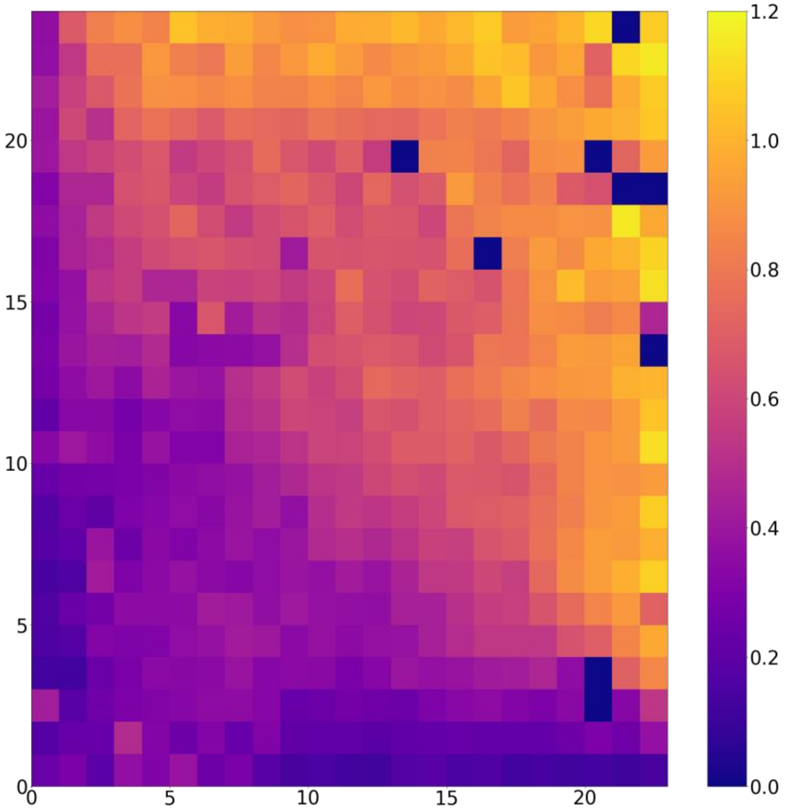
Data: COSMOS2015 catalog. $5 \cdot 10^5$ galaxies, >30 photometric bands, 0<spectr_z<9

Methods: MLPQNA photo-z algorithm + SOM

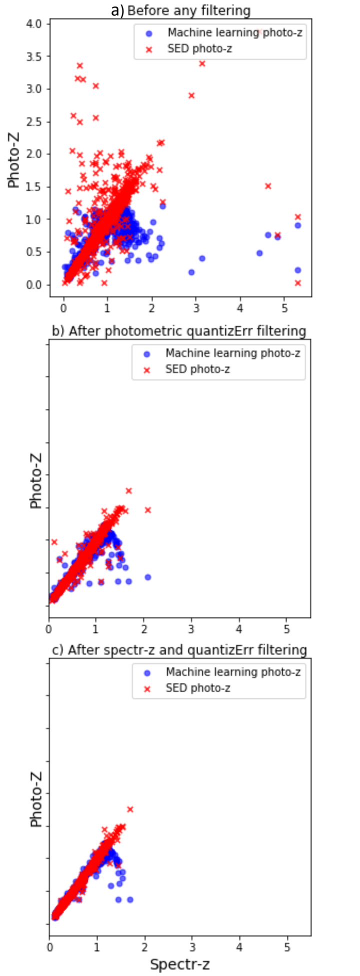
Main idea: Use SOM to detect spectroscopic and photometric outliers for which we don't have enough knowledge base

Results: selection of subsample of catalog with σ(Δz) up to 2 times better than for the whole catalog and % of outliers up to 5 times lower

Further plans: investigation of the detected outliers; experiments with other outlier detection approaches



SOM with post-labelled mean spectr-z



Scatter plots for independent DEIMOS subset

Open Science Grid

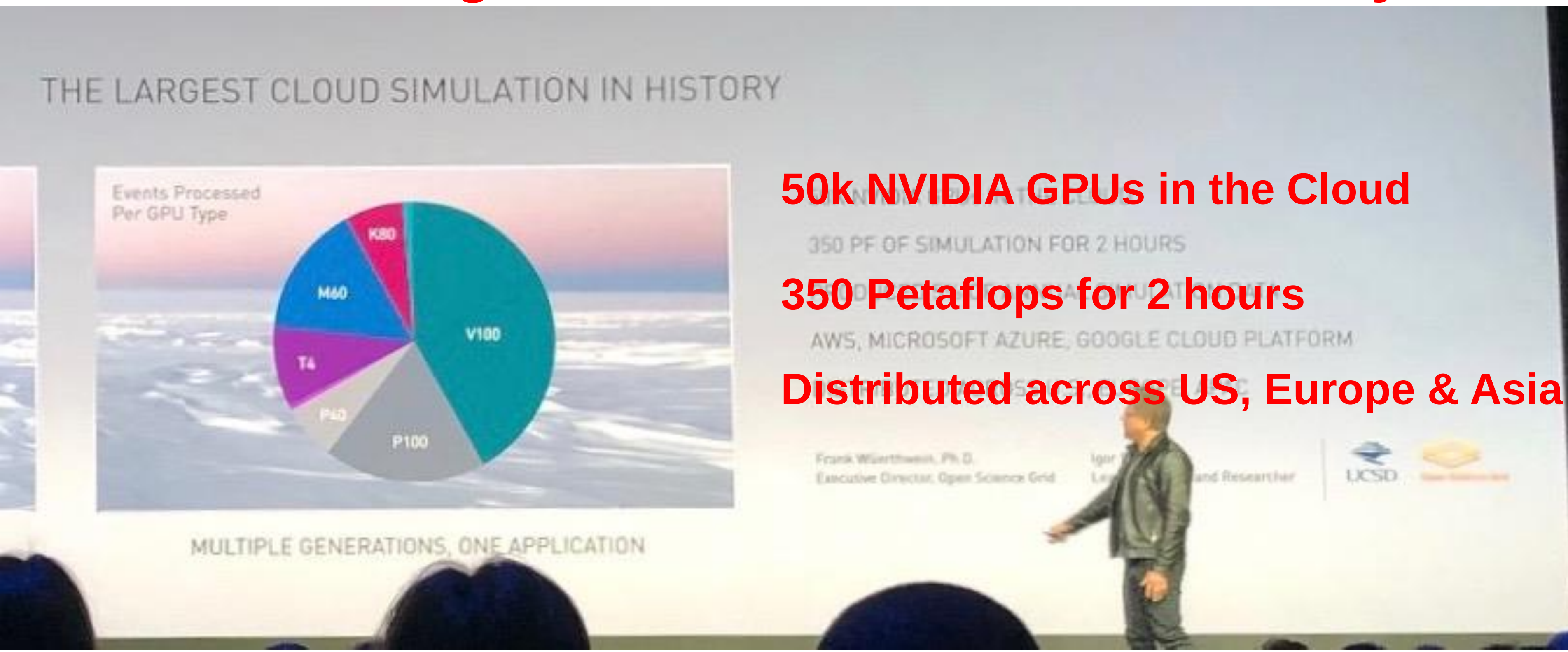
CENTER FOR HIGH THROUGHPUT COMPUTING

INTERNET2

Multi-Cloud Burst for MMA

NSF UCSD

The Largest Cloud Simulation in History



50k NVIDIA GPUs in the Cloud

350 Petaflops for 2 hours

Distributed across US, Europe & Asia

On Nov 16 2019 we bought all GPU capacity that was for sale in Amazon Web Services, Microsoft Azure, and Google Cloud Platform worldwide

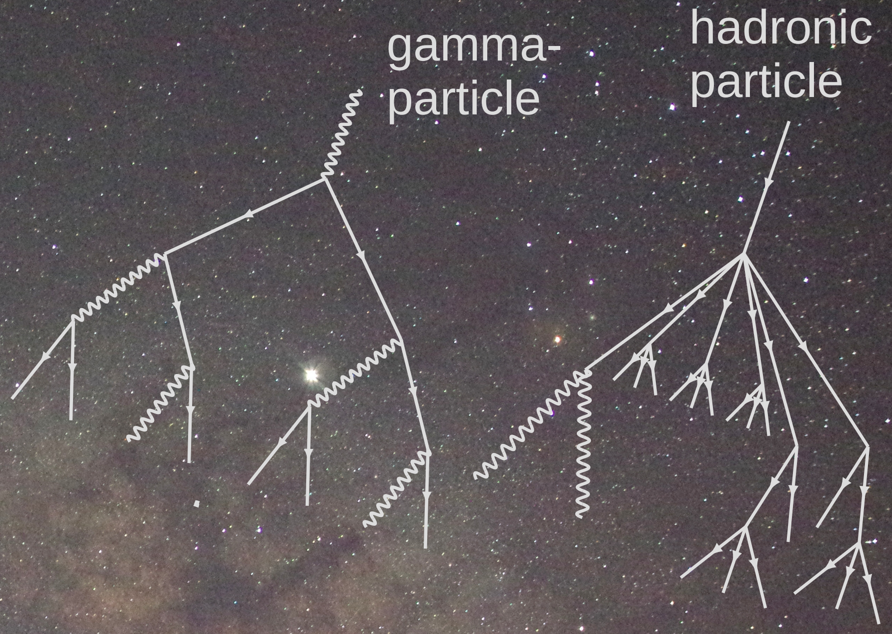
SDSC SAN DIEGO SUPERCOMPUTER CENTER

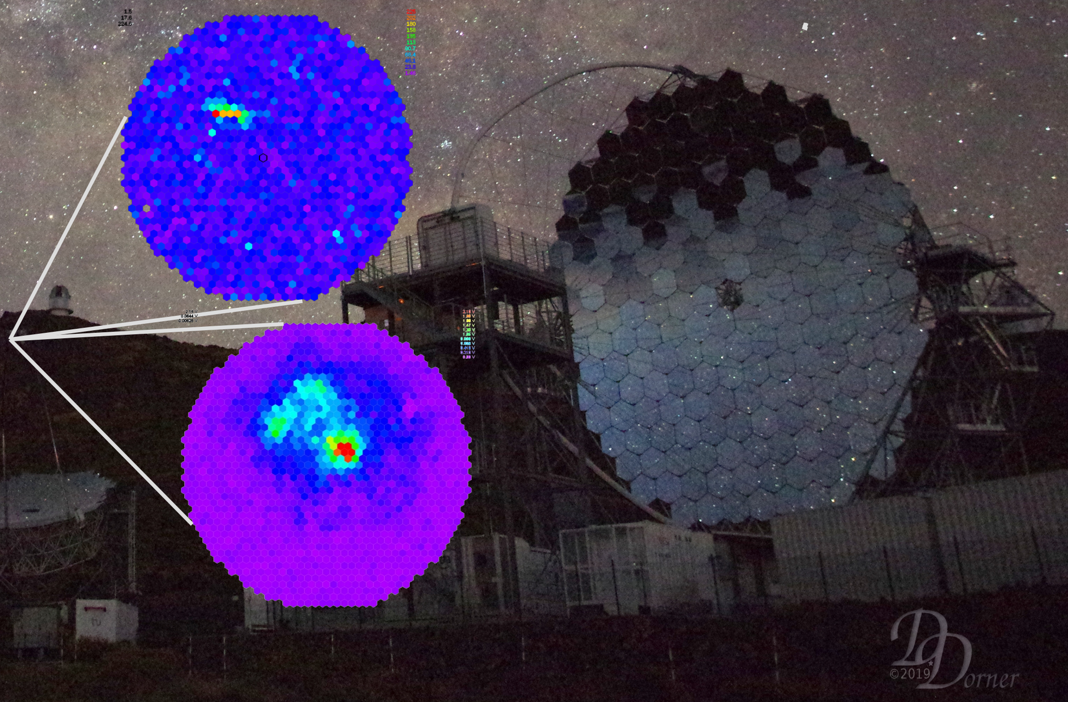
ICECUBE SOUTH POLE NEUTRINO OBSERVATORY

PRP PACIFIC RESEARCH PLATFORM

Machine Learning and Big Data in Cherenkov Astronomy

- **Challenge:** Huge amount of data
- Atmosphere is part of the detector
  - ➔ Simulate different sets of parameter: Zenith distance, ambient light, atmosphere, clouds, calima
  - ➔ **Challenge:** Computing intensive
- Data analysis
  - ➔ Classification and regression problems: Particle type, shower origin, energy
  - ➔ **Challenge:** Mismatch between real and simulated events
- Multi-wavelength context
  - ➔ Machine learning to predict flares
  - ➔ **Challenge:** Simulations describing real light curves





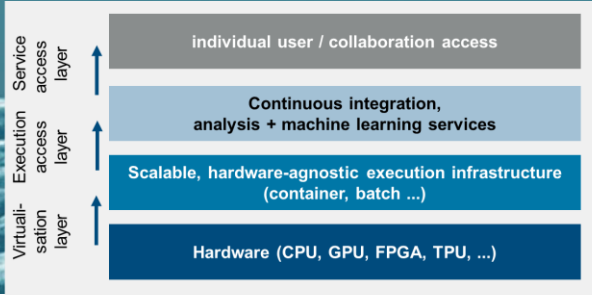
The PAHN-PaN Consortium

➔ see poster #16

Particle, Astroparticle, Hadron & Nuclear Physics accelerate the NFDI

For the PAHN-PaN Consortium:
Kilian Schwarz (GSI)

Bad Honnef, January 14 2020
Heraeus Seminar



Particle, astroparticle, and hadron & nuclear physics

- Decade-long experience in operating self-developed global big data management infrastructure (WLCG, the world's largest grid).

PAHN-PaN goals

- Innovative, industry-standard solutions for *FAIR* Exabyte data management and scientific services.

Synergies, solutions and services (see Cross Cutting Topics)

- Knowlegde and technology transfer to/from entire NFDI..

Task area 1: **Developing workflows and tools for data management**

Task area 2: **FAIR data lifecycle concepts and open data**

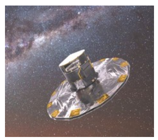
Task area 3: **Data analysis procedures and services**

Task area 4: **Real-time data analysis and selection**

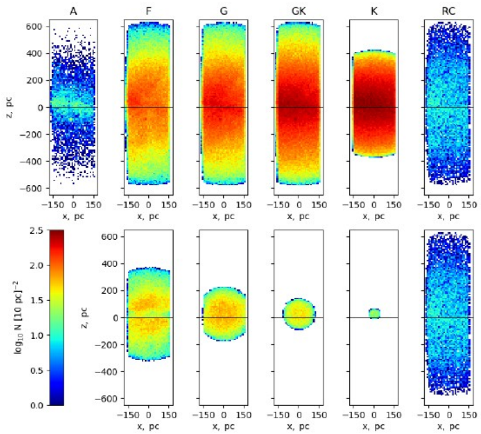
DFG Deutsche Forschungsgemeinschaft

Improving the Milky Way model with Gaia

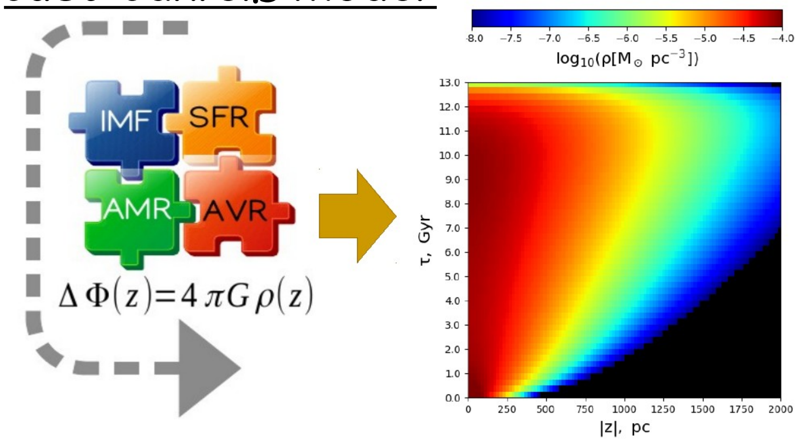
Data:
Gaia DR2, ~10^9 stars:
Complete local samples, ~10^6 stars



Method:
Bayesian approach,
MCMC posterior sampling



Just-Jahreiß model:

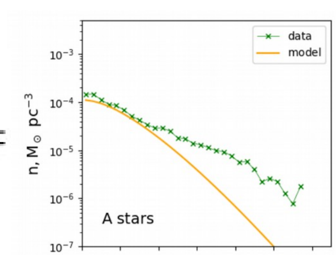


Task: Quickly simulate spatial distribution, kinematics and photometry of 10^6 local stars

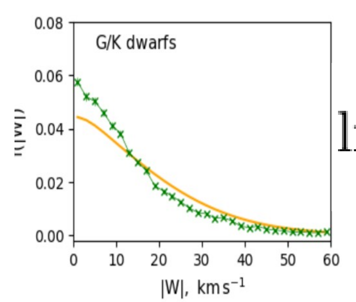
Solution: Wise data selection (complete samples), simplifying assumptions (AMR fixed for a single MCMC run), appropriate bin sizes (10^4 bins to simulate), 10 parameters to fit:

$$\boldsymbol{\theta} = \{\Sigma_d, \Sigma_t, \Sigma_{dh}, \sigma_e, \alpha_{\mathrm{AVR}}, \sigma_t, \alpha_{\mathrm{SFR}}, \beta_{\mathrm{SFR}}, \tau_p, \Sigma_p\}$$

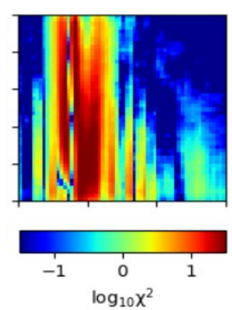
$\ln \mathbb{P} =$



$\ln \mathcal{L}_z +$



$\ln \mathcal{L}_{\mathrm{W}} +$

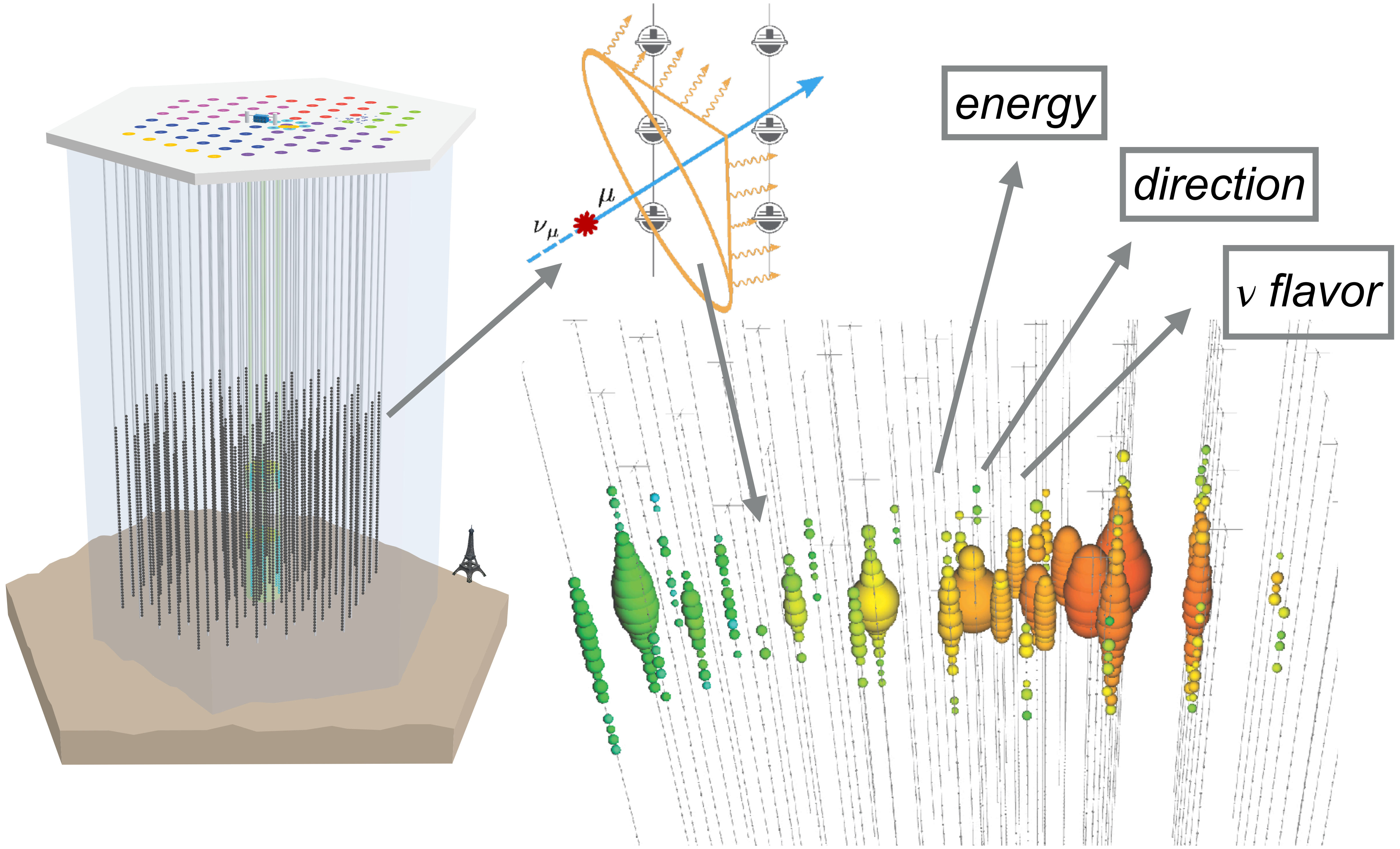


$\ln \mathcal{L}_h + \ln \dfrac{\mathcal{P}}{\mathcal{P}_0}$

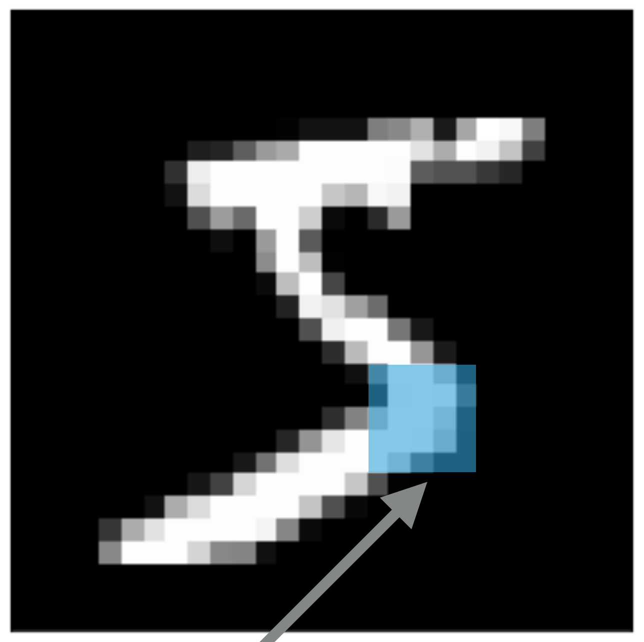
Preliminary results:
Hard to reproduce all observed data features within the assumed framework (wrong IMF? Model reached its predictive limit?) More in Sysoliatina and Just, in prep.

From 2D to 3D to Graphs: Representing Detector Geometries in Neural Networks.

A. Trettin for the IceCube Collaboration

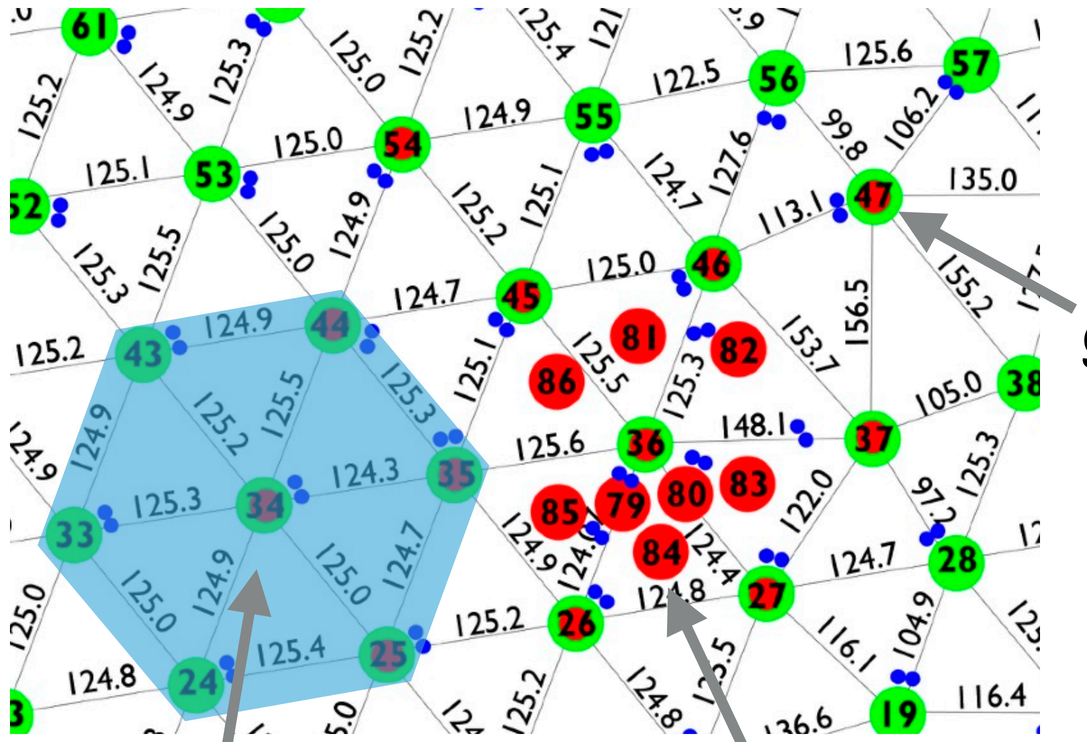


what a CNN expects…

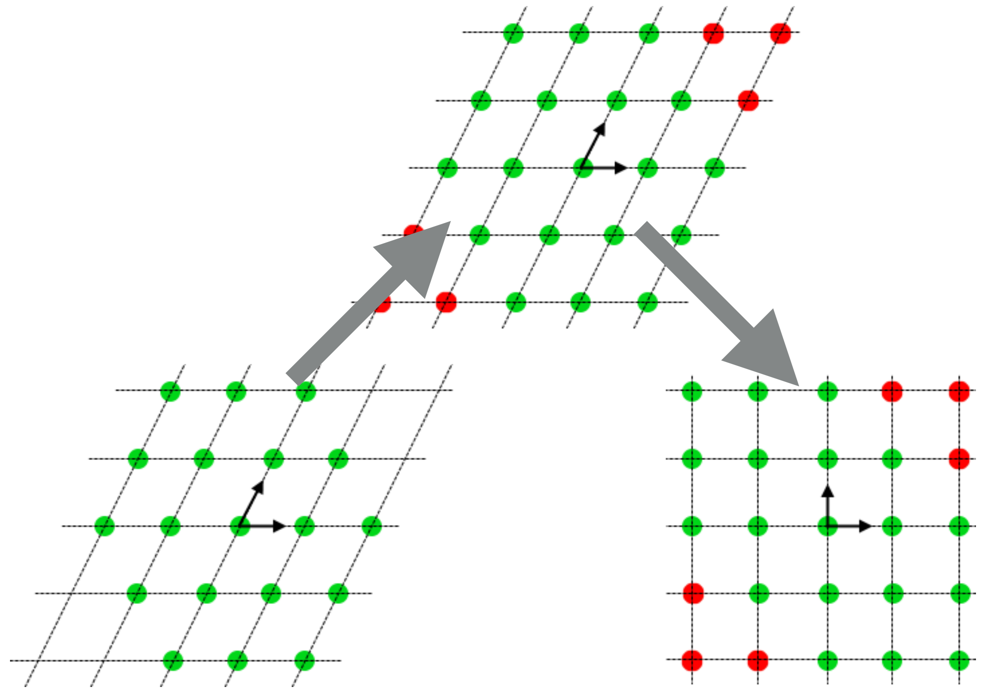


vs.

what we have to handle:

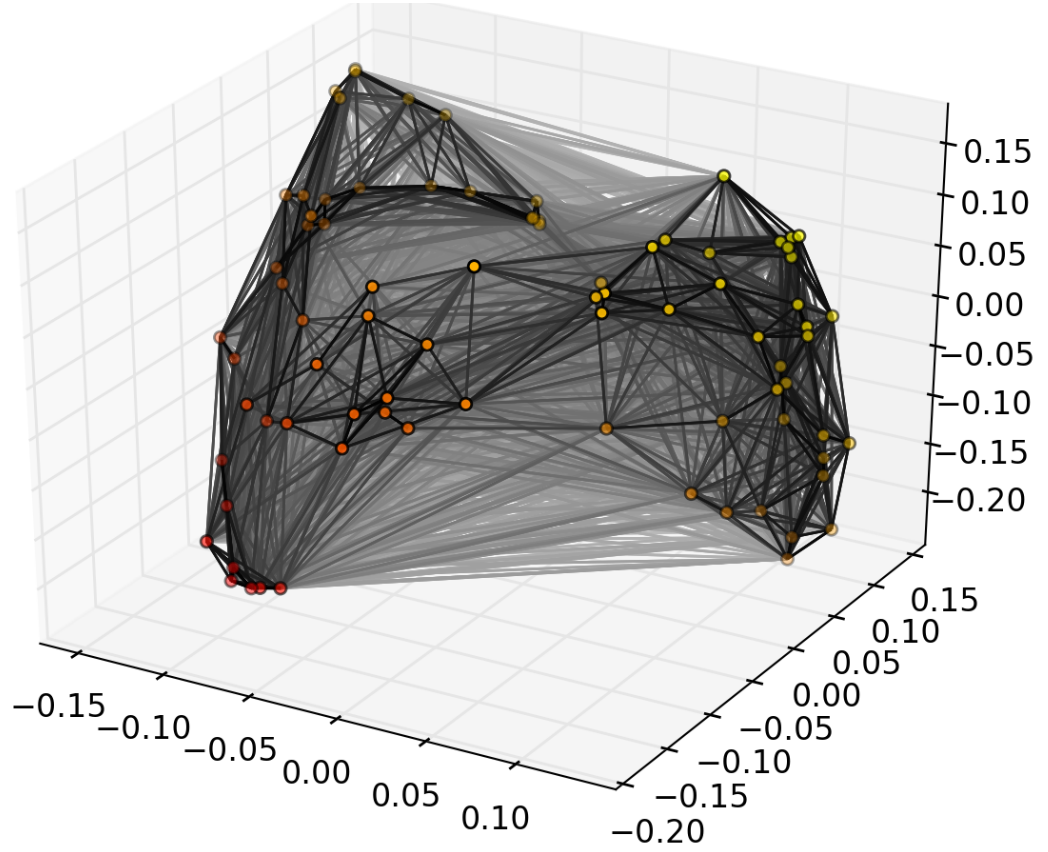


either just *make* it square…



or

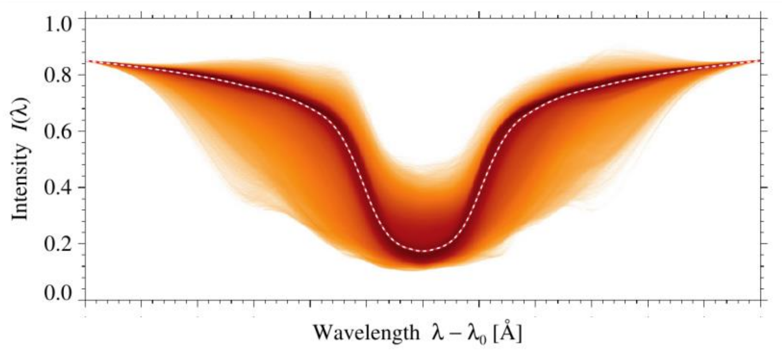
…let the network learn geometry!



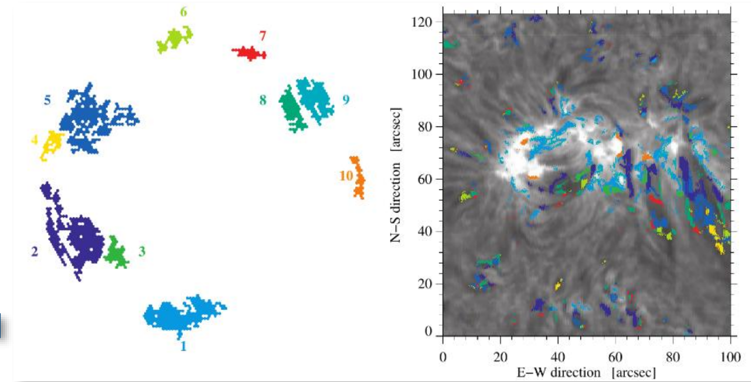
HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

Classification of High-resolution Solar Hα Spectra

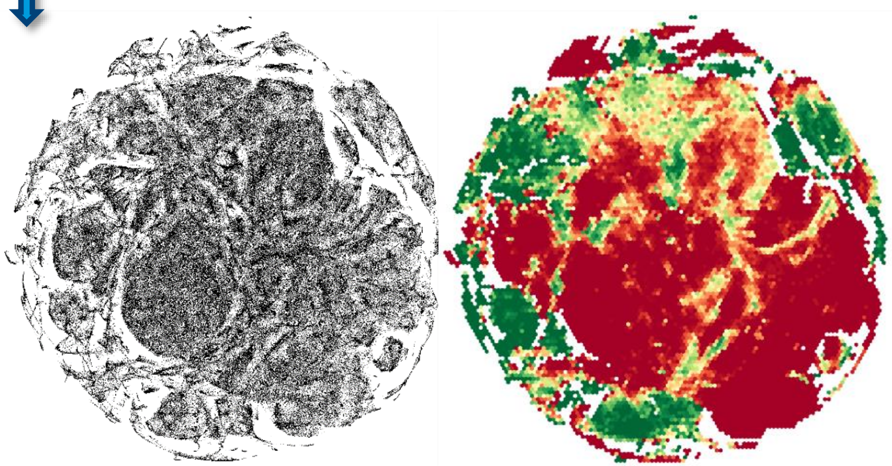
3D data → Science in every pixel



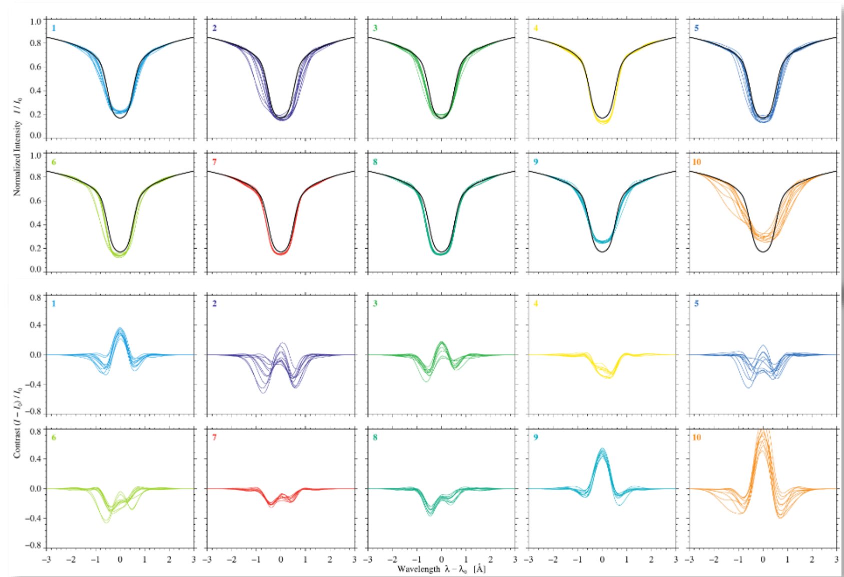
Back mapping



t-SNE → Appropriate tool to classify spectra



Profiles for selected classes

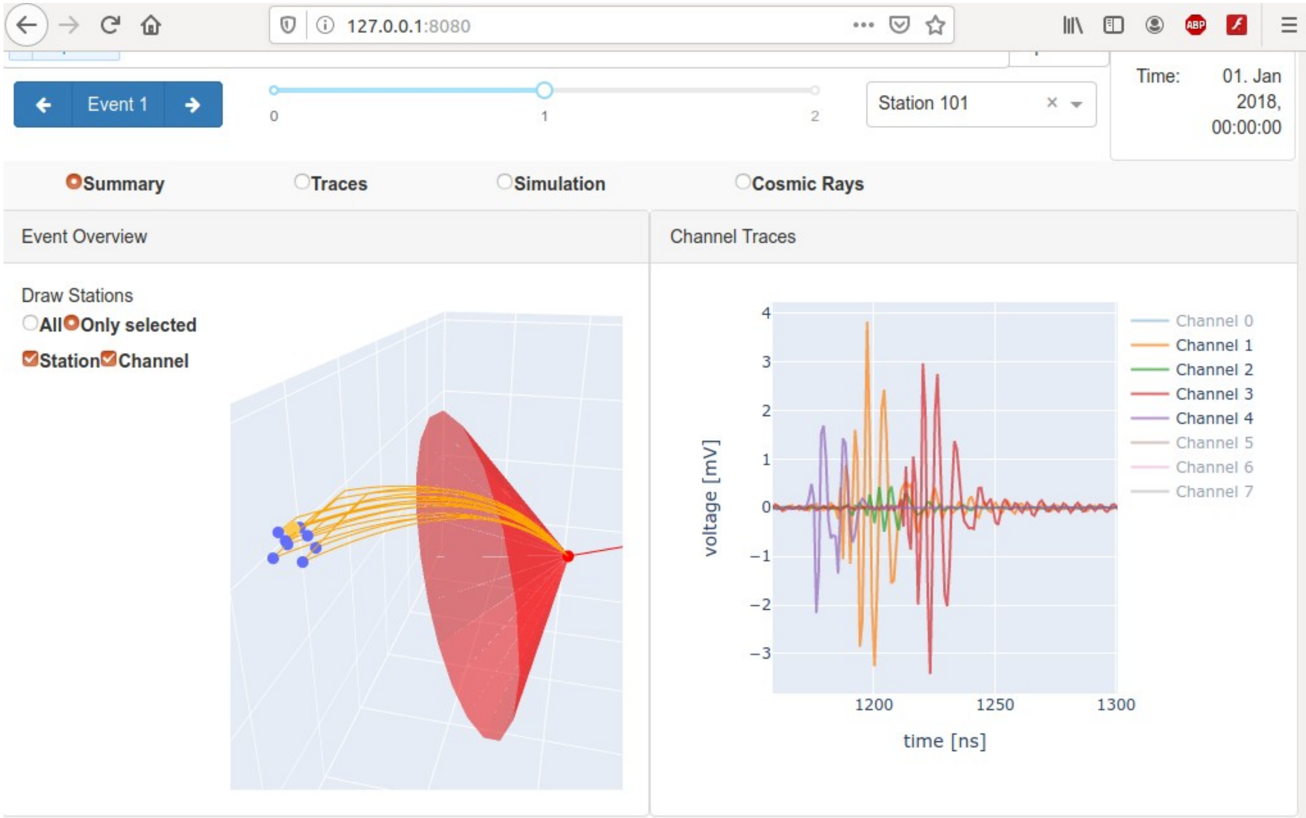


M. Verma, G. Matijevič, C. Denker, A. Diercke, C. Kuckein, E. Dineva, H. Balthasar, I. Kontogiannis, and P. S. Pal

NuRadioReco and NuRadioMC

A Software Framework for the Radio Detector Community

- Radio signals from neutrinos in ice
- First discovery-scale detector to be built this year
- Need for simulation and reconstruction framework
- Detector layout likely to change
- Radio already used for cosmic rays



Our Goal:

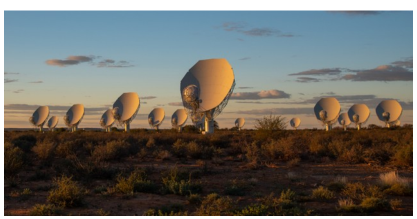
Build a general radio-detector framework

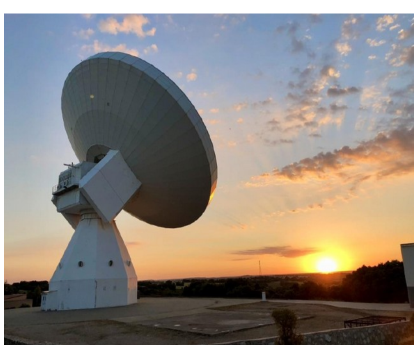
Prototypes for the Next Generation of Computing Backends in Radio-Astronomy

T. Winchen, Max Planck Institue for Radioastronomy

Max-Planck-Institut für Radioastronomie

Different Telescopes – One Modular Backend Software Design





Backend Design Summary

Adaptability:

Rapid adaption to individual observations and new/exotic science cases, e.g. real time transients with machine learning, detection of extreme energy particles, …

\+ **Commensality:**

Simultaneous processing for multiple disciplines and also offline data processing

\+ **Simplification**:

Minimize in-house solutions, maximize reuse of components and prefer COTS computing hardware.

\+ **Standardization:**

Use established open source software and industry operation standards

---

= **Maximize science / €**

Reduce development time (=money)

BoostNumpy

BoostNumpy: Big Data Processing in C++ with Python convenience

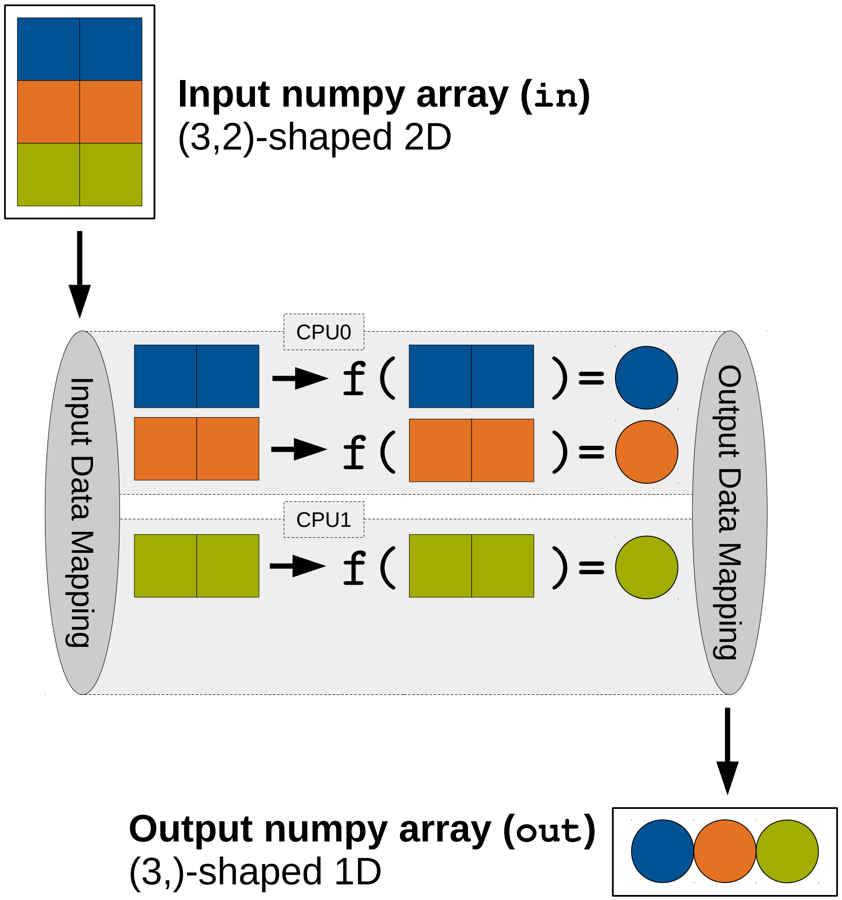
github.com/martwo/BoostNumpy

Question

- How to process big data fast and in a convenient way?

Answer

- Processing **big data** stored in **numpy arrays** using **C++** functions from within a **Python** script
- Benefit from both worlds: **speed** & **usability**
- **No data copying** required
- Automatic **Parallelization** on multi-CPU computers possible
- **Very easy connection of C++ & Python** by the developer via boost & MPL
  → It's a one-liner!



Martin Wolf, ECP, TUM SFB 1258 Neutrinos Dark Matter Messengers

Small Problems with Big Data in Astronomy

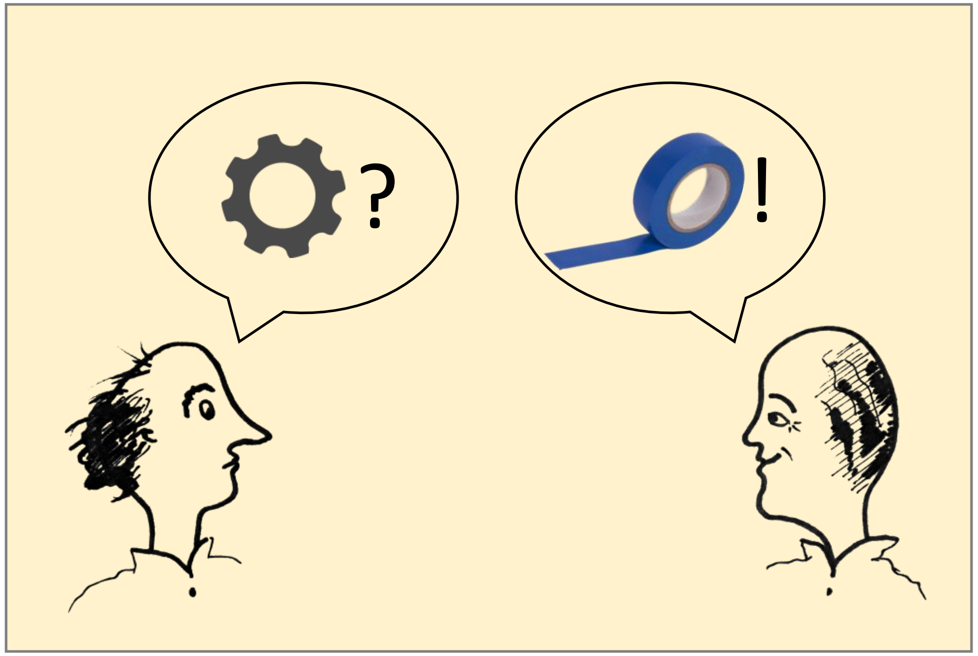
Oleksandra Razim[1], Kseniia Sysoliatina[2]
[1]Department of Physics, Strada Vicinale Cupa Cintia, 21, 80126, University Federico II, Napoli, Italy,
[2] Astronomisches Rechen-Institut, Mönchhofstr. 12-14, 69120 Heidelberg, Germany

Catalog search

- Virtual Observatory
- Naming conventions
- …

Data interface

- ADQL
- User scenarios
- …



Formatting

- FITS
- Guidelines
- …

Reproducibility

- Jupyter Notebook
- Git
- …