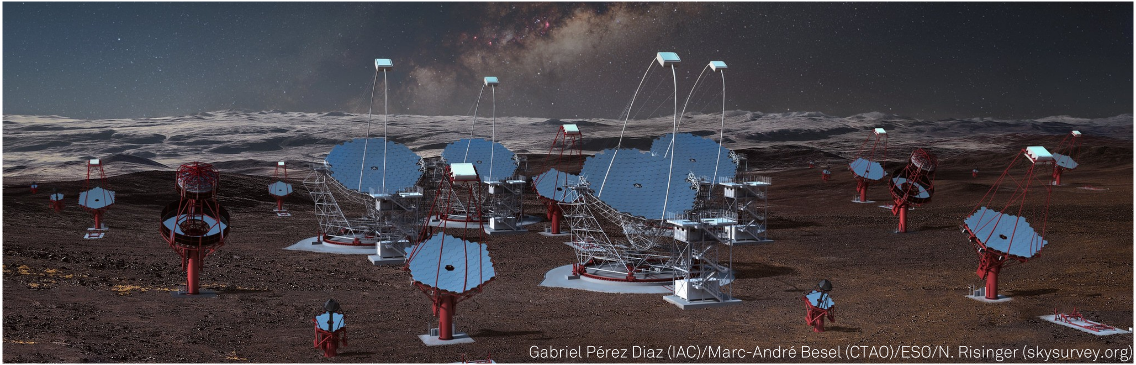
technische universität dortmund

SFB 876 Providing Information by Resource-Constrained Data Analysis

exp. physik 5 astroteilchenphysik



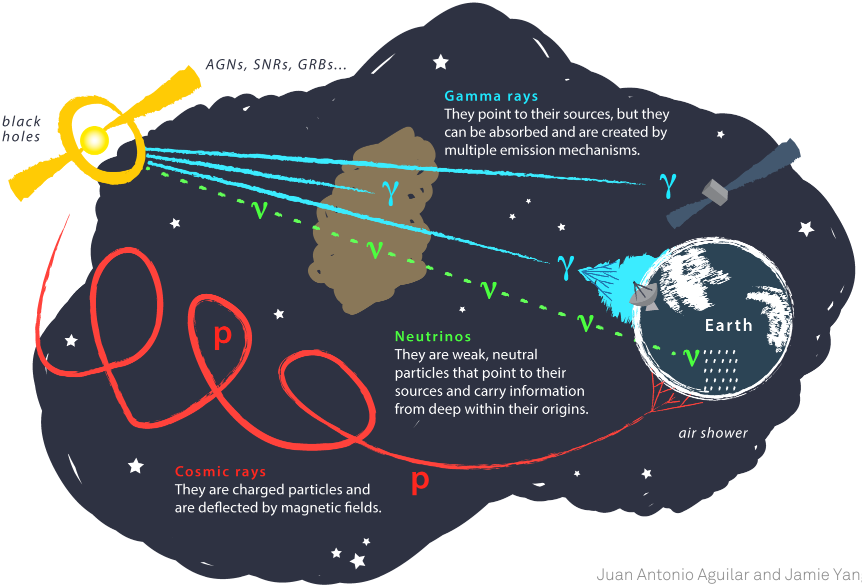
Gabriel Pérez Diaz (IAC)/Marc-André Besel (CTAO)/ESO/N. Risinger (skysurvey.org)

Event Reconstruction for the Cherenkov Telescope Array

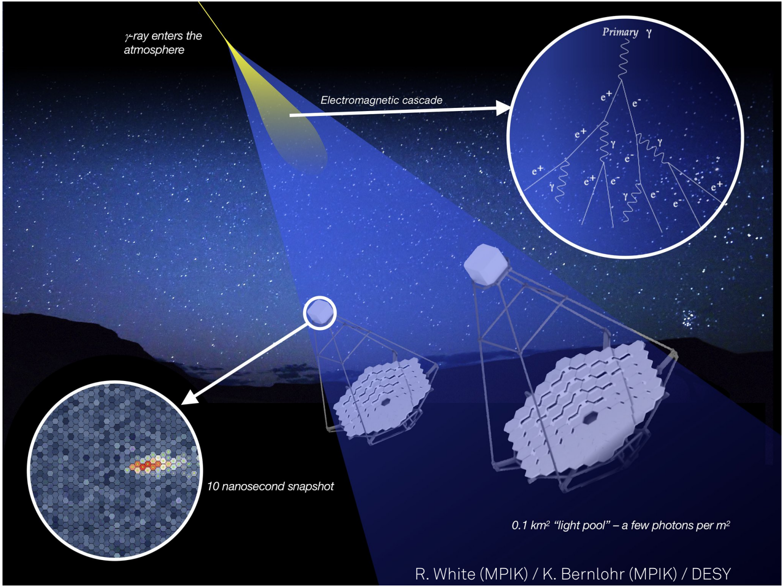
Maximilian Nöthe

Dortmund Data Science Center Kolloquium – 2022-01-11

Astroparticle Physics – TU Dortmund

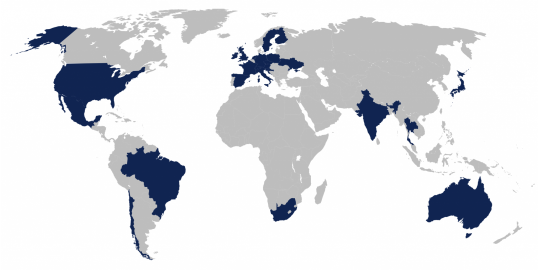


Juan Antonio Aguilar and Jamie Yang. IceCube/WIPAC


γ-ray enters the atmosphere
Electromagnetic cascade
Primary γ
e+
e-
γ
10 nanosecond snapshot
0.1 km² "light pool" – a few photons per m²
R. White (MPIK) / K. Bernlohr (MPIK) / DESY

cta cherenkov telescope array

- CTA will be the next generation gamma-ray astronomy observatory
- The first ground-based gamma-ray instrument to operate as an open observatory
- ~100 telescopes of four types at two sites, one in each hemisphere
- Energy range from **20 GeV** to **300 TeV**
- Prototypes of all telescopes are operating
- Construction of first phase to be finished in 2025



1500 Members
from 150 institutes
in 25 countries

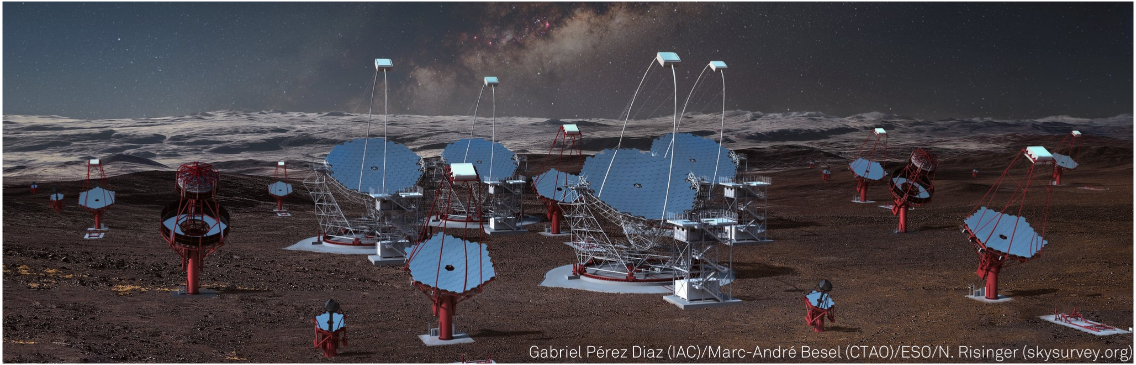
Planned Array in the Northern Hemisphere

Observatorio del Roque de los Muchachos, La Palma, Canary Islands, Spain

- Currently planned for first stage: 4 Large Sized Telescopes (**23 m**), 9 Medium Sized Telescopes (**12 m**)
- Goal: 4 LSTs, 15 MSTs

Planned Array in the Southern Hemisphere

Paranal Observatory, Atacama Desert, Chile



Gabriel Pérez Diaz (IAC)/Marc-André Besel (CTAO)/ESO/N. Risinger (skysurvey.org)

- Currently planned for first stage: 14 Medium Sized Telescopes (**12 m**), 37 Small Sized Telescopes (**4 m**)
- Goal: 4 LSTs, 25 MSTs, 70 SSTs

The LST Prototype

- Inaugurated in 2018
- Currently in commissioning
- Lots of observations already taken → “real” data
- Energy range: **20 GeV** to **50 TeV** (Dominitating CTA sensitivity between **20 GeV** and **200 GeV**)

The LST Prototype

- Inaugurated in 2018
- Currently in commissioning
- Lots of observations already taken
  → “real” data
- Energy range: **20 GeV** to **50 TeV**
  (Dominitating CTA sensitivity between **20 GeV** and **200 GeV**)
- Observations stopped since 2021-09-26 due to volcano eruption

M. Nöthe

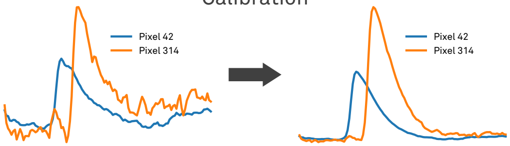
How does CTA data look like?

- Cherenkov telescopes feature extremely fast, sensitive cameras
- Detect single to hundreds of photons in a time frame of $\approx$ **40 ns**
- Recorded data consists of time series for each pixel
- LST single telescope losslessly compressed raw data rate: **10 000 Events/s** $\Rightarrow \approx$ **3 TB h^{-1}**
- Typical good observation time: **2 000 h/year**
- CTA goal: maximum of **20 PB/year** $\Rightarrow$ Lossy data volume reduction needed

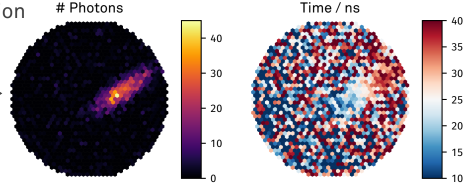
First Analysis Goal: Estimate the key properties of the primary particle in each event

Energy	1d-Regression
Particle type	Classification
Direction of origin	2d-Regression

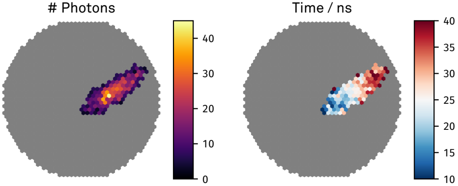
Calibration



Pulse Integration



Cleaning

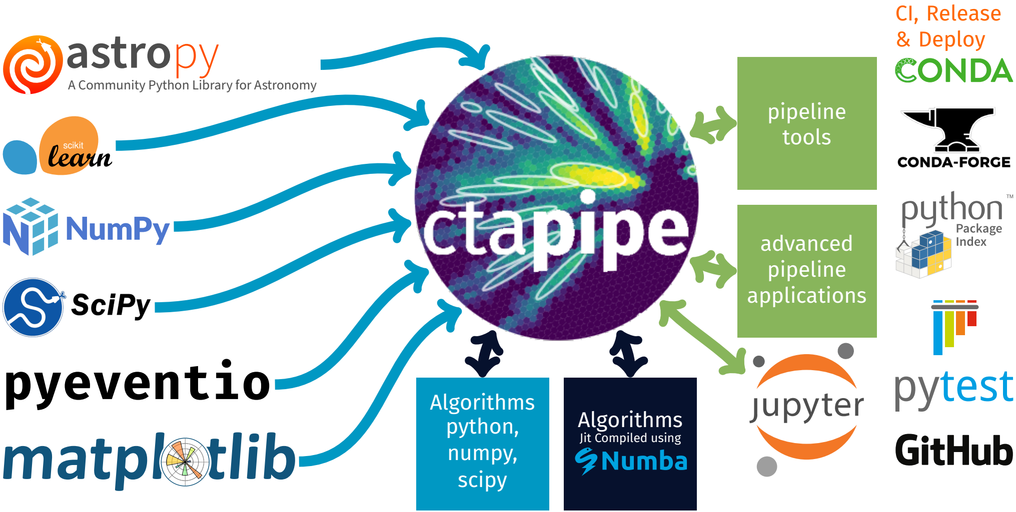


Feature Extraction

event	width	length	…	size
0	0.15	0.52	…	1253.1
2	0.05	0.12	…	321.3
5	0.08	0.19	…	512.7

Parameter Estimation

event	energy	gammaness	ra	dec	time
–	GeV	–	deg	deg	MJD
0	1500	0.82	83.6	22.1	59024.63123
2	400	0.73	83.5	21.9	59024.64183
5	680	0.92	83.7	22.0	59024.67093



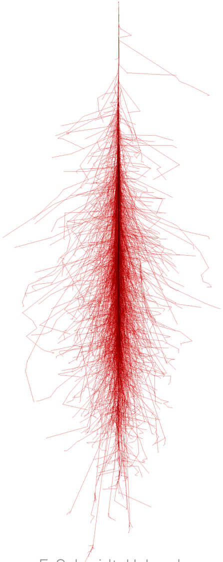
https://github.com/cta-observatory/ctapipe

Side note: Labeled Data

- Supervised learning for the event property estimation
- Mostly decision tree based ensemble methods

⇒ Labeled datasets are needed for training

- Air shower physics and detector electronics are Monte-Carlo simulated for known primaries
- Brings its own challenges with respect to
  - Data storage
  - Computation time
  - Agreement between simulations and actual observations
- Several Multi-PB simulation datasets used for the planning and optimization of CTA



F. Schmidt, U. Leeds

After the Event Property Estimation: Solving the inverse Problem

- The full measurement process of a Cherenkov Telescope can be described using

$$e(\hat{E}, \hat{\alpha}, \hat{\delta}, t) = \int R(\hat{E}, \hat{\alpha}, \hat{\delta} | E, \alpha, \delta, t) \cdot I(E, \alpha, \delta, t) \, \mathrm{d}E \, \mathrm{d}\Omega + b(\hat{E}, \hat{\alpha}, \hat{\delta}, t)$$

with

$\hat{E}, E$	estimated / true gamma-ray energy
$\hat{\alpha}, \hat{\delta}, \alpha, \delta$	estimated / true gamma-ray direction in celestial coordinates
t	Time
$e(\hat{E}, \hat{\alpha}, \hat{\delta}, t)$	Observed event distribution
$R(\hat{E}, \hat{\alpha}, \hat{\delta} \| E, \alpha, \delta, t)$	Instrument Response Function
$I(E, \alpha, \delta, t)$	True gamma-ray signal distribution
$b(\hat{E}, \hat{\alpha}, \hat{\delta}, t)$	Background from cosmic rays and other sources

Goal: Estimate $\boldsymbol{I}$ from $\boldsymbol{e}$, $\boldsymbol{R}$ and $\boldsymbol{b}$ $\Rightarrow$ Inverse Problem, Unfolding

More labeled data needed to calculate $\boldsymbol{R}$

$\boldsymbol{R}$ changes rapidly (20 minutes) during observations due to environmental conditions, pointing direction, moon light, …

Summary

- The Cherenkov Telescope Array will be the next, large step in gamma-ray astronomy
- The scale and tasks of the project pose unique challenges to data processing and physics analysis
- Event reconstruction is the central part of the analysis pipeline before data is given out to astronomers
- Data given out to users of the observatory will consist of reconstructed event lists with corresponding instrument response functions
- Further analysis of this data requires solving the inverse problem, e. g. using unfolding techniques