

SFB 876 Providing Information by Resource-**Constrained Data Analysis**





Project C3

Multi-level statistical analysis of high-frequency spatio-temporal process data

Prof. Dr. Katharina Morik, Prof. Dr. Dr. Wolfgang Rhode, Dr. Tim Ruhe

Challenges Ahead

- Next generation of experiments now being planned and built
- Detector arrays instead of single telescopes
- Dramatic increase in data volumes
- Growing complexity of data

CTA

Prob

Faster reaction times needed for multi-messenger observations

> IceCube-Gen2 100 T B/ day 100 T B/ day

SKA 100 PB/day



Upcoming Experiments

IceCube Gen2



Square Kilometer Array



Spectral Reconstruction

Time-dependent aggregation in DSEA⁺ facilitates **time series analyses** to detect concept drifts







- Design objective function to incorporate ordinal nature of class labels
- Improvements also increase sensitivity towards BSM physics (C4) (C5)

Towards Tailored Deep Learning Methods

Geometric Deep Learning (A6)

- Extension to better suited architectures
- Matches the complex data measured by physics detectors

Exploitation of A Priori Knowledge

- Physicists have comprehensive prior knowledge about their experiments
- Develop methods to include symmetries, constraints, and laws of nature into Deep Learning architectures





Generative Networks

- Event reconstruction
- Fast approximate simulation
- Extraction of physics parameters from observations, e.g. ice properties

Data Modelling and Simulation

Next Generation CORSIKA

- Cherenkov light production and photon raytracing amount to 80% 90% of runtime (a 2 TeV shower produces 10⁹ to 10¹¹ photons)
- **Offload to GPUs** \Rightarrow factor 3 expected for decrease in computing time for CTA
- Main challenge: keeping simulations deterministic



Active Sampling for Simulation Control

Challenge:

SKA



- Many simulated examples are similar
- Only examples with **relevant information** should be selected/simulated

Strategy:

- Iteratively update the training set with a Query By Commitee (QBC) approach
- Reweight examples according to their usefulness
- Sample/Generate more useful examples in succeeding iteration
- Adopt and refine approach of (B3)



