**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
- Faster reaction times needed for multi-messenger observations

**Upcoming Experiments**
- CTA: 100 TeV/day
- IceCube Gen2: 100 TeV/day
- SKA: 100 PB/day

**Cherenkov Telescope Array**
- 100 telescopes planned for the Atacama Desert, Chile
- 20 telescopes for La Palma, Spain
- Several prototypes exist
- First 23 m telescope nearly finished
- Latest simulated data set is 1.4 PB

**Square Kilometer Array**
- 130,000 dipole antennas and 200 dishes with 15 m in diameter
- 1 km² collection area
- Located in Australia and South Africa
- Raw data rate will amount to several PBs (100x total WWW traffic in 2017)
- Sky survey 1000 times faster than existing systems
- Continuous stream of data

---

**Geometric Deep Learning**
- 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

**Efficient CNNs**
- Hexagonal kernels
- FPGA implementation

---

**Spectral Reconstruction**
- Time-dependent aggregation in DSEA facilitates time series analyses to detect concept drifts
- Enhance DSEA to unfold continuous spectra from a regression task
- Design objective function to incorporate ordinal nature of class labels
- Improvements also increase sensitivity towards BSM physics

**Towards Tailored Deep Learning Methods**
- Time-dependent aggregation in DSEA facilitates time series analyses to detect concept drifts
- Enhance DSEA to unfold continuous spectra from a regression task
- Design objective function to incorporate ordinal nature of class labels
- Improvements also increase sensitivity towards BSM physics

---

**CTA (Brügge et al. 2017)**
- Transfer FACT experience to CTA
- Prototype already meets current CTA real-time requirements
- Horizontal scalability through Apache Spark for future requirements
- Recording starts by the end of 2018

**SKA**
- From streams of events to a continuous data stream

---

**Data Modelling and Simulation**
- Next Generation CORSIKA
  - Cherenkov light production and photon raytracing amount to 80% – 90% of runtime (a 2 TeV shower produces 10^9 to 10^11 photons)
  - Offload to GPUs – factor 3 expected for decrease in computing time for CTA
- Main challenge: keeping simulations deterministic

---

**Active Sampling for Simulation Control**
- Many simulated examples are similar
- Only examples with relevant information should be selected/simulated

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