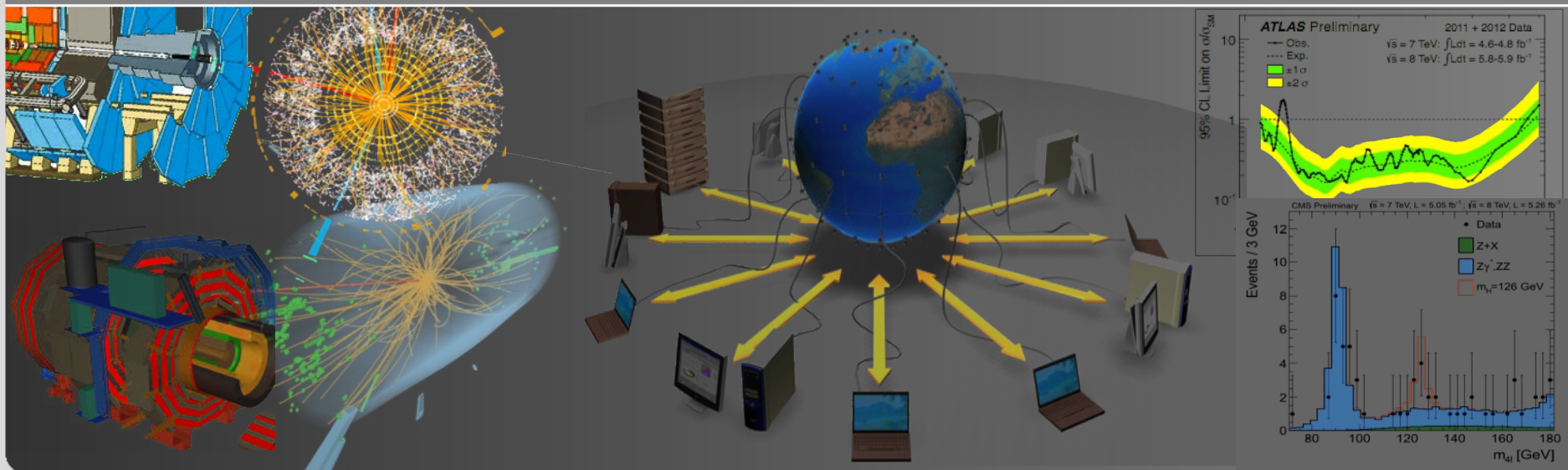


Machine Learning in Particle Physics

Günter Quast, Torbern Ferber, Ulrich Husemann, Markus Klute

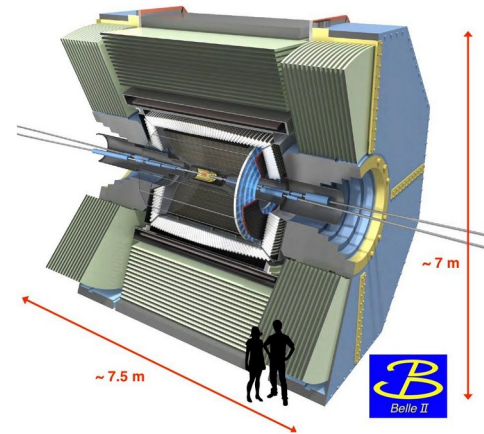
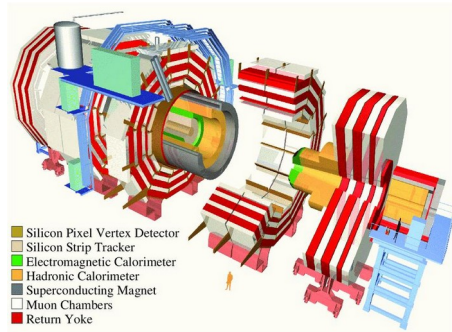
Fakultät für Physik
Institut für Experimentelle Teilchenphysik

KCETA-MathSEE Workshop

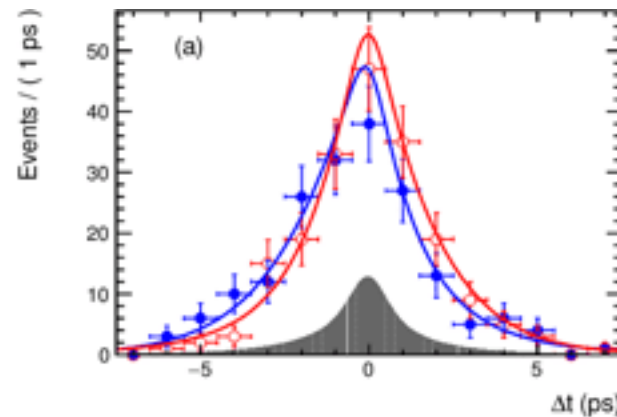
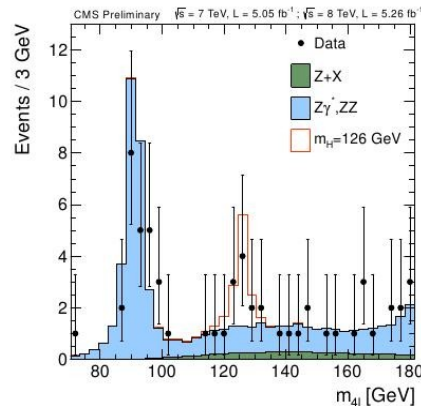
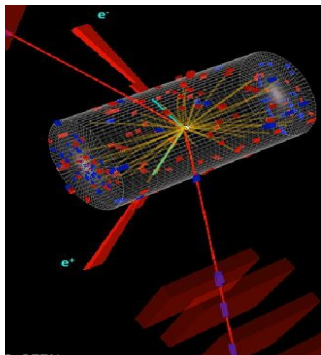


The Challenge

Particle Detectors produce huge amounts of data – $O(\text{PB/sec})$



Interesting Physics Events (=“images” of a particle collision)
are (in most cases extremely) **rare**



Many steps of
data reconstruction, reduction and selection
from raw data to physics

Application Areas of ML

- On-line **classification** of Particles and Events
using FPGA, GPU and CPU Farms
Challenge: performance and speed (decision within $\sim 500\text{ns}$)
- Exploring speed-up of simulations using
Generative Adversarial Networks (**GAN**)
Challenge: need good approximation of detailed simulation !
- Traditionally using a wide range of various ML techniques
(called “Multi-Variate Analysis”, **MVA**) for
 - identification of “physics objects” (= element in a particle collision)
(e.g. b-, top- and tau-Tagging, electron and muon identification, ...)
 - multivariate regression techniques for improved
reconstruction (e.g. particle masses and missing transverse energy)
 - simple binary classification (e.g. signal vs. background)**Well established** check via redundancy or in signal-free control regions
- Classification of rare or so-far unobserved events
testing all kinds of modern deep NN approaches
Challenges: training on simulated samples only
→ propagation of simulation uncertainties through network
and reliable confidence regions for classifiers

Example: ML in Higgs to fermions

- Search for 1 out of 10^{10} :

$$\sigma_{\text{incl}}(pp) \approx 10^{11} \text{ pb}$$

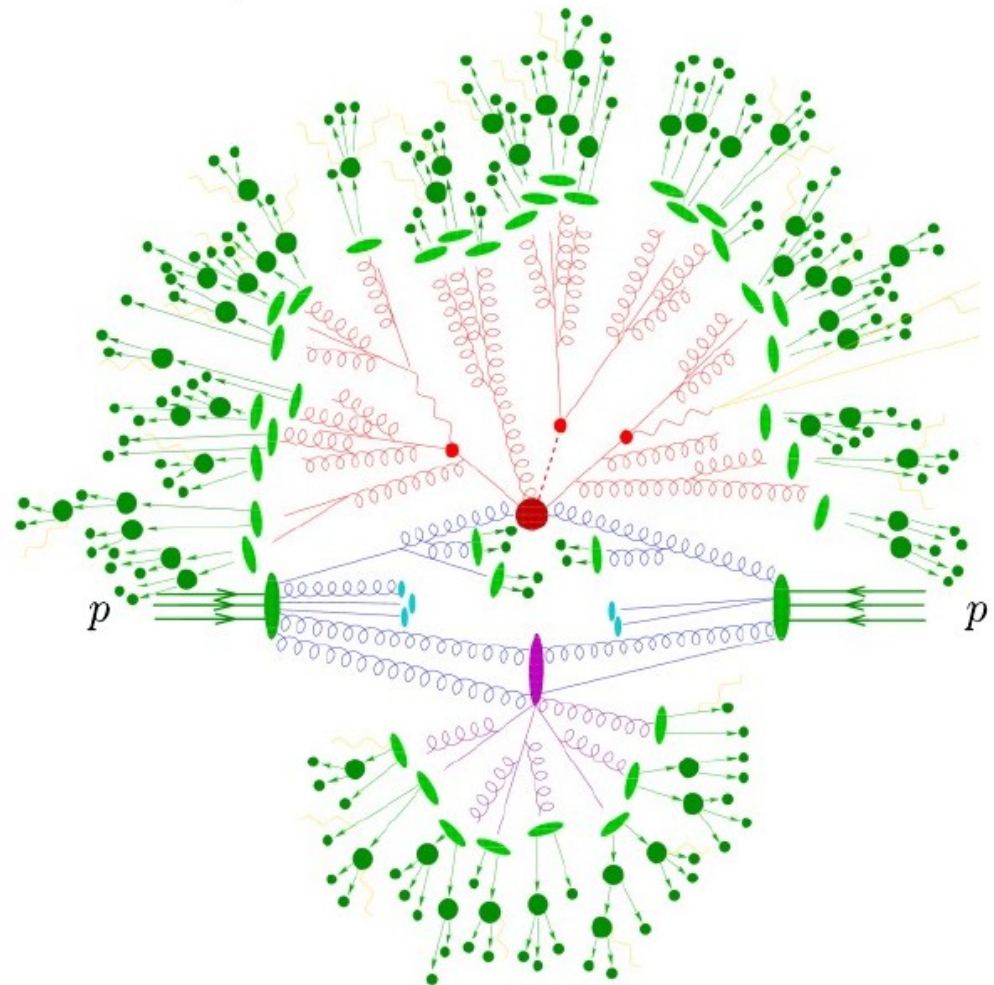
$$\sigma_{\text{incl}}(t\bar{t}) \approx 840 \text{ pb}$$

$$\sigma_{\text{incl}}(H) \approx 50 \text{ pb}$$

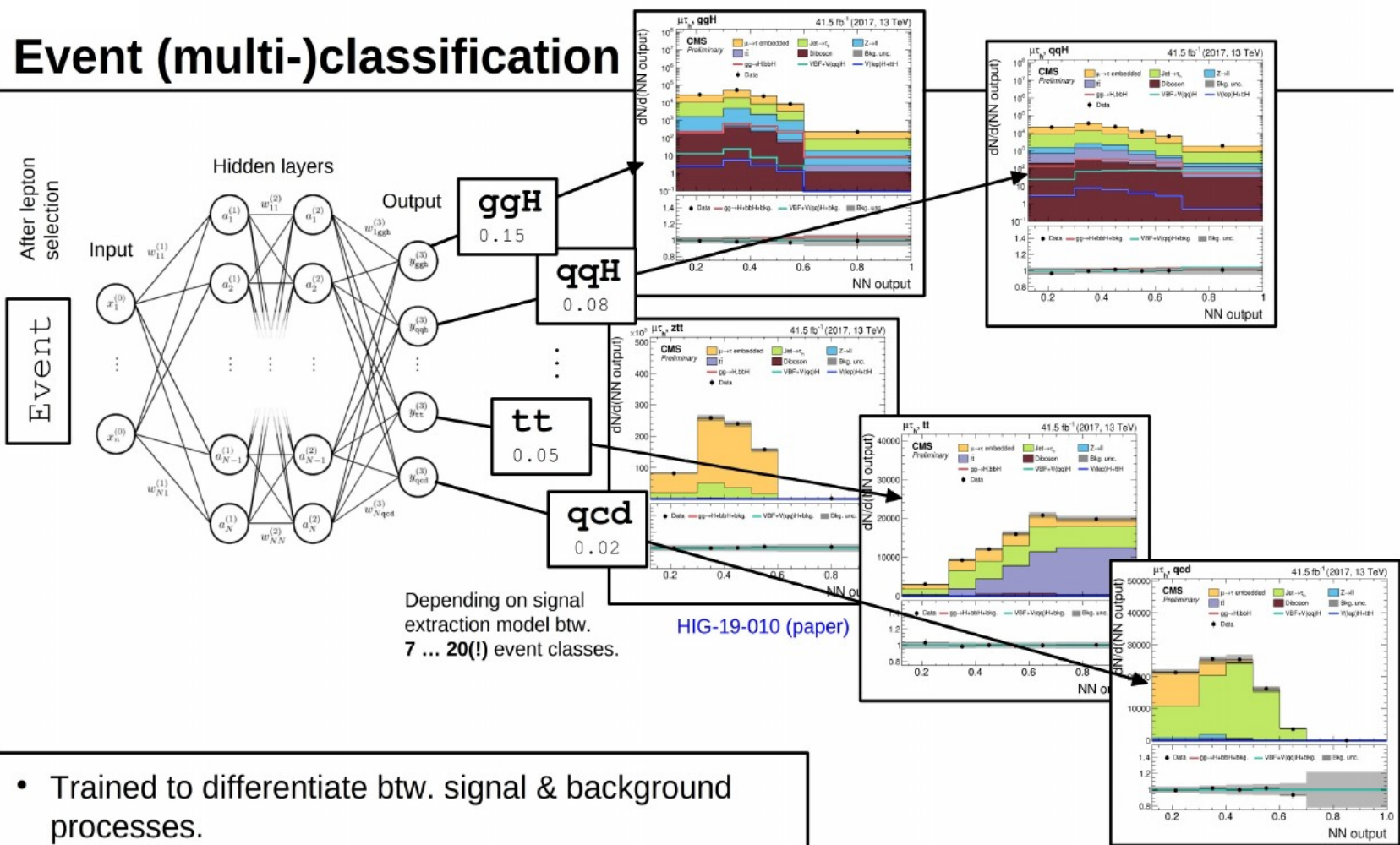


* for us finding the Higgs it was
48 years = 1,513,728,000 sec

- Complex environment full of *features*.
- Overlap of signal(s) & background(s) in *feature space*.



Event (multi-)classification



- Trained to differentiate btw. signal & background processes.
- Output $\rightarrow$ tuple of scores ($\sim$ Bayesian probabilities) for the event to belong to a given process.
- Highest score defines class the event is associated to.

NB: Corresponding to 72 ... 240 1D distributions as input to the stat. inference.

Analyses in Particle Physics without ML hardly possible !

Over many years we have made the effort to understand and control how NNs learn:

Using a **Taylor expansions of the NN output** function we can tell apart what are the most influential features for the NN output function. We can even tell apart when they are learned and from what data. This gives us many insights and strict control over what the NN is actually learning and finally doing.

Most (important) work is invested in:

- Selection of the input variables to the NN training.
- Scrutiny of the input data to the NN.
 - Once the input space is sufficiently under control the analysis is basically done.
-

List of Topics from recent theses

Bachelor-Arbeiten (mit ML-Thematik im Titel) der letzten zwei Jahre:

- Studien zu Graph Neural Networks in $tt + bb$ -Prozessen am CMS-Experiment (Halenke, Timo)
- Studien zu neuronalen Netzen im Rahmen einer leptonischen Mono-top-Analyse (Klingel, Alina)
- Classification of ttH Events using Lorentz Layer Networks (Schaub, Simeon David)
- Multiklassifikation mit Bayesian Neural Networks in der $ttH(bb)$ -Analyse (Guthmann, Dorian)
- Untersuchungen zur Domänenanpassung bei neuronalen Netzen im Rahmen einer $ttbar$ -Analyse mit offenen ATLAS-Daten des Jahres 2012 (Lippmann, Tim)
- Illustration of neural network learning with uncertainties (Christian Winter)
- Studien zu erweiterten Netzwerkarchitekturen im Rahmen einer $tt+X$ Analyse am CMS-Experiment (Lenschen, Paul Jakob)
- Weiterführende Studien zu Convolutional Neural Networks im Rahmen einer ttH -Analyse am CMS-Experiment (Vautz, Julia Caterina)
- Untersuchung von Optimierungsstrategien für Bayesian Neural Networks im Rahmen der $ttH(bb)$ -Analyse am CMS-Experiment am CERN (Cung, Yee-Ying Christina)
- Verbesserung der Rekonstruktion des Z-Bosons mit neuronalen Netzen im Kontext einer $ttbar+X$ -Analyse am CMS-Experiment (Hsu, Thomas)
- Rekonstruktion des Higgs-Bosons in assoziierter Produktion mit einem Top-Quark-Antiquark-Paar mithilfe Neuroner Netze bei CMS (Heim, Marlene)
- Illustration of the Neural Network Learning Process during Training (Heine, Greta Sophie)
- Untersuchung von Convolutional Neural Networks im Rahmen einer ttH -Analyse am CMS-Experiment (Seyffert, Yan)
- Untersuchung von Boosting-Algorithmen für Neuronale Netze im Rahmen einer ttH -Analyse am CMS-Experiment (Golks, Niklas)

List of Topics from recent theses (2)

Master-Arbeiten (mit ML-Thematik im Titel) der letzten drei Jahre:

- Deep Continuum Suppression with Predictive Uncertainties at the Belle II Experiment (Sowa, Lars)
- ANN-based Jet Assignment in $t\bar{t}+X$ Cross-section Measurements at the CMS Experiment (Armbruster, Lukas)
- Optimization of neural networks in the presence of systematic uncertainties (Risto, Gessi)
- Treating Uncertainties with Bayesian Neural Networks in the measurement of $t\bar{t}H(bb)$ production (Shadskiy, Nikita)
- Standard Model $H \rightarrow \tau\tau$ Analysis with a Neural Network Trained on a Mix of Simulation and Data Samples (Scham, Moritz)
- B-tagging with Deep Neural Networks (Boeckh, Tobias)
- Studies of the usage of neural networks in particle physics analyses (Jörger, Simon)
- Reducing background modeling uncertainties in the CMS analysis of the associated production of a Higgs boson and a top quark-antiquark pair with adversarial neural networks (Schindler, Jörg)
- Recoil Calibration as a Neural Network Task (Kopf, Tanja)
- Domain Adaptation Studies in Deep Neural Networks for Heavy-Flavor Jet Identification Algorithms with the CMS Experiment (Walter, David)
- Multivariate Klassifizierung mit künstlichen neuronalen Netzen zur Trennung von Unterprozessen der Top-Quark-Antiquark-Paar-Produktion (Hilser, Lukas)
- Continuum Suppression with Deep Learning techniques for the Belle II Experiment (Weyland, Dennis)
- A Novel Strategy for the Standard Model $H \rightarrow \tau\tau$ Analysis with Emphasis on Minimizing Systematic Uncertainties in Presence of Modern Multi-Variate Methods (Wunsch, Stefan)
- Multivariate Analysemethoden zur Vorhersage der fehlenden transversalen Energie am CMS Experiment (Zäh, Nicola)
- Studies of neural network architectures in the Search for NMSSM inspired di-Higgs events in the $\tau\tau+bb$ final state (Schmieder, Ralf)

BMBF-Ausschreibung vom 4.6.2021:

„Software und Algorithmen“

zur Erforschung von Universum und Materie (ErUM)

mit Schwerpunkt auf Künstliche Intelligenz und Maschinelles Lernen

*Die **Ziele der Fördermaßnahme** umfassen*

- die Vernetzung relevanter Akteure und Strukturen mit Blick auf die Digitalisierung,*
- den Ausbau von digitalen Kompetenzen in der Forschung und*
- die Stärkung des Transfers und der Kommunikation – innerhalb von ErUM und über ErUM hinaus –*

im Themenfeld „Software und Algorithmen“ des Aktionsplans ErUM-Data.

Durch die Fördermaßnahme sollen eine koordinierte Herangehensweise und eine breite Verständigung der verschiedenen Fachdisziplinen erreicht werden, sowie übergeordnete Lösungen ermöglicht werden.

Zwei Antragssizzen aus der Teilchenphysik am KIT eingereicht

- Confident Inference (**COIN**)
- MASchinelles Lernen und Neuronale Netze für EchtzeitTriggersysteme mit FPGAs an Großgeräten der Teilchenphysik (**MANTA**)

Antragsskizzen ErUM-Data

■ ErUM-Data Proposal: **Confident Inference (COIN)**

Projektpartner: TU Dortmund, FAU Erlangen, KIT, Univ. Hamburg,
Univ. Heidelberg, Univ. Göttingen, RWTH Aachen,

Inferenz: Rückgängigmachen von Störeffekten, die Information über Naturphänomene überdecken
(durch Messverfahren mathematisch/theoretische Näherungen).

In Teilchen- und Astroteilchenphysik mit Hilfe einer Kombination von Daten und Simulationen,
die auf fundamentaler Theorie basieren.

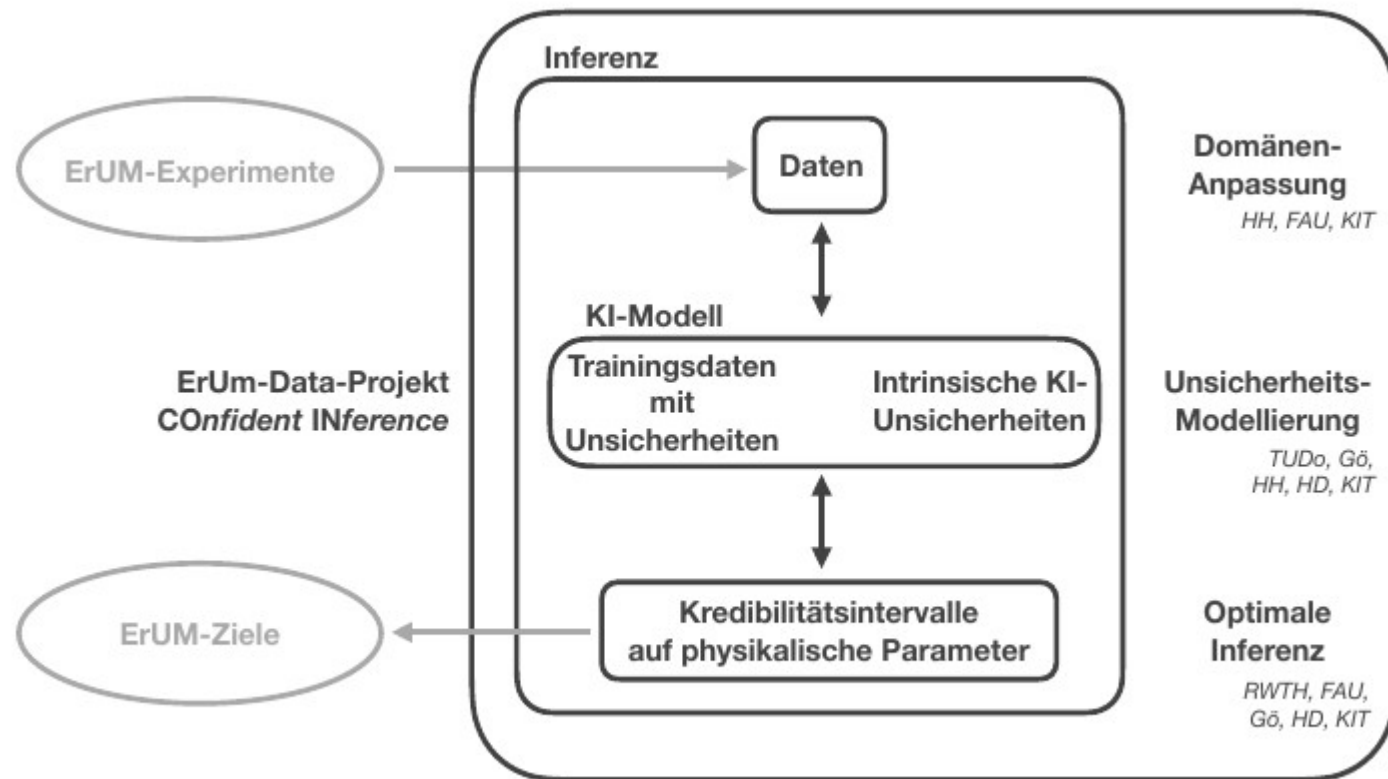
Erfordert das bestmögliche Verständnis aller statistischen und systematischen Störeffekte
(Simulations- und Theorieunsicherheiten)

bisher: behandelt durch Variationen der Modelle im Rahmen der Unsicherheiten.

neu: Anwendung moderner Methoden des maschinellen Lernens (ML, insb. DL):

- DomänenAnpassung:
 - Simulationsmodelle an Messdaten angleichen,
 - Unsicherheiten in den Trainingsdaten im Training berücksichtigen, um Sensitivität zu reduzieren
- Unsicherheits-Modellierung:
generische Einfügung von Unsicherheiten in NNs (z.B. Bayesianische NNs (BNNs), Generative Adversarial Networks (GANs))
- Optimale Inferenz:
 - hohe Zahl von Parametern erlaubt weitgehende Erfassung der Phasenräume;
 - invertierbare NNs (Normalizing Flows) und deren Varianten (z.B. conditional Invertible NNs, cINNs) mit Erhaltung der Wahrscheinlichkeiten
 - Berücksichtigung systematischer Unsicherheiten in der Kostenfunktion des NN

Antragsskizzen ErUM-Data Coin (2)



■ ErUM-Data Proposal:

MAschinelles Lernen und Neuronale Netze für EchtzeitTriggersysteme mit FPGAs an Großgeräten der Teilchenphysik (MANTA)

Koordination KIT mit Partnern Univ. Hamburg, Univ. Mainz, Univ. Dresden, TU München, Univ. Heidelberg sowie assoziierte Partner

Verarbeitung von Daten in Echtzeit ist bei Projekten der Teilchenphysik essentiell, um komplexe Entscheidungen auf Trigger-Niveau treffen zu können.

Für zukünftige Experimente können ML-Methoden auf FPGAs ausgeführt werden.

Dazu müssen innovative Anpassungen von Triggeralgorithmen und an die Hardware angepasste Werkzeuge entwickelt werden:

- Algorithmenentwicklung
- System-Konfiguration
- Entwicklung optimierter Werkzeuge und FPGA Firmware-Modle
- Tests von Prototypen am Experiment und Integrationsaspekte

Antragsskizzen ErUM-Data MANTA (2)



Abbildung 1: Überblick über die Herangehensweise im Projekt von Echtzeit-Trigger Ansätzen zu geplanten initialen Trigger-Systemen für die jeweiligen Experimente.

Themen von Besonderem Interesse

- Strict Interpretation of the NN output as a confidence interval of the “true value” and methods to mitigate systematic effects in training data (part of “COIN”)
 - Bayesian Networks as a promising Ansatz
 - Inclusion of systematics in cost function
 - detailed monitoring of the learning process
 - Generative Adversarial Network (GAN) to reduce sensitivity to input variables affected by systematics
- Implementaion of ML on FPGA (part of “MANTA”)
- Gain experience with Graph Neural Networks
- Application and development of technologies to use more complex NN topologies for multi-classification task, e.g. Parametric NNs or alternative, parallel sub-NNs for multi-input, multi-output networks (ideal for deployment on GPU clusters)
- Explore methods of “anomaly detection” to search for new effects (i.e. “new physics”) in data, e.g. Variational Recurrent NN, GAN with Auto-Encoder, ...
(more ideas and guicance needed, see <https://arxiv.org/abs/2101.08320>)