ACAT'19, Track 2 summary

Data Analysis - Algorithms and Tools

Sofia Vallecorsa, Michael Kagan, Andrey Ustyuzhanin

2019-03-15, Saas-Fee, Switzerland

Track2: Data Analysis - Algorithms and Tools

- 36 talks
- 39 posters

Track sub-sections

Tracking, online (Monday)	Simulation and Event reconstruction (Tuesday)
Statistics, Uncertainty (Wednesday)	New Physics, Cosmic (Thursday)

1. Tracking and online event processing

- Tracking represents one of the most time-consuming parts in reconstruction
- Very rich session:
  - **Parallelized Kalman-Filter-Based Reconstruction of Particle Tracks on Many-Core Architectures with the CMS Detector**
  - **HEP.TrkX Charged Particle Tracking using Graph Neural Networks**
  - **hls4ml: deploying deep learning on FPGAs for trigger and data acquisition**
  - **ConformalTracking: a geometry agnostic tracking library**
  - **A 3D Track Finder for the Belle II CDC L1 Trigger**
  - **Charged Particle Tracking as a QUBO problem solved with Quantum Annealing-Inspired Optimization**
  - **Real-time reconstruction of long-lived particles at LHCb using FPGAs**
  - **A hybrid deep learning approach to vertexing**
  - **Belle2VR – An Interactive Virtual Reality Visualization of GEANT4 Event Histories**

Deep Learning on FPGAs for Trigger and Data Acquisition

J. Ngadiuba

Efficient NN design for FPGAs

FPGAs provide huge flexibility

Performance depends on how well you take advantage of this

Constraints:

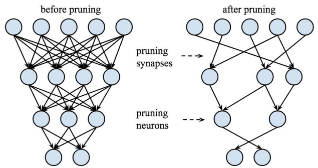
- Input bandwidth
- FPGA resources
- Latency

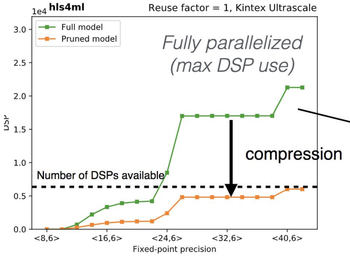
With `hls4ml` package we have studied/optimized the FPGA design through:

- **compression:** reduce number of synapses or neurons — NN TRAINING
- **quantization:** reduces the precision of the calculations (inputs, weights, biases)
- **parallelization**: tune how much to parallelize to make the inference faster/slower versus FPGA resources — FPGA PROJECT DESIGNING

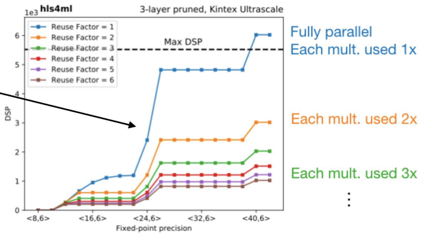
Resource usage: DSPs

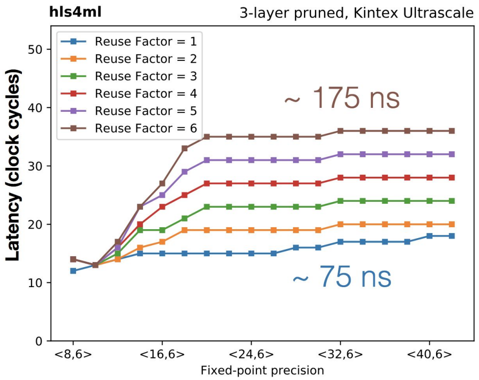
- DSPs (used for multiplication) are often limiting resource
  - maximum use when fully parallelized
  - DSPs have a max size for input (e.g. 27x18 bits), so number of DSPs per multiplication changes with precision





70% compression ~ 70% fewer DSPs





Longer latency

Each mult. used 6x

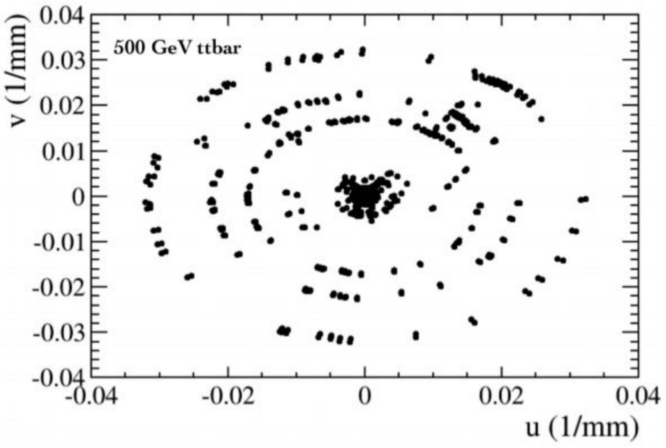
Each mult. used 3x

Fully parallel
Each mult. used 1x

More resources

Tracking-1

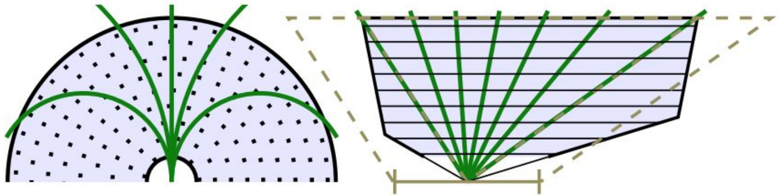
geometry agnostic tracking library by **Marko Petric**: Conformal map applies a geometry transform that maps circles in the x,y plane passing through the origin point into straight lines in the u,v plane. Designed for CLIC, but also works with different detectors e.g. FCCee CLD



[Github](#)

3D Track Finder for the Belle II by **Sebastian Skambraks**:

- The novel triggering techniques copes with the severe background conditions coming along with the upgrade of the instantaneous luminosity by a factor of 40
- precise drift-time information of the central drift chamber + neural network trigger
- 3D finder (single hit representations in the Hough plane are trained using Monte Carlo)
- This 3D finder enables an improvement of the track finding efficiency by including the stereo sense wires as input.

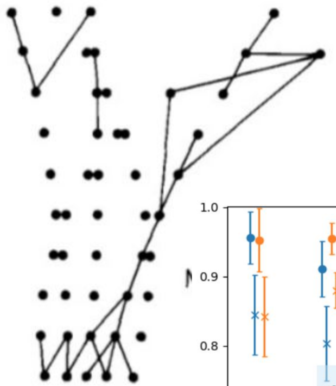


Sectors in p_T (left) and in ϑ (right).

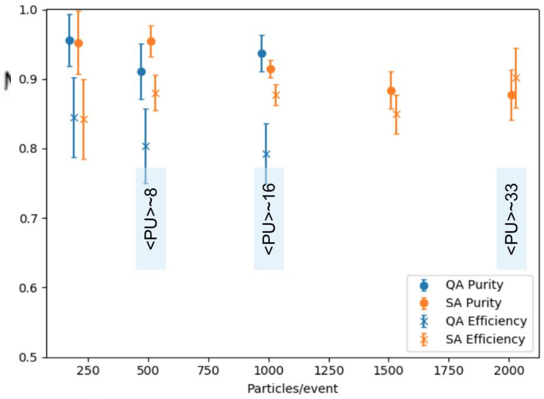
Tracking-2

Quantum Annealing tracking by **J-R. Vlimant**:

Quadratic Unconstrained Binary Optimization (QUBO) can be mapped to an Ising Hamiltonian with change of variable {0,1}↔{-1,1}

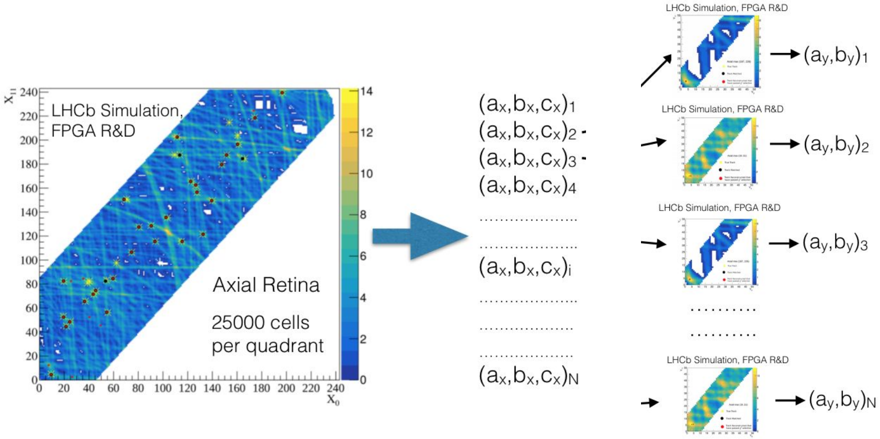


$$E = \boxed{-\frac{1}{2}\sum_{i,j} w_{ij} s_i s_j}$$



Real-time reconstruction of long-lived particles at LHCb using FPGAs by **Michael Morello**.

Study of the performances of a future innovative real-time tracking system based on FPGAs, R&D developed in the context of the LHCb Upgrade Ib (LHC Run 4) dedicated to reconstructing particle downstream of the magnet in the forward tracking detector.



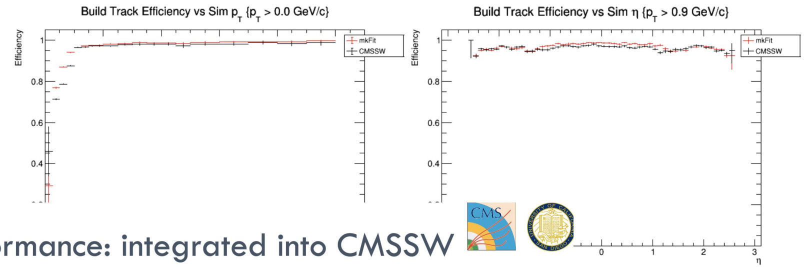
Parallelised Kalman Filter

mkFit efficiency: mkFit validation

14 Mario Masciovecchio (UCSD), 11 March 2019

- $t\bar{t}$ (PU=70)
  - Algorithm-level efficiency, for long tracks

→ mkFit is at least as efficient as **CMSSW**



Parallelized Kalman-Filter-based Reconstruction of Particle Tracks on Many-Core Architectures with the CMS detector

11 March 2019

G. Cerati[4], P. Elmer[3], B. Gravelle[5], M. Kortelainen[4], S. Krutelyov[1], S. Lantz[2], <u>**M. Masciovecchio**</u>[1], K. McDermott[2], B. Norris[5], A. Reinsvold Hall[4], D. Riley[2], M. Tadel[1], P. Wittich[2], F. Würthwein[1], A. Yagil[1]

1. UCSD 2. Cornell 3. Princeton 4. FNAL 5. Oregon

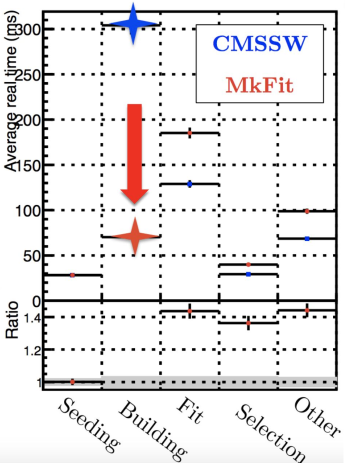
ACAT 2019, Saas Fee (11-15 March 2019)

Time performance: integrated into CMSSW

17 Mario Masciovecchio (UCSD), 11 March 2019

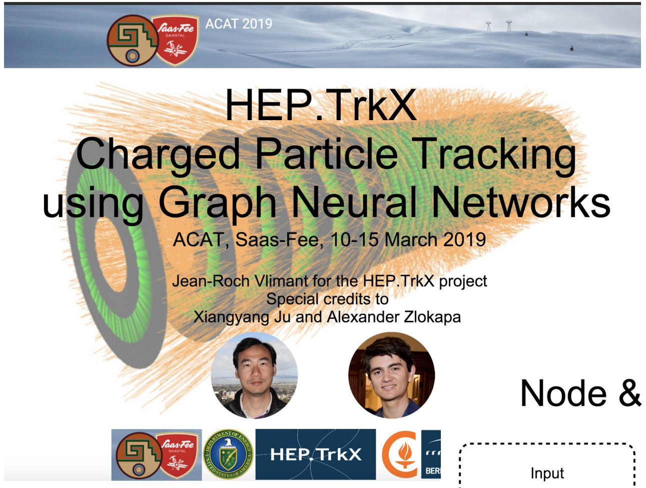
- Time performance of mkFit when integrated into CMSSW, vs. CMSSW
  - For corresponding tracking step
  - N(threads) =1; N(streams) =1
  - Using $t\bar{t}$ (PU=50)
  - Test on SKL-SP Gold
  - mkFit compiled with AVX512

→ **Track building** is **4.4x faster** (mkFit)
  - mkFit time currently accounts for data-format conversions
    - **~40%** ⇒ Actually ≳ **7x** faster
  - mkFit gets faster with # threads
  - ➢ Can only improve from here



[slides]

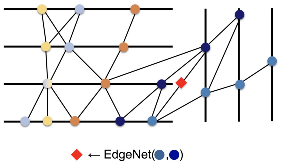
Tracking with Graph NN



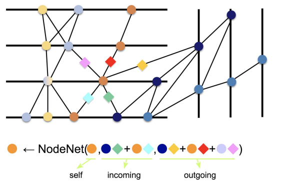
Neural Networks

- **Input Network**
  - Transforms from hit features (r,φ , z) to the node latent representation (N for 8 to 128)
  - Dense : 3→...→N
- **Edge Network**
  - Predicts an edge weight from the node latent representation at both ends
  - Dense : N+N→...→1
- **Node Network**
  - Predicts a node latent representation from the current node representation, weighted sum of node latent representation from incoming edge, and weighted sum
  - Dense : N+N+N→...→N

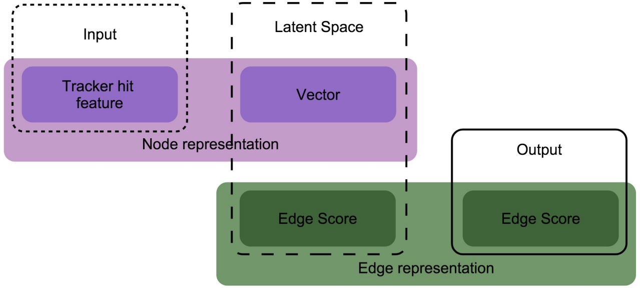
Edge Network



Node Network

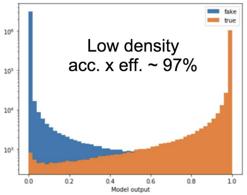


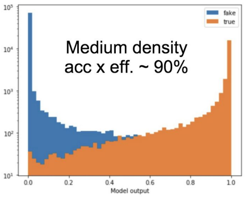
Node & Edge Representations

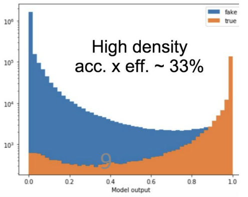


Latent edge representation taken to be the classification score instead of some latent vector representation

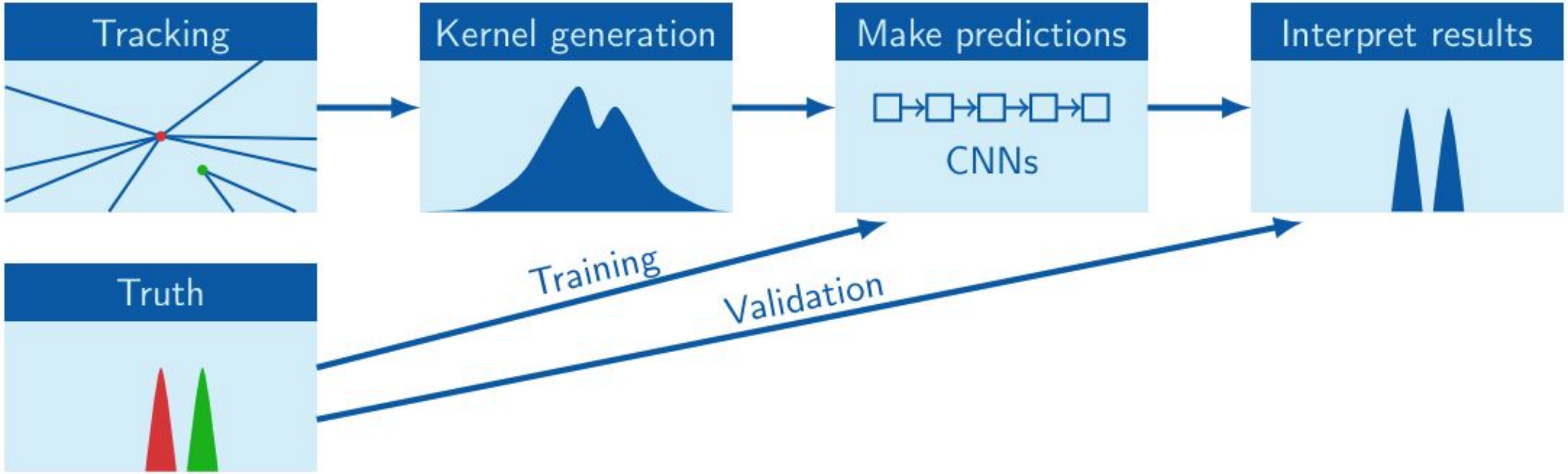
Performance





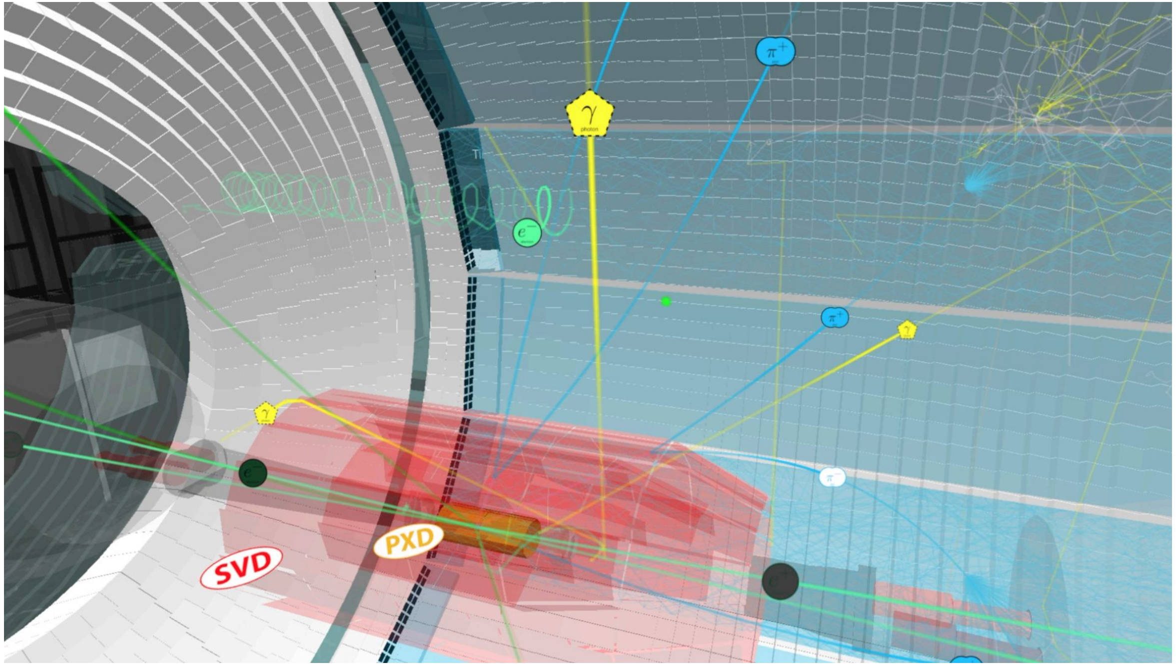


Hybrid Deep learning vertexing by H. Schreiner



- Proof-of-Principle established: 5-layer CNN finds primary vertices with efficiencies and false positive rates similar to traditional algorithms.
- Efficiency is tunable; increasing the efficiency also increases the false positive rate due to Asymmetric cost function

Belle2VR: An interactive virtual reality visualization of GEANT4 event histories (Leo Piilonen)



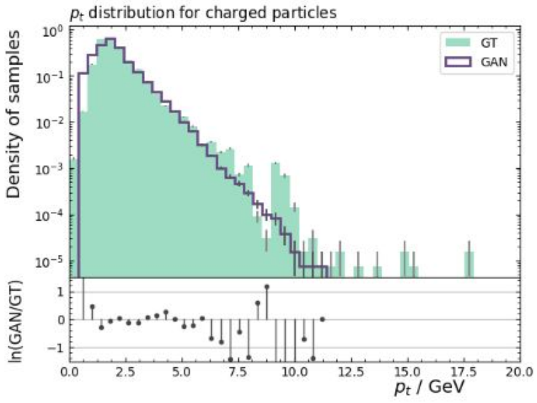
https://vimeo.com/220004044

2. Simulation and Event Reconstruction

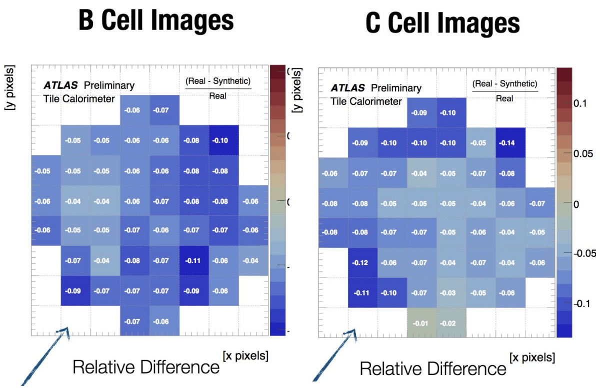
- **Recurrent GANs for particle-based simulation at the LHC**
- **Fast Data-Driven simulation of Cherenkov Detectors Using Generative Adversarial Networks.**
- **Physics inspired feature engineering with Lorentz Boost Networks**
- **Reinforced Jet Grooming**
- **Constructing mass-decorrelated hadronic decay taggers in ATLAS**
- **Towards the Increase in Granularity for the Main Hadronic ATLAS Calorimeter: Exploiting Deep Learning Methods**
- **Energy reconstruction of the ATLAS Tile Calorimeter under high pile-up conditions using the Wiener filter**
- **Electromagnetic calorimeter reconstruction in Belle II**
- **Selective background Monte Carlo simulation at Belle II**

GAN-based Simulation

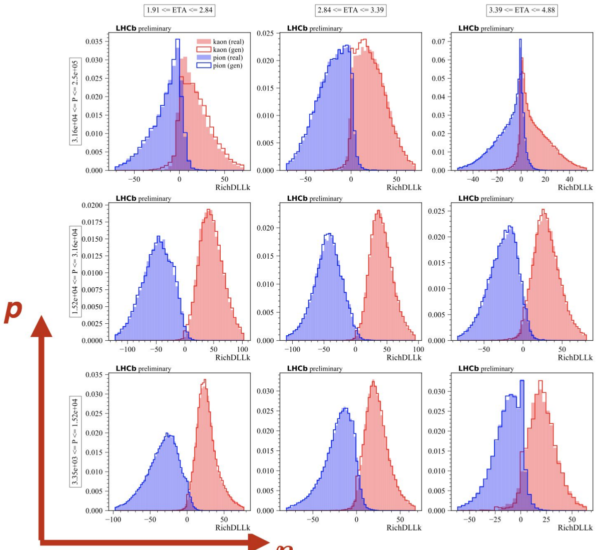
T. Nguyen showed how to use conditional GANs to generate list of particles mimicking particle-flow candidates. Can be used directly by reconstruction algorithms and to generate pile-up. Different layers are pre-trained, then inserted into the GAN generator.



DCGAN were also used **(P. Gaspar**) to reproduce showers within the high granularity upgrade for the ATLAS calorimeter. In order to insure high quality images, GAN output is run through an additional CNN classifier. Results show a 11-14% agreement with Geant4 simulation



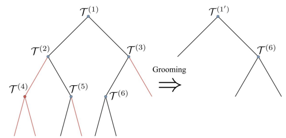
Another GAN application was presented by **A. Maevskiy** to the simulation of the LHCb RICH detector. Cranmer distance is used to train a fully connected layers GAN and stabilize the gradients

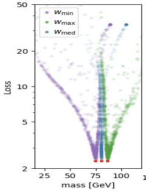


Reconstruction: Machine Learning and Jets

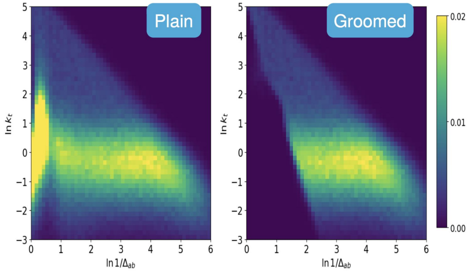
Large array of literature on ML for jets, and we saw new tools developed to include physics knowledge, constrain classifiers, or learn to groom them!

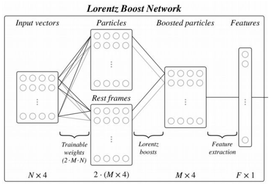
F. Dreyer introduced a grooming procedure for iteratively removing soft-wide angle radiation from jet, can be learned using a reinforcement learning paradigm

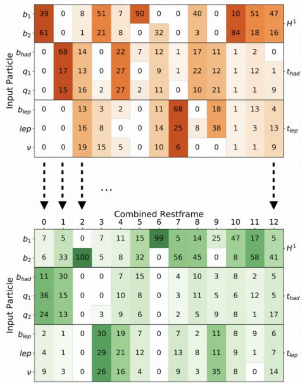




$$R(m, a_t, \Delta, z) = R_M(m) + \frac{1}{N_{\mathrm{SD}}} R_{\mathrm{SD}}(a_t, \Delta, z)$$

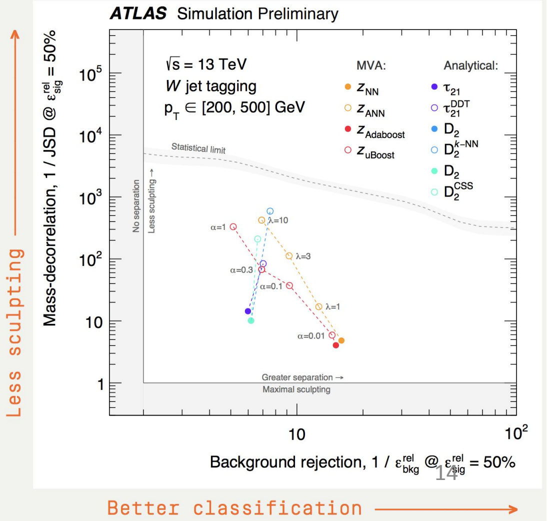






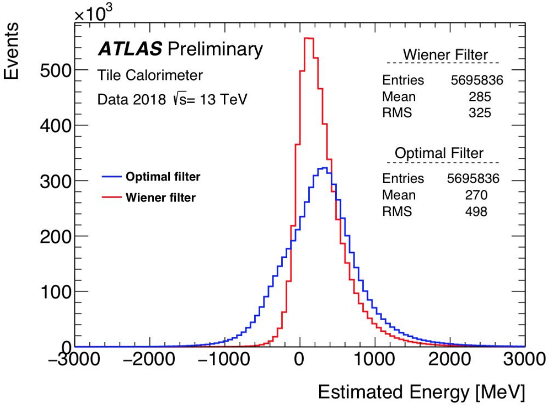
Y. Rath discussed Lorentz Boost Networks, physics inspired architectures Able to autonomously create characteristic features from particle four momenta with physics interpretability

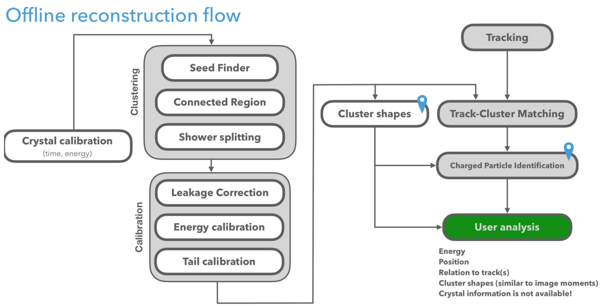
A. Sogaard compared several methods for constraining sculpting effects on data from models, and examined the performance vs sculpting tradeoff



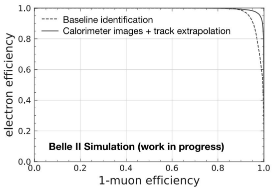
Energy Reconstruction: Calorimeters

D. Gonçalves presents the Wiener Filter - algorithm for energy estimation in the ATLAS Tile Calorimeter under high pile-up conditions.

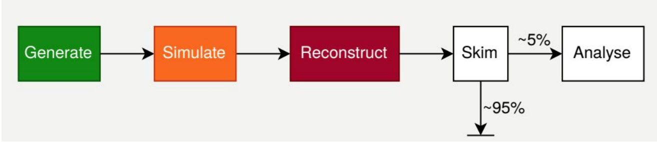


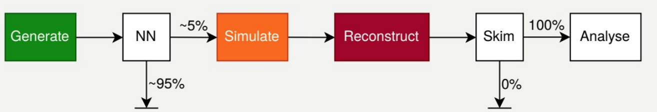


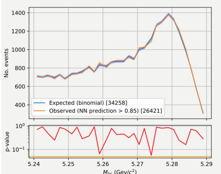
S. Ferber discussed the Belle II ML offline reconstruction algorithm tested on the data taken in 2018. Demonstrated improvements in energy and position reconstruction.



M. Ritter presented method of predicting in the early stages of the simulation process the likelihood of relevancy of an individual event to the target study using convolutional neural networks.







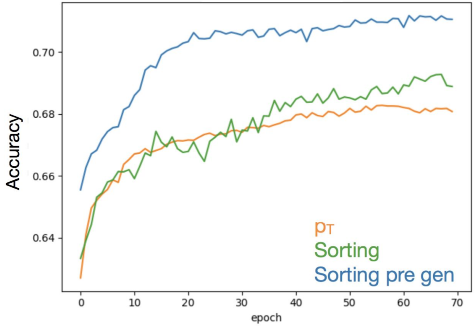
3. Statistics and Uncertainty

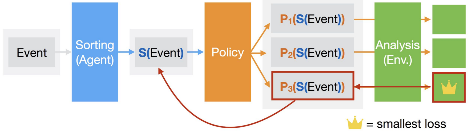
- **Incorporation of Systematic Uncertainties in the Training of Multivariate Methods**
- **Global fits of BSM physics models with Gambit**
- **INFERNO: Inference-Aware Neural Optimisation**
- **Uncertainty reduction by gradient descent**
- **Full Event Interpretation at Belle II**
- **Adversarial Neural Network-based data-simulation corrections for jet-tagging at CMS**
- **Reinforced Sorting Networks for Particle Physics Analyses**
- **Variational Autoencoders for New Physics Mining at the Large Hadron Collider**
- **Excursion Set Estimation using Sequential Entropy Reduction for Efficient Searches for New Physics at the LHC**
- **Machine Learning on sWeighted data**
- **Variational Dropout Sparsification for Particle Identification speed-up**

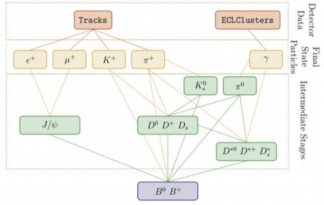
Analyzing Whole Events

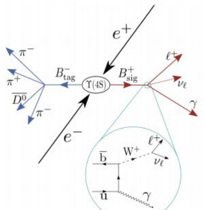
New mechanism explored for analyzing events wholistically!

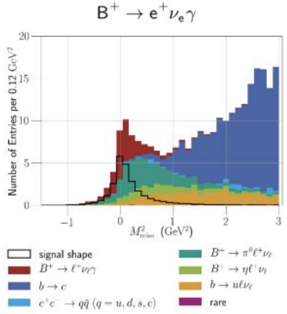
Which reconstructed objects should be assigned to which truth-level object for event reconstruction? ***D. Noll*** discussed doing this with reinforcement learning for sorting!





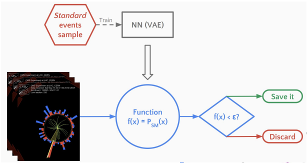


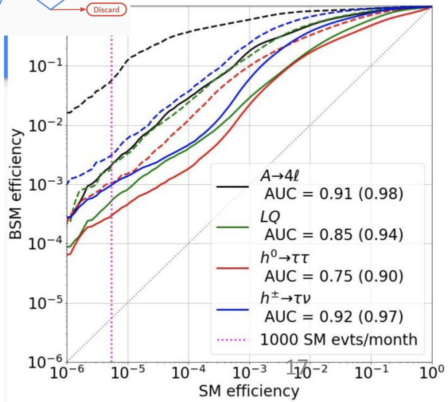




W. Sutcliffe utilized a group of O(200) decay channel classifiers to reconstruct O(1000) decay chains for tag-side B-reconstruction at Belle II. Saw factor of 3 gains in efficiency when applied to Belle Data!

O. Cerri explored how the latent space of a variational autoencoder learned to reconstruct the SM may help identify anomalous events, as a model-independent trigger

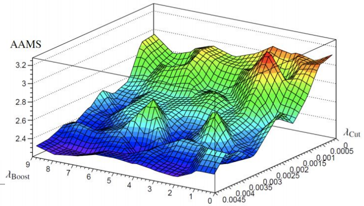


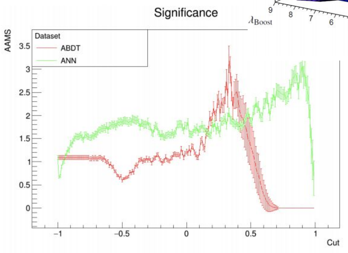


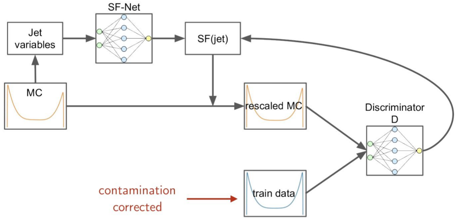
Systematic Uncertainties

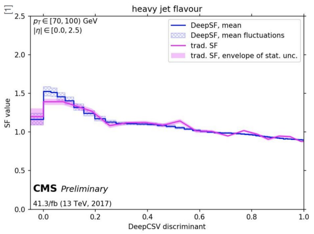
How to deal with systematic uncertainties (i.e. known unknowns about the data) when building a model?

For BDTs, ***T. Alef*** showed how to augment the AdaBoost algorithm to try to enforce invariance to data shifts to systematics at training time



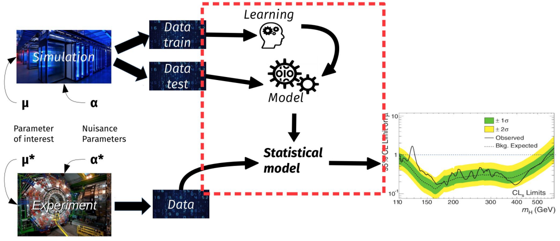


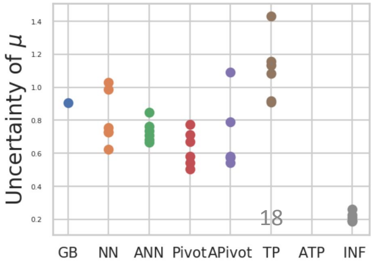




B. Fischer showed how adversarial technique, used in the past to enforce invariance to systematics, can be used to derive correction factors for simulations with DeepSF

V. Estrade compared techniques, including Adversarial NNs, gradient constraints, and INFERNO to study sensitivity to systematics in parameter estimation

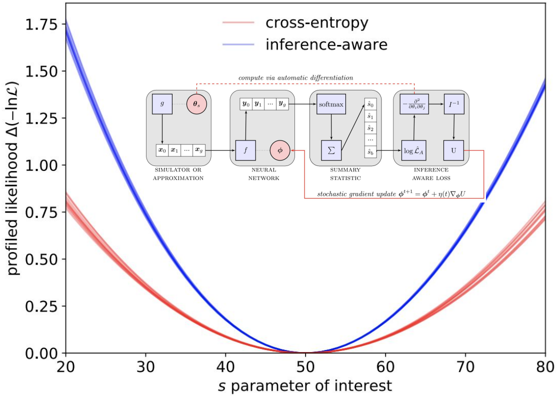


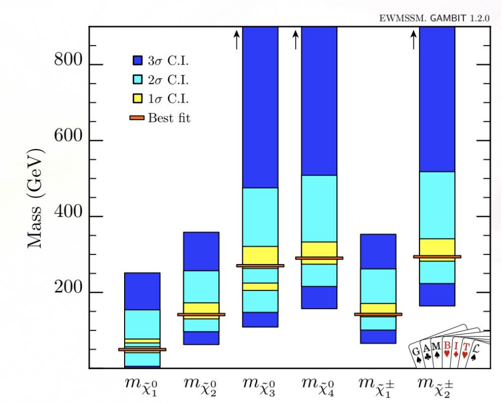


Parameter Inference

How best to extract a parameter of interest or determining a confidence interval?

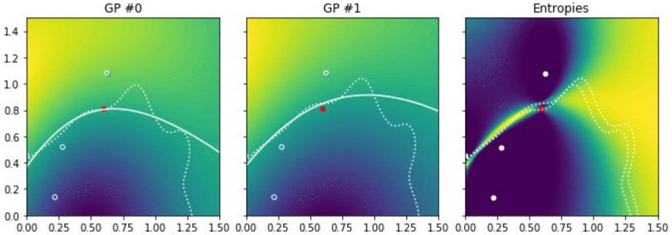
INFERNO, presented by ***P. de Castro***, aims to learn non-linear summary statistics by directly minimizing an approximation of the expected profiled (or marginalised) interval width accounting for the effect of nuisance parameters



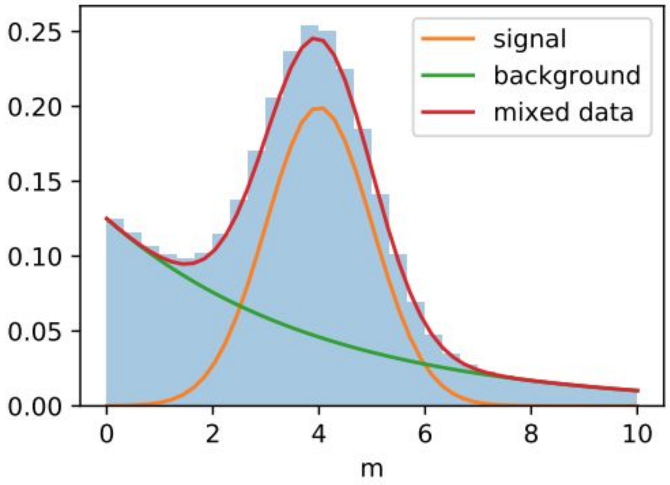


Gambit, presented by ***A. Buckley***, performs global BSM fits to a wide array of analyses, including Collider, DM, Flavour Physics, and Precision EW data.

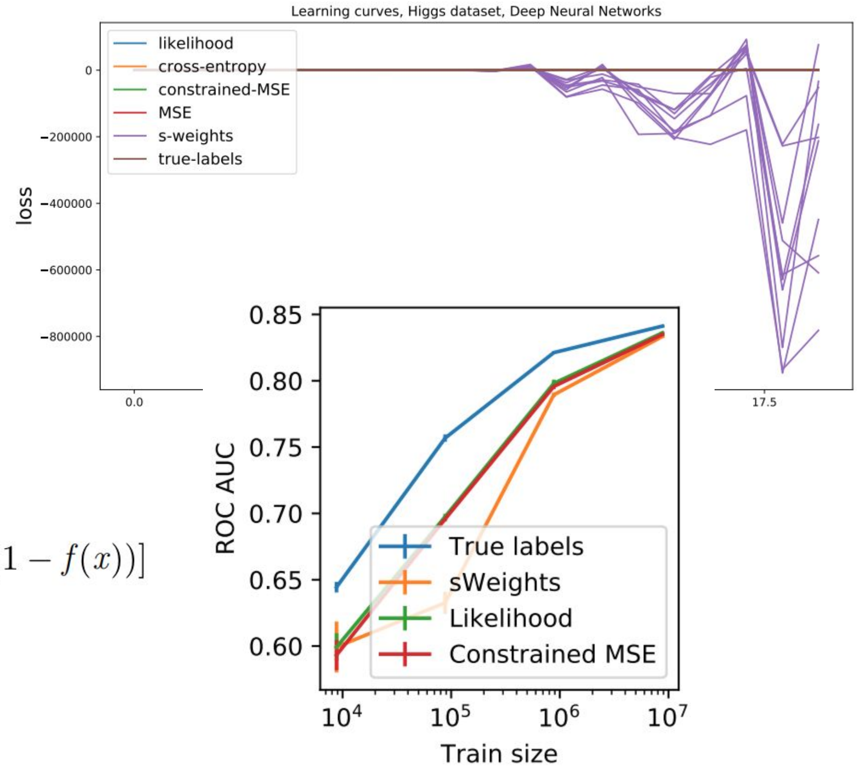
L. Heinrich used active learning to determine where to best sample new theory points for finding limit contours, based on looking for points that will maximize the expected information gain



Machine Learning on sWeighted data, N. Kazeev



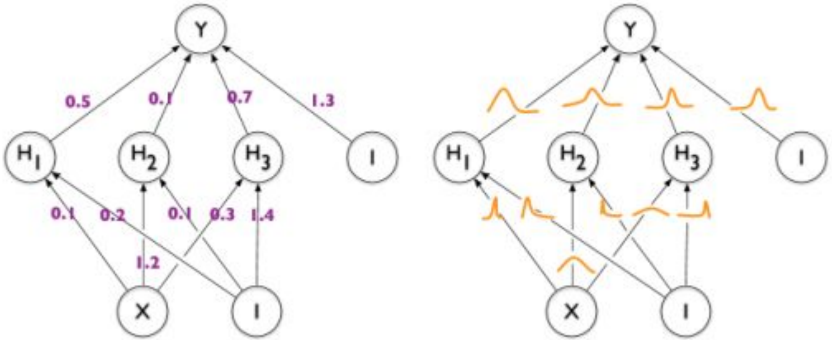
$$-\log\left[p\left(\text{signal}|m\right)\cdot f(x)+p\left(\text{background}|m\right)\cdot\left(1-f(x)\right)\right]$$



Variational Dropout Sparsification for Particle Identification speed-up

A.Ryzhikov suggests using Bayesian NN that gives hints which NN weights could be removed => inference speed increased.

Significant improvement for LHCb PID!



Method	# Neurons	Electron	Ghost	Kaon	Muon	Pion	Proton	Speed-Up
6xDNN	45-48	0.9855	0.9485	0.9148	0.9844	0.9346	0.9178	x1
1xDNN	150	0.9863	0.9570	0.9145	0.9889	0.9463	0.9167	x1
Ternary	Auto	0.9843	0.9435	0.9154	0.9834	0.9352	0.9110	x5
1xDNN	30	0.9871	0.9557	0.9158	0.9893	0.9427	0.9125	x5
BDNN	Auto	**0.9881**	0.9548	**0.9244**	**0.9896**	**0.9509**	**0.9228**	**x16**

github

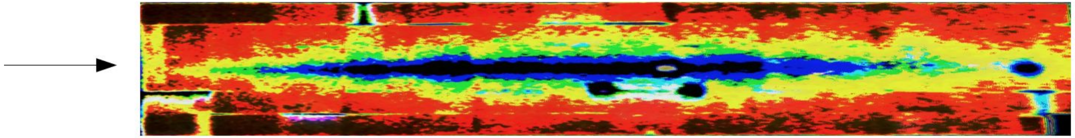
4. New Physics, Cosmic

- **Weak signal extraction using matrix decomposition**
- **Particle Identification in PICO Using Semi-supervised Learning**
- **Accelerating dark matter search in emulsion SHiP detector by Deep Learning**
- **Electromagnetic-shower generation with Graphical GANs**
- **Submanifold Sparse Convolutional Networks for Sparse, Locally Dense Particle Image Analysis**
- **Air shower reconstruction with hexagonal convolutional neural networks**
- **Deep Learning based Algorithms in Astroparticle Physics**

Weak signal extraction using matrix decomposition

Steven Prohira

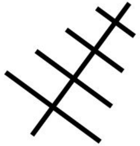
shower created in the target (HDPE)



(actual beam profile)

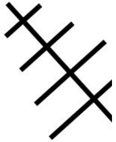
10 GeV electron beam, N=10^9

incident radio is reflected

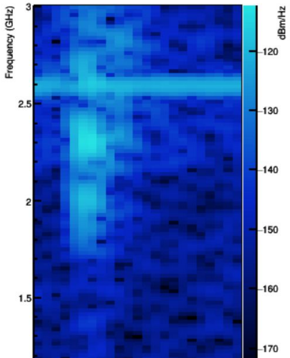


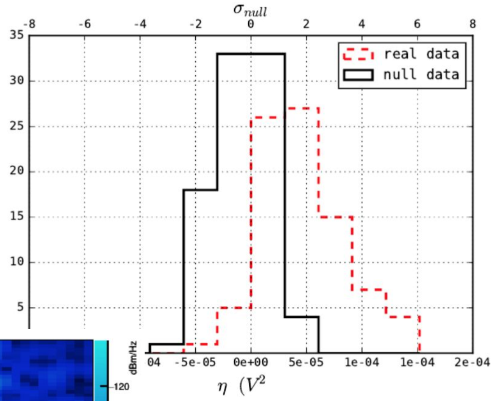
transmitter

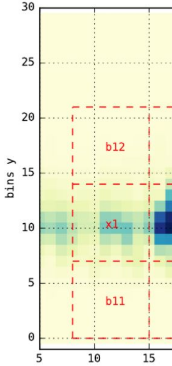
Towards detecting ultra-high-energy neutrinos with radar



receiver







blished a signal excess at 2.4 sigma via sideband sul

slides

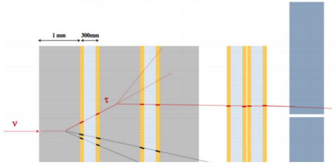
Dark matter search experiments

Search for WIMP interactions requires a bit of ML

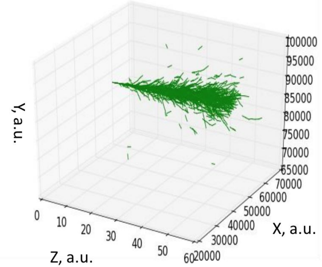
For the PICO experiment **G. Cao** presented comparison of traditional approach - frequency analysis in Fourier space with semi-supervised discriminator.

S. Shirobokov presented ML-based search for SNL traces in SHiP experiment

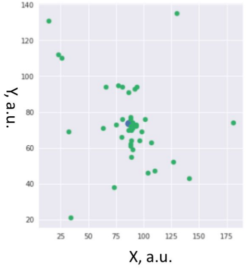
iSHiP structure

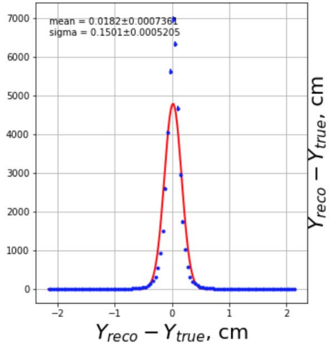


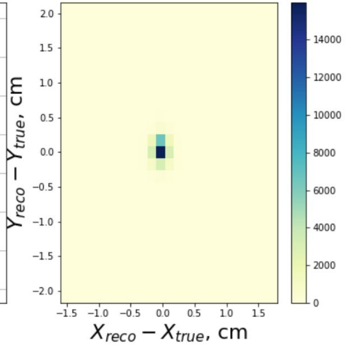
View of a shower in emulsion



View of a shower in TT



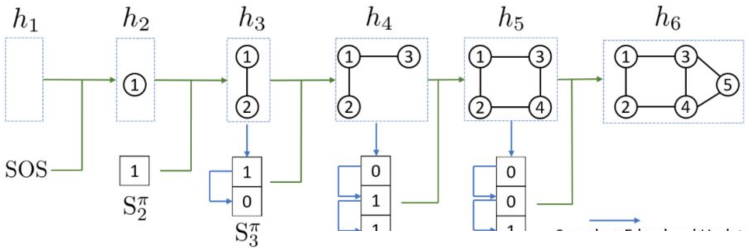


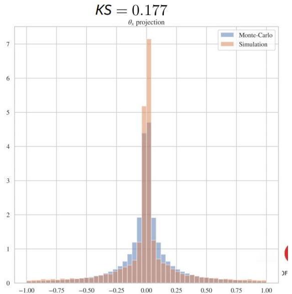


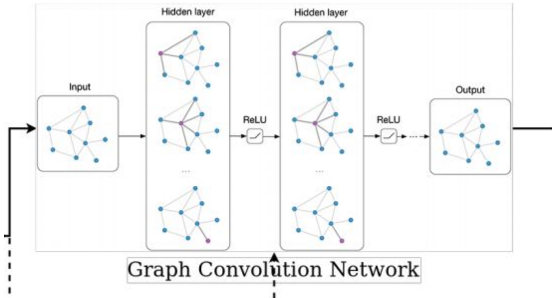
Electromagnetic showers

From DM to cosmic

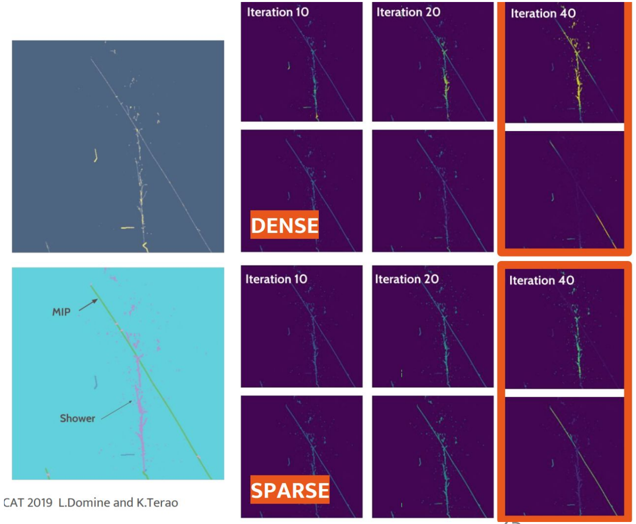
V. Belavin presented Graphical generative models - dual generation of Graph and signal within it for SHiP experiment.







L.Domine presented Sparse convolution-based approach for MicroBooNE (LArTPC) experiment. Trains faster, requires way less memory.



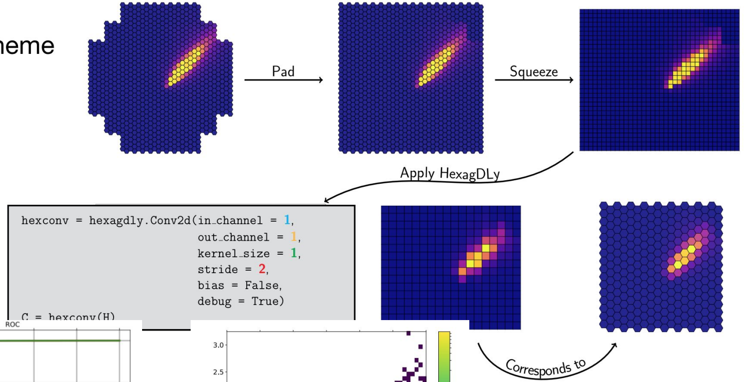
CAT 2019 L.Domine and K.Terao

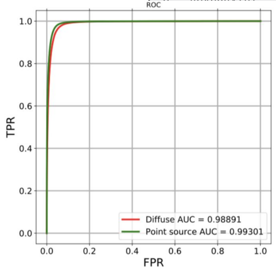
https://osf.io/9b3cv/

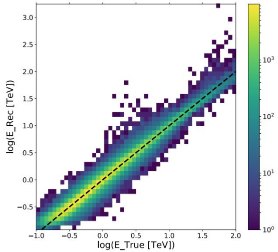
Cosmic particles meet DL

C. Steppa presented the way to adapt convolutional layers to hexagonal sensors of HESS and CTA telescopes for air showers reconstruction.

- Addressing scheme
- Kernels
- Convolution
- **Example**

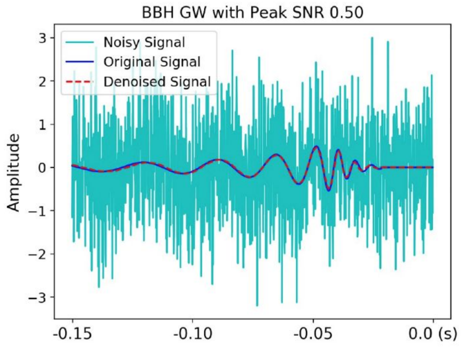


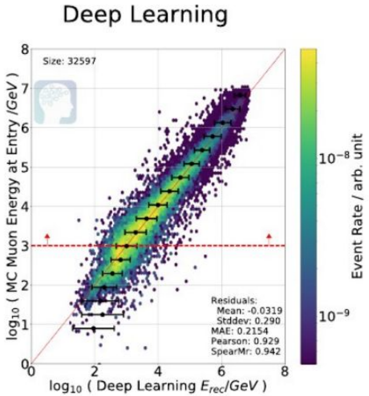




J.Glombitza presented fireworks of ML-based techniques for variety of astrophysics experiments:

- Neutrino reconstruction for Ice Cube
- Reconstruct binary black hole signal with denoising autoencoders
- LArTPC images segmentation for MicroBooNE
- Simulation Refinement





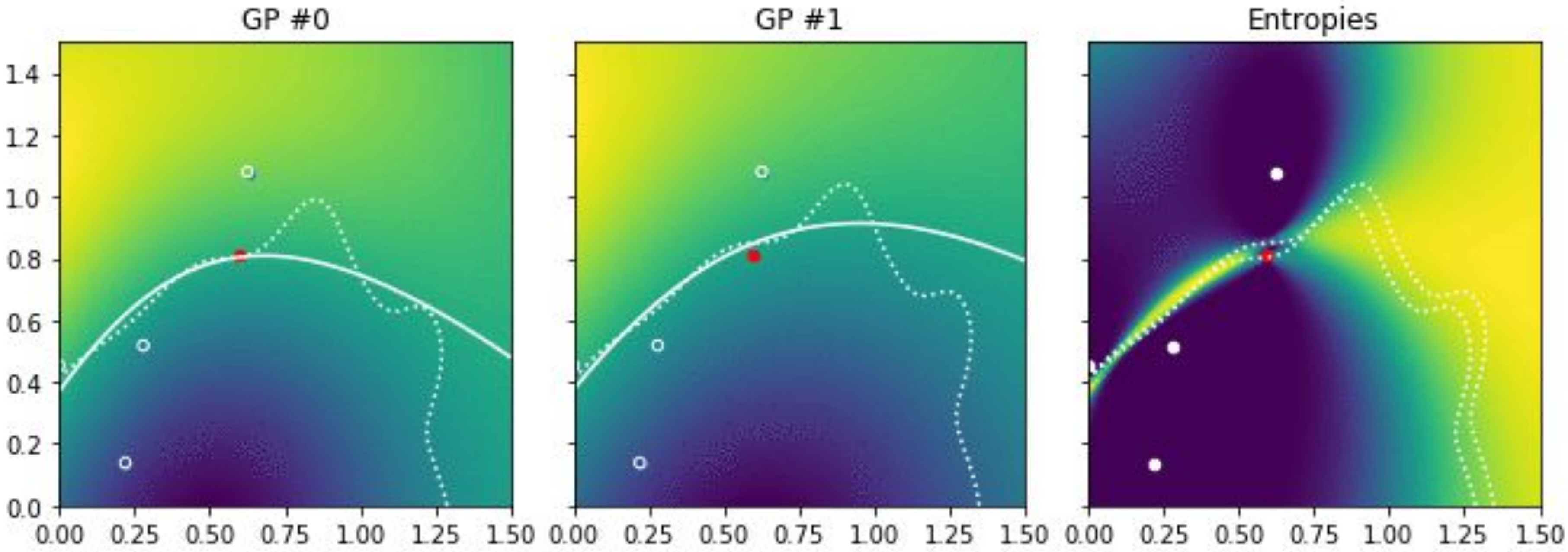
Summary

- Enlightening session and consistent representation of the current state of the art! (ML at HEP)
- Good focus on
  - Model robustness (uncertainty estimation)
  - Wholistic approach for event analysis
- Bayesian approach is heavily underrepresented wrt Frequentist (GAN)
- ML/DL Panel to be transcribed and made available
- **Thanks to all the presenters and attendees!**
- Apologies for all the work that could not be properly mentioned in a 20-minutes summary

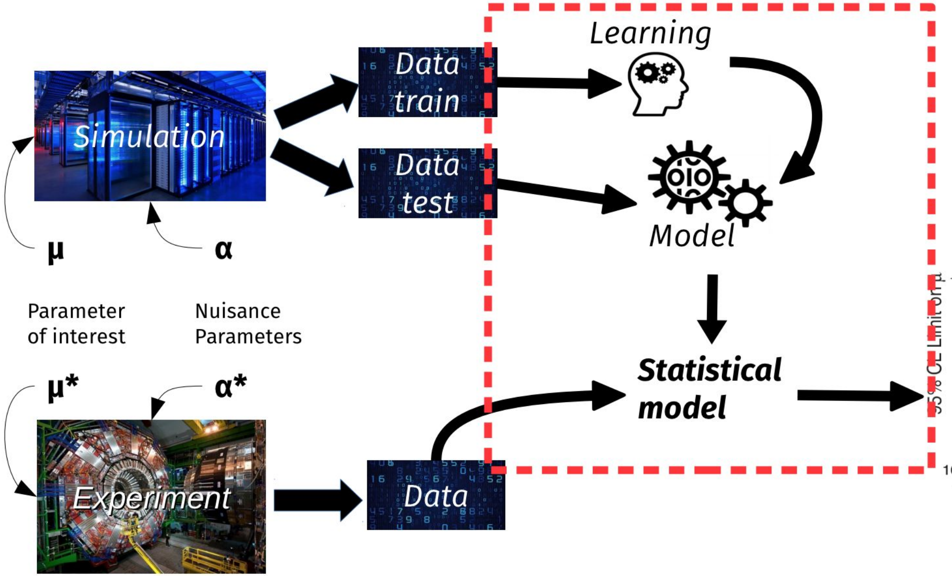
Backup

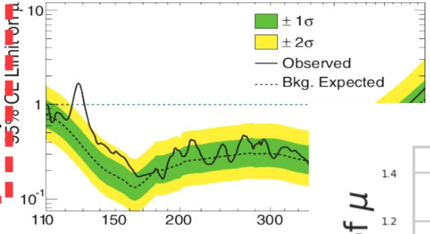
Active Learning for Excursion Set Estimation

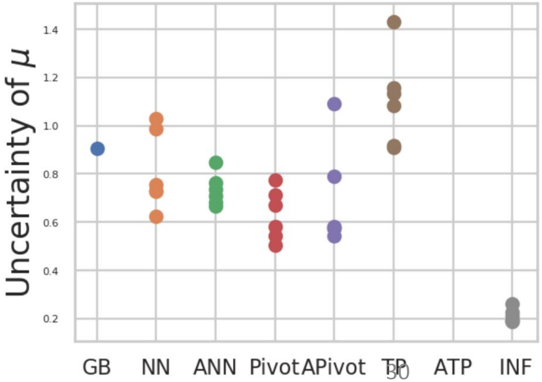
Kyle Cranmer, Lukas Heinrich, Gilles Louppe



Uncertainty reduction by gradient descent,







***Victor Estrade*[1]**, *Cécile Germain*[1], *Isabelle Guyon*[1], *David Rousseau*[2]

[slides]

[1] LRI, University of Paris-Sud, University of Paris-Saclay