Deep Learning Ideas for LHC Data Analysis

Gregor Kasieczka

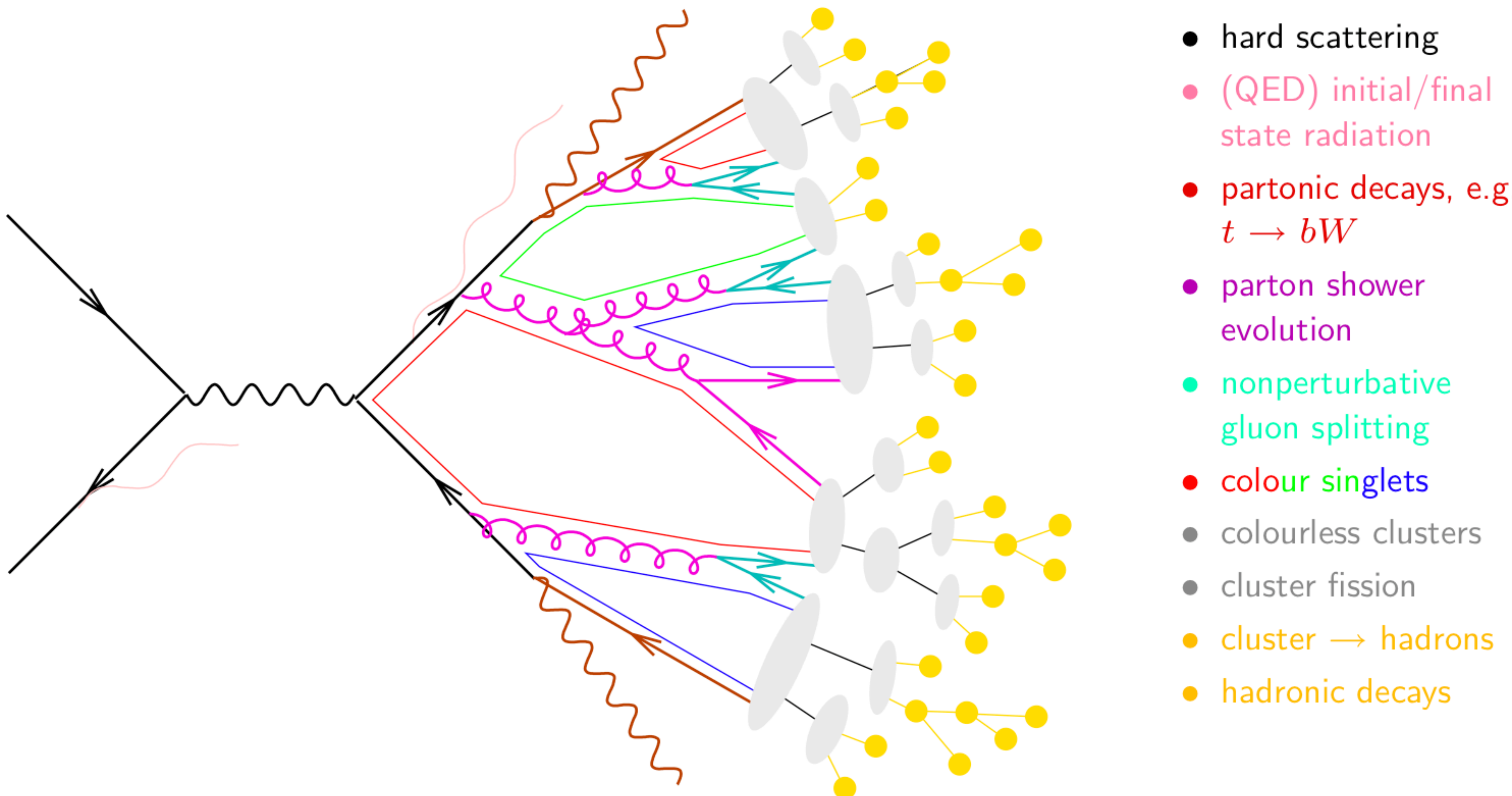
LHC Discussion: Machine Learning

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Emmy Noether-Programm
Deutsche Forschungsgemeinschaft
DFG

Bundesministerium für Bildung und Forschung

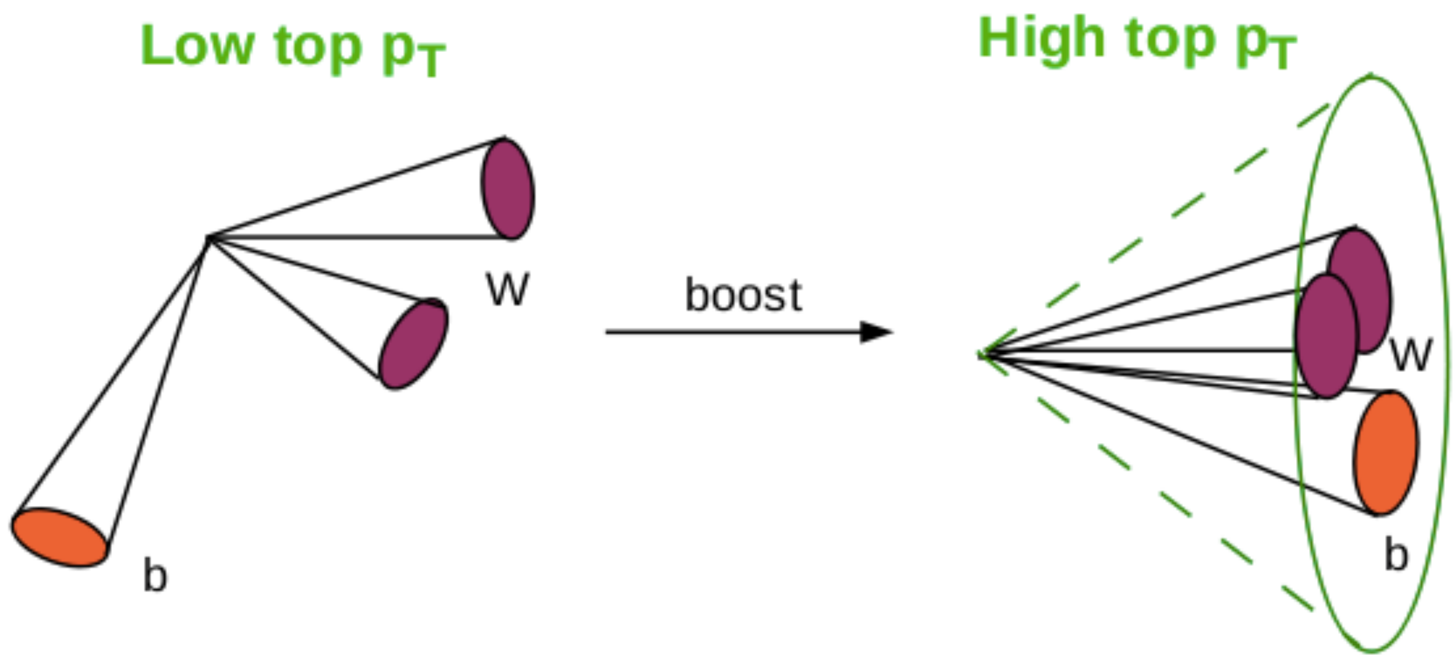
Physics at the LHC



We want to infer underlying physics from measurements in the detector. How can deep neural networks assist us?

http://www.quantumdiaries.org/

Heavy Resonance Tagging



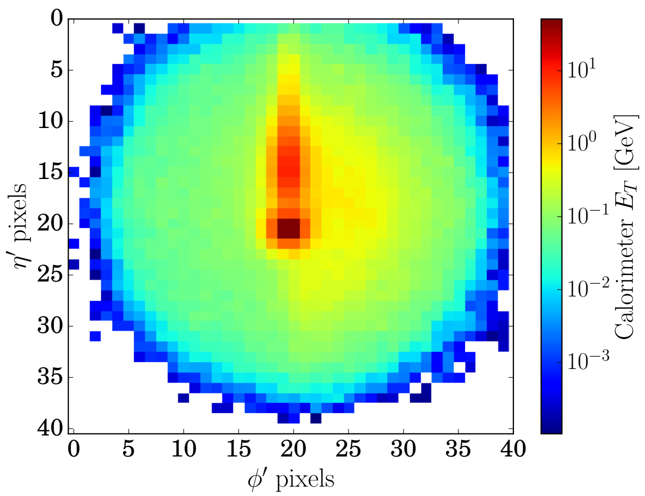
- Hadronically decaying top/Higgs/W/Z
- Contained in one (large-R) jet
- How to distinguish from light quark/gluon jets (and from each other)
- For new physics searches (and SM studies)

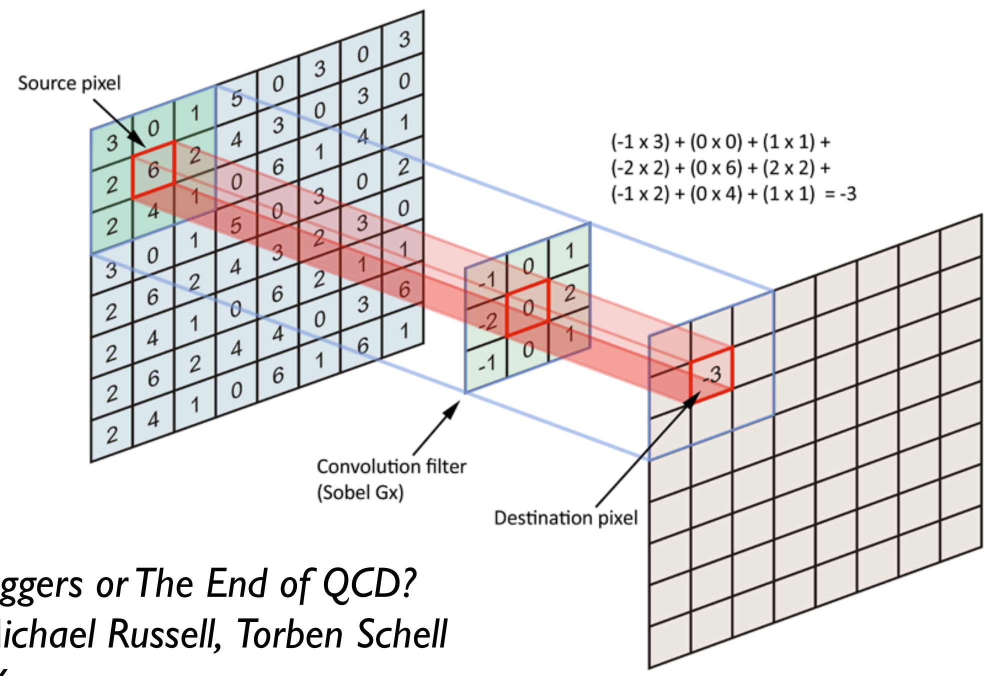
Towards an Understanding of the Correlations in Jet Substructure
D Adams et al (BOOST 2013 Participants), Eur.Phys.J. C75
Top Tagging, T Plehn, M Spannowksy, J.Phys. G39 (2012) 083001
Boosted Top Tagging Method Overview, GK, Proc. Top2017

Some Classical solutions:
(aka jet substructure)

- Mass
  Calculate using a grooming algorithm (eg mMDT/softdrop or pruning)
- Centers of hard radiation
  n-subjettiness or energy correlation functions
- Flavour
  b tagging of large-R jets or subjets
- Soft substructure
  Color connection
- Inclusive reconstruction
  HEPTopTagger V2, HOTVR
- Other substructure variables
  Shower deconstruction, template tagger, …

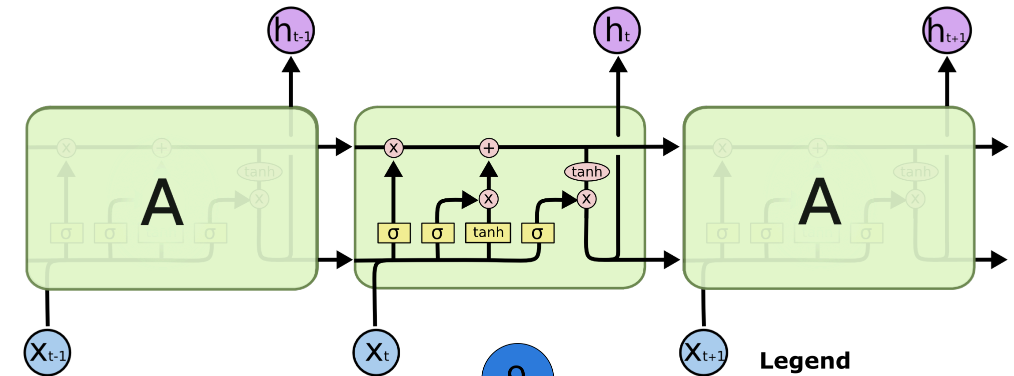
Images

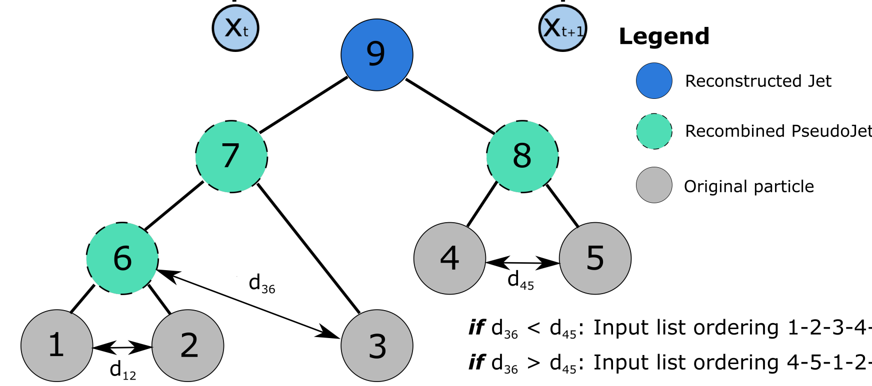




Deep-learning Top Taggers or The End of QCD?
GK, Tilman Plehn, Michael Russell, Torben Schell
JHEP 05 (2017) 006
Deep learning in color: towards automated quark/gluon jet discrimination
PT Komiske, EM Metodiev, MD Schwartz
JHEP 01 (2017) 110
Jet-Images: Computer Vision Inspired Techniques for Jet Tagging
J Cogan, M Kagan, E Strauss, A Schwartzman
arXiv:1407.5675
Jet-Images -- Deep Learning Edition
Ld Oliveira, M Kagan, L Mackey, B Nachman, A Schwartzman
JHEP 1607 069
Quark and gluon tagging with Jet Images in ATLAS, ATL-PHYS-PUB-2017-017

Recurrent

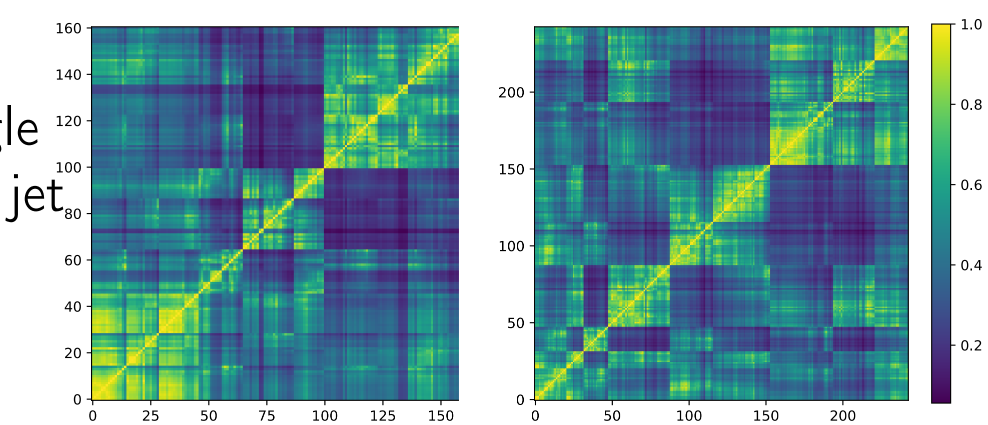




Long Short-Term Memory (LSTM) networks with jet constituents for boosted top tagging at the LHC
S Egan, W Fedorko, A Lister, J Pearkes, C Gay
arXiv: 1711.09059
QCD-Aware Recursive Neural Networks for Jet Physics G Louppe, K Cho, C Becot, K Cranmer
arXiv:1702.00748

Other

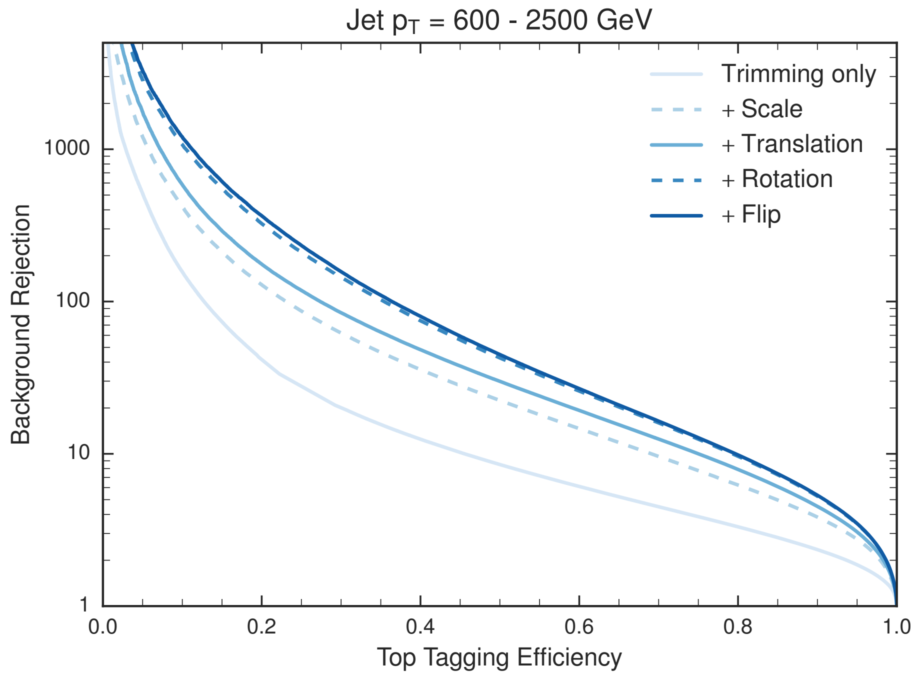
Single QCD jet

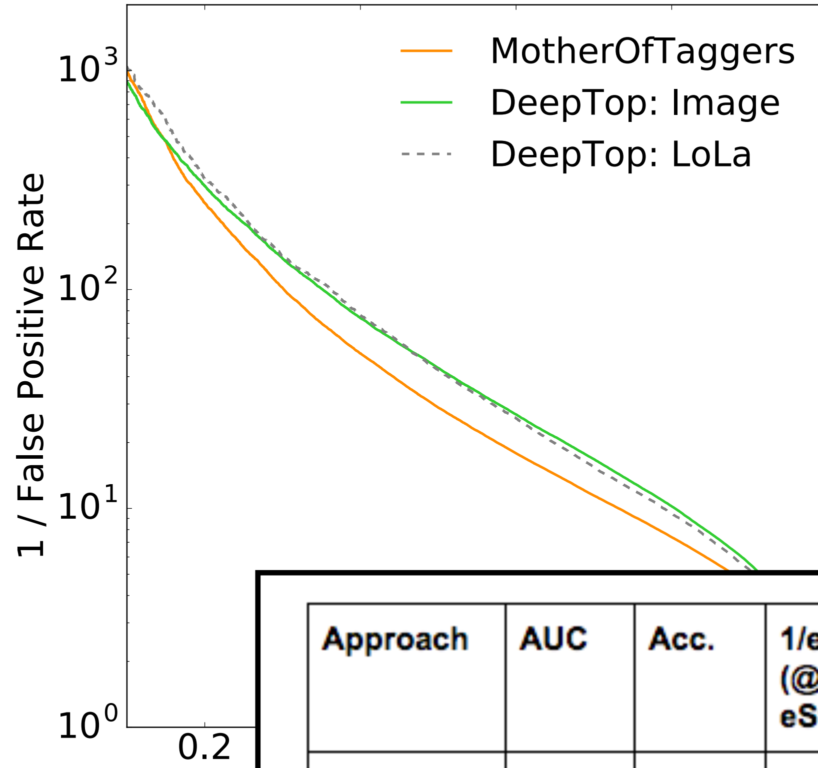


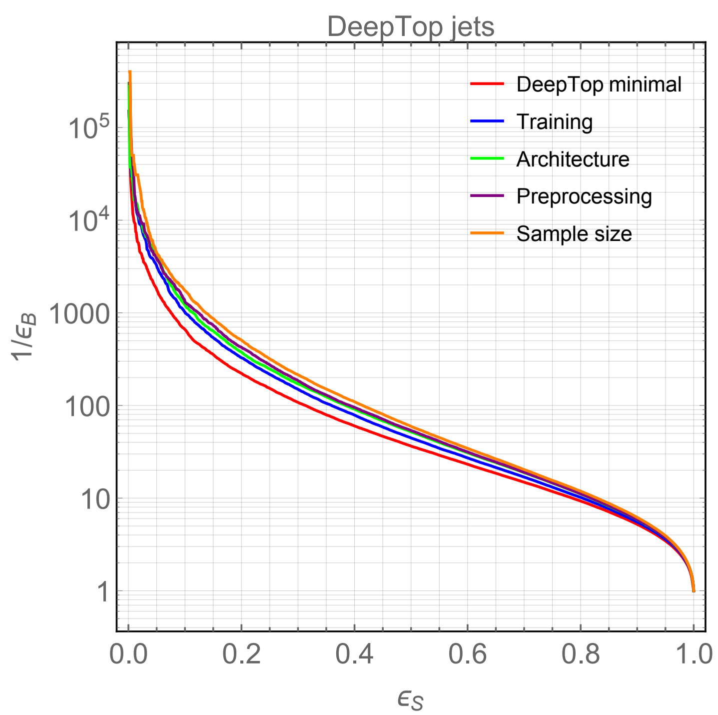
Single *W* jet

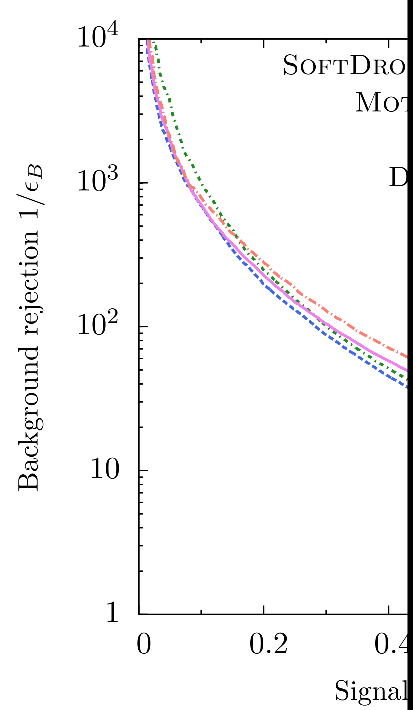
Neural Message Passing for Jet Physics I Henrion et al
Procs. of the Deep Learning for Physical Sciences Workshop at NIPS (2017)
Deep-learning Top Taggers & No End to QCD A Butter, GK, T Plehn, M Russell
1707.08966

Performance Overview





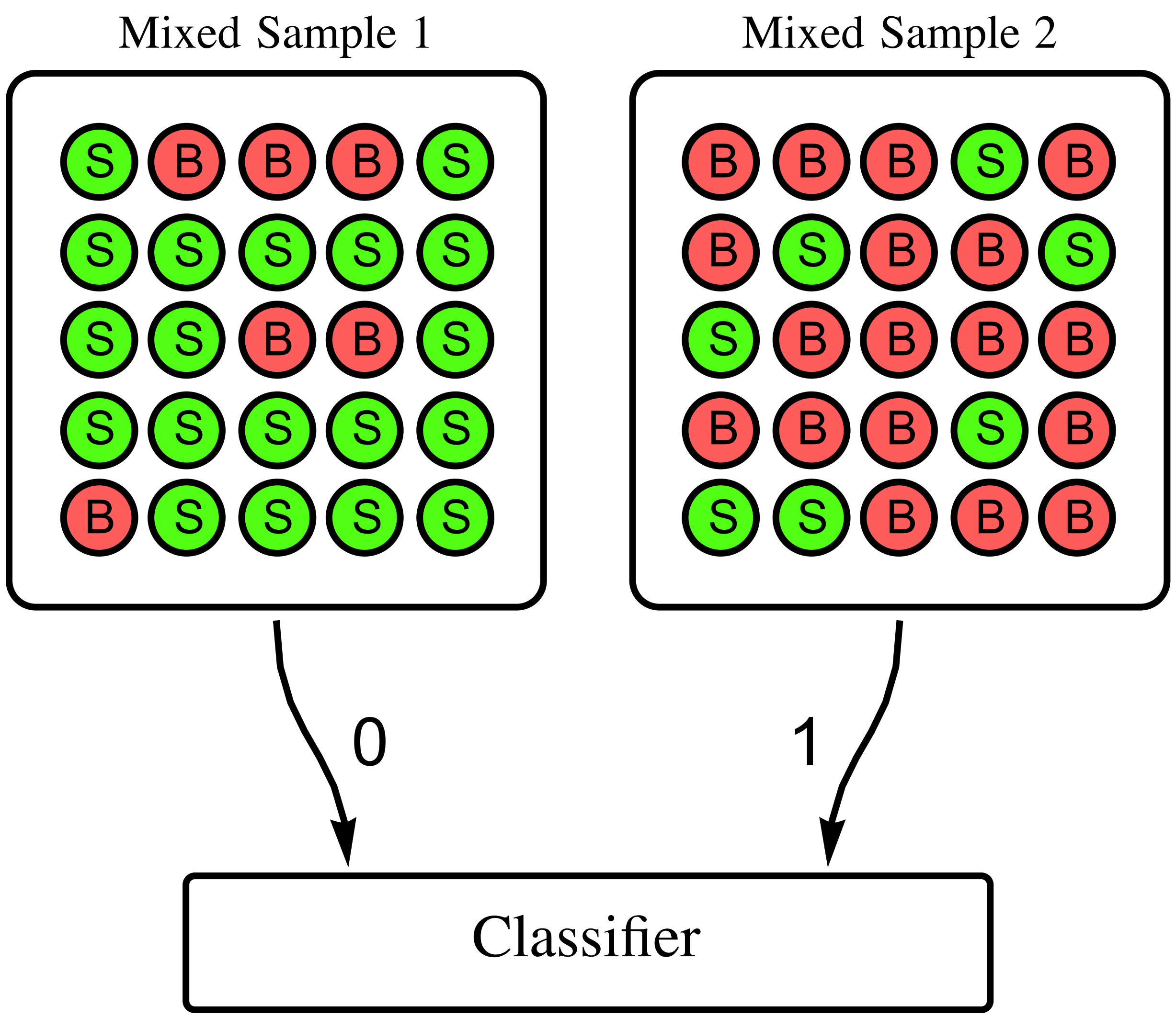


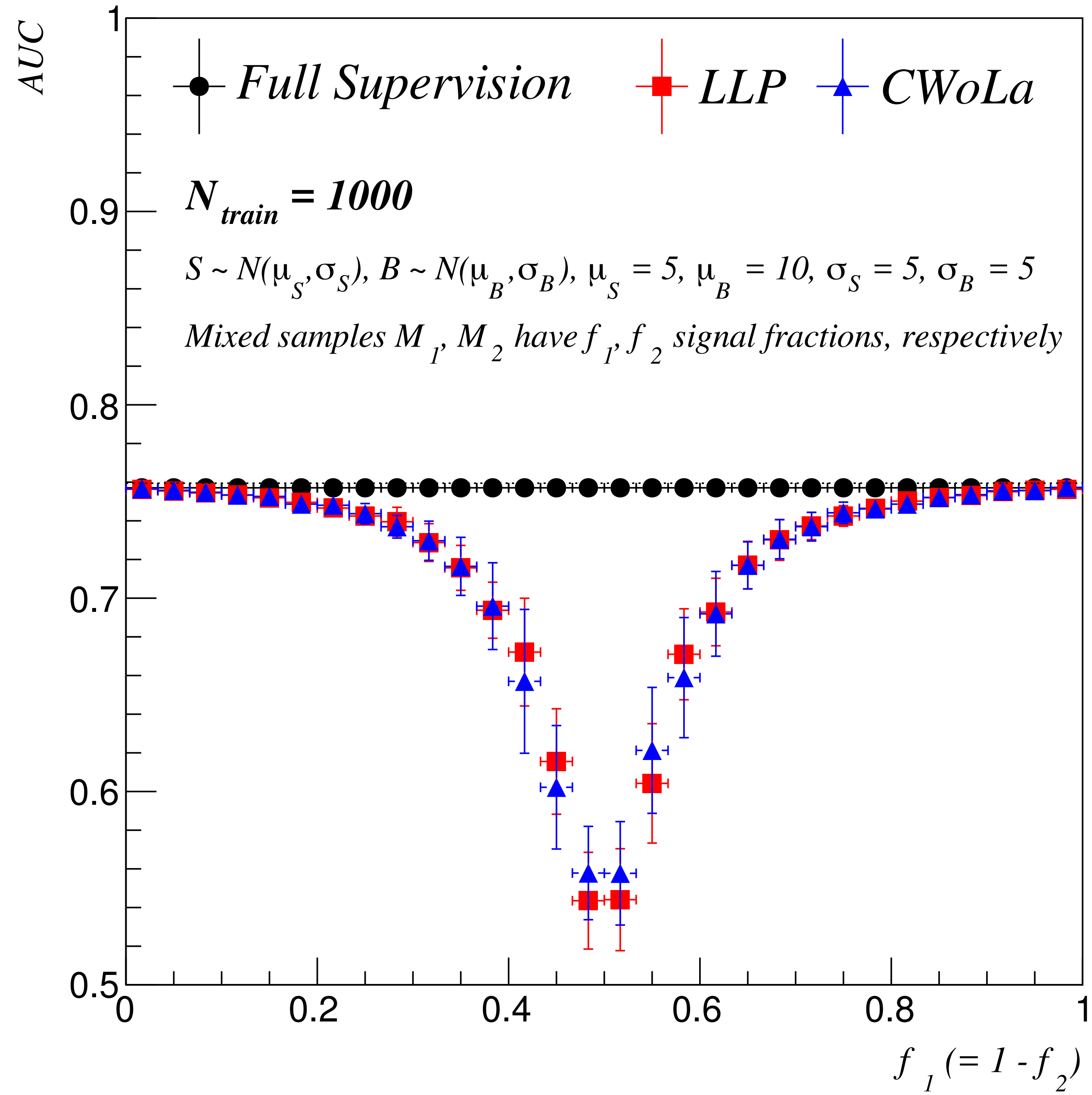


Approach	AUC	Acc.	1/eB (@ eS=0.3)	Training Time	Contact	Comments
LoLa	0.979	0.928			GK / SImon Leiss	Preliminary number, based on LoLa
LBN	0.979	0.928	728		Marcel Rieger	Preliminary number
CNN	0.981	0.93	780		David Shih	Model from *Pulling Out All the Tops with Computer Vision and Deep Learning (1803.00107)*
P-CNN (1D CNN)	0.980	0.930	782		Huilin Qu, Loukas Gouskos	Preliminary, use kinematic info only (https://indico.physics.lbl.gov/indico/event/546/contributions/1270/)

Our top tagging reference sample:
https://goo.gl/XGYju3

Learning from Data

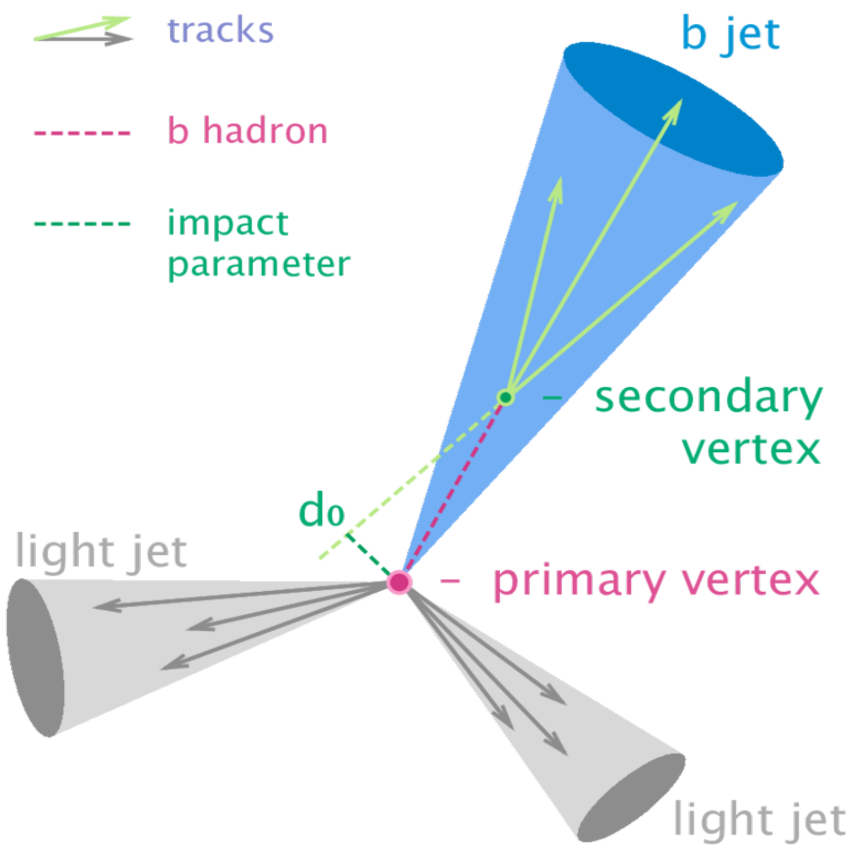


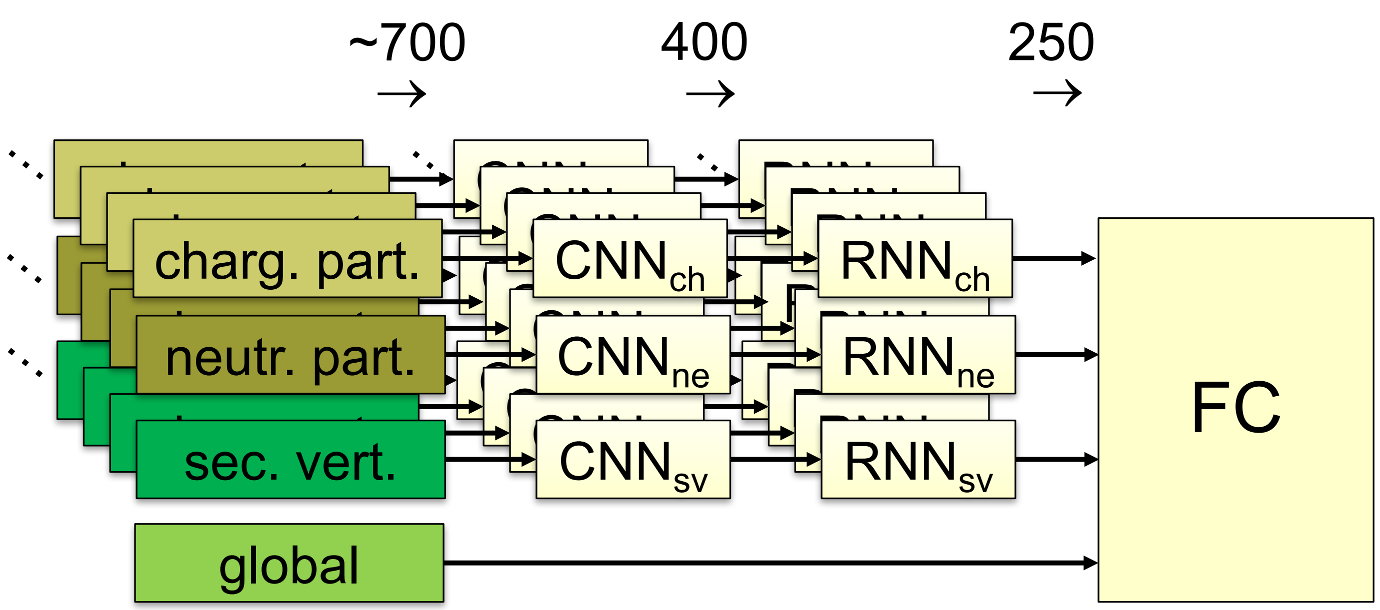


Distinguishing mixed samples is equivalent to signal/background classification!

Weakly Supervised Classification in High Energy Physics
LM Dery, B Nachman, F Rubbo, A Schwartzman, 1702.00414
Learning to Classify from Impure Samples
PT Komiske, EM Metodiev, B Nachman, MD Schwartz, 1801.10158
Classification without labels: Learning from mixed samples in high energy physics, EM Metodiev, B Nachman, J Thaler, 1708.02949

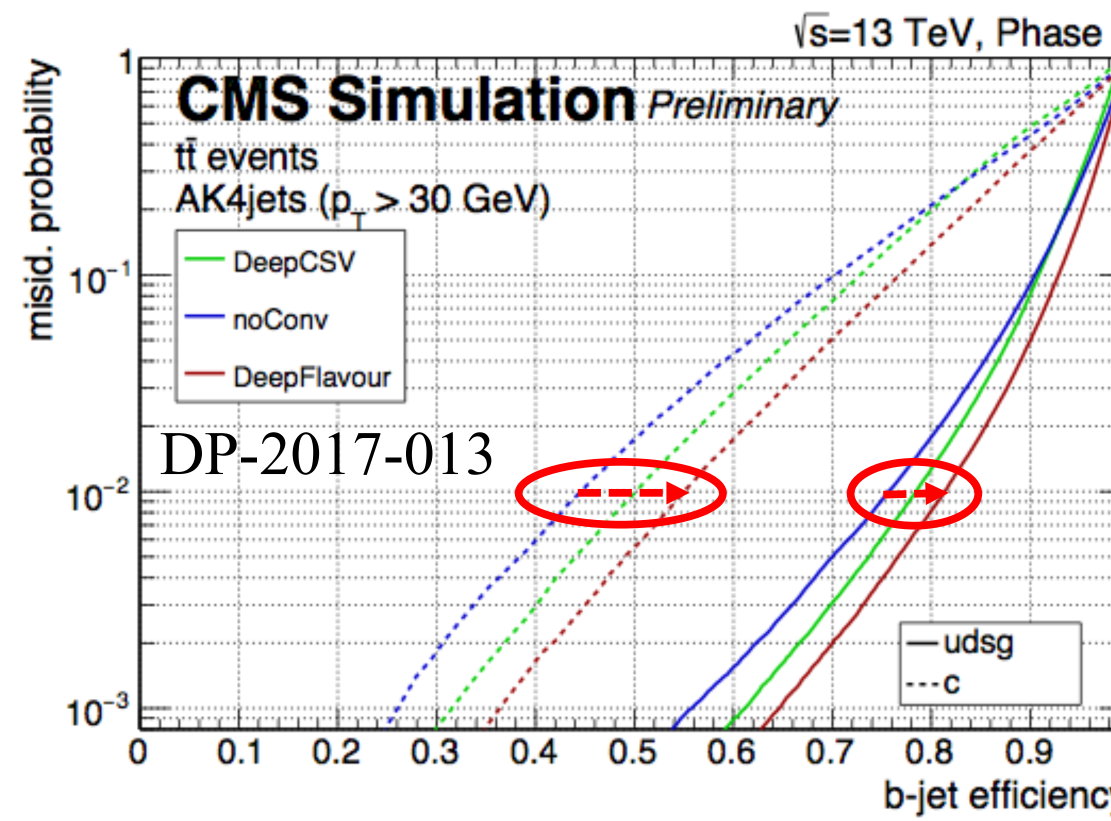
b Quark Identification





~ 700 inputs and 250.000 model parameters

- *DeepCSV*: Standard variables in deep neural network
- *DeepFlavour*: Complex architecture on per-particle quantities

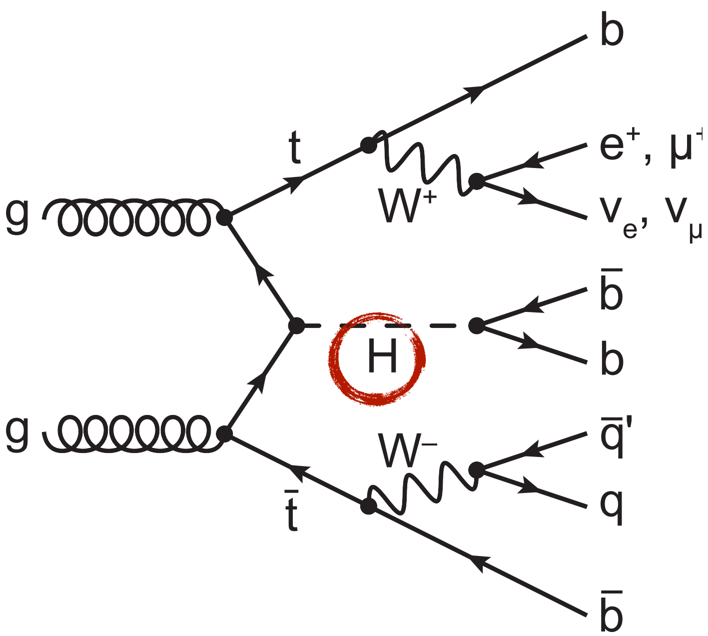


Blue: generic DNN (650 inputs)
Green: CMS tagger (~65 human made inputs)
Red: Physics inspired DNN (650 inputs)

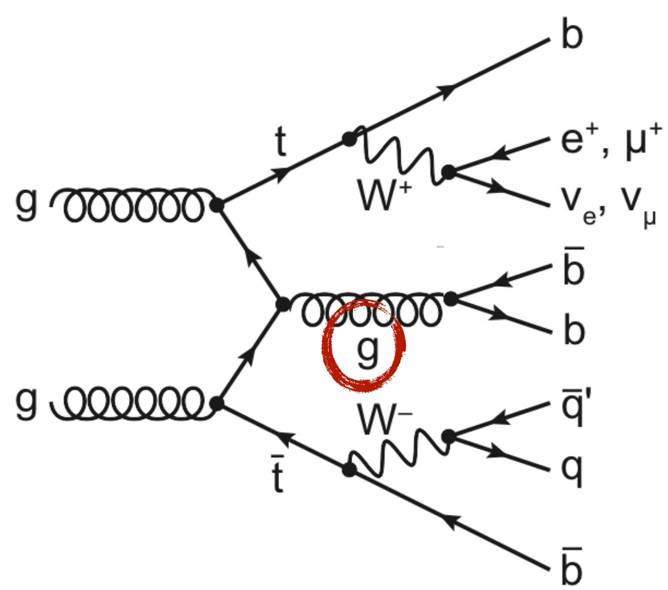
Machine Learning for Jet Physics in CMS
Markus Store (for CMS)
Jets in ML Workshop, Berkeley, 2017

CMS DP 2017-013

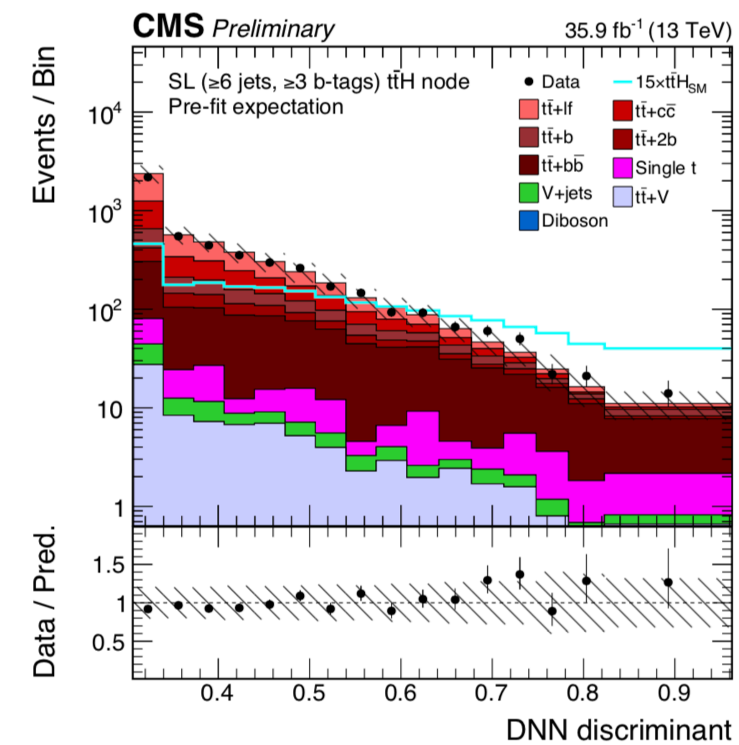
Event Classification Example



signal



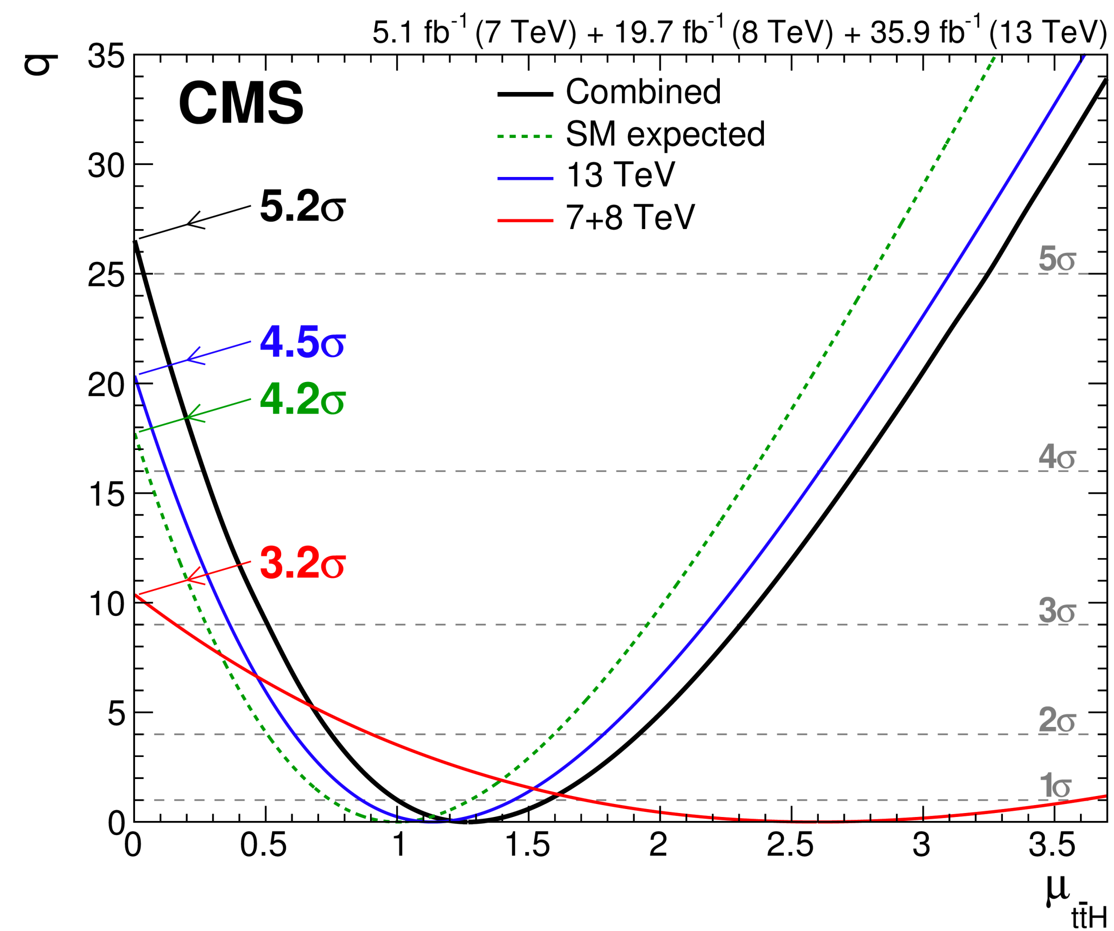
difficult background



Distinguish signal and background sub-processes per jet-multiplicity category

Fully connected network based on

- *object kinematics*
- *event shapes*
  (including Matrix Element Method output)
- *b-tagging information*



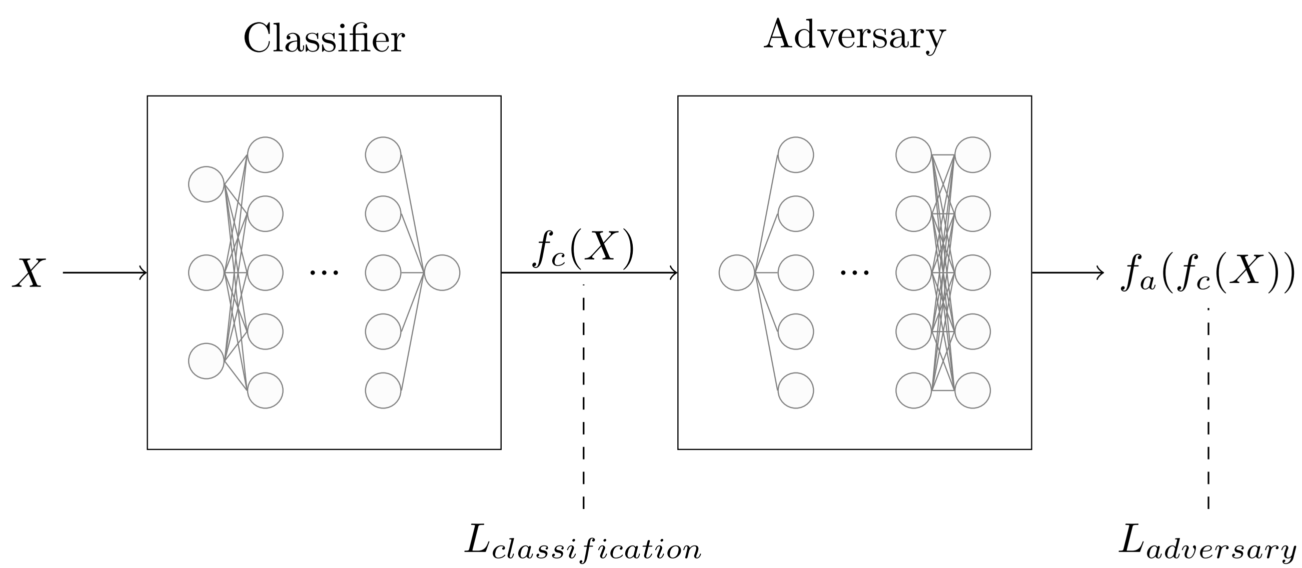
Observation of ttH Production CMS PRL 120 (2018) 231801

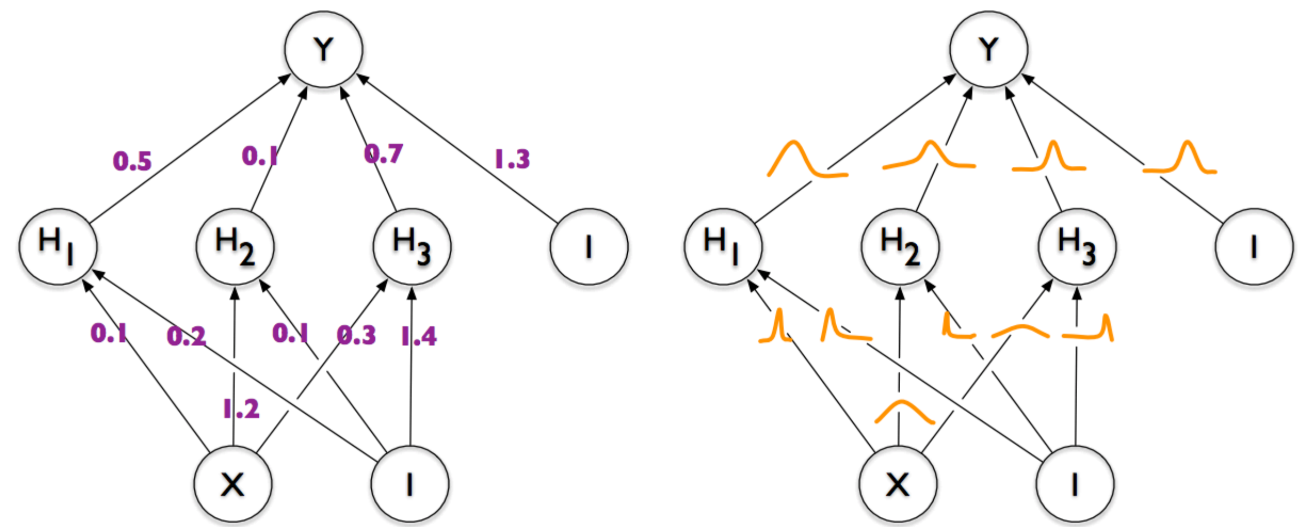
Search for ttH production in the H → bb decay channel with leptonic tt decays in proton-proton collisions at sqrt(s) = 13 TeV with the CMS detector CMS Collaboration, PAS HIG-17-026

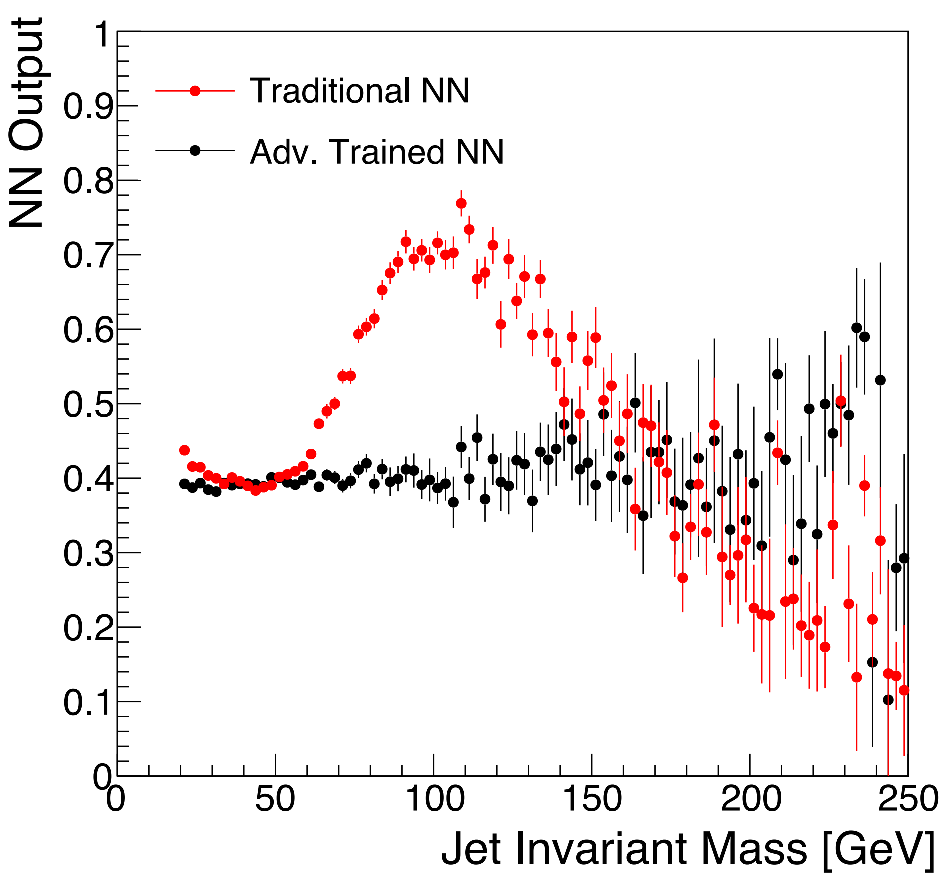
Jet-Parton Assignment in ttH Events using Deep Learning
M Erdmann, B Fischer, M Rieger, JINST 12 (2017) P08020

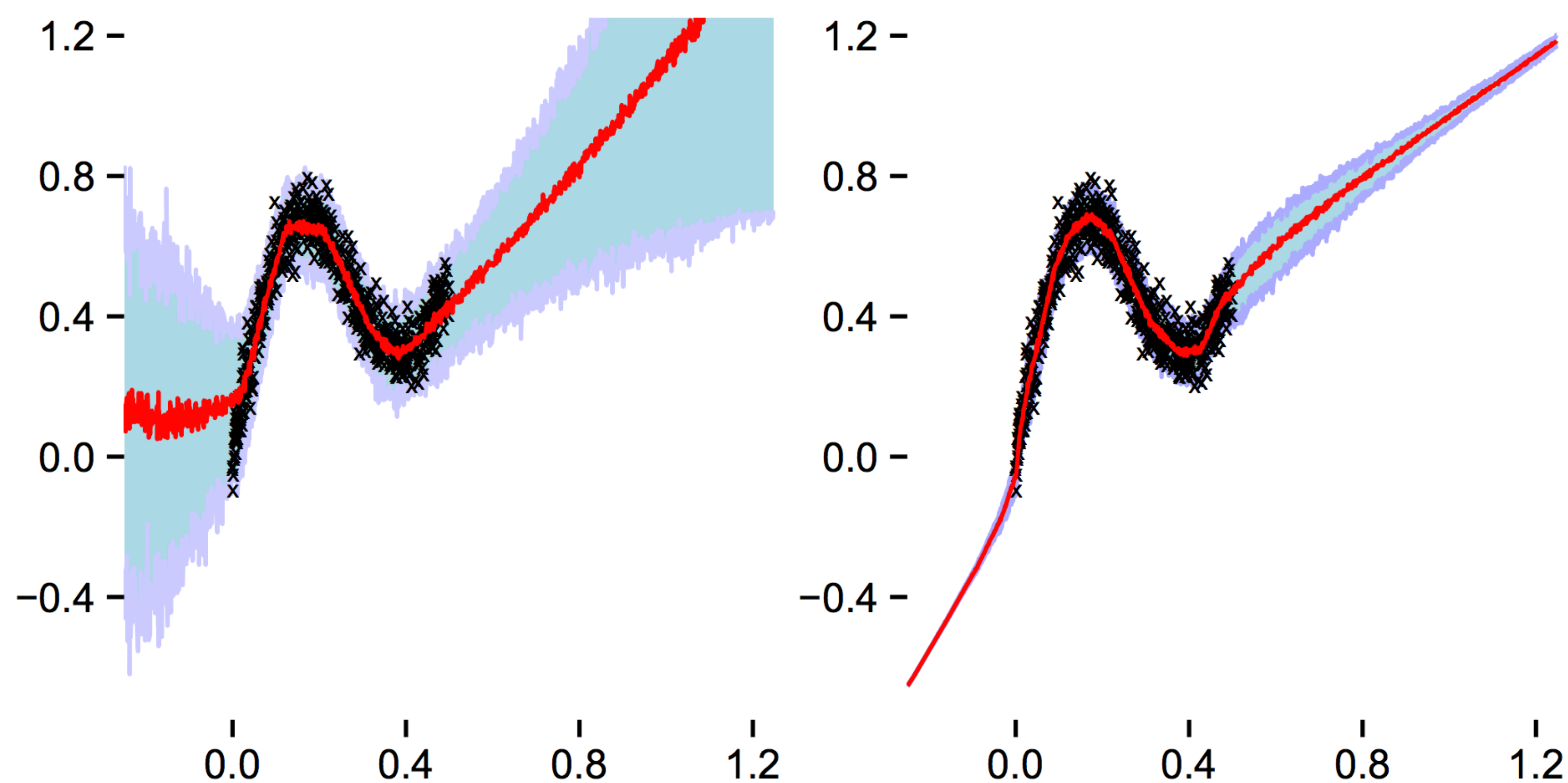
Calibration/Correlations/Uncertainties

Decorrelated Jet Substructure Tagging using Adversarial Neural Networks
C Shimmin, P Sadowski, P Baldi, E Weik, D Whiteson, E Goul, A Søgaard 1703.03507
Weight Uncertainty in Neural Networks
C Blundell et al, ICML Proc's 2015





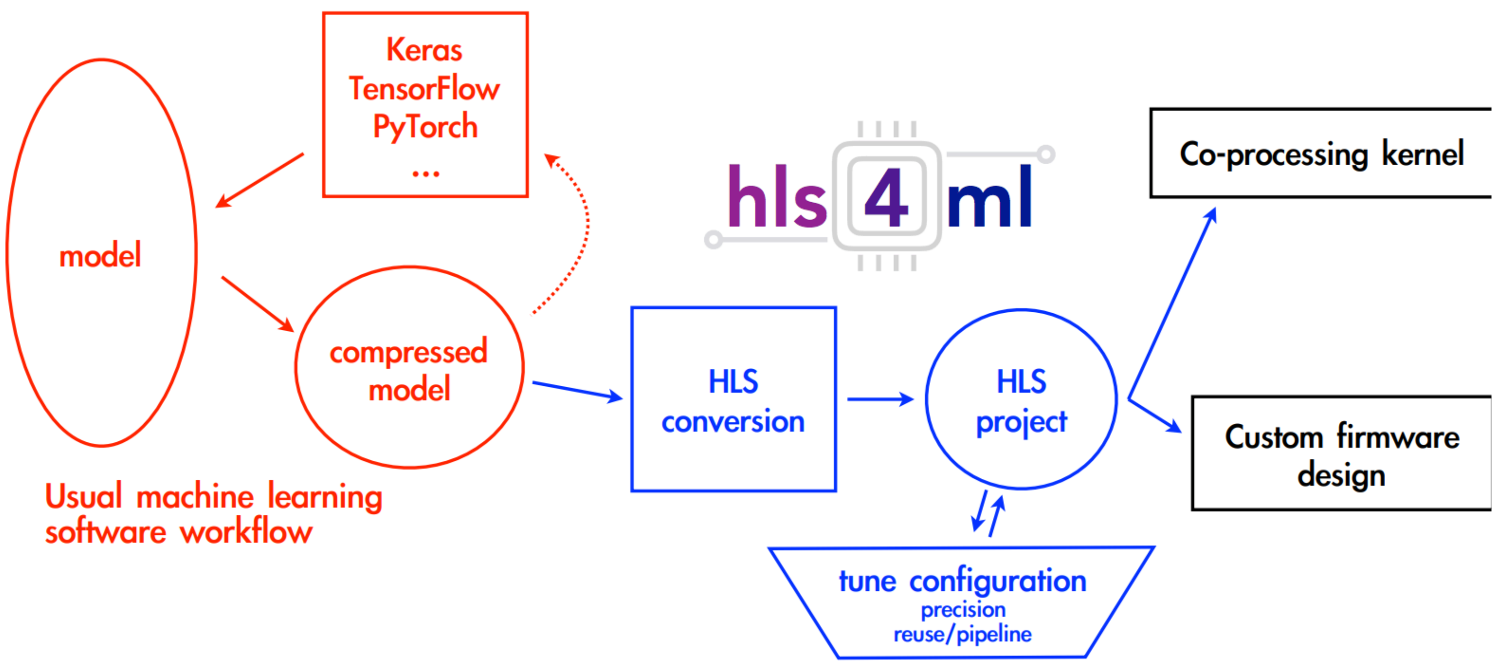




- **Trade-off discrimination power and stability**

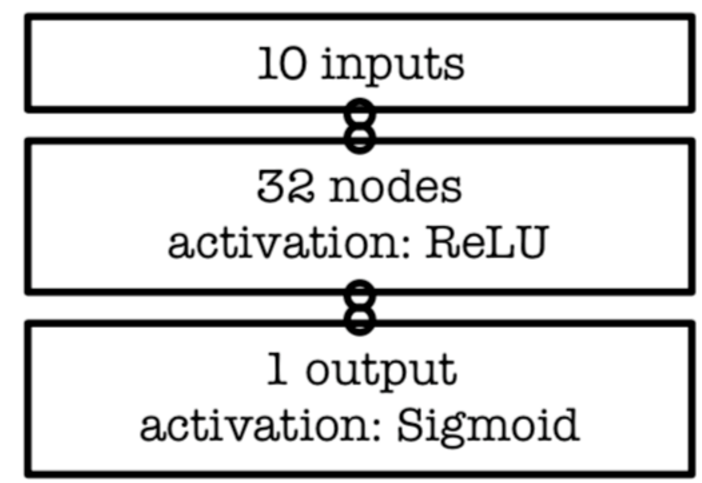
FPGA DNN Triggers

Upcoming talk in Joint Instrumentation Seminar (Date TBC)



Observables

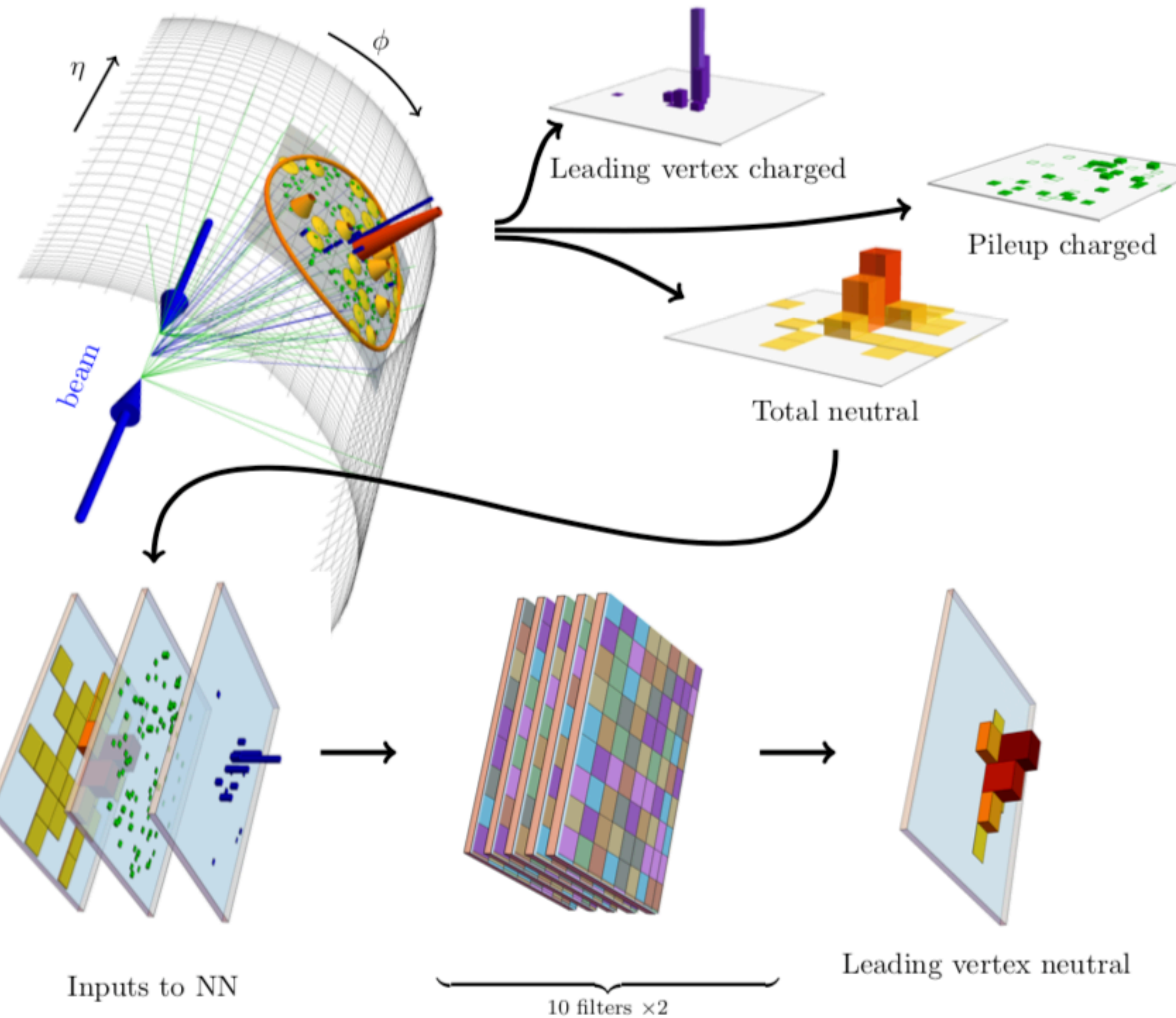
m_{mMDT}
$N_2^{\beta=1,2}$
$M_2^{\beta=1,2}$
$C_1^{\beta=0,1,2}$
$C_2^{\beta=1,2}$
$D_2^{\beta=1,2}$
$D_2^{(\alpha,\beta)=(1,1),(1,2)}$
$\sum z \log z$
Multiplicity

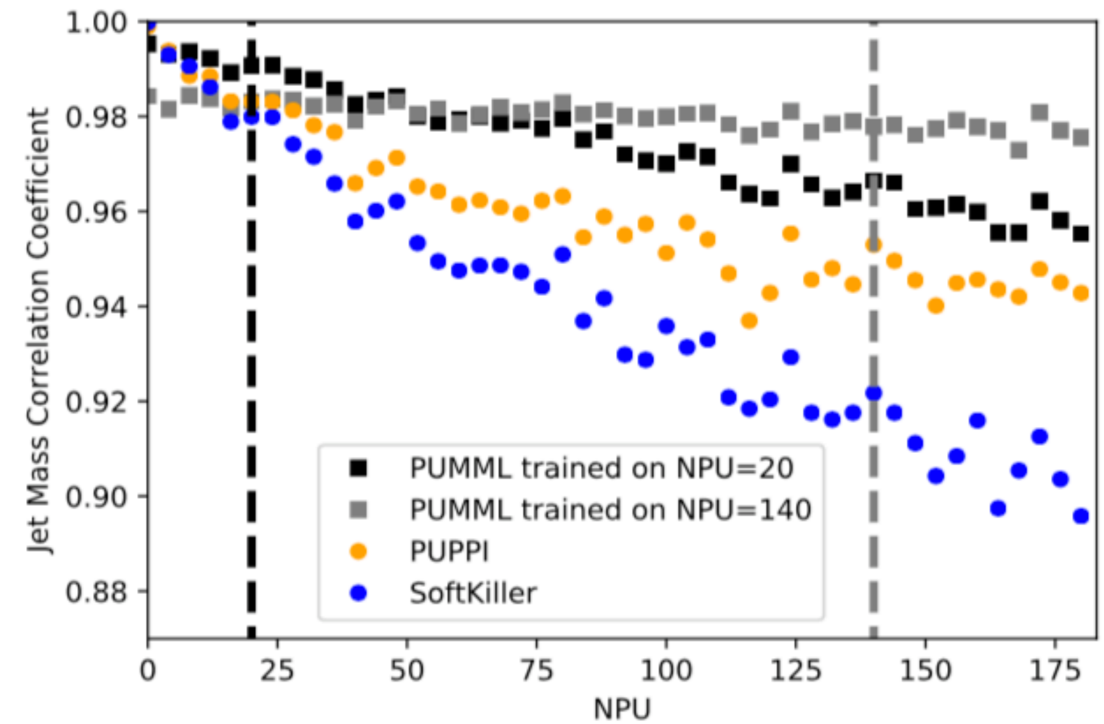


Fast inference of deep neural networks in FPGAs for particle physics
J Duarte et al
1804.06913

- Framework to translate NNs to FPGAs for fast (L1 trigger) execution
- Latency of 75-150 ns

Pile-Up



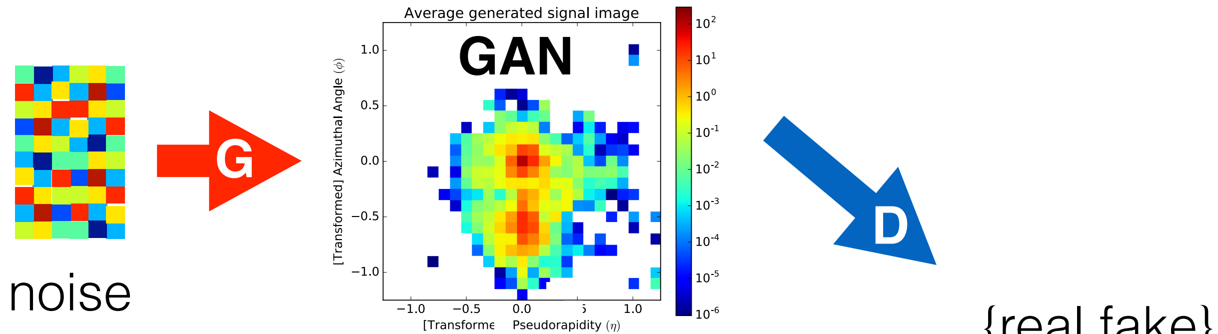


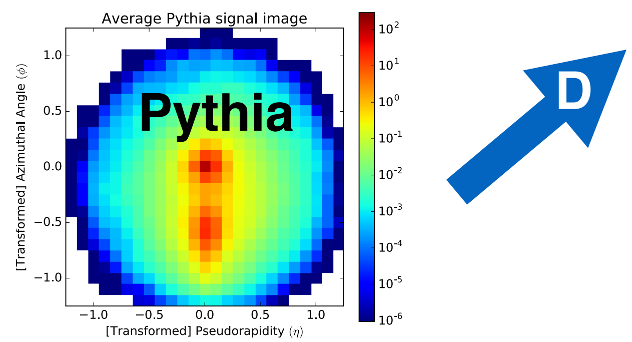
- Predict pile-up distribution per-event
- Alternative (CMS): Can we improve PF/PUPPI with ML?

Pileup Mitigation with Machine Learning (PUMML)
PT Komiske et al
1707.08600

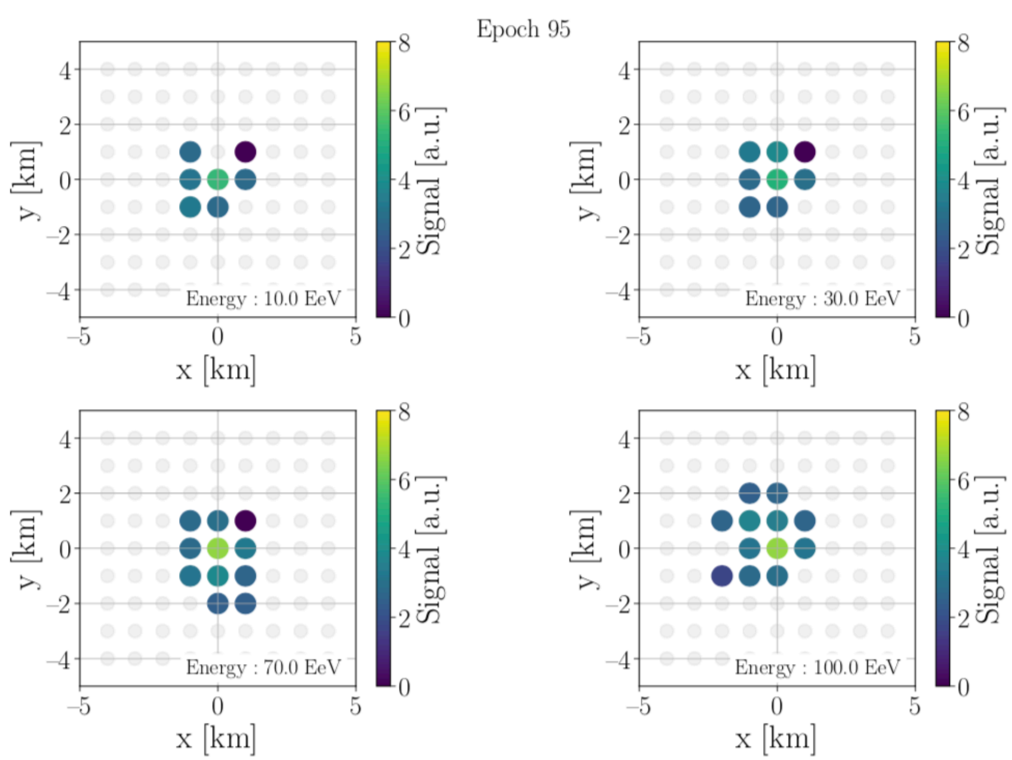
Generative Adversarial Networks

Generative Adversarial Networks (GAN):
*A two-network game where one **maps noise to images** and one **classifies images as fake or real**.*

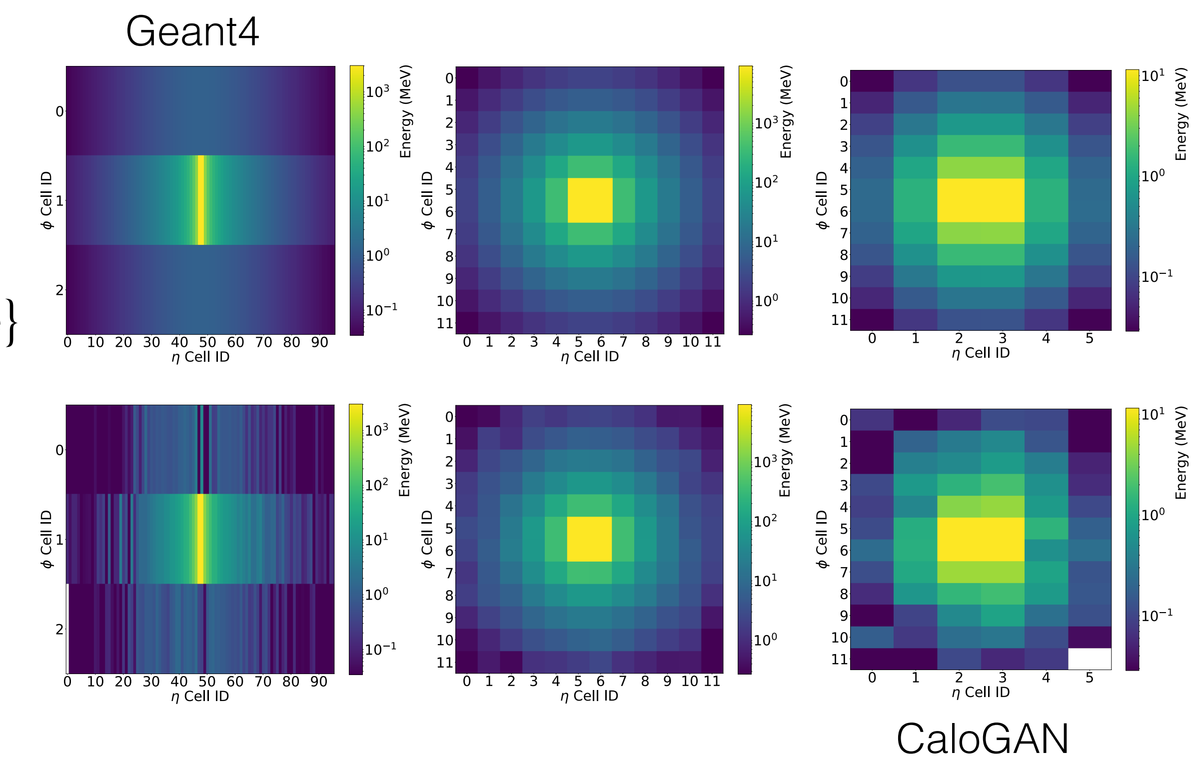




When **D** is maximally confused, **G** will be a good generator



Simulation of cosmic air-showers captured by Cherenkov detector. Use to improve energy reconstruction!

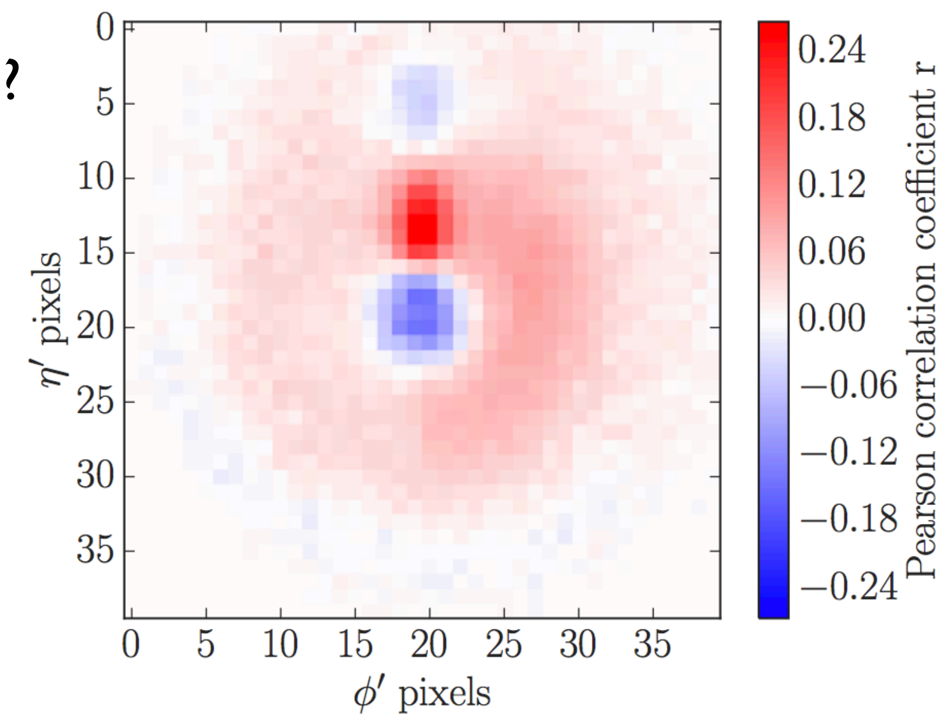


Alternative simulation of calorimeter events. Several orders of magnitudes faster than Geant4

L. de Oliveira, M Paganini, B Nachman:
Learning Particle Physics by Example: Location-Aware Generative Adversarial Networks for Physics Synthesis,, Comput Softw Big Sci *(2017) 1:4,*
Accelerating Science with Generative Adversarial Networks: An Application to 3D Particle Showers in Multi-Layer Calorimeters, PRL 120, 042003
CaloGAN: Simulating 3D High Energy Particle Showers in Multi-Layer Electromagnetic Calorimeters with Generative Adversarial Networks, PR D 97, 014021
Generating and refining particle detector simulations using the Wasserstein distance in adversarial networks M Erdmann et al, 1802.03325

Opening the Black Box

- What is necessary to trust a new algorithm?
  - Learn what the network learns!
    - Visualise decision process
    - Encode physics in the network
    - Correlation with known variables
    - *We want to understand, not just describe!*
  - Information theoretic approaches

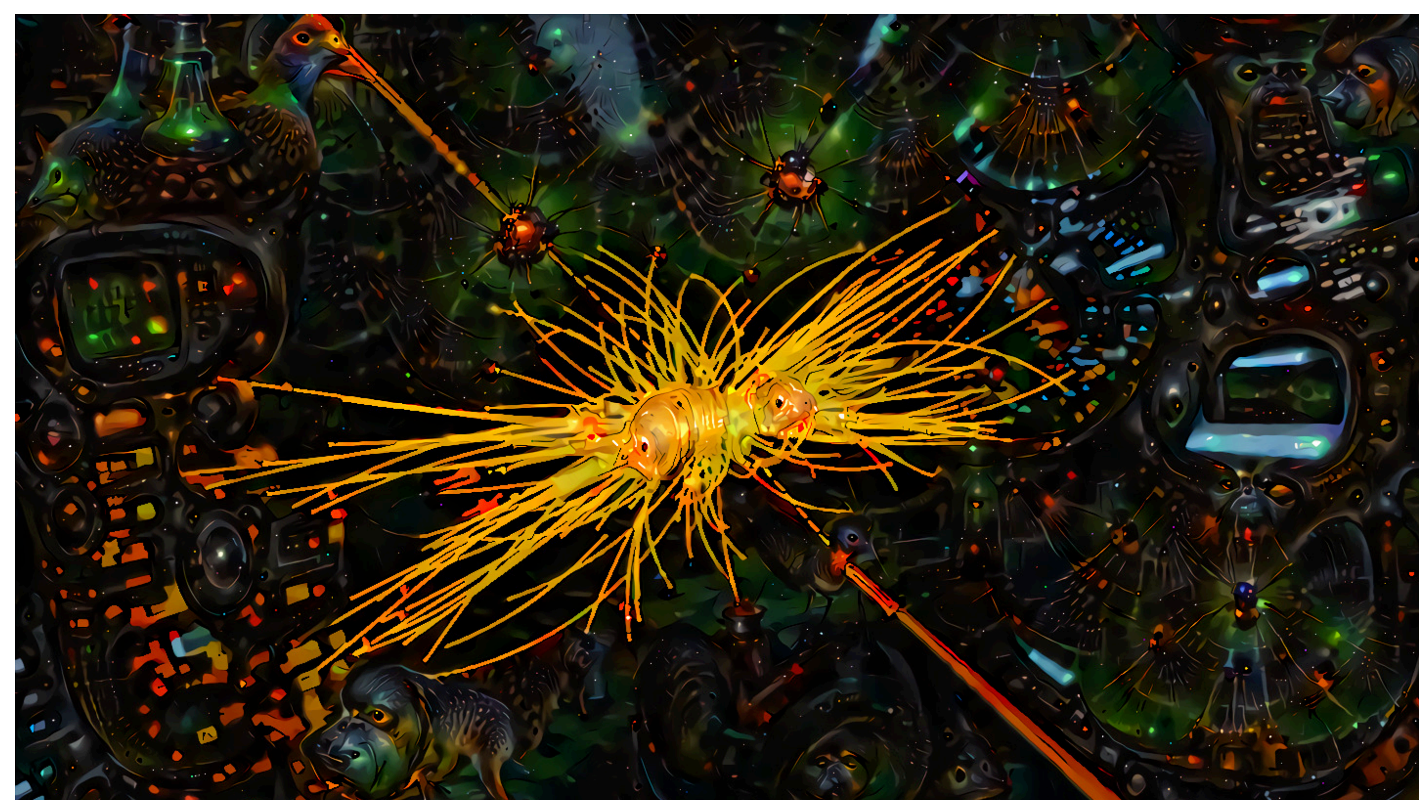


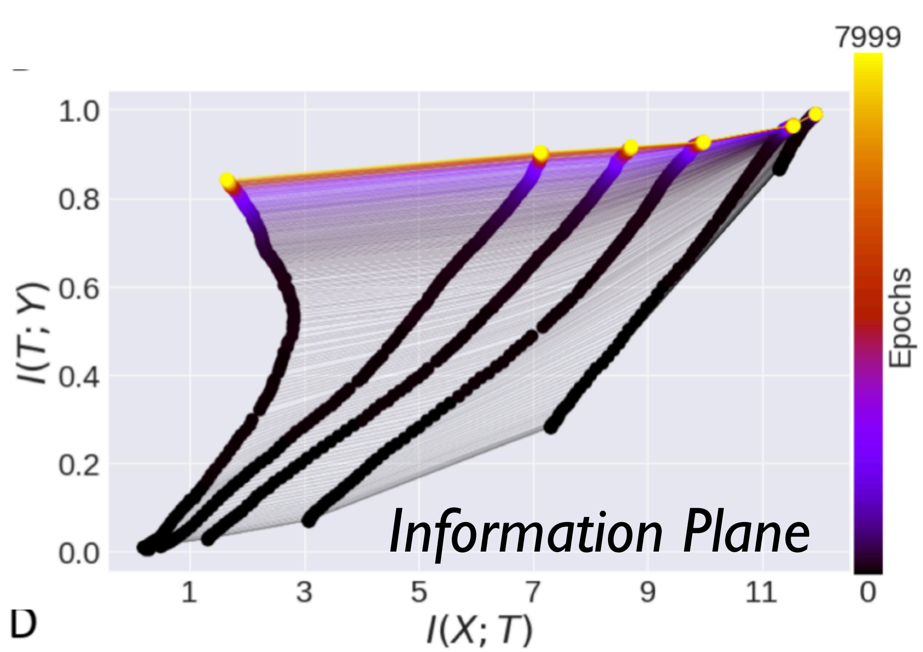
On the Information Bottleneck Theory of Deep Learning, Saxe et al, ICLR Proc 2018

Deep-learning Top Taggers or The End of QCD? GK, Tilman Plehn, Michael Russell, Torben Schell JHEP 05 (2017) 006

Weight Uncertainty in Neural Networks C Blundell et al, ICML Proc's 2015

DeepDream: Slighly modify image to increase classification score. Highlight the features the network learned





Closing

- Presented (some) use cases
  - Object identification
    - Heavy resonances
    - Jet flavour and gluon
  - Event classification
  - Calibration / Uncertainties
  - Trigger
  - Pile-Up
  - Simulation
  - Understanding

- *Many ideas available*
- *So far mostly on “pheno” level*
- *Gain from actual use in experiments*

Thank you!