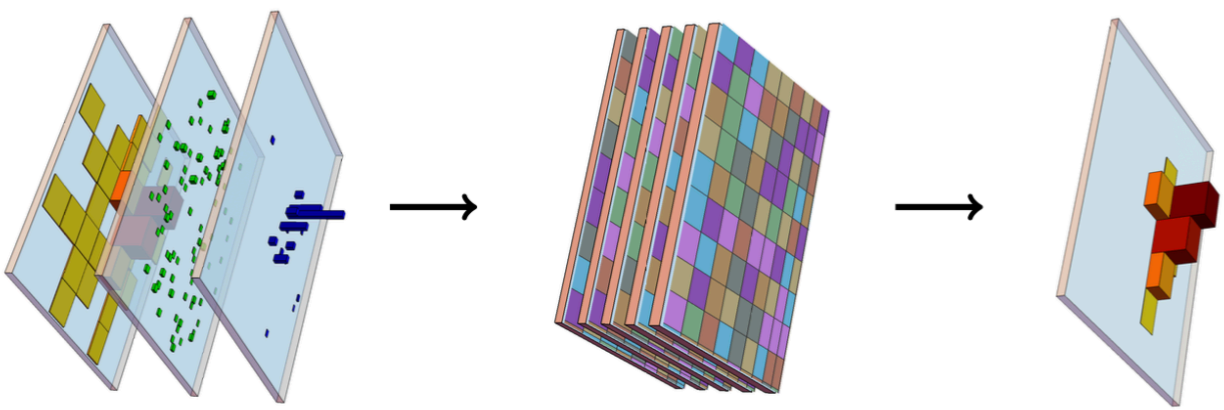
Jet Physics with Machine Learning

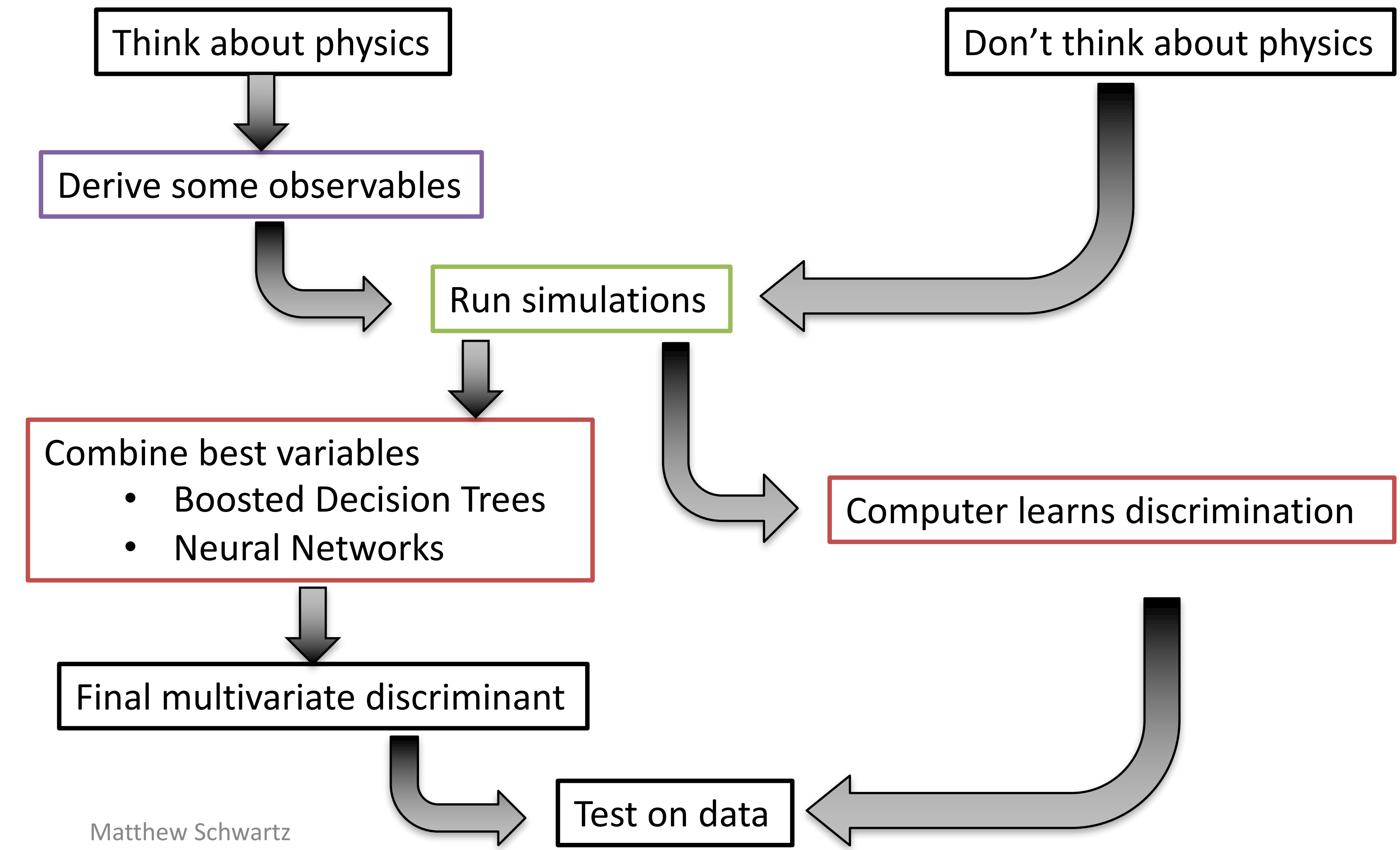
Chicagoland Mini-Workshop on
LHC Run II Analysis
January 15, 2019

Matthew Schwartz
Harvard University



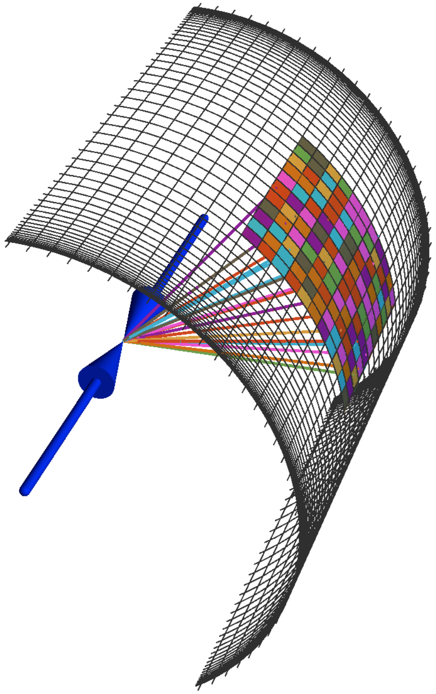
Modern Machine Learning for Particle Physics

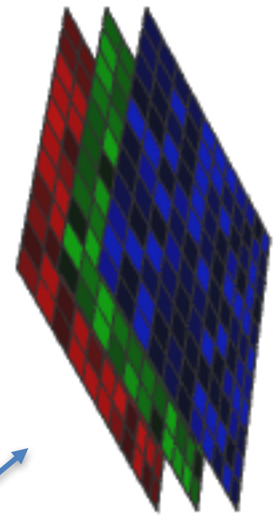
Traditional approach | Modern machine learning



Convolutional Neural Networks for quark/gluon jet discrimination

Komiske, Metodiev, MDS (arXiv:1612.01551)

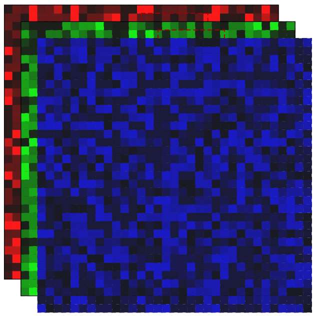




preprocess

- Center
- Crop
- Normalize
- Zero
- Standardize

NN inputs

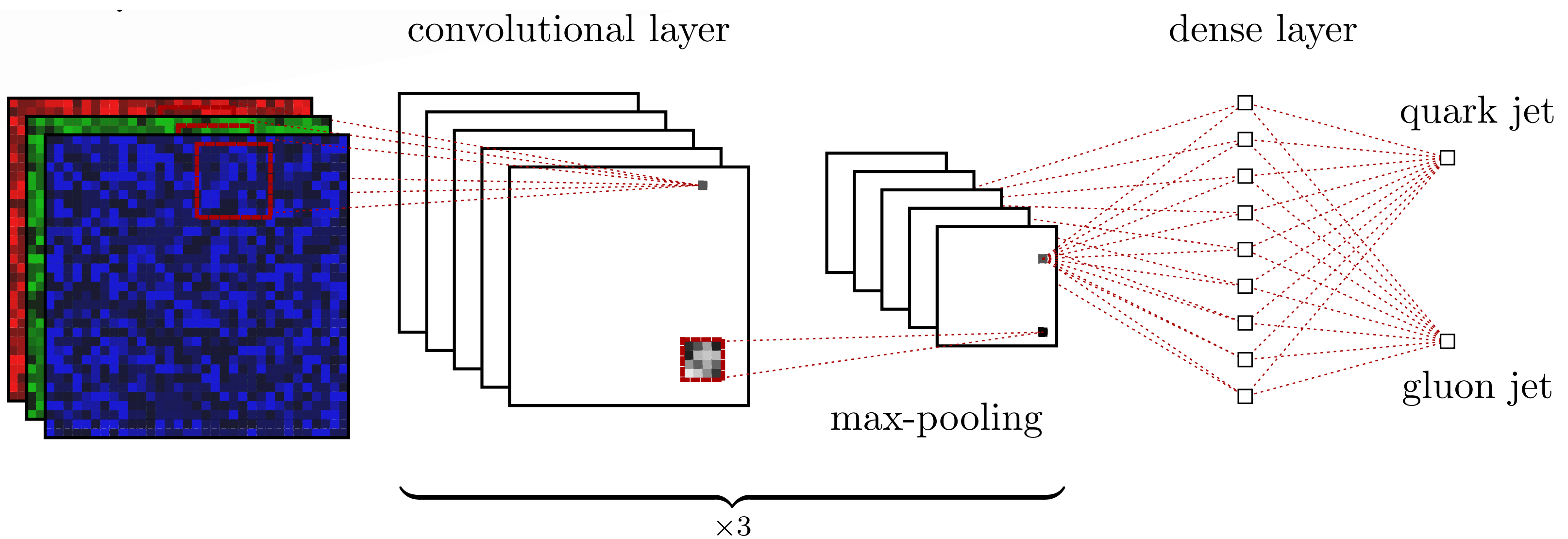


Three input layers

- Red = energy of charged particles
- Green = energy of neutral particles
- Blue = number of charged particles

CNN architecture

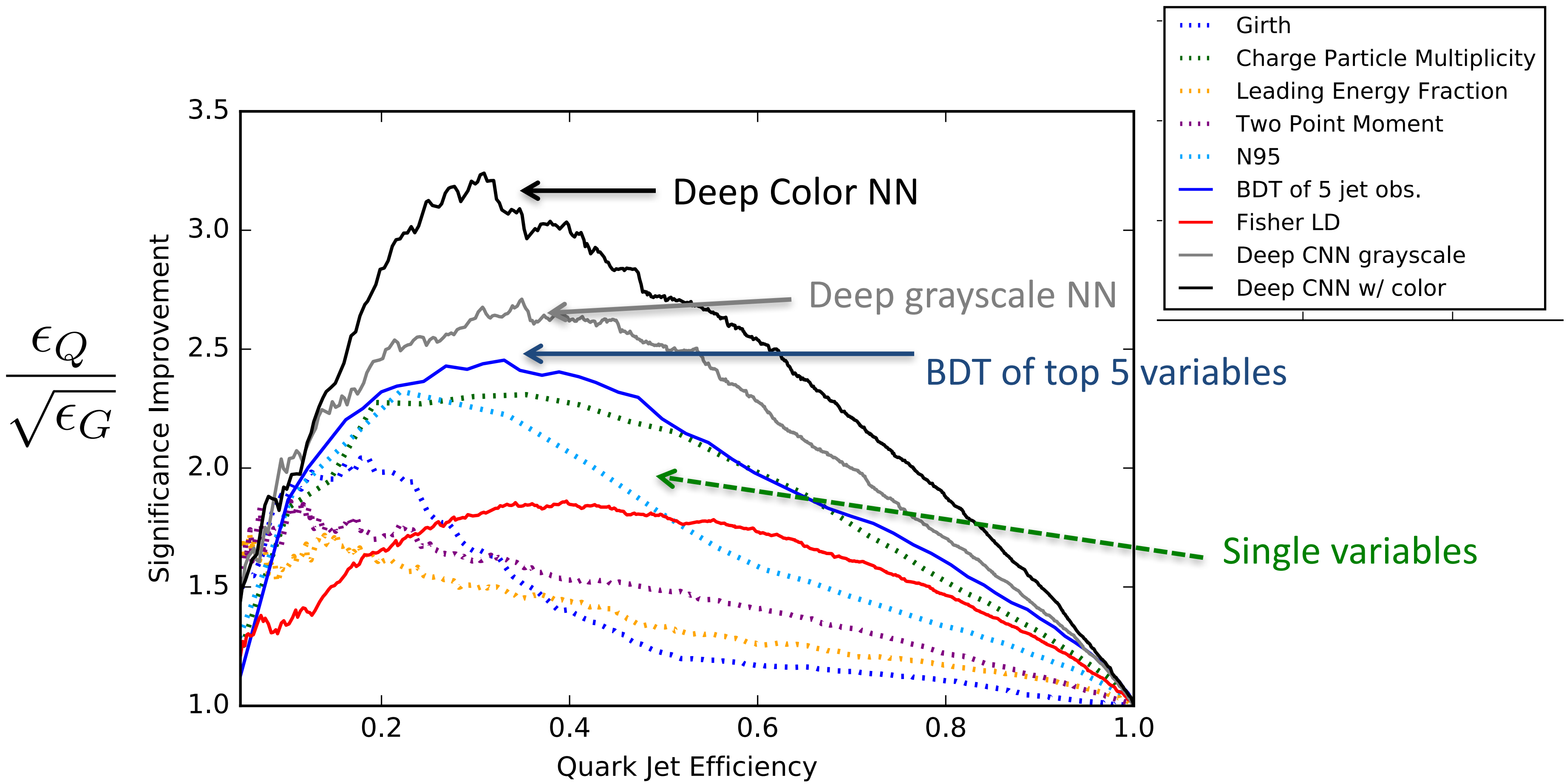
Komiske, Metodiev, MDS (arXiv:1612.01551)



- Convolution layers
  - 32×32 (image) ➔ 8×8 ➔ 4×4 ➔ 2×2 ➔ 1
- Final layer is densely connected to all final filters
- Output nodes connected to all notes in final hidden layer

Quark/Gluon CNN results

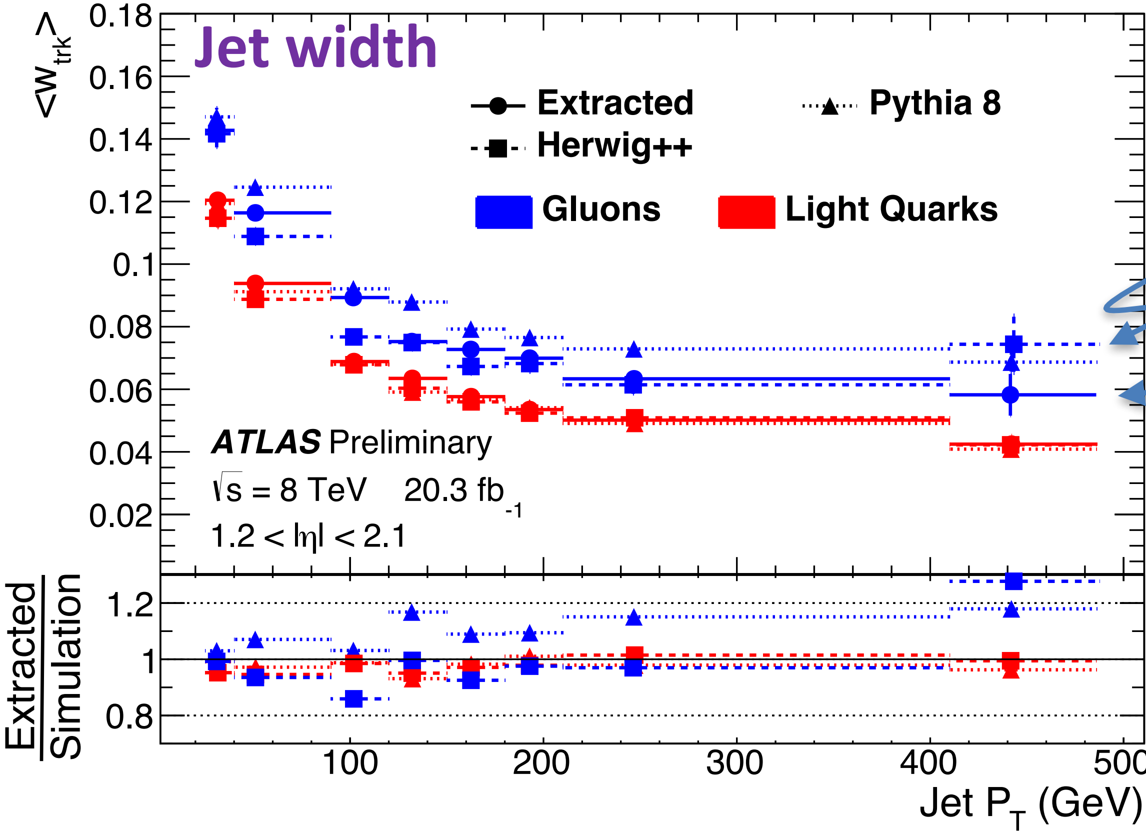
Komiske, Metodiev, MDS (arXiv:1612.01551)



Works really well – especially considering we don't put in any physics!

LHC data on quark and gluon jets

ATLAS (arXiv:1405.6583)
ATLAS (ATLAS-CONF-2016-034)



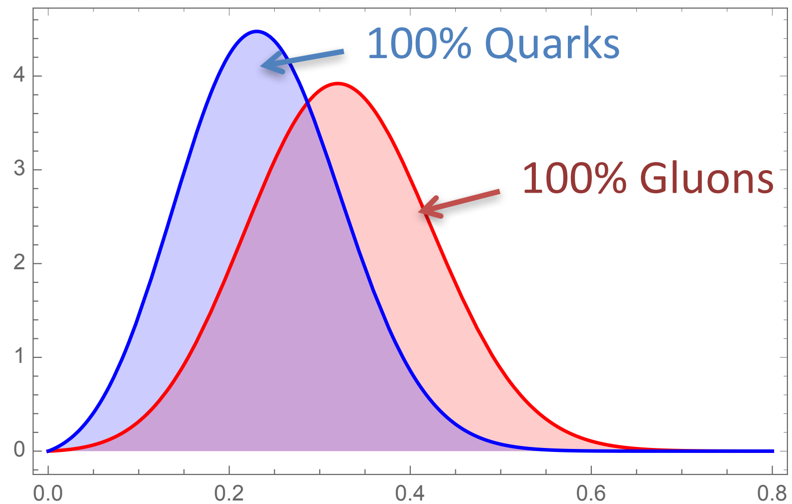
Pythia/Herwig gluon jets

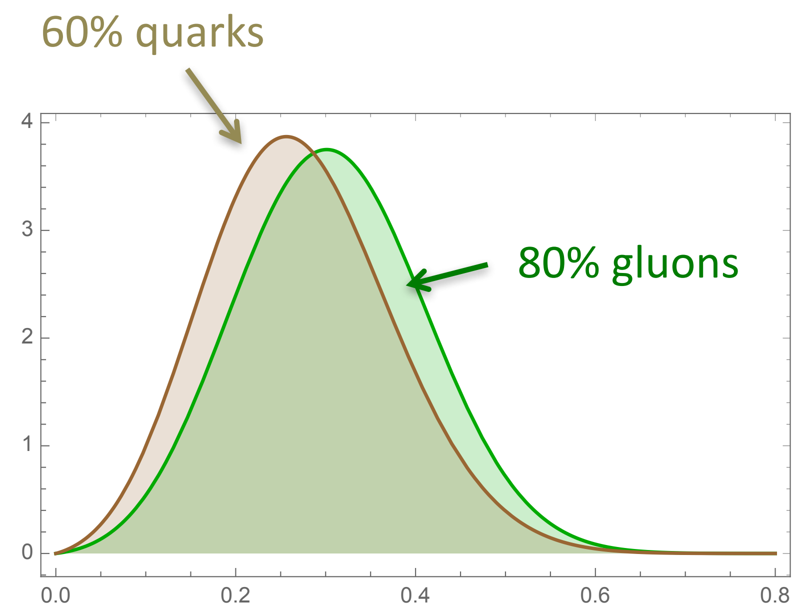
Gluon jets in data

- Simulations don't agree with each other
- Data does not agree with any simulation

Training Q/G discriminants on data

- Don't have samples of pure quark and gluon jets in data
- Do we need them?





$Q(x) = 2\ f_1(x) - f_2(x)$

$G(x) = -0.5\ f_1(x) + 1.5\ f_2(x)$

Sample 1: $f_1\ (x) = 0.6\ Q(x) + 0.4\ G(x)$

Sample 2: $f_2(x) = 0.2\ Q(x) + 0.8\ G(x)$

- Any two independent samples will do, in principle
- Inversion requires high statistics, impossible for multidimensional inputs

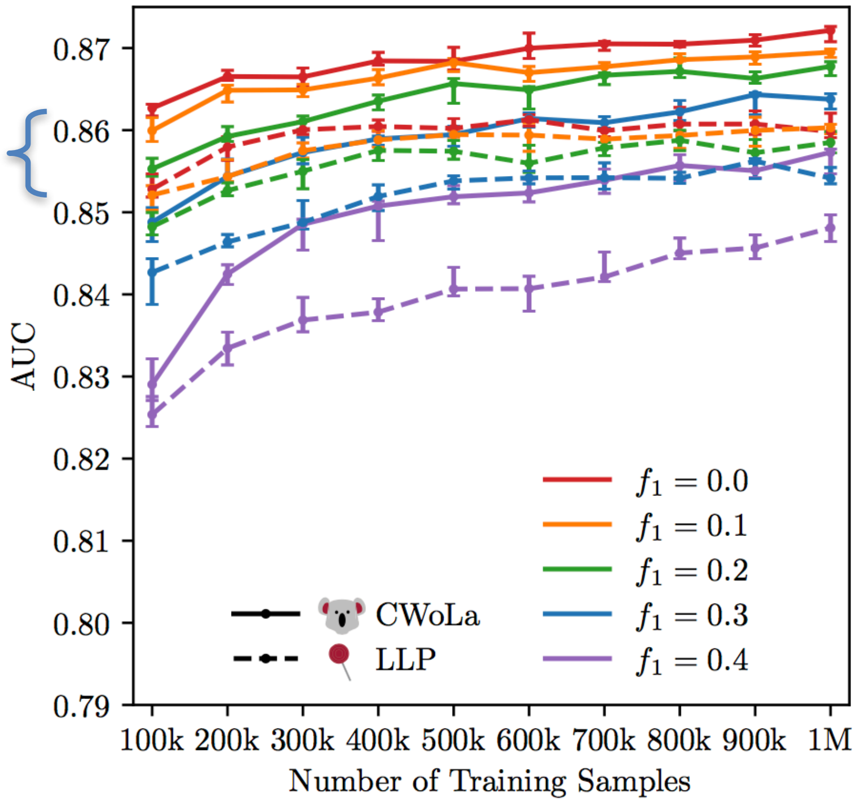
Weak supervision: don't try to unmix samples, just learn discrimination

Cohen, Freytsis, Ostdiek (arXiv:1706.09451)
Metodiev, Nachman, Thaler (arXiv:1708.02949)

Jet images + weak supervision

MDS, Komiske, Metodiev, Nachman (arXiv:1801.10158)

Only ~1% spread

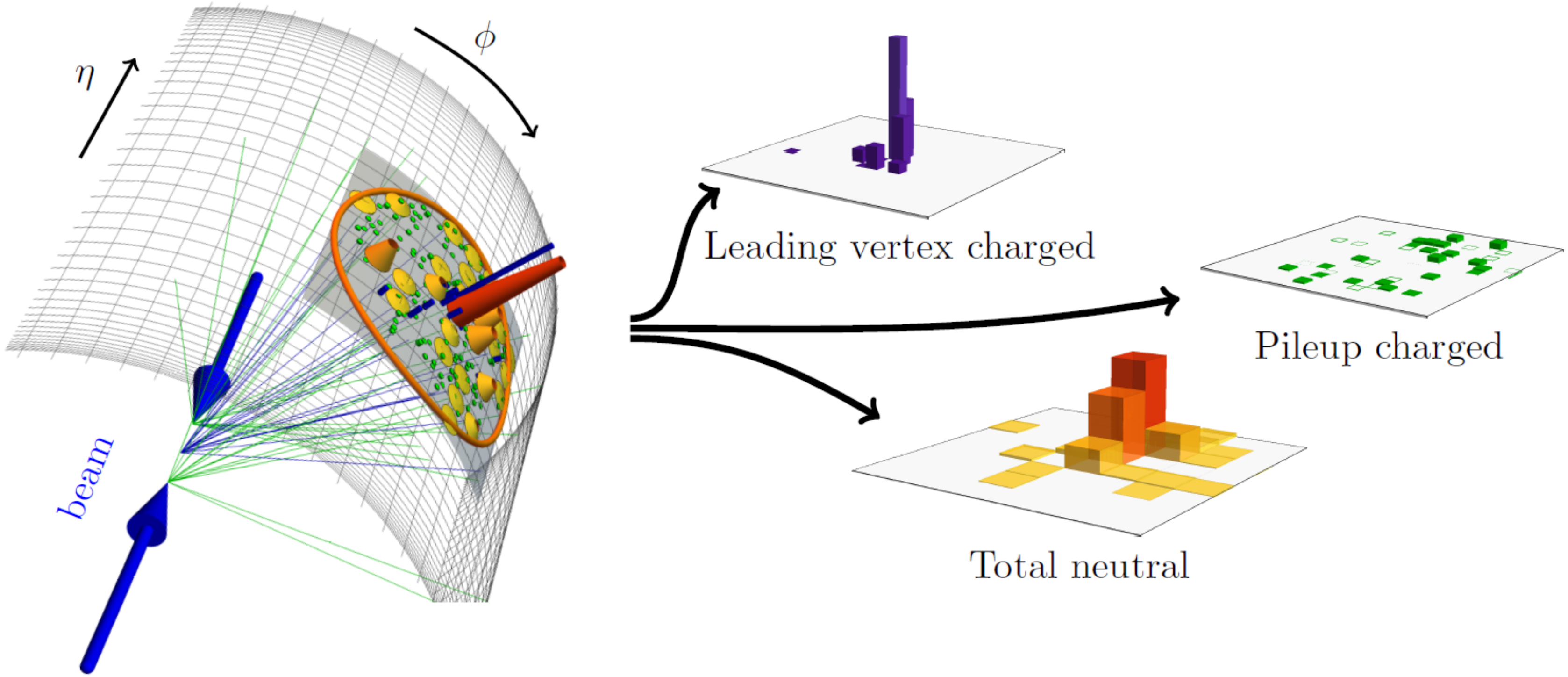


Fully supervised

Weak supervision (Mixed samples)

- Learning from mixed samples without labels as good learning from pure samples
- Labels not needed even for complex inputs

Pileup removal as regression problem



Can measure

1. Leading vertex charged particles
2. Pileup charged particles
3. Total neutral particles

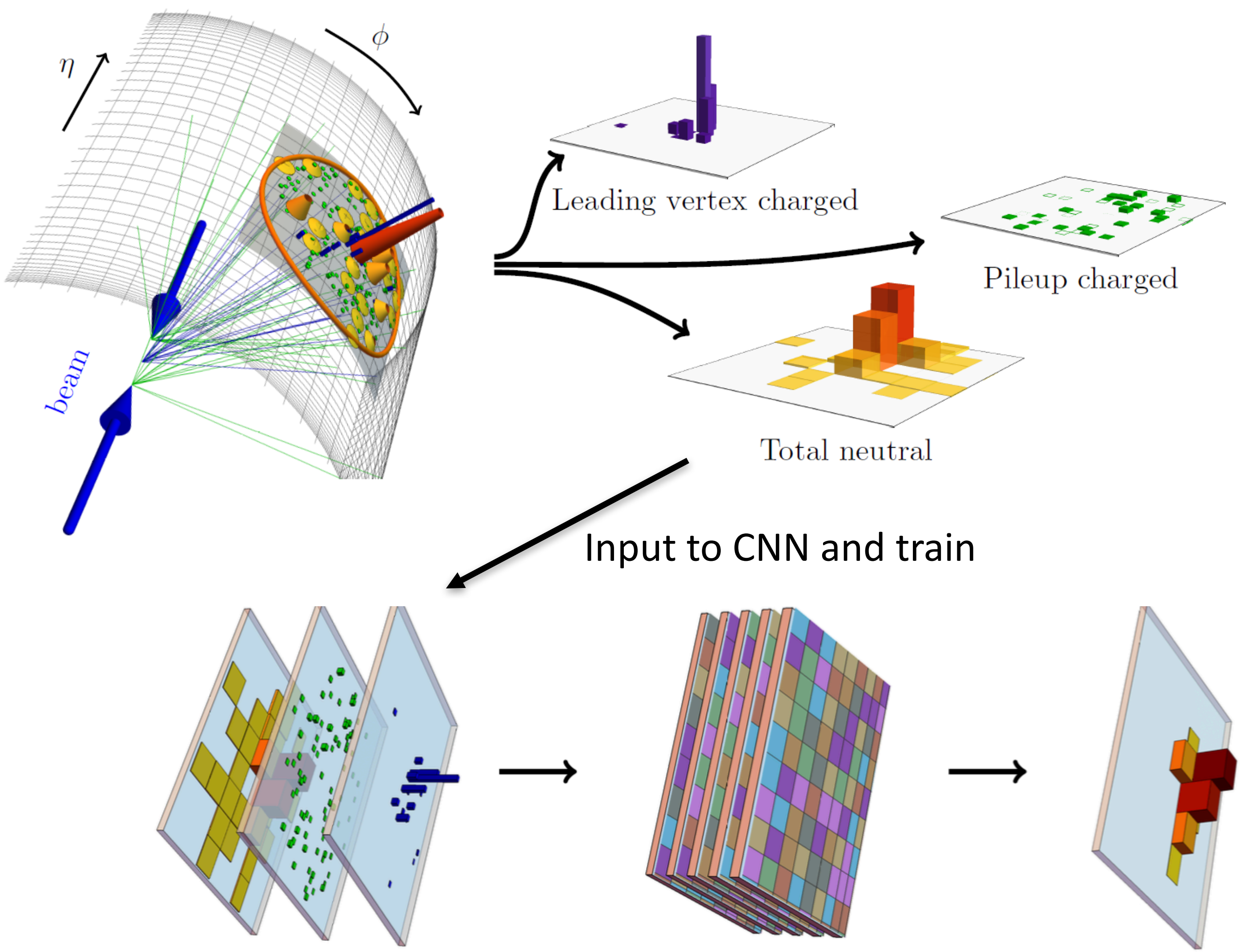
?

Leading vertex neutral particles

CNNs for Pileup Removal

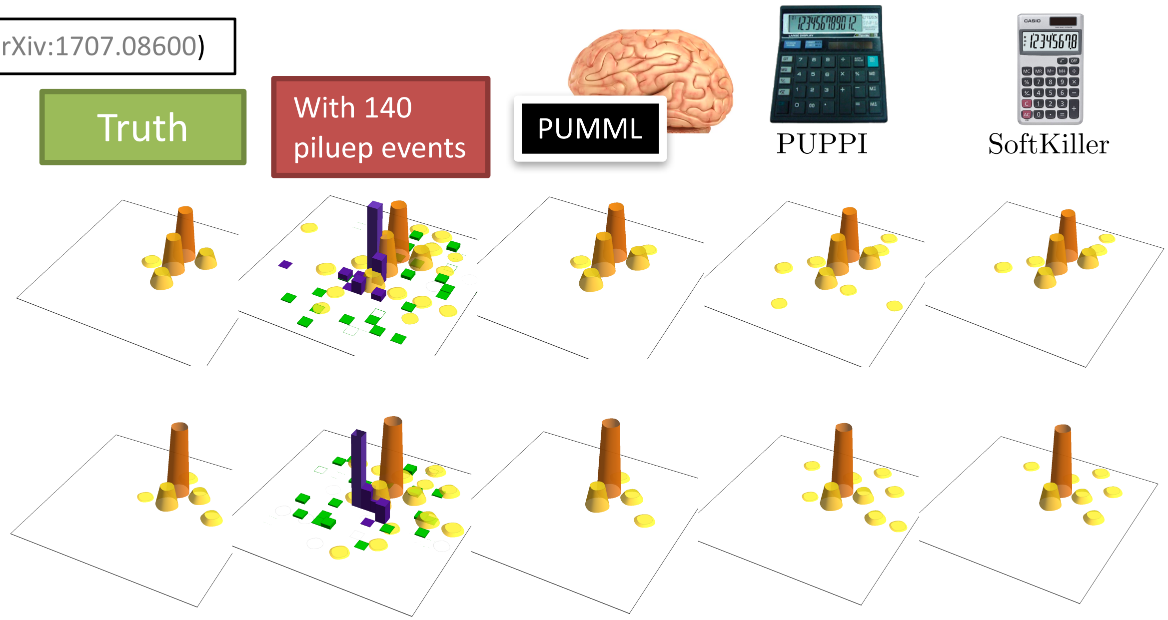
Komiske, Metodiev, Nachman, MDS (arXiv:1707.08600)

- Separate observable energy deposits into 3 images



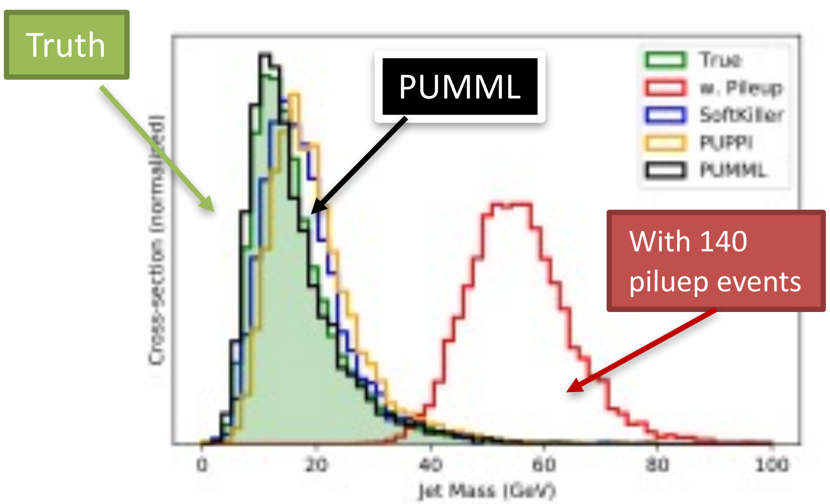
PileUp Mitigation with Machine Learning (PUMML)

Komiske, Metodiev, Nachman, MDS (arXiv:1707.08600)

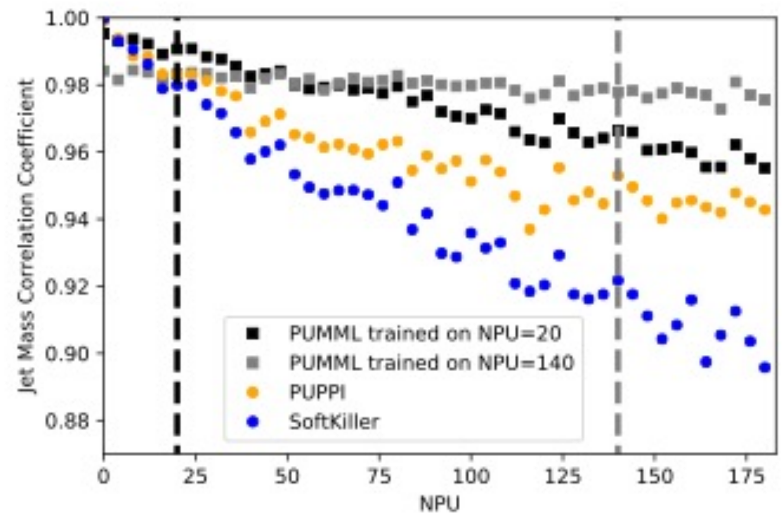


- Excellent leading vertex (truth) reconstruction

- Excellent observable reconstruction



- Excellent stability for variable piluep #



Other ML applications

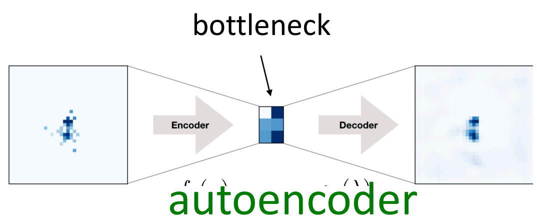
Check out ML-for-Jets conference @ Fermilab (Nov 14-16, 2018):
https://indico.cern.ch/event/745718/timetable

- Model-independent BSM searches

Collis, Howe, Nachman (arXiv:1805.02664)

Heimel et al. (arXiv:1808.08979)

Farina, Nakai, Shih (arXiv:1808.08992)



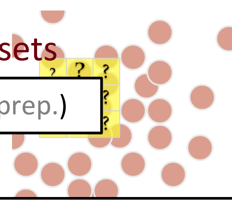
- Jet charge

Fraser, MDS (arXiv:1803.08066)

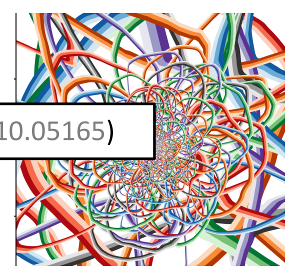
- Strange quark tagging

Nakai, Shih, Thomas (in prep.)

- Jet representations as unordered sets
  - Particle Clouds
  - Energy flow networks

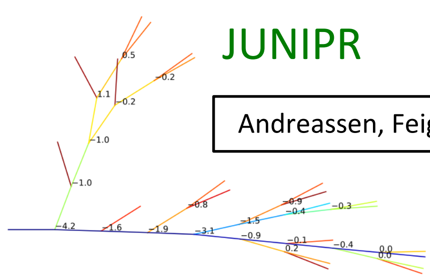


Gouskos, Qu (in prep.)



Komiske, Metodiev, Thaler (arXiv:1810.05165)

- Interpretable network respresentations

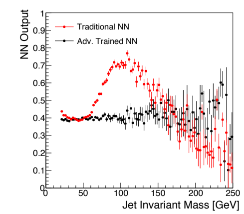


JUNIPR

Andreassen, Feige, Frye, MDS. (arXiv:1804.09720)

- Adversarial networks

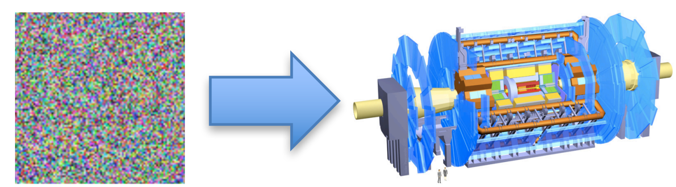
Louppe, Kagan, Cranmer (arXiv:1611.01046)
Shimmin et al. (arXiv:1703.03507)



- Decorrelates NN output from mass

- Generative adversarial networks (GANs)

Oliveira, Paganini, Nachman. (arXiv:1707.08966)



- Phase space sampling

Bothmann, del Debbio (arXiv:1808.07802)

Klimek, Perelstein (arXiv:1810.11509)

Conclusions

- Machine learning is a rapidly growing, exciting area of high energy physics

- Early successes
  - Discrimination:
    - Quark vs Gluon, Boosted Tops, W's, jet charge
  - Data driven discrimination
    - Weak supervision: do not need labeled samples
  - Bump hunting
    - Train on sidebands, look for anomalies
  - Efficiency improvements
    - CaloGAN, PS generators: simulate in microseconds

- Future
  - More testing on data
  - Learning physics: can we learn something we didn't already know?
  - Symbolic ML: can it do QFT?
  - Killer ap for reinforcement learning in HEP?