University of Zurich[UZH]

Physik-Institut

Discriminating GGF and VBF in Higgs production with ML

Nicolas Greiner

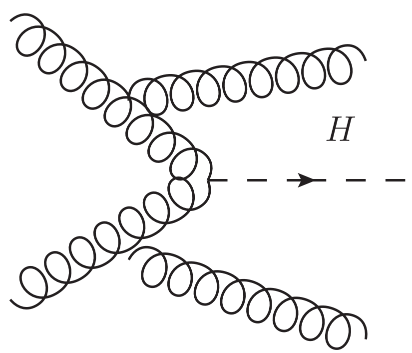
In collaboration with G.Luisoni, J.Winter

FNSNF

Fonds national suisse
Schweizerischer Nationalfonds
Fondo nazionale svizzero
Swiss National Science Foundation

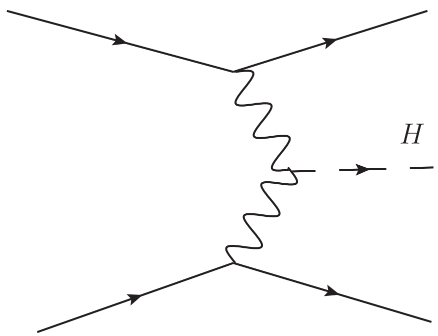
Objective

GGF



VS

VBF



- Gluon fusion dominant contribution (loop induced)
- Vector boson fusion probes couplings of Higgs to heavy gauge bosons
- Treat VBF as signal, GGF as background
- Fraction of VBF can be enhanced by applying ‘VBF-cuts’

Computational setup

- Amplitudes for **HEFT-GGF** and **VBF** computed with **GoSam** + **Sherpa** via BLHA
  [Cullen, v.Deurzen, Heinrich, Luisoni, Mastrolia, Mirabella, Ossola, Peraro, Schlenk, v.Soden-Fraunhofen, Tramontano, NG '14]
  [Gleisberg, Hoeche, Krauss, Schoenherr, Schuhmann]
  - Virtual amplitudes: **GoSam** with **Ninja** [v.Deurzen, Luisoni, Mastrolia, Mirabella, Ossola, Peraro '14]
    Scalar integrals with **OneLoop** [v.Hameren '11]
  - Tree amplitudes and integration with **Sherpa** and **Comix** [Gleisberg, Hoeche '08]
- Phenomenological analysis via generation of **ROOT NTUPLES**
  - **GGF Events: H + 1 / 2 / 3** jets for 8, 13, 14 and 100 TeV
    VBF Events: H + 2 / 3 jets for 13 TeV
  - For kt / anti-kt with R = 0.1 (or larger)
  - Allows for fast analysis, change of **scale, pdf, cuts, jet-tagging**

Physical Setup / Cuts

General settings (GGF HEFT, VBF)

- Scale choice: $\mu_F = \mu_R = \frac{\hat{H}'_T}{2} = \frac{1}{2}\left(\sqrt{m_{\mathrm{H}}^2 + p_{T,\mathrm{H}}^2} + \sum_i |p_{T,i}|\right)$
- PDF: CT14nlo

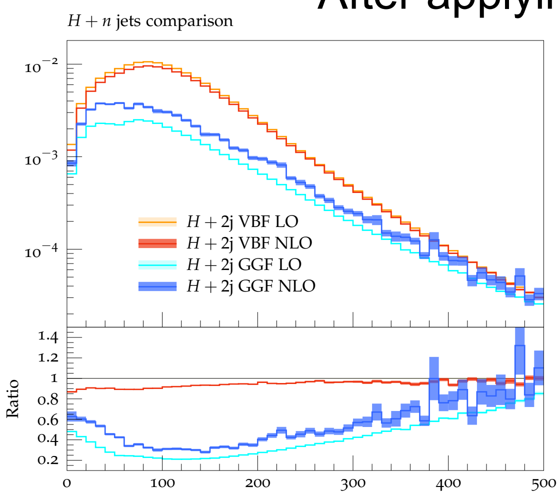
Cuts

- **Baseline cuts:** anti-kt with $R = 0.4, \quad p_T > 30 \text{ GeV}, \quad |\eta| < 4.4$
- **VBF cuts:** $m_{j1,j2} > 400 \text{ GeV}, \quad |\Delta y_{j1,j2}| > 2.8$

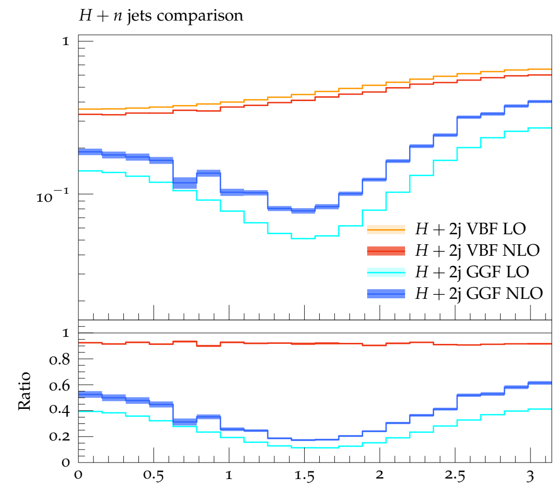
Starting point: LO Ntuples for H+2/3 jets for 13 TeV

Differential distributions for GGF and VBF

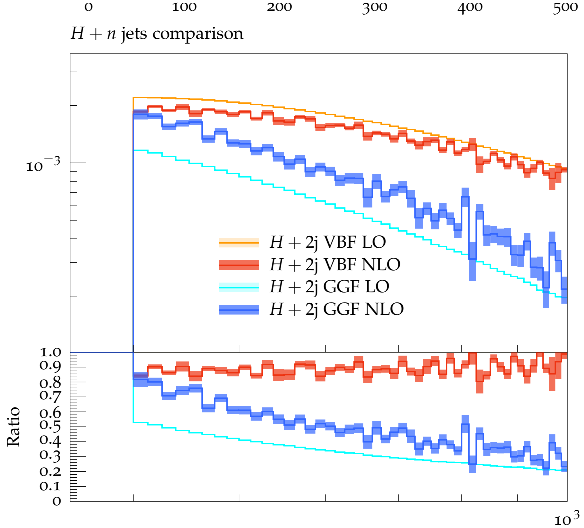
After applying VBF cuts on GGF (**preliminary plots**)



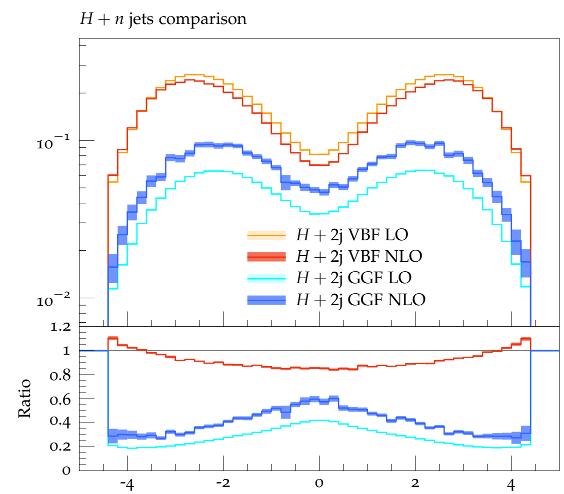
$p_{T,H}$



$\Delta\Phi_{jj}$



m_{jj}



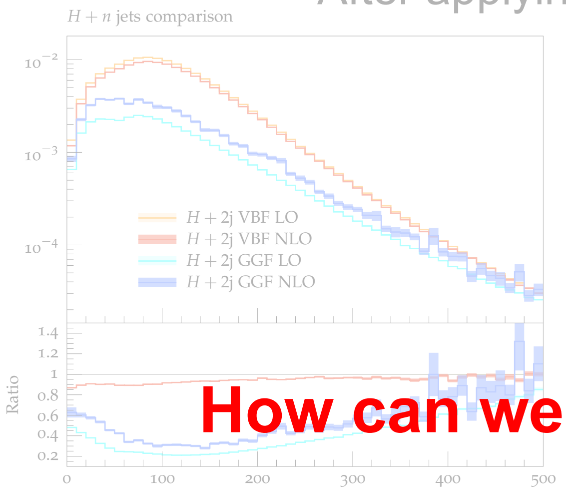
y_{j1}

University of Zurich[UZH]

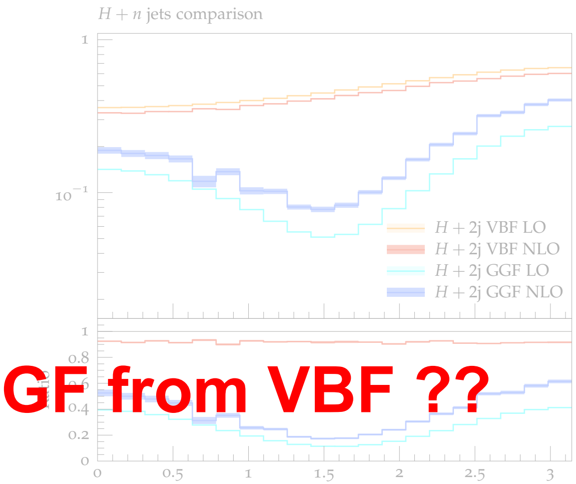
Physik-Institut

Differential distributions for GGF and VBF

After applying VBF cuts on GGF (**preliminary plots**)

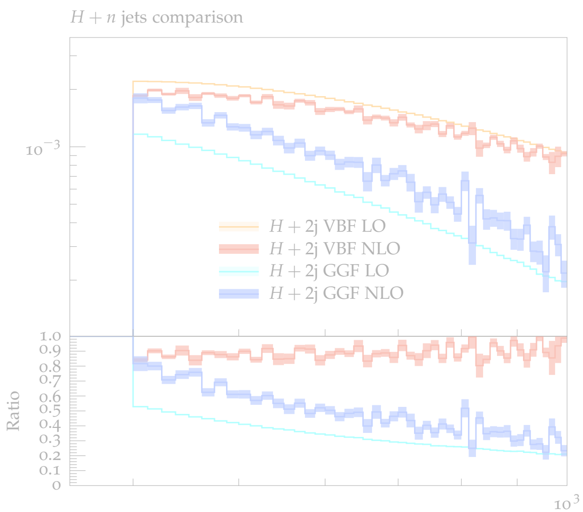


$p_{T,H}$

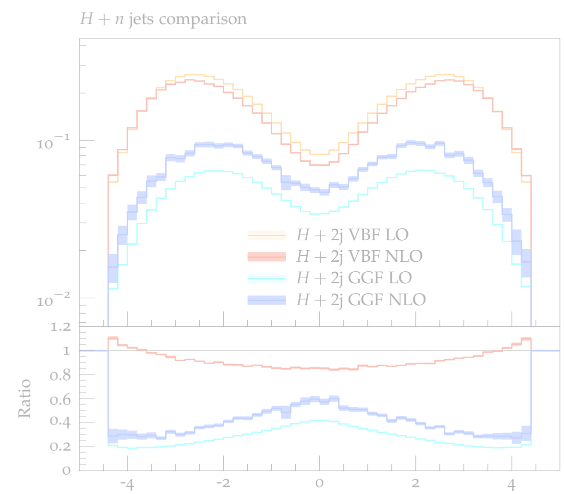


$\Delta\Phi_{jj}$

How can we separate GGF from VBF ??



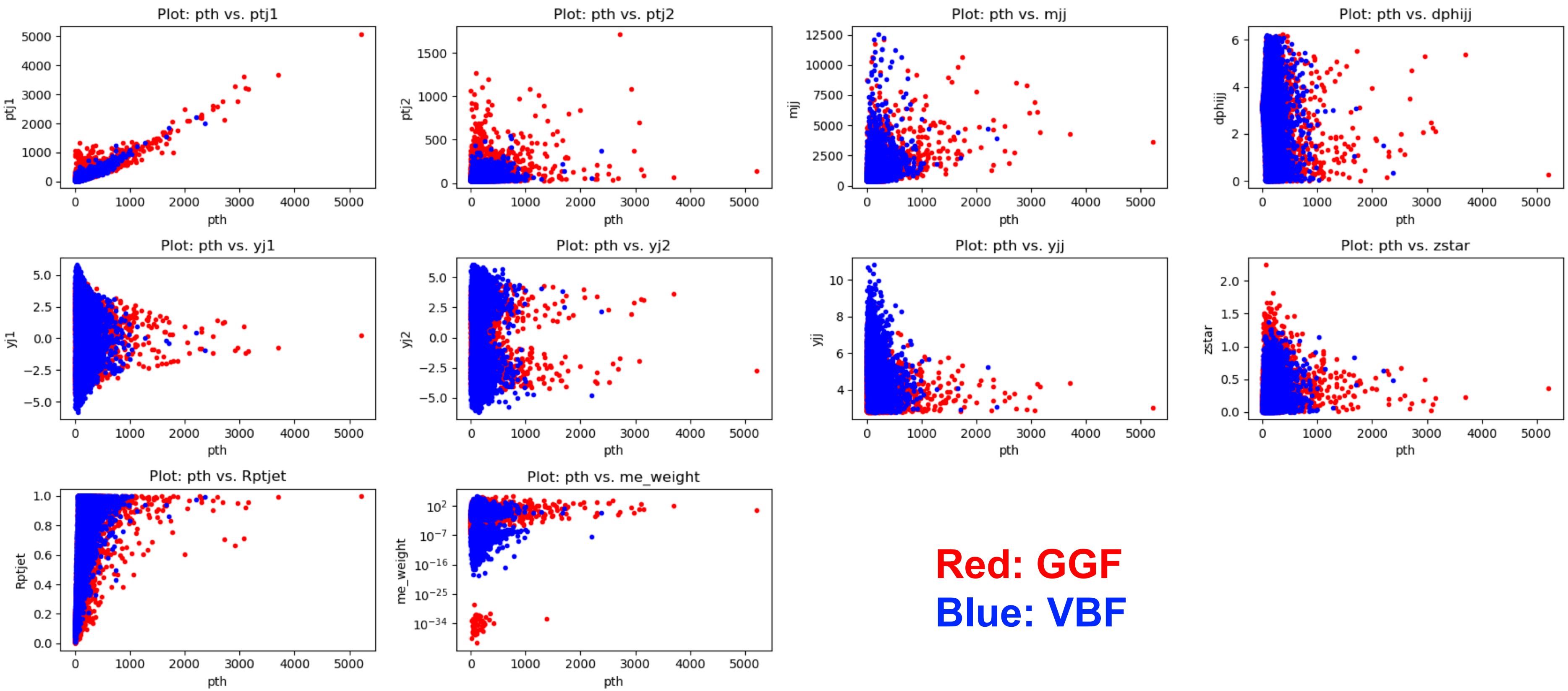
m_{jj}



y_{j1}

2-D plots for H+2j

Example: **Higgs pT** vs. other observables



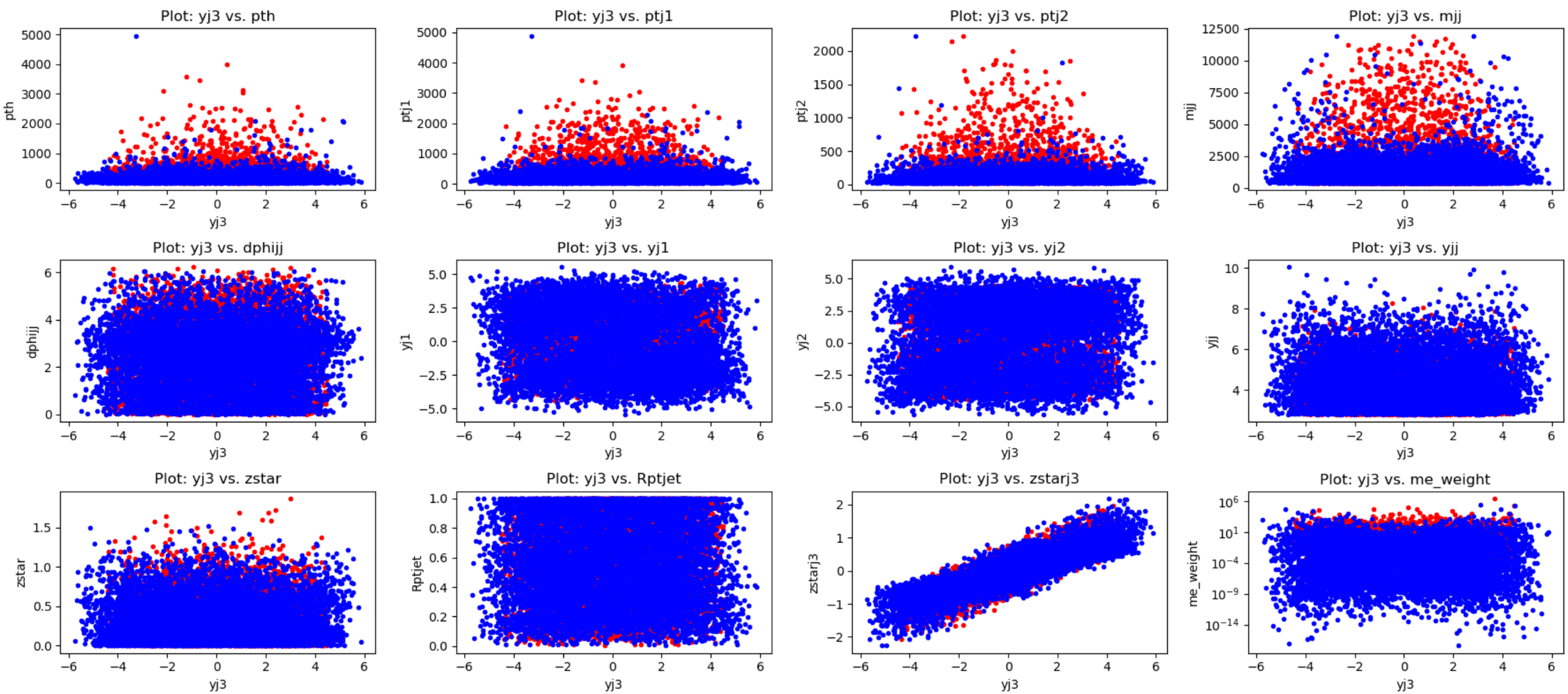
University of Zurich[UZH]

Physik-Institut

2-D plots for H+3j

Red: GGF
Blue: VBF

Example: **Rapidity of 3rd jet** vs other observables



Tensorflow

- Apply **Google's Tensorflow** machine learning algorithms to distinguish between GGF and VBF

- Use ntuples for GGF and VBF as training/ evaluation sets

- Input variables:

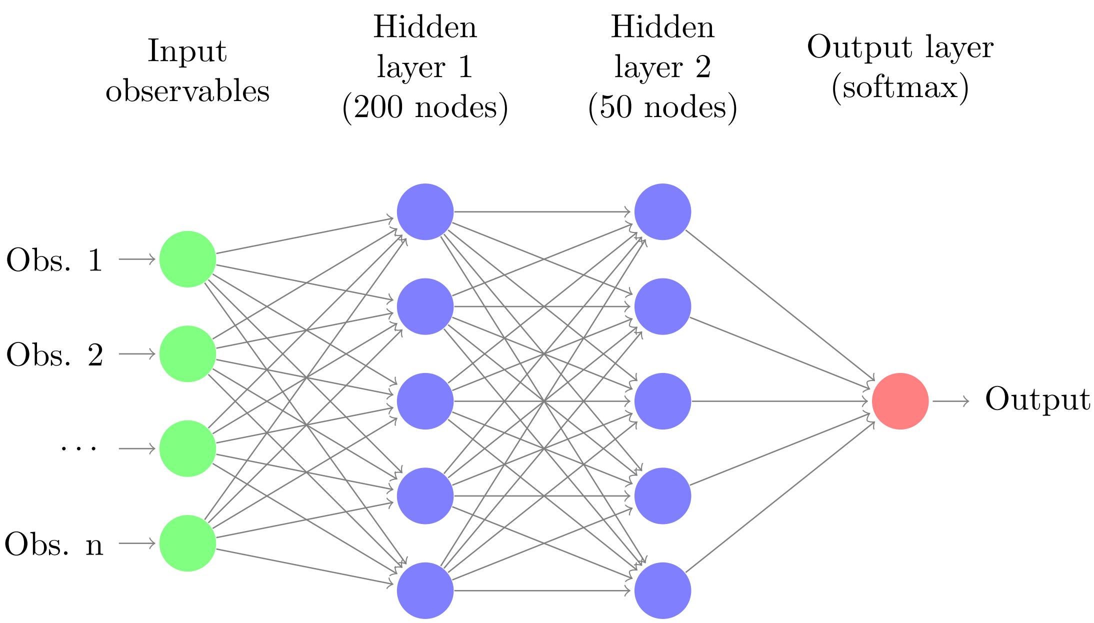
H+2: $p_{T,H}, p_{T,j1,j2}, m_{jj}, \Delta\Phi_{jj}, y_{j1,j2}, \Delta y_{jj}$

$$z^* = \frac{y_H - \left(\frac{y_{j1}+y_{j2}}{2}\right)}{\Delta y_{jj}}, \; R_{p_{T,j}} = \frac{p_{T,j1j2}}{p_{T,j1} + p_{T,j2}}, \; |\mathcal{M}|^2$$

H+3: Additionally:

$$z^*_{j3} = \frac{y_{j3} - \left(\frac{y_{j1}+y_{j2}}{2}\right)}{\Delta y_{j1,j2}}$$

Neural Net with Tensorflow



Output:

GGF: [1,0]
VBF : [0, 1]

$$\mathrm{softmax}(x)_i = \frac{\exp(x_i)}{\sum_j \exp(x_j)}$$

Normalizes result in (0,1)
-> probability

- Internal nodes: linear activation function (relu: rectified linear unit) $f(x) = x^+ = \max(0, x)$
- Adaption of grid: Adam Optimizer
- Learning rate: 0.0001
- Dropout in hidden layer 2: 0.9 (10% of nodes are not used) (increases speed and avoids overtraining)

Training and Testing of NN

- Test/training sample with ~50% GGF and ~50% VBF (approx. 16k events for each process)

- Determine **accuracy** (fraction of correct predictions) every n step by comparing NN output with test sample

- Goal: Minimize **cost function**
  Cross entropy (log-likelihood):

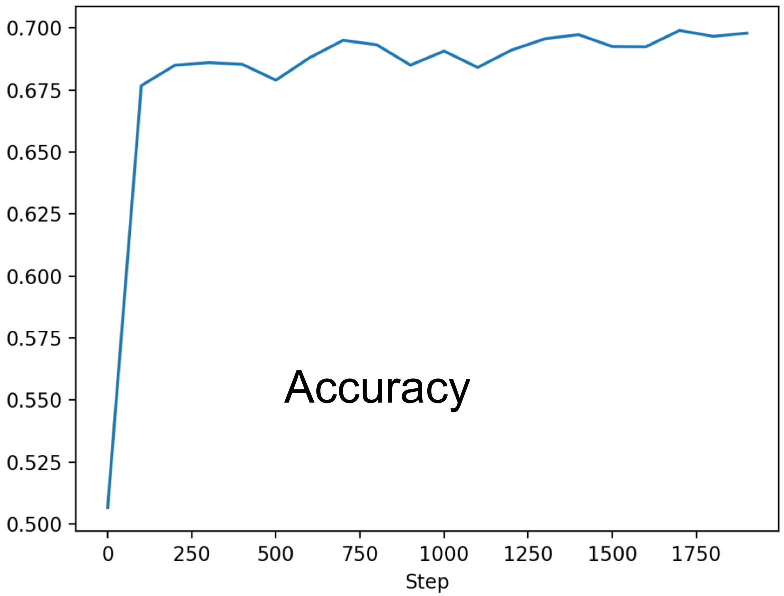
$$H(y', y) := -\frac{1}{n}\sum_{i}^{n} y'_i \ln(y_i)$$

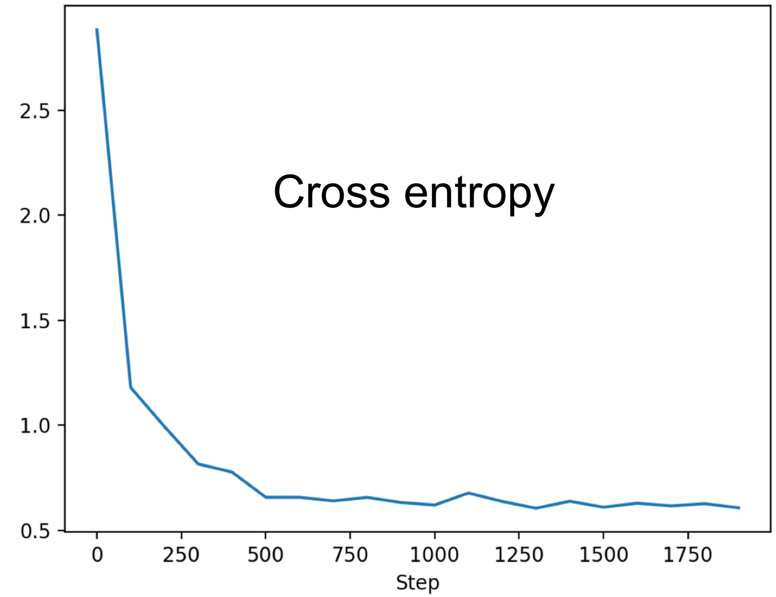
y' : test sample (real/correct)
y : NN output

- 2000 training steps to minimize cost function

- Random shuffle of input events

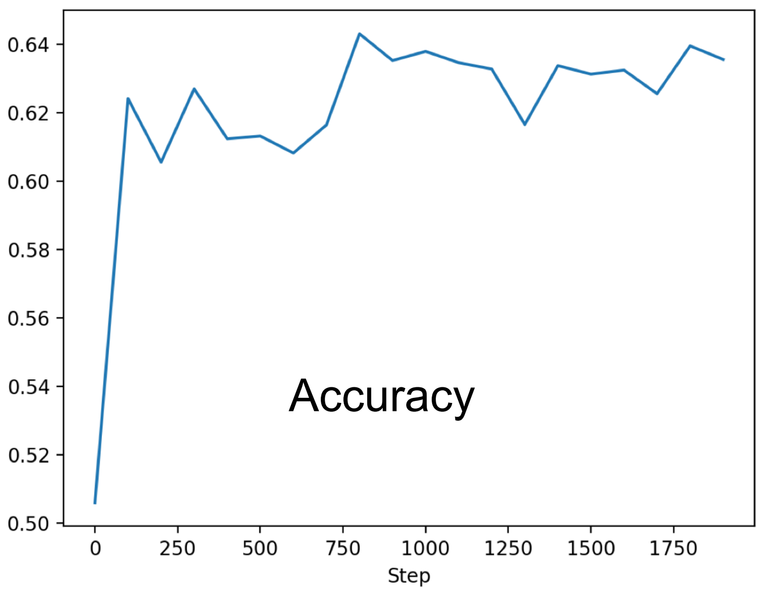
Accuracy and cost function for H+2

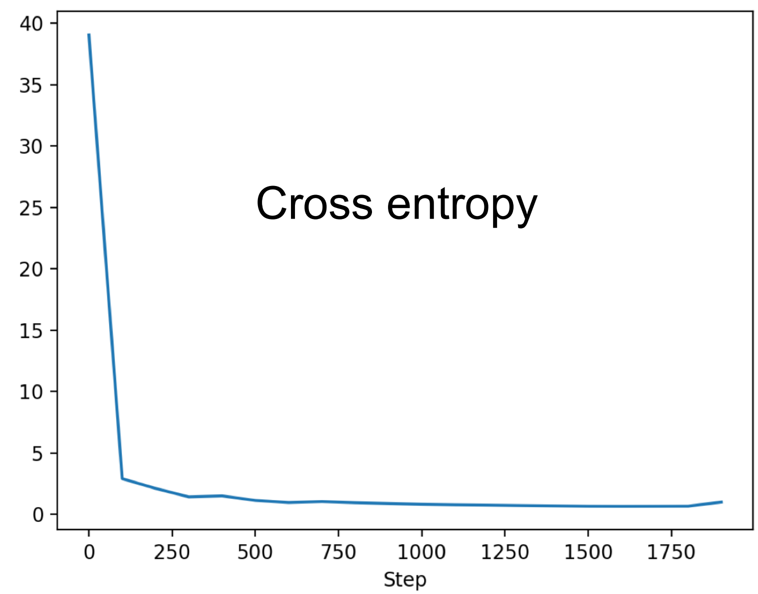




- Accuracy reaches ~70 % for H+2j
- **Note:** Accuracy highly dependent on composition of event sample (fraction of GGF vs VBF)

Accuracy and cost function for H+3

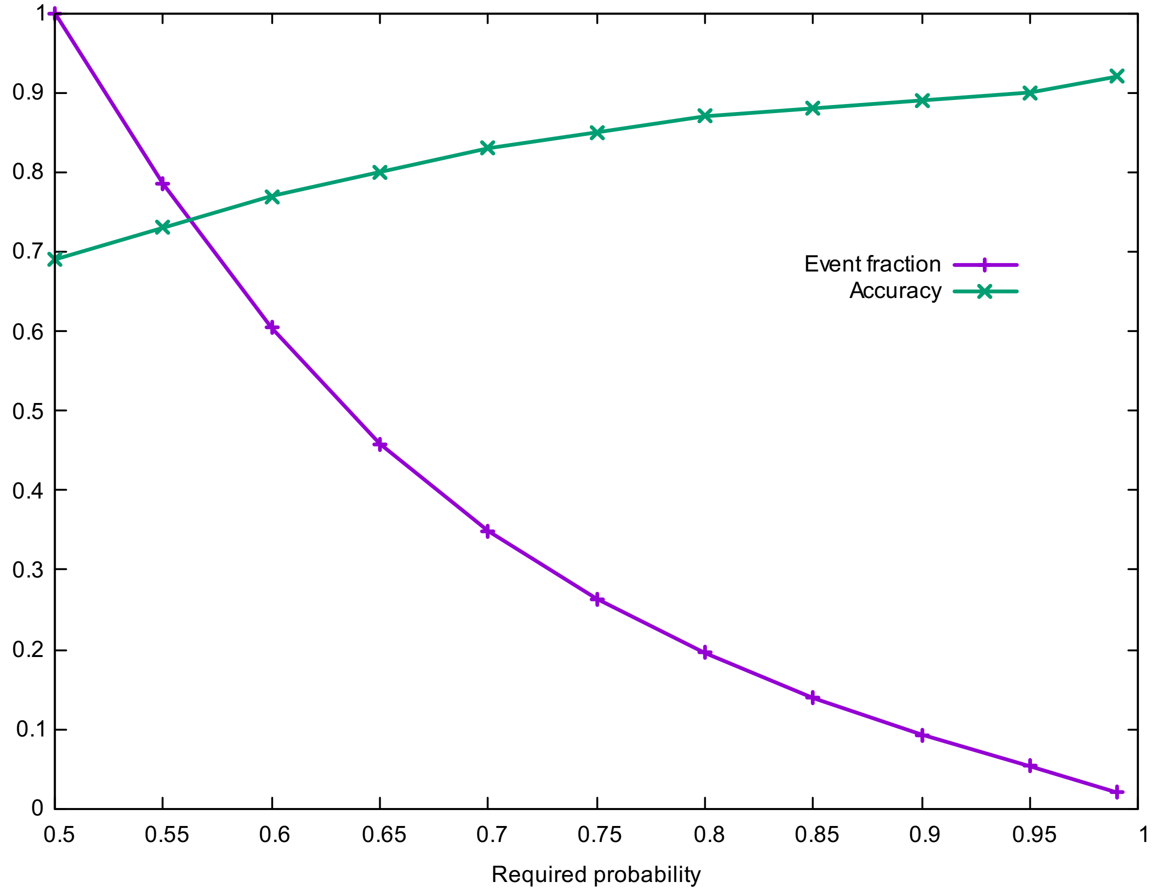




- Accuracy slightly worse than for H+2, ~65%
- In general (H+2/3): Results largely independent on technical details of the NN and on several input parameters

Cut on confidence

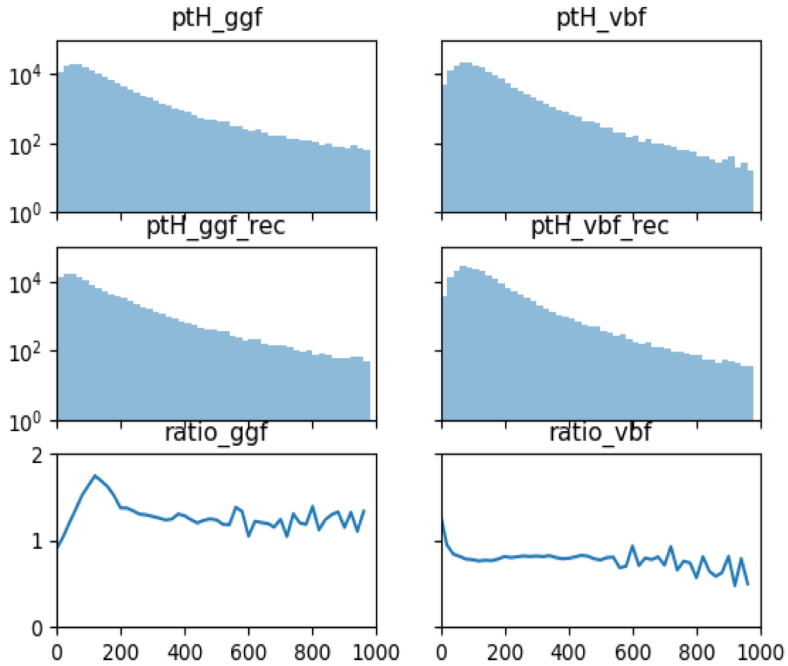
- Output of NN can be interpreted as confidence, how sure the NN is in its decision: [0.49, 0.51] vs [0.99, 0.01]
  -> Investigate cut on confidence level



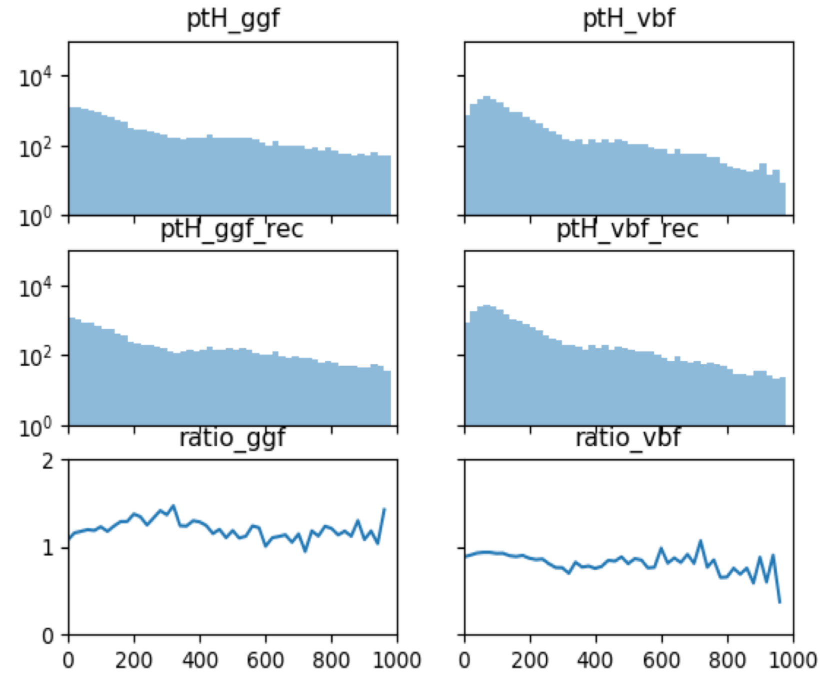
Reconstructing differential distributions

- Mix GGF and VBF and let the NN reconstruct the distributions
- For confidence levels of 0.5 (left) and 0.9 (right) for H+2

Higgs pT



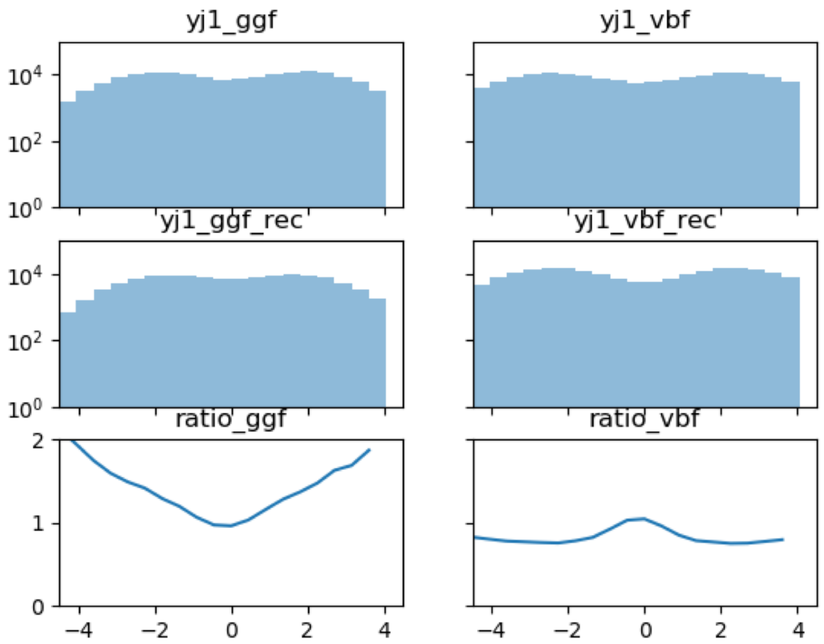
$p > 0.5$

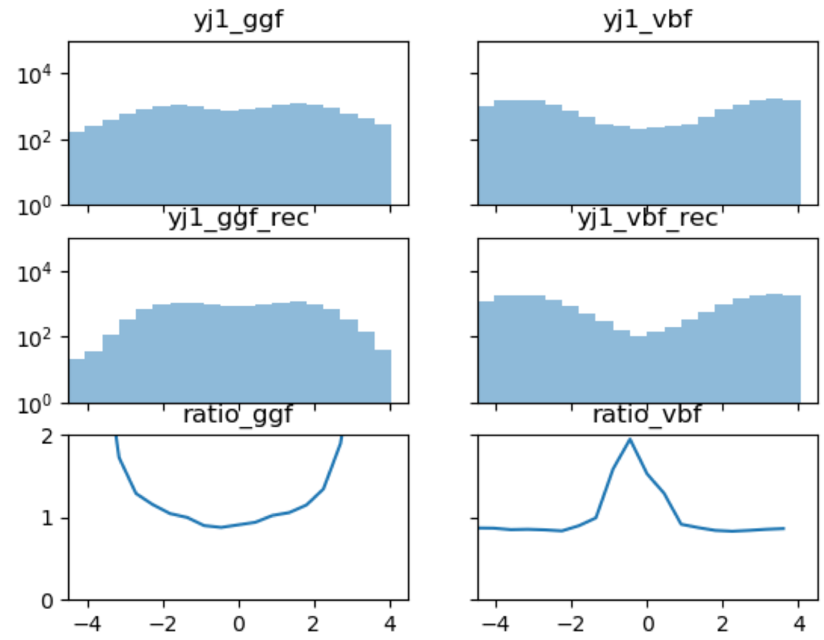


$p > 0.9$

Reconstructing differential distributions

Rapidity of 1st jet





-> Reconstruction different for various observables

Conclusions and Outlook

- Objective: Trying to separate GGF and VBF in the best possible way
- Use ntuples as input and feed them into NN (classification problem)
- What/ how much is a good accuracy ?
- Use weighted cross entropy (GGF classified as VBF gets higher weight): Effect seems to be marginal
- Usefulness of partonic events ?
- Possible development: Change input variables from observables to ‘detector pixels’
  Advantage: Input variables independent of number of particles (also showered events possible)