# Boosting Higgs Pair Production in the $b\bar{b}b\bar{b}$ Final State with Multivariate Techniques

**Çiğdem İşsever**

UNIVERSITY OF OXFORD

based on Behr, Bortoletto, Frost, Issever, Hartland, Rojo, arxiv:1512.08928

IOP Institute of Physics

Joint annual HEPP and APP conference

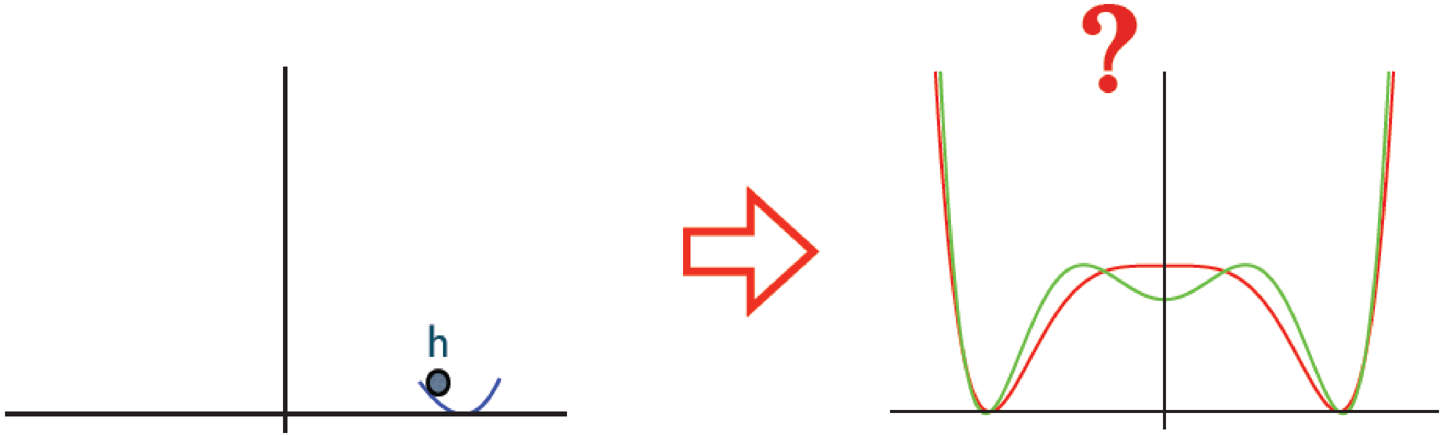
21–23 March 2016, University of Sussex, Brighton, UK

# Probing electroweak symmetry breaking

- Higgs mechanism
  - Postulated shape of potential completely ad-hoc, no first principles
  - It works, but highly unsatisfactory
  - Other EWSB mechanisms conceivable

Current measurements

Di-Higgs production




Arkani-Hamed, Han, Mangano, Wang, arxiv:1511.06495

# Alternative EWSB Mechanisms

Arkani-Hamed, Han, Mangano, Wang, arxiv:1511.06495

- Higgs mechanism:

$$V(h) = m_h^2 h^\dagger h + \frac{1}{2}\lambda\left(h^\dagger h\right)^2 \text{ with } m_h^2 < 0 \text{ and } \lambda > 0$$

- Alternative: negative quadric ↔ positive sextic

$$V(h) \to m_h^2 h^\dagger h + \frac{1}{2}\lambda\left(h^\dagger h\right)^2 + \frac{1}{3!\,\Lambda^2}\left(h^\dagger h\right)^3 \text{ with } \lambda < 0$$

- Alternative: Coleman-Weinberg

$$V(h) \to \frac{1}{2}\lambda\left(h^\dagger h\right)^2 \log\left[\frac{(h^\dagger h)}{m^2}\right]$$

Each leads to different EWSB mechanism,

- with crucial implications for the hierarchy problem,
- the structure of quantum field theory,
- and New Physics at the EW scale

# Alternative EWSB Mechanisms

- Higgs mechanism:

$$\lambda_{HHH}^{SM} = 3\left(m_H^2/v\right)$$

Leading differences show up in the cubic Higgs self-coupling

- Alternative: negative quadric ↔ positive sextic

$$\lambda_{HHH} = 7\left(m_H^2/v\right) = (7/3)\lambda_{HHH}^{SM}$$

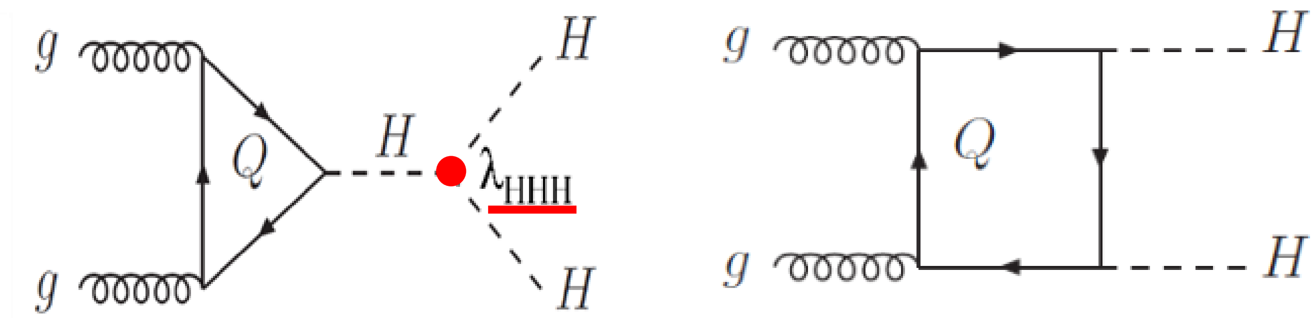
- Alternative: Coleman-Weinberg

$$\lambda_{HHH} = (5/3)\lambda_{HHH}^{SM}$$

Arkani-Hamed, Han, Mangano, Wang, arxiv:1511.06495

# Importance of Higgs Pair Production

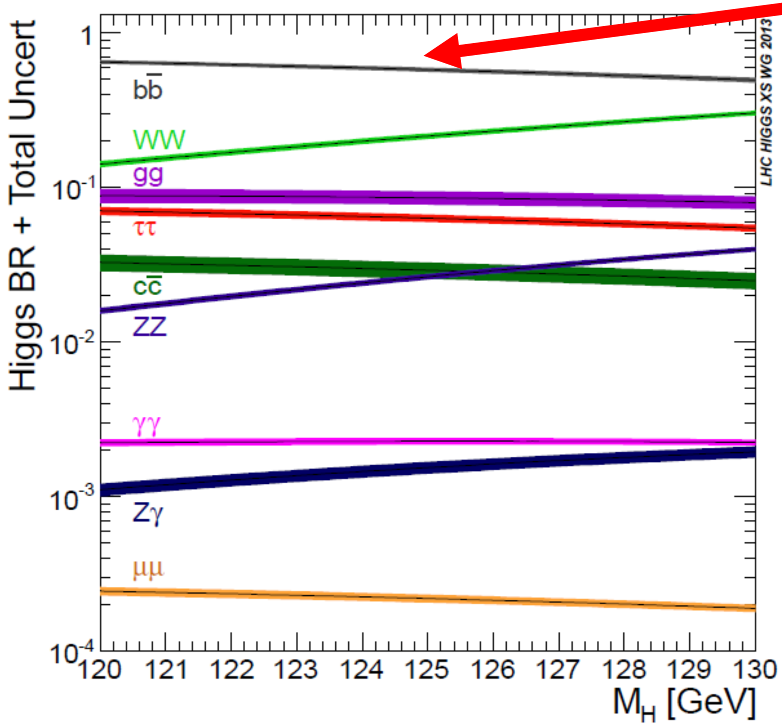
- Uniquely sensitive to Higgs self-coupling, $\lambda_{HHH}$




- In SM, **hh rates are small:** in the leading gluon-fusion production mode, NNLO+NNLL **cross-section at 14 TeV is ~40 fb.** These are further suppressed by the branching fractions**!**

# Pros and Cons of $b\bar{b}b\bar{b}$ final state

- **Pros**: **largest branching ratio**: BR($H \rightarrow bb$) ~ 0.57



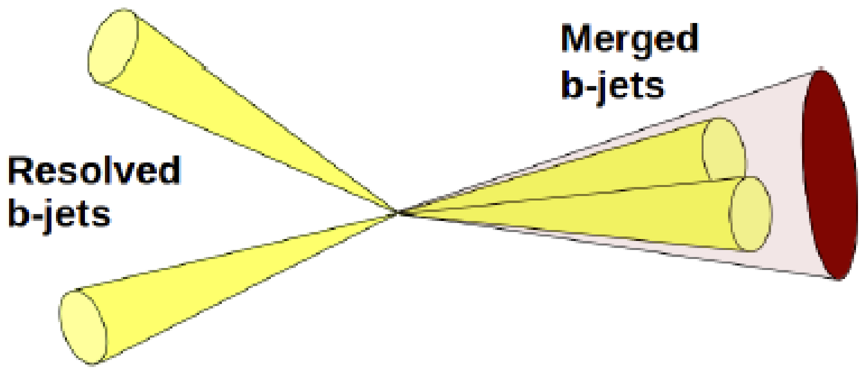

arXiv:1307.1347 and PDG

- **Cons:** Huge QCD multi-jets background
  - Previous studies: $S/\sqrt{B}$ ~ 2 @ HL-LHC (no PU, missing relevant backgrounds)

# Analysis Strategy

[arxiv:1512.08928], submitted to EPJC

- **Multivariate** + loose cut based analysis techniques
- Use all di-Higgs decay topologies



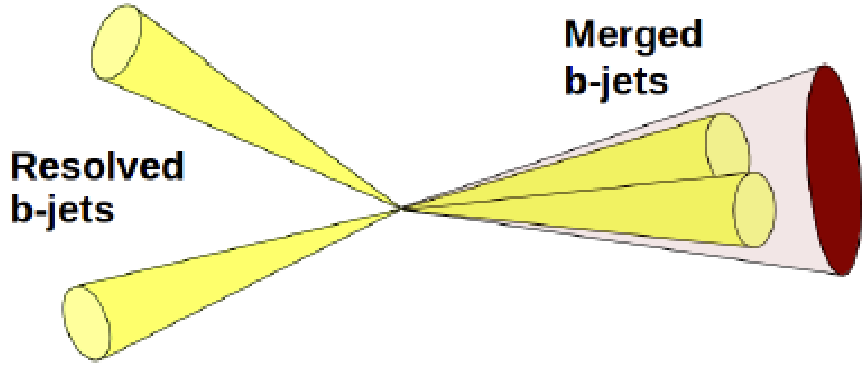

- Model pile-up (PU) effects: 80 (baseline) and 150
- Include all relevant backgrounds, esp. 2b2j

| Process | Generator | $N_{\rm evt}$ | $\sigma_{\rm LO}$ (pb) | $K$-factor |
|---|---|---|---|---|
| $pp \to hh \to 4b$ | MadGraph5_aMC@NLO | 1M | $6.2 \cdot 10^{-3}$ | 2.4 (NNLO+NNLL) |
| $pp \to b\bar{b}b\bar{b}$ | SHERPA | 3M | $1.1 \cdot 10^{3}$ | 1.6 (NLO) |
| $pp \to b\bar{b}jj$ | SHERPA | 3M | $2.7 \cdot 10^{5}$ | 1.3 (NLO) |
| $pp \to jjjj$ | SHERPA | 3M | $9.7 \cdot 10^{6}$ | 0.6 (NLO) |
| $pp \to t\bar{t} \to b\bar{b}jjjj$ | SHERPA | 3M | $2.5 \cdot 10^{3}$ | 1.4 (NNLO+NNLL) |

# Higgs identification (tagging)

- Separation of b-jets shrinks: $\Delta R(b, \overline{b}) \sim 2m^{H}/p_{T}^{H}$

  → Single large-R jet for $p_T^H > 250\ \mathrm{GeV}$




**Resolved Higgs**

- 2 AKT4 jets
- $p_T > 40$ GeV and $|\eta| < 2.5$

**Merged Higgs**

- 1 AKT10 jet
- $p_T > 200$ GeV, $|\eta| < 2.5$
- Higgs tagging: BDRS mass-drop tagger [arxiv:0802:2470]

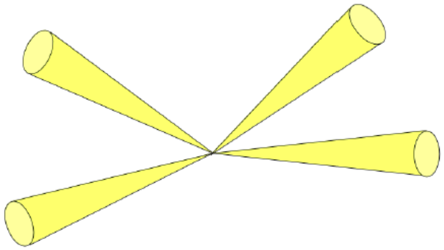
**b-tagging efficiencies**

- b-jet: $f_b$=0.8
- c-jet: $f_c$=0.1
- light-jet: $f_l$=0.01

**Large-R jet b-tagged if two associated b-tagged small-R jets**

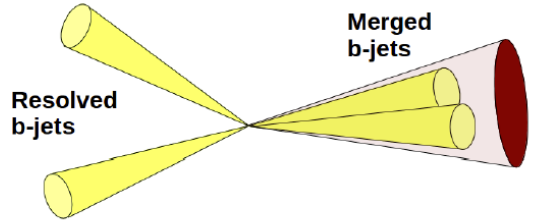
# Higgs Topologies

## Resolved



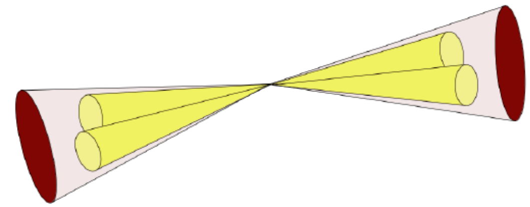
- **≥ 4 b-tagged small-R jets**
- Di-Higgs reconstructed from 4 leading jets
- Minimize mass difference between dijet systems

## Intermediate




- **=1 large-R jet (b- and Higgs tagged)**
- **≥ 2 b-tagged AKT4 jets**
- **ΔR>1.2 w.r.t large-R jet**
- Higgs reconstructed from 2 leading small-R jets
- Minimize mass difference between dijet system and large-R jet

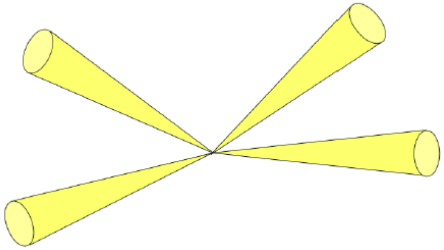
## Boosted



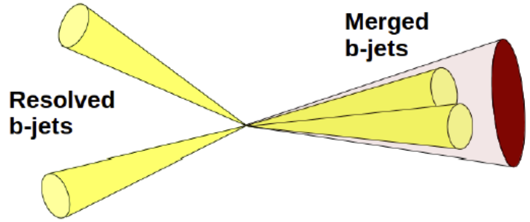
- **≥ 2 b- and Higgs tagged large-R jets**
- Leading 2 jets taken as Higgs candidates

# Higgs Topologies

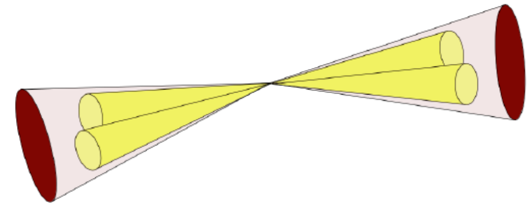
**Resolved**



**Intermediate**




**Boosted**



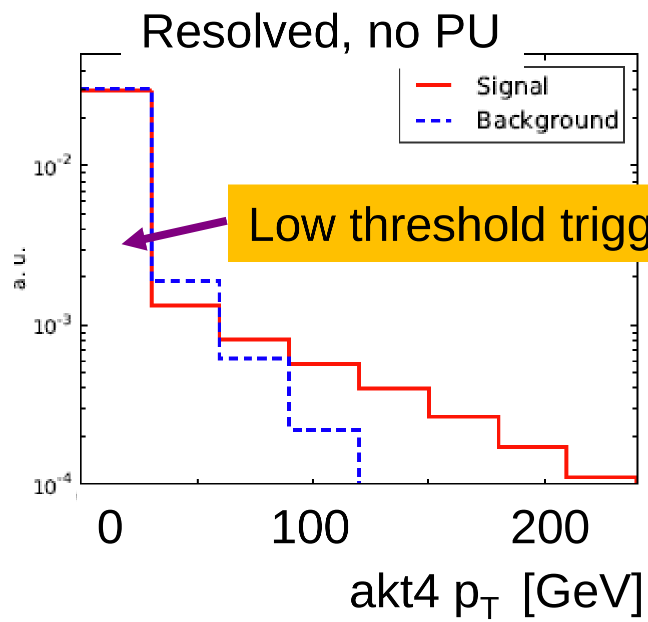
**Rank categories by S/√B to make them exclusive:**

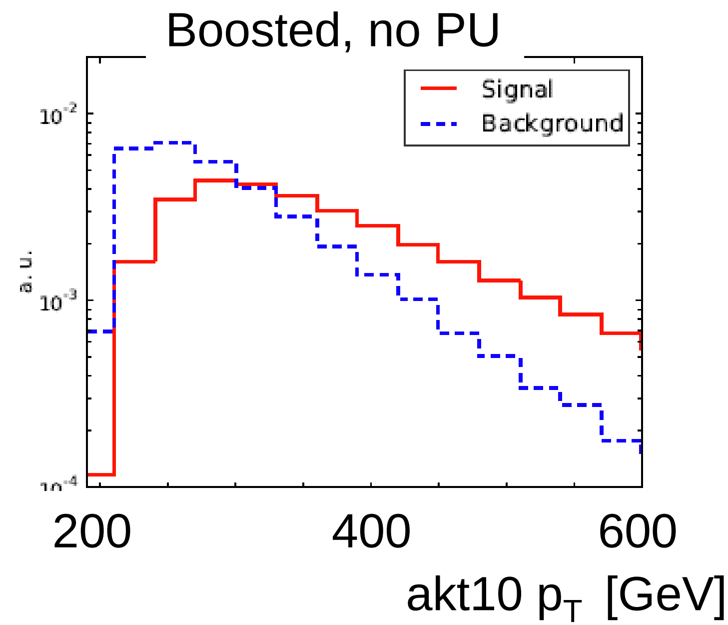
**Resolved < Intermediate < Boosted**

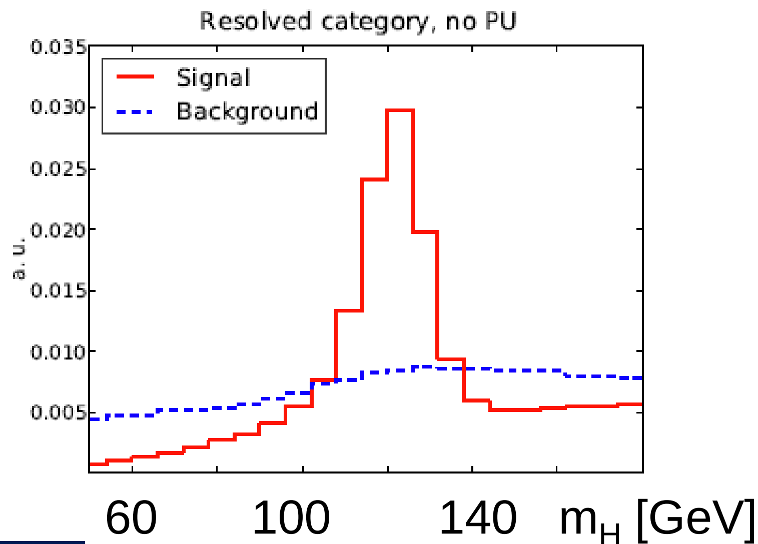
**Loose mass cut:** $|\mathbf{m_H} - 125\ \mathbf{GeV}| < 40\ \mathbf{GeV}$

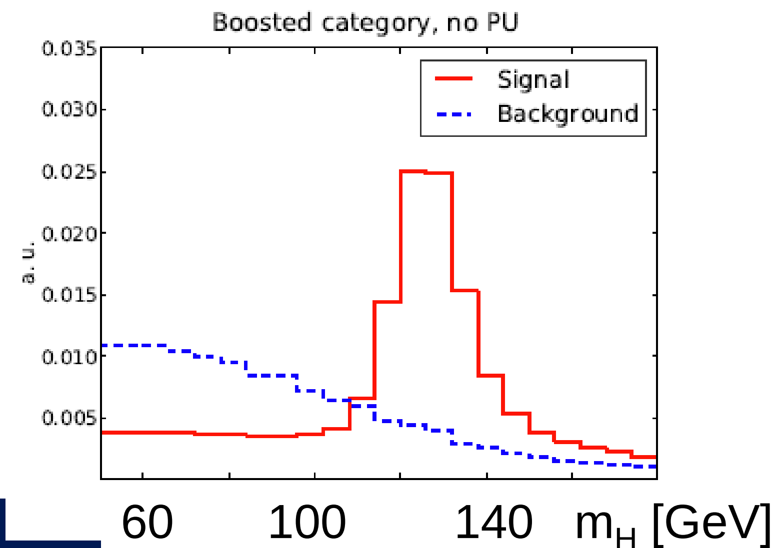
# Pre-MVA distributions

Signal
Background



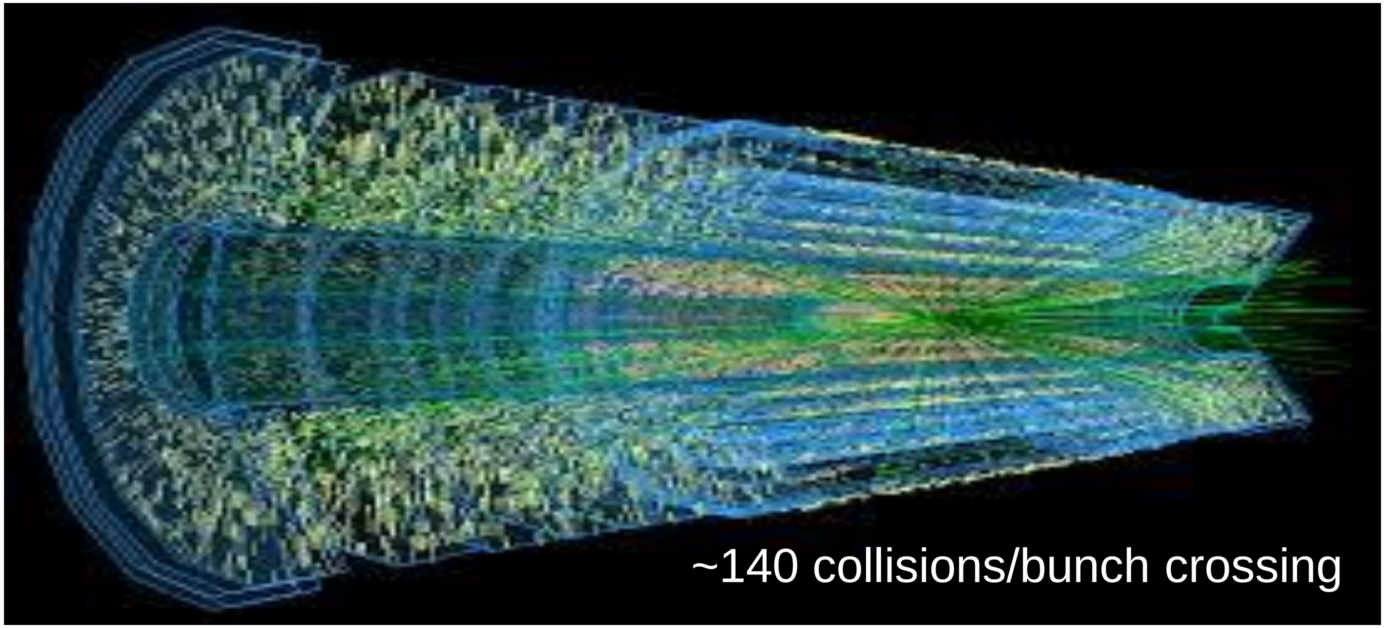



# Pile-Up modelling

- Up to 150 interactions per bunch-crossing @ HL-LHC
- Impact on signal significance:
  - Additional (hard) jets in reconstruction
  - Affects mass and substructure of large-*R* jets
- **Not taken into account in previous studies**

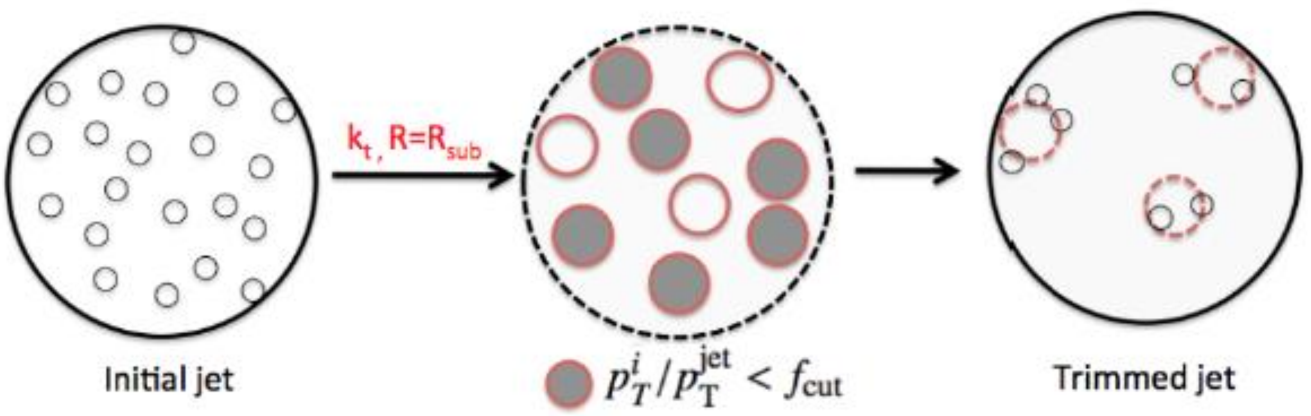


# Pile-Up modelling and mitigation

- Superimpose ***n*****(PU) = 80 Minimum Bias events** on each signal/background event
  - Similar results for PU = 150

- Pile-up mitigation
  - Event level: Remove soft particles using **SoftKiller** [arXiv:1407.0408]
  - Jet level: Apply trimming [arXiv:0912.1342] (large-*R* jets only)

JHEP09 (2013) 076




$R_{sub}=0.2$
$f_{cut}=5\%$

# Impact of Pile-up on Higgs mass reconstruction

| Resolved category | | | |
|---|---|---|---|
| | | $\langle m_h^{\text{reco}} \rangle - m_h$ | $\sigma_{m_h}$ |
| no PU | leading $h$ | -3.8 GeV | $(8.5 \pm 0.2)$ GeV |
| | subleading $h$ | -5.8 GeV | $(9.1 \pm 0.3)$ GeV |
| PU80 | leading $h$ | +33 GeV | $(8.8 \pm 1.5)$ GeV |
| | subleading $h$ | +31 GeV | $(11.7 \pm 3.3)$ GeV |
| PU80+SK | leading $h$ | +3.9 GeV | $(10.7 \pm 0.3)$ GeV |
| | subleading $h$ | +2.1 GeV | $(10.5 \pm 0.3)$ GeV |

No PU:
~ 9 GeV

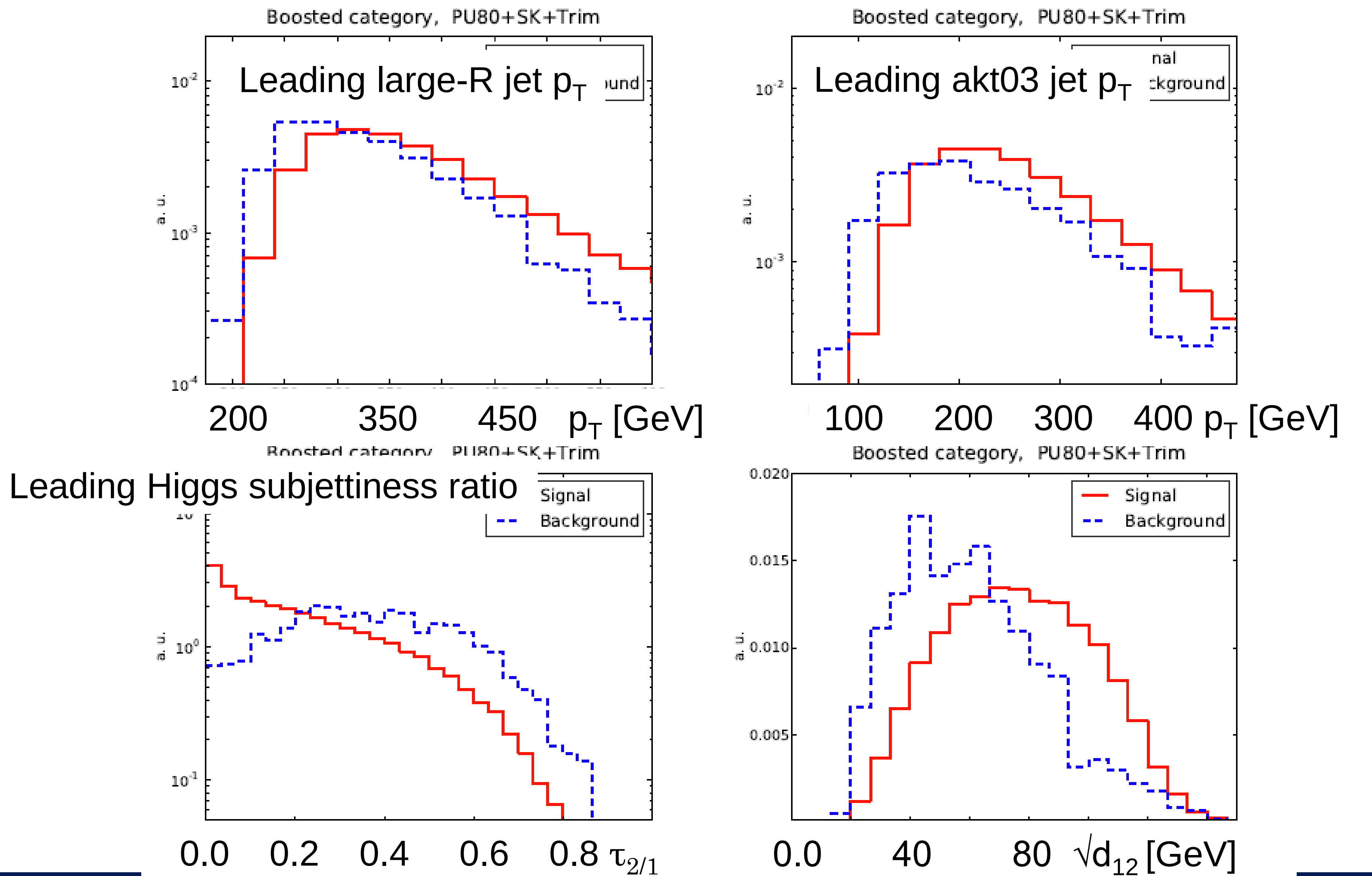
80 PU:
~ 11 GeV

| Boosted category | | | |
|---|---|---|---|
| | | $\langle m_h^{\text{reco}} \rangle - m_h$ | $\sigma_{m_h}$ |
| no PU | leading $h$ | +2.0 GeV | $(8.2 \pm 0.5)$ GeV |
| | subleading $h$ | +1.0 GeV | $(8.8 \pm 0.5)$ GeV |
| PU80+SK+Trim | leading $h$ | -2.2 GeV | $(8.7 \pm 0.7)$ GeV |
| | subleading $h$ | -4.9 GeV | $(9.0 \pm 0.8)$ GeV |

No PU:
~ 9 GeV

80 PU:
~ 9 GeV

# Boosted with 80 PU + Softkiller + Trimming



# Pre-MVA 2b2j background -- BIG

| HL-LHC, Resolved category, PU+SK with $n_{PU} = 80$ | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Cross-section [fb] | | | | | | $S/B$ | | $S/\sqrt{B}$ | |
| | $hh4b$ | total bkg | $4b$ | $2b2j$ | $4j$ | $t\bar{t}$ | tot | $4b$ | tot | $4b$ |
| C1a | 11 | $4.4 \cdot 10^8$ | $1.5 \cdot 10^5$ | $3.0 \cdot 10^7$ | $4.1 \cdot 10^8$ | $2.6 \cdot 10^5$ | $2.4 \cdot 10^{-8}$ | $7.2 \cdot 10^{-5}$ | 0.03 | 1.5 |
| C1b | 11 | $4.4 \cdot 10^8$ | $1.5 \cdot 10^5$ | $3.0 \cdot 10^7$ | $4.1 \cdot 10^8$ | $2.6 \cdot 10^5$ | $2.4 \cdot 10^{-8}$ | $7.2 \cdot 10^{-5}$ | 0.03 | 1.5 |
| C1c | 3 | $1.1 \cdot 10^8$ | $4.2 \cdot 10^4$ | $7.7 \cdot 10^6$ | $9.9 \cdot 10^7$ | $1.1 \cdot 10^5$ | $2.8 \cdot 10^{-8}$ | $7.4 \cdot 10^{-5}$ | 0.02 | 0.8 |
| C2 | 0.6 | $9.0 \cdot 10^3$ | $3.5 \cdot 10^3$ | $5.1 \cdot 10^3$ | $3.1 \cdot 10^2$ | 50 | $6.5 \cdot 10^{-5}$ | $1.7 \cdot 10^{-4}$ | 0.4 | 0.5 |

| 4b | 2b2j | 4j |
|---|---|---|
| $3.5 \cdot 10^3$ | $5.1 \cdot 10^3$ | $3.1 \cdot 10^2$ |

*b*-quark radiation in parton shower

→ non-negligible fraction of 2*b*2*j* events with ≥ 2 *b*-jets

Additional light jets from parton shower

→ Not all of the 4 leading jets in 4*b* events are *b*-jets

**Good control of light and charm jet fake rates impotant!**

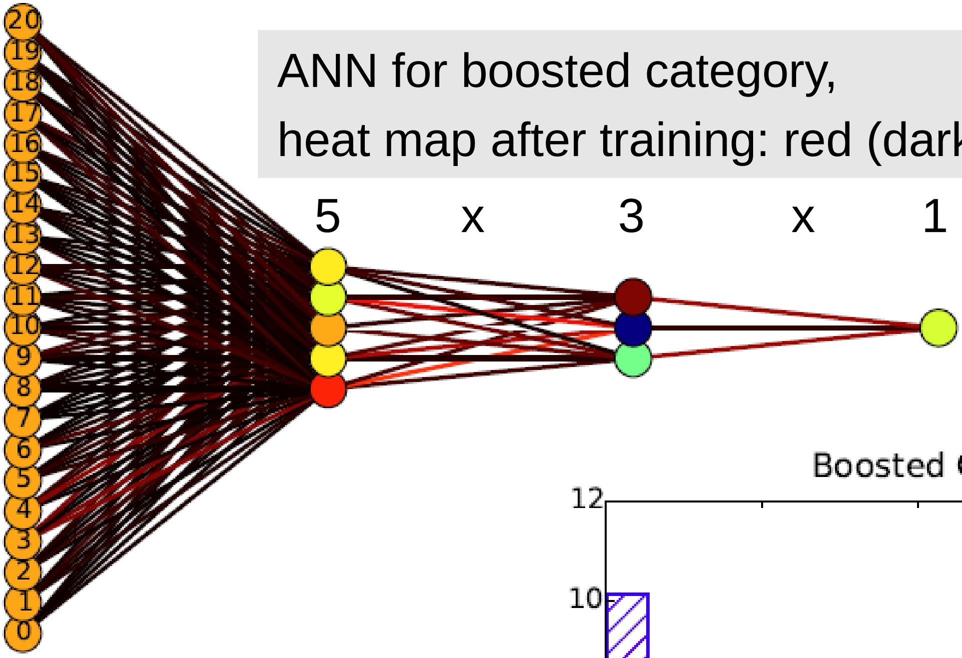
# Multivariate Analysis Techniques

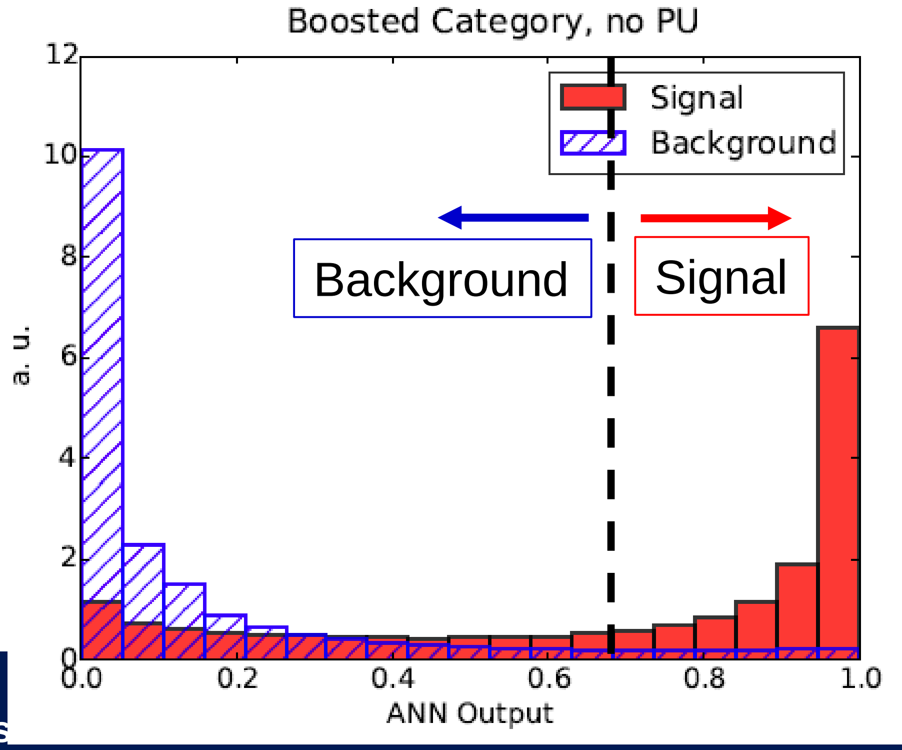
## Multi-layer feed-forward artifical neural net (ANN) - *perceptron*

Input variables

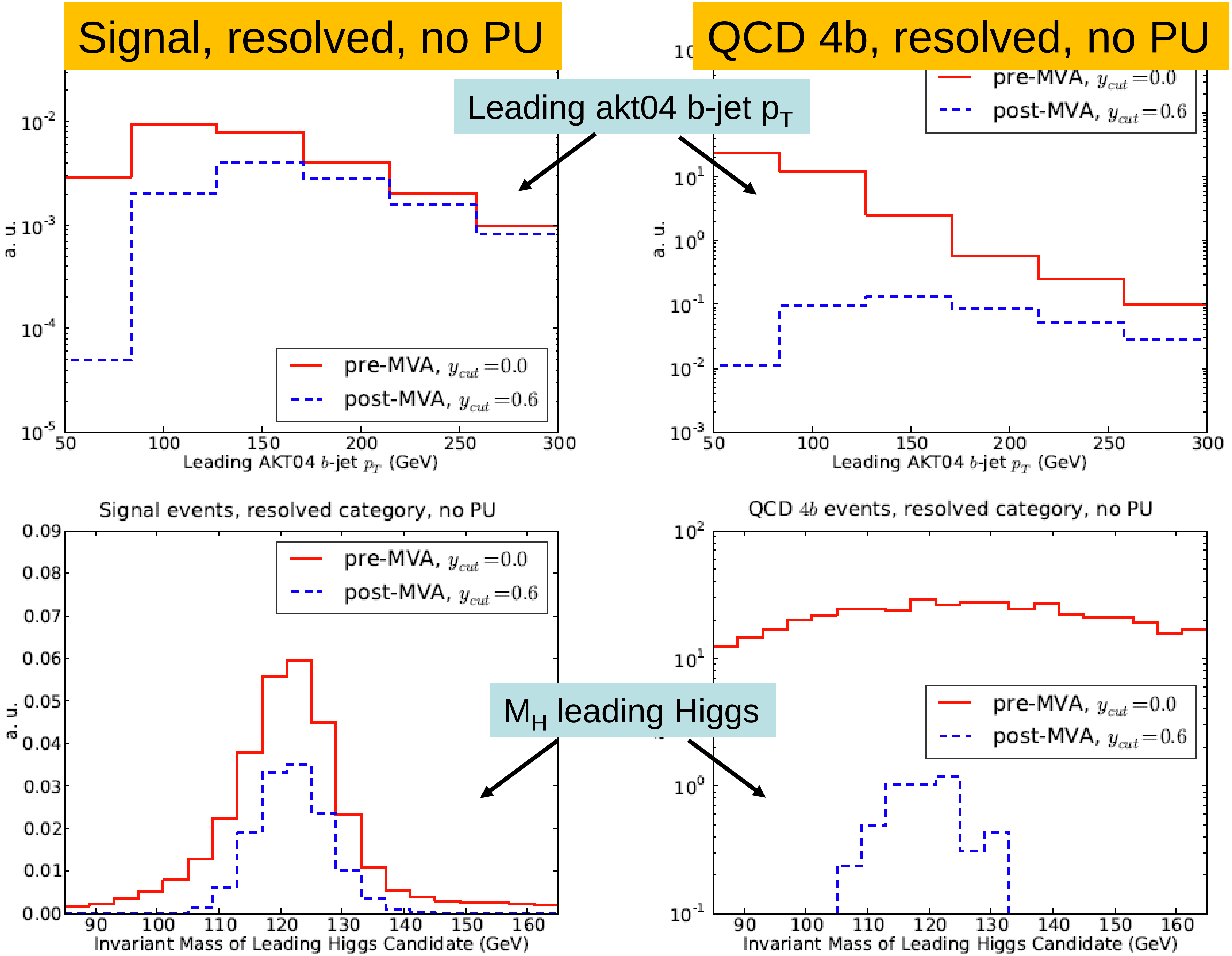
- 13 resolved
- 17 intermediate
- 21 boosted

ANN for boosted category,
heat map after training: red (dark) high (low) weight



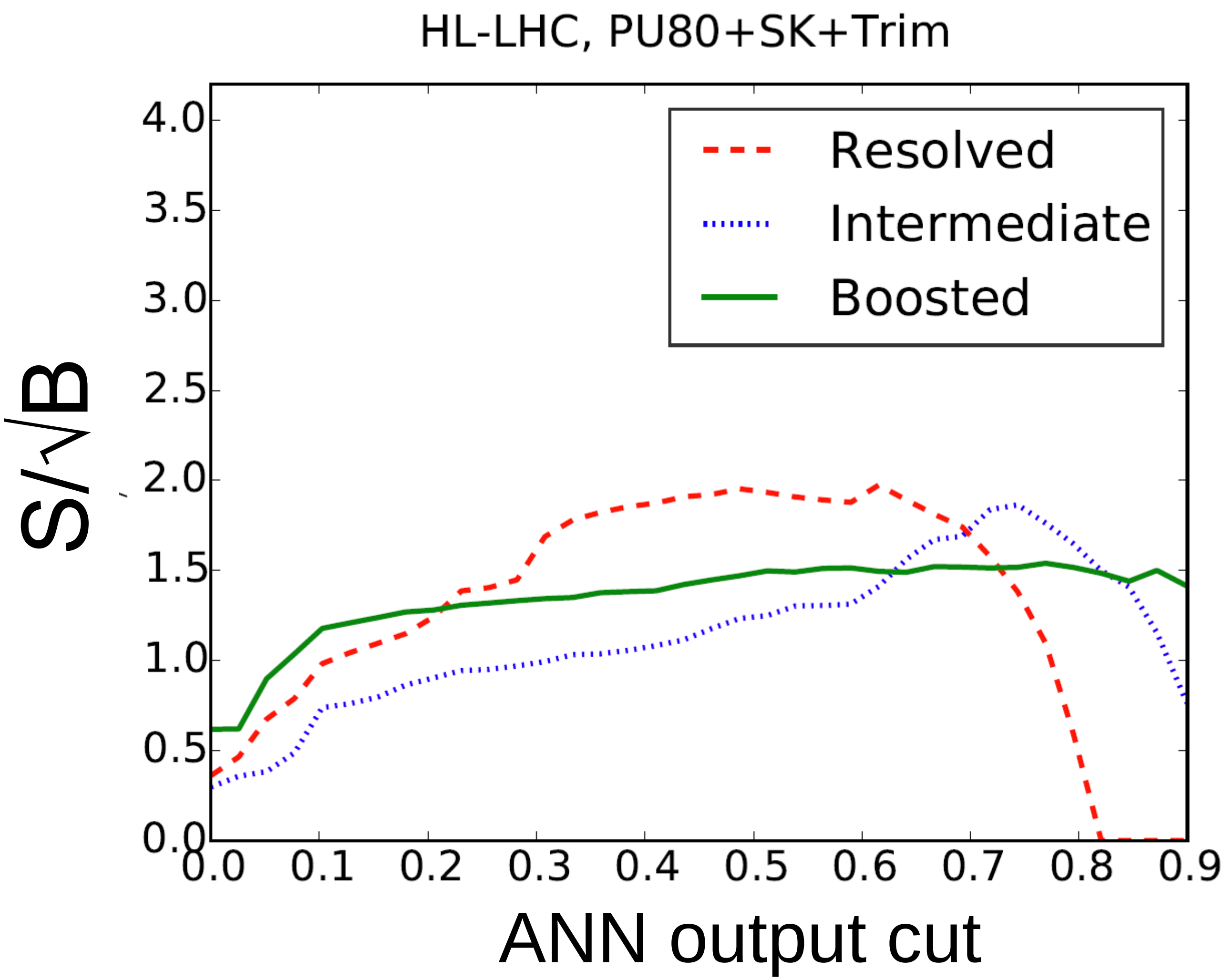




Trained on signal & background
Separately for the 3 categories



Signal, resolved, no PU
QCD 4b, resolved, no PU
Leading akt04 b-jet $p_T$
pre-MVA, $y_{cut}=0.0$
post-MVA, $y_{cut}=0.6$
a. u.
Leading AKT04 $b$-jet $p_T$ (GeV)
Signal events, resolved category, no PU
QCD $4b$ events, resolved category, no PU
$M_H$ leading Higgs
Invariant Mass of Leading Higgs Candidate (GeV)

# S/√B post-MVA results



# Post-MVA results for HL-LHC (3 ab$^{-1}$)

| Category | | signal $N_{ev}$ | background $N_{ev}^{tot}$ | $N_{ev}^{4b}$ | $S/\sqrt{B_{tot}}$ | $S/\sqrt{B_{4b}}$ | $S/B_{tot}$ | $S/B_{4b}$ |
|---|---|---|---|---|---|---|---|---|
| Boosted | no PU | 290 | $1.2 \cdot 10^4$ | $8.0 \cdot 10^3$ | 2.7 | 3.2 | 0.03 | 0.04 |
| | PU80+SK+Trim | 290 | $3.7 \cdot 10^4$ | $1.2 \cdot 10^4$ | 1.5 | 2.7 | 0.01 | 0.02 |
| Intermediate | no PU | 130 | $3.1 \cdot 10^3$ | $1.5 \cdot 10^3$ | 2.3 | 3.3 | 0.04 | 0.08 |
| | PU80+SK+Trim | 140 | $5.6 \cdot 10^3$ | $2.4 \cdot 10^3$ | 1.9 | 2.9 | 0.03 | 0.06 |
| Resolved | no PU | 630 | $1.1 \cdot 10^5$ | $5.8 \cdot 10^4$ | 1.9 | 2.7 | 0.01 | 0.01 |
| | PU80+SK | 640 | $1.0 \cdot 10^5$ | $7.0 \cdot 10^4$ | 2.0 | 2.6 | 0.01 | 0.01 |
| **Combined** | no PU | | | | 4.0 | 5.3 | | |
| | PU80+SK+Trim | | | | 3.1 | 4.7 | | |

$$\left(\frac{S}{\sqrt{B}}\right)_{\rm tot} \simeq 3.1\ (1.0)\,, \quad \mathcal{L} = 3000\ (300)\ {\rm fb}^{-1}$$

$$\left(\frac{S}{\sqrt{B_{4b}}}\right)_{\rm tot} \simeq 4.7\ (1.5)$$

1) Combination of ALL categories important
2) 2b2j background has significant impact

# Post-MVA results for HL-LHC (3 $ab^{-1}$)

| Category | | signal $N_{ev}$ | background $N_{ev}^{tot}$ | background $N_{ev}^{4b}$ | $S/\sqrt{B_{tot}}$ | $S/\sqrt{B_{4b}}$ | $S/B_{tot}$ | $S/B_{4b}$ |
|---|---|---|---|---|---|---|---|---|
| Boosted | no PU | 290 | $1.2 \cdot 10^4$ | $8.0 \cdot 10^3$ | 2.7 | 3.2 | 0.03 | 0.04 |
| | PU80+SK+Trim | 290 | $3.7 \cdot 10^4$ | $1.2 \cdot 10^4$ | 1.5 | 2.7 | 0.01 | 0.02 |
| Intermediate | no PU | 130 | $3.1 \cdot 10^3$ | $1.5 \cdot 10^3$ | 2.3 | 3.3 | 0.04 | 0.08 |
| | PU80+SK+Trim | 140 | $5.6 \cdot 10^3$ | $2.4 \cdot 10^3$ | 1.9 | 2.9 | 0.03 | 0.06 |
| Resolved | no PU | 630 | $1.1 \cdot 10^5$ | $5.8 \cdot 10^4$ | 1.9 | 2.7 | 0.01 | 0.01 |
| | PU80+SK | 640 | $1.0 \cdot 10^5$ | $7.0 \cdot 10^4$ | 2.0 | 2.6 | 0.01 | 0.01 |
| **Combined** | no PU | | | | 4.0 | 5.3 | | |
| | PU80+SK+Trim | | | | 3.1 | 4.7 | | |

$$\left(\frac{S}{\sqrt{B}}\right)_{\text{tot}} \simeq 3.1\ (1.0)\,, \quad \mathcal{L} = 3000\ (300)\,\text{fb}^{-1}$$

Observation with just 4b channel seems possible!

$$\left(\frac{S}{\sqrt{B_{4b}}}\right)_{\text{tot}} \simeq 4.7\ (1.5)$$

# Conclusions

- Di-Higgs production **corner stone of LHC program**
- Presented new feasibility study of 2H → bbbb
  - MVA based and combination of all di-Higgs topologies
  - **1st time:**
    - pile-up modelled (PU = 80 and PU = 150)
      - **Results are stable** under PU with mitigation
    - 2b2j and 4j backgrounds included
      - Light and charm quark **fake rate reduction important**
    - **Low pt trigger thresholds important** for resolved regime
  - 4b final state alone enough to claim evidence for 2H process at HL-LHC
  - **If experimental performance improved** 2H process observation with just 200 – 300 $fb^{-1}$ possible.

# Outlook

- Include $Z \rightarrow bb$, ttH and HZ backgrounds
- Add more statistics to QCD backgrounds
- Estimation of accuracy of Higgs self-coupling measurement at LHC, HL-LHC, 100 TeV FCC collider


Backup Slides

# Pre-MVA cut-flow

- **C1a**: check that we have at least two large-$R$ jets (in the boosted case), one large-$R$ jet and at least 2 small-$R$ jets (in the intermediate case) and at least four small-$R$ jets (in the resolved case).

  In addition, require that these jets satisfy the corresponding $p_T$ thresholds; $p_T \geq 200$ GeV for large-$R$ jets and $p_T \geq 40$ GeV for small-$R$ jets, as well as the associated rapidity acceptance constraints.

- **C1b**: the two leading large-$R$ jets must be mass-drop tagged in the boosted category. In the intermediate category, the large-$R$ jet must also be mass-drop tagged.

- **C1c**: after the two Higgs candidates have been reconstructed, their invariant masses are required to lie within a window around $m_H$, in particular between 85 and 165 GeV, Eq. (3.6).

- **C2**: the $b$-tagging conditions are imposed (see Sect. 3.2), and the event is categorised exclusively into one of the three topologies, according to the hierarchy determined in Sect. 3.3.

# Pre-MVA 2b2j background -- BIG

| HL-LHC, Boosted category, PU+SK+Trim with $n_{PU} = 80$ | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Cross-section [fb] | | | | | | $S/B$ | | $S/\sqrt{B}$ | |
| | $hh4b$ | total bkg | $4b$ | $2b2j$ | $4j$ | $t\bar{t}$ | tot | $4b$ | tot | $4b$ |
| C1a | 3.5 | $4.1 \cdot 10^7$ | $1.0 \cdot 10^4$ | $2.7 \cdot 10^6$ | $3.8 \cdot 10^7$ | $2.0 \cdot 10^4$ | $8.6 \cdot 10^{-8}$ | $3.4 \cdot 10^{-4}$ | 0.03 | 1.9 |
| C1b | 2.5 | $3.2 \cdot 10^7$ | $6.8 \cdot 10^3$ | $1.9 \cdot 10^6$ | $3.0 \cdot 10^7$ | $1.9 \cdot 10^4$ | $7.8 \cdot 10^{-8}$ | $3.6 \cdot 10^{-4}$ | 0.02 | 1.6 |
| C1c | 0.8 | $2.2 \cdot 10^6$ | $5.4 \cdot 10^2$ | $1.4 \cdot 10^5$ | $2.0 \cdot 10^6$ | $4.8 \cdot 10^3$ | $3.8 \cdot 10^{-7}$ | $1.6 \cdot 10^{-3}$ | 0.03 | 2.0 |
| C2 | 0.14 | $1.5 \cdot 10^2$ | 40 | 86 | 22 | 1.8 | $9.0 \cdot 10^{-4}$ | $3.5 \cdot 10^{-3}$ | 0.6 | 1.2 |

# Background and signal relative fractions

| | | $n_0^{(\text{b-jet})}$ | $n_1^{(\text{b-jet})}$ | $n_2^{(\text{b-jet})}$ | $n_3^{(\text{b-jet})}$ | $n_4^{(\text{b-jet})}$ | $\text{EFF}_{\text{b-tag}}$ |
|---|---|---|---|---|---|---|---|
| Signal | $hh \to 4b$ | 0.1% | 3% | 25% | 53% | 20% | 8.5% |
| Background | QCD $4b$ | 1% | 8% | 27% | 44% | 20% | 8.4% |
| | QCD $2b2j$ | 9% | 42% | 49% | 1% | 0.1% | 0.04% |
| | QCD $4j$ | 96% | 3.5% | 0.5% | 0.01% | $3 \cdot 10^{-4}\%$ | $2 \cdot 10^{-4}\%$ |

**Table 5.** The relative fractions $n_j^{(\text{b-jet})}$ of events for the resolved selection for which out of the four leading small-$R$ jets of the event, $j$ jets contain at least one $b$-quark with $p_T^b \geq 15$ GeV. This information is provided for the di-Higgs signal events and for the three QCD background samples. The last column indicates the overall selection efficiency as defined in Eq. (4.1)

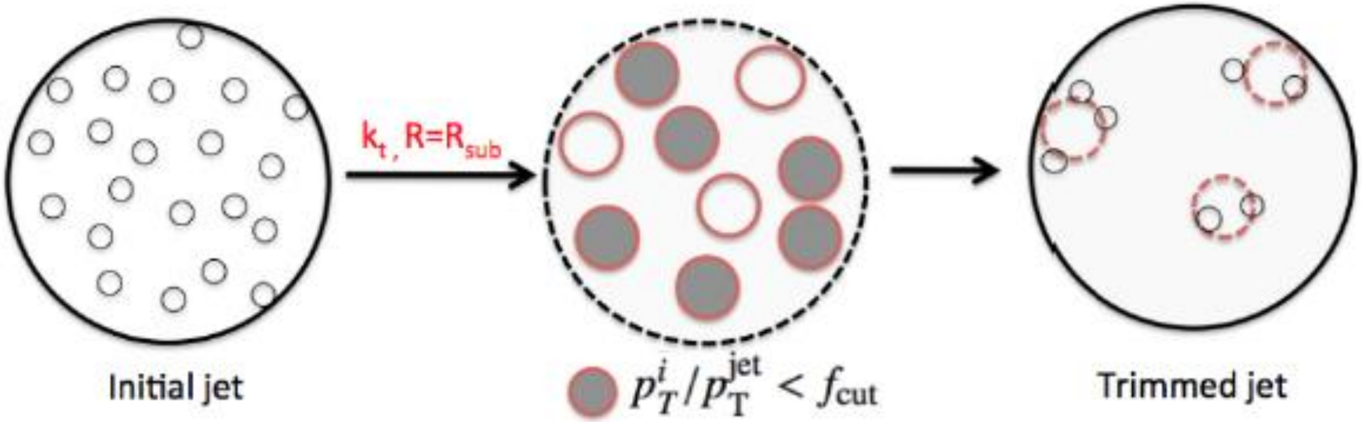
# Pile-up simulation and mitigation

- Up to 150 interactions per bunch-crossing @ HL-LHC
- Impact on signal significance:
  - Additional (hard) jets in reconstruction
  - Affects mass and substructure of large-$R$ jets
- **Not taken into account in previous studies**

Strategy:

- Superimpose $n_{\mathrm{PU}} = 80$ Minimum Bias events on each signal/background event
- Pile-up mitigation
  - Event level: Remove soft particles using SoftKiller [arXiv:1407.0408]
  - Jet level: Apply trimming [arXiv:0912.1342] (large-$R$ jets only)

JHEP09 (2013) 076



# Multivariate techniques

Given a set of $N_{var}$ kinematic variables $\{k_i\}$ associated to MC event $i$, and a set of ANN weight parameters $\{\omega\}$, the ANN output $y_i$ interpreted as **probability that this event originates from signal process**

$$y_i = P(y'_i = 1|\{k\}_i, \{\omega\}) \, ,$$

With $y'_i$ the true MC classification: $y'_i=1$ for signal, $y'_i=0$ for background

The **general classification probability** including background events is

$$P(y'_i|\{k\}_i, \{\omega\}) = y_i^{y'_i}(1 - y_i)^{1-y'_i}$$

Thus the **error function to be minimised during the training** is the **cross-entropy:**

$$E(\{\omega\}) \equiv -\log\left(\prod_i^{N_{\rm ev}} P(y'_i|\{k\}_i, \{\omega\})\right)$$
$$= \sum_i^{N_{\rm ev}} \left[y'_i \log y_i + (1 - y'_i) \log (1 - y_i)\right]$$

ANN training performed with **Genetic Algorithms** using **cross-validation stopping**

# Post-MVA ROC curve and number of events

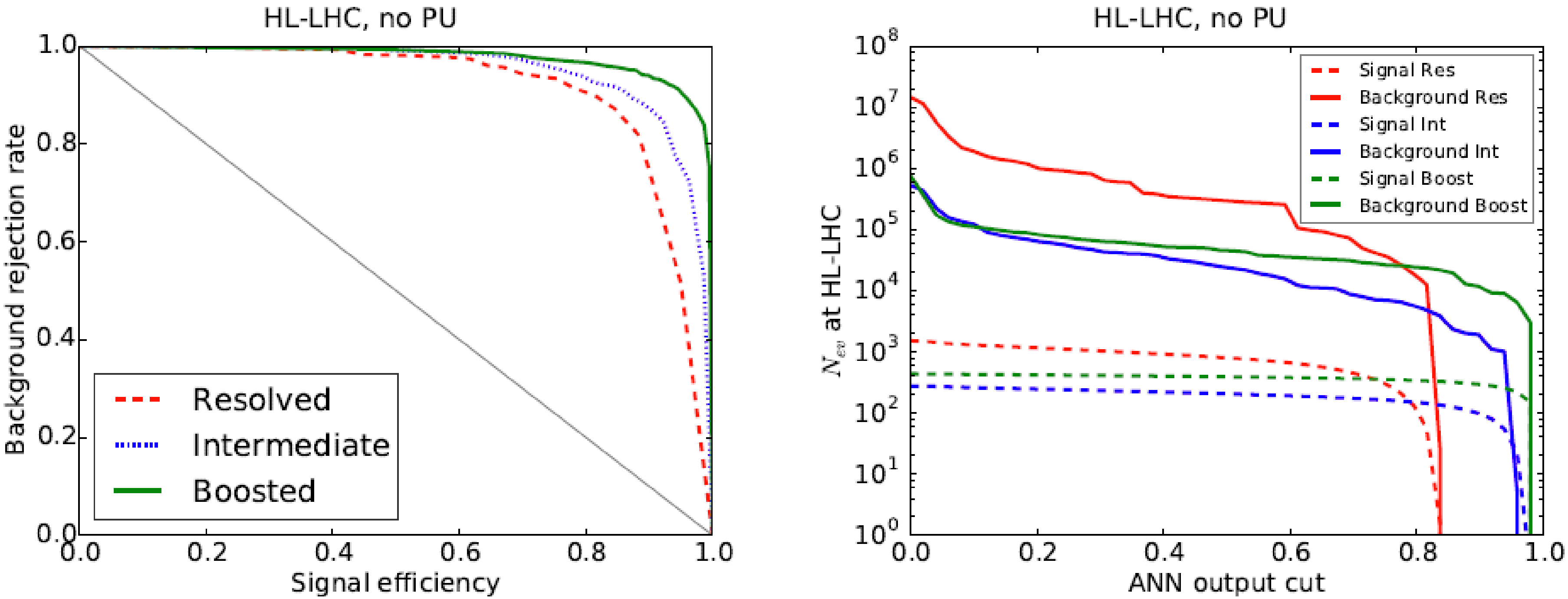
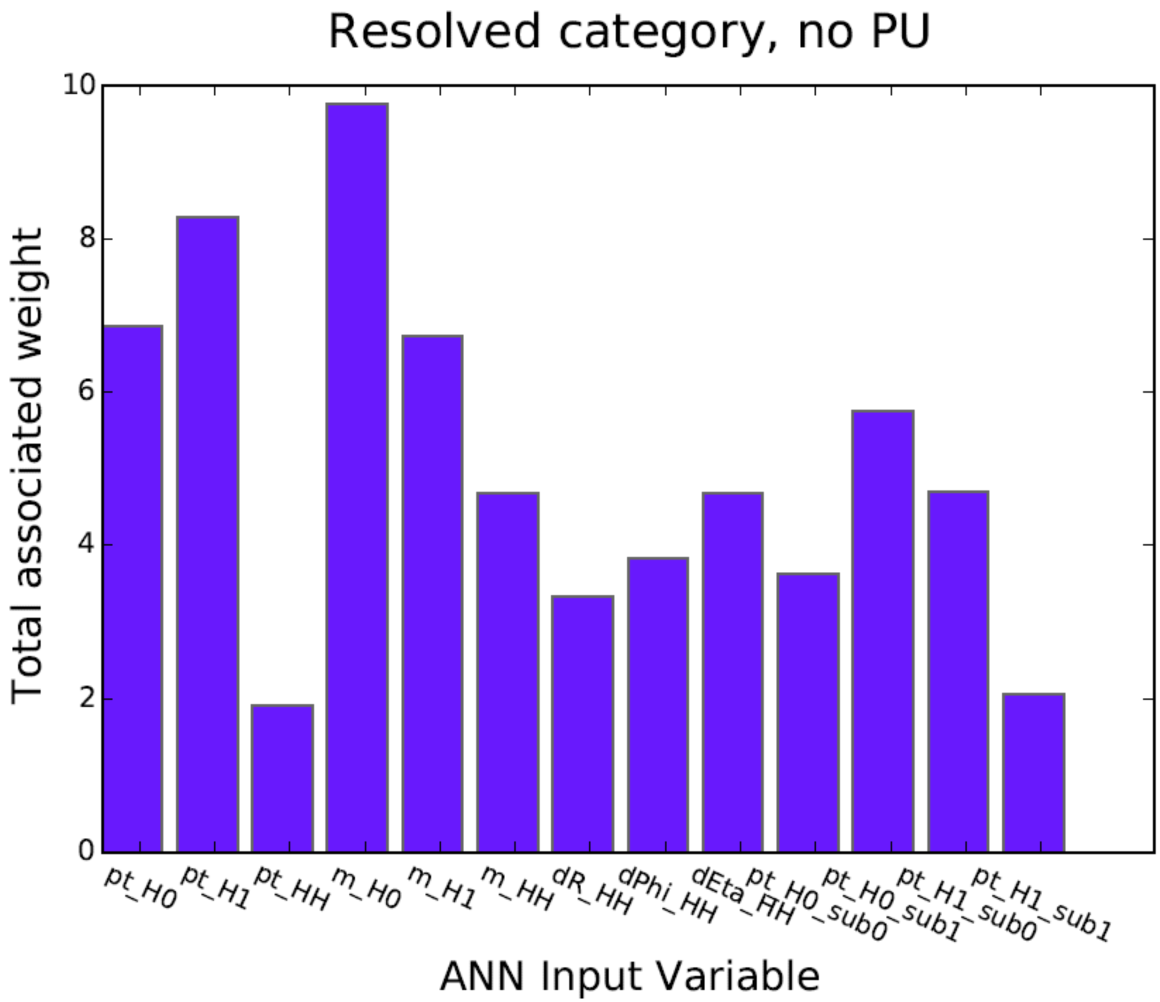
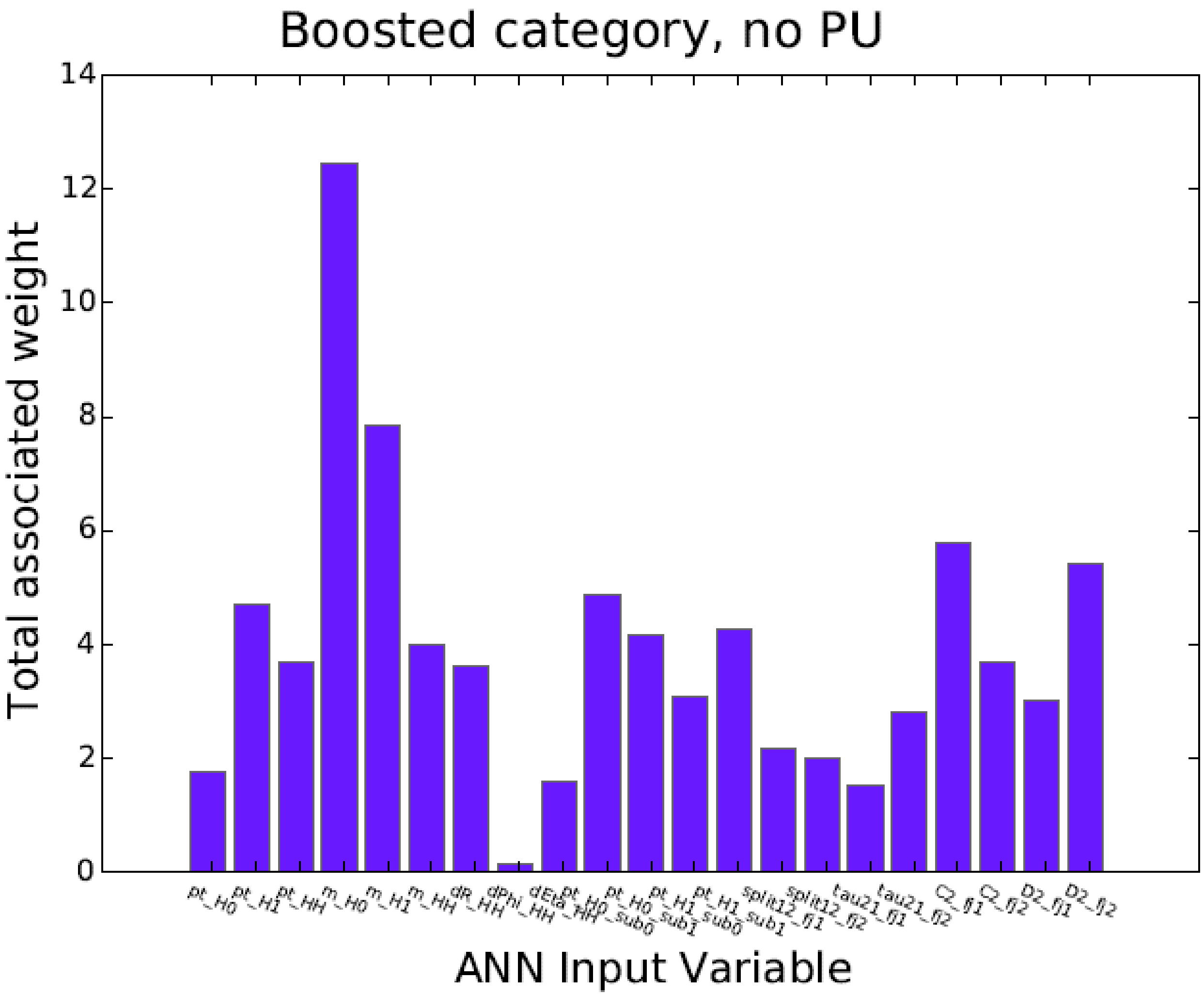



**Figure 17.** Left: ROC curve for the background rejection rate as a function of the signal selection efficiency, as the cut $y_{\rm cut}$ in the ANN output is varied. Right: Number of signal (dashed) and background (solid) events expected at the HL-LHC as a function of the $y_{\rm cut}$.

# MVA Analysis



# MVA



Boosted category, no PU
Total associated weight
0
2
4
6
8
10
12
14
pt_H0
pt_H1
pt_HH
m_H0
m_H1
m_HH
dR_HH
dPhi_HH
dEta_HH
pt_H0_sub0
pt_H0_sub1
pt_H1_sub0
pt_H1_sub1
split12_fj1
split12_fj2
tau21_fj1
tau21_fj2
C2_fj1
C2_fj2
D2_fj1
D2_fj2
ANN Input Variable

# Pre-MVA distributions