BOOSTING $HH \rightarrow (b\bar{b})(b\bar{b})$ WITH MULTIVARIATE TECHNIQUES

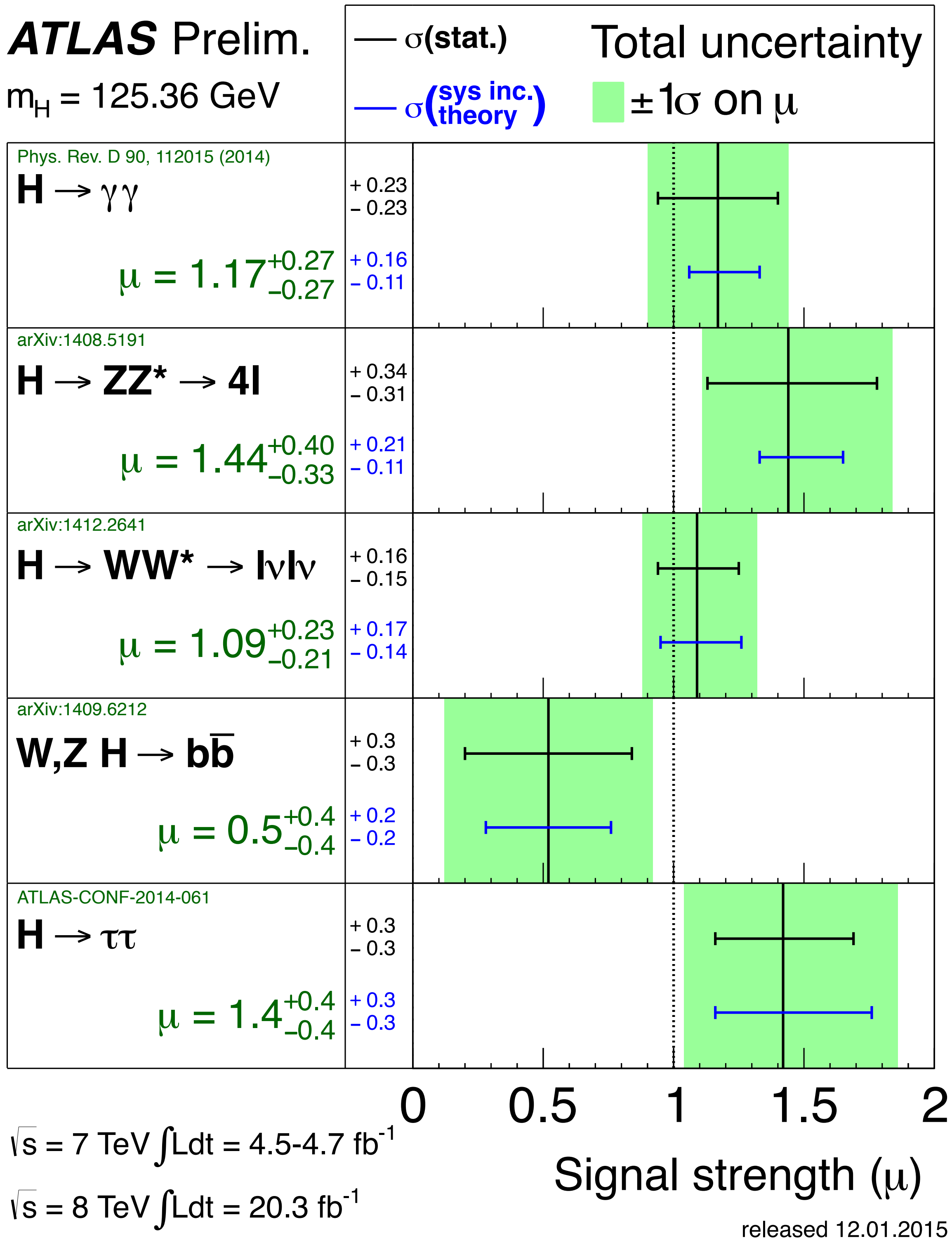
BEHR, BORTOLETTO, FROST, ISSEVER, NH, ROJO [1512.08928]

Nathan Hartland
University of Oxford

DIS 2016

DESY Hamburg, 13/04/16

HIGGS PHYSICS AT THE LHC



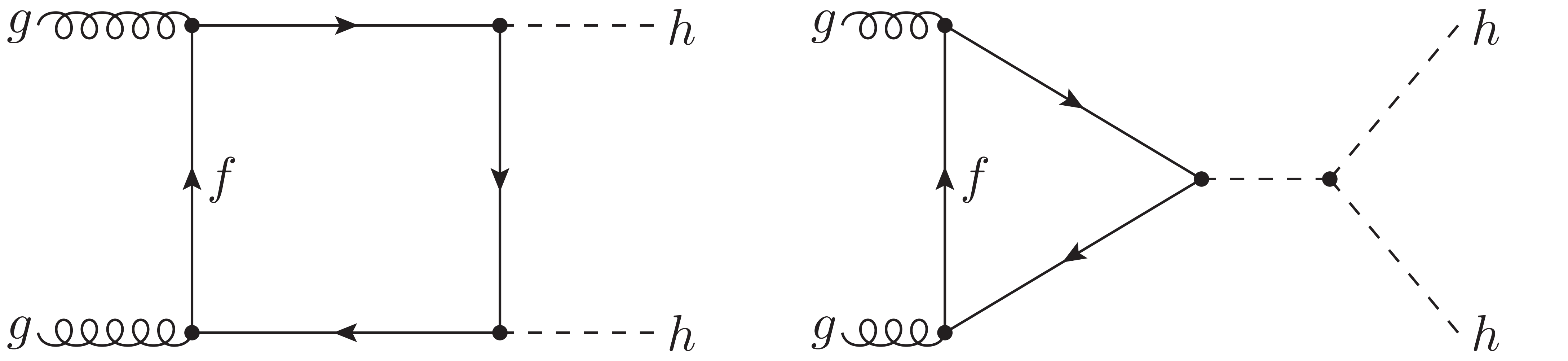
Resonance at ~126 GeV is pretty consistent with expectations for SM Higgs.

So far, we have only explored the **minimum** of the EWSB potential

$$V = \frac{1}{2} m_H^2 H^2 + v\lambda_3 H^3 + \frac{1}{4}\lambda_4 H^4$$

To understand the full potential, we need to measure (*at least*) double-Higgs production

DIHIGGS AT THE LHC



Clear difficulty: HH production cross-section at LHC is *tiny*

14TeV NNLO $\sigma_{HH} \simeq 40$ fb *factor of ~10^3 smaller than single H*

Compounded by usual H reconstruction problems

Final State	BF
bbbb	33%
bbWW	13%
bb\tau\tau	3.5%
WWWW	5.3%

*Lots of cross-section decaying to **challenging** fully hadronic final state*

Recent ATLAS bound:

$$\sigma(pp \to hh \to b\bar{b}b\bar{b}) < 1.22 \text{ pb}$$

(Dirk Duschinger's Talk, ATLAS-CONF-2016-017)

(Data is king: 3000 fb^{-1} HL-LHC)

$HH \to (b\bar{b})(b\bar{b})$ BACKGROUNDS

Fully hadronic (gg)Higgs channel: experimentally challenging on a lot of fronts

Primary challenge: *Overwhelming QCD background*

Signal	Cross-section
HH	$4.0 \cdot 10^{-2}$ pb

Background	Cross-section
ZH	$7.7 \cdot 10^{-1}$ pb
ttH	$4.6 \cdot 10^{-1}$ pb
bbH	$6.1 \cdot 10^{-1}$ pb

Background	Cross-section
bbbb	$1.8 \cdot 10^{3}$ pb
bbjj	$3.5 \cdot 10^{5}$ pb
jjjj	$5.8 \cdot 10^{6}$ pb
tt(bbjjjj)	$3.5 \cdot 10^{3}$ pb

$HH \to (b\bar{b})(b\bar{b})$ BACKGROUNDS

Background	Cross-section
bbbb	$1.8\cdot10^3$ pb
bbjj	$3.5\cdot10^5$ pb

Contributions from light jets usually discounted by b-tagging arguments. Assume (optimistically):

$$\epsilon_{\text{tag}} = 0.8 \qquad \epsilon_{\text{mistag}} = 0.01$$

Then tagged 4b cross section ~ $\epsilon_{\text{tag}}^4 \cdot \sigma_{bbbb} = 7.3 \cdot 10^2 \text{ pb}$

and tagged 2b2j cross section ~ $\epsilon_{\text{tag}}^2 \epsilon_{\text{mistag}}^2 \cdot \sigma_{bbjj} = 22 \text{ pb}$

Consider instead the fraction of events containing n b-jets

$(R = 0.4 \text{ ak}_T \quad p_T > 20\text{GeV})$

	0 b-jet	1 b-jet	2 b-jet	3 b-jet	4 b-jet	ϵ_{sel}
bbbb	1%	8%	27%	44%	20%	**8.4%**
bbjj	9%	42%	49%	1%	0.1%	**0.04%**

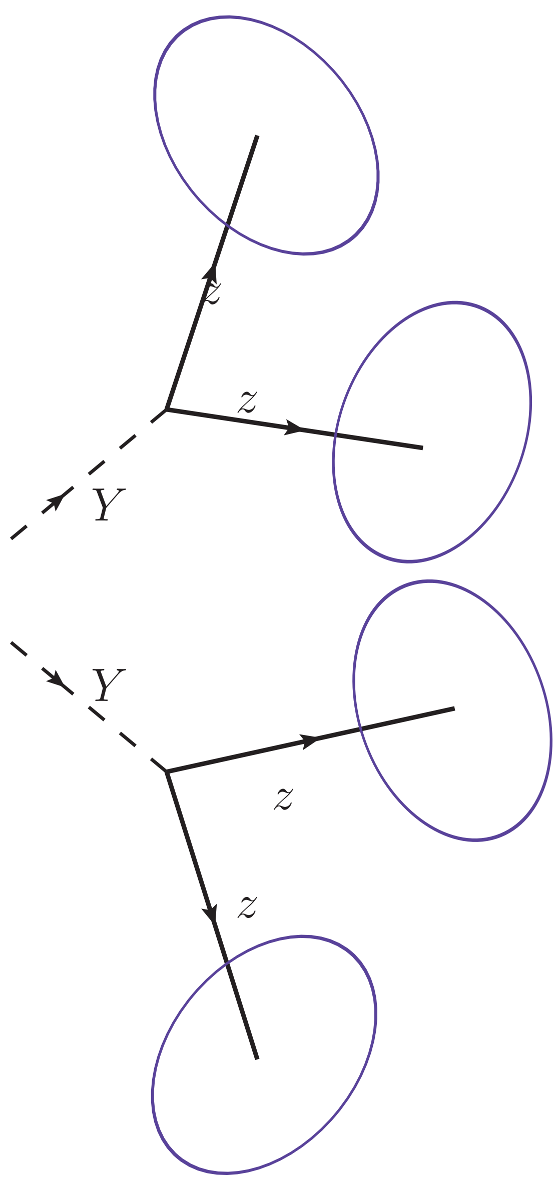
$$\epsilon_{\text{sel},bbbb} \cdot \sigma_{bbbb} \simeq 150\text{pb}$$

$$\epsilon_{\text{sel},bbjj} \cdot \sigma_{bbjj} \simeq 140\text{pb}$$

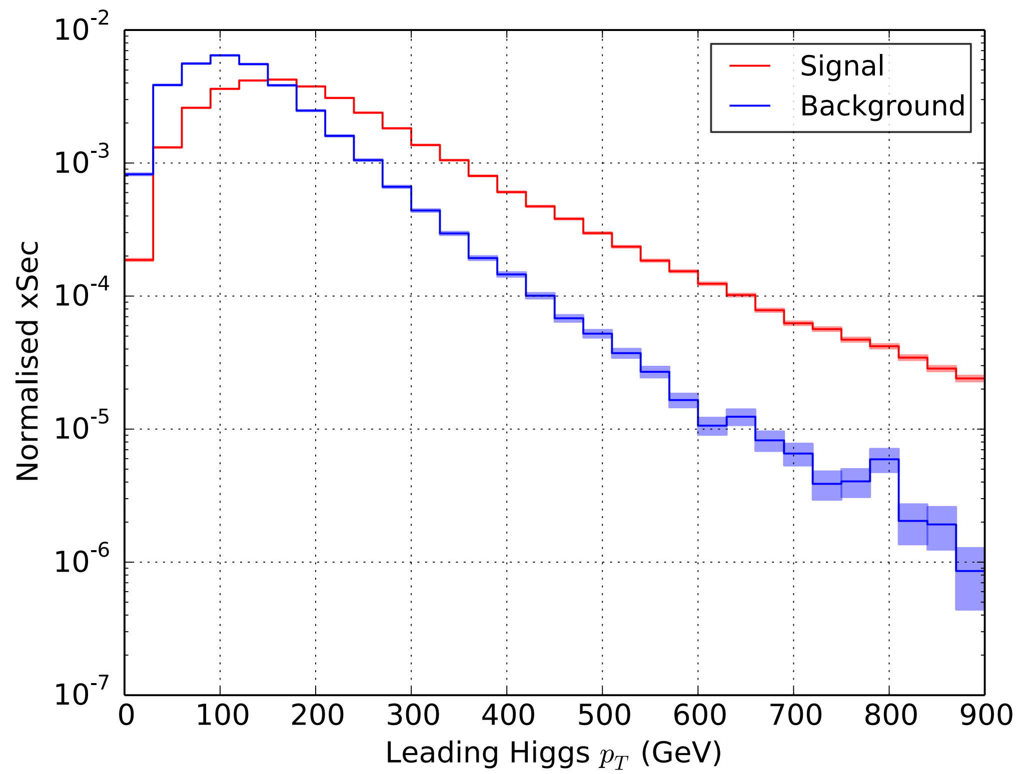
Contributions e.g bbjj are not negligible w.r.t 'irreducible' 4b background

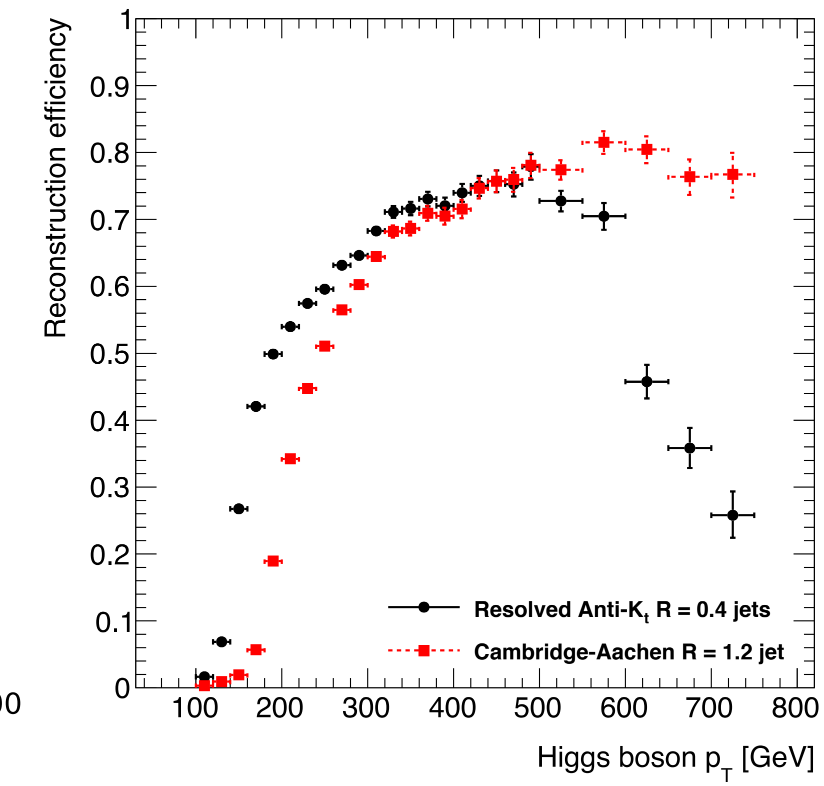
$HH \to (b\bar{b})(b\bar{b})$ ANALYSIS TOPOLOGIES

Resolved Reconstruct Higgs decay products in four separate (small-R) jets



(Gouzevitch et al 1303.6636)





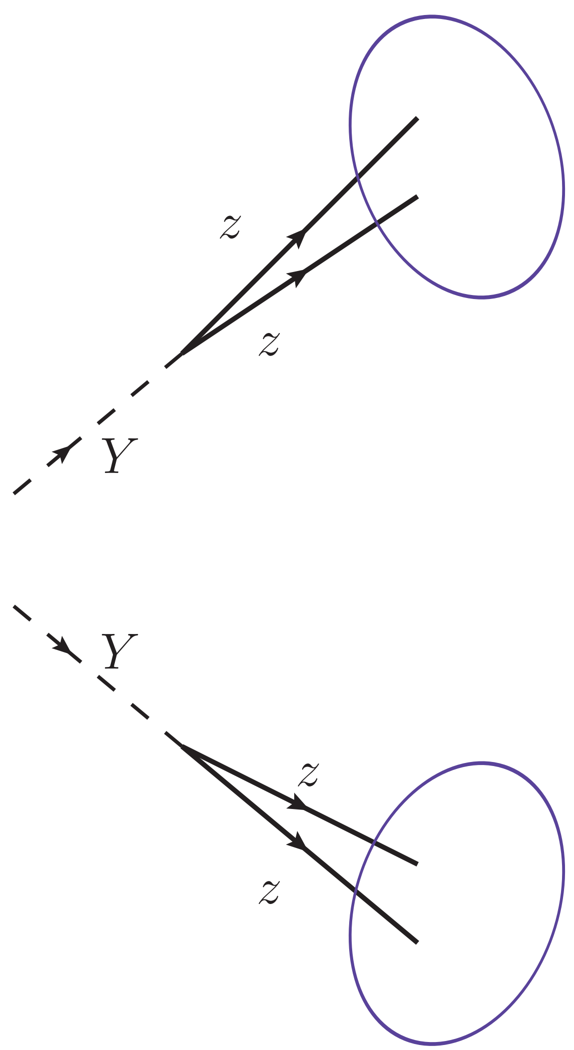
(Wardrope et al 1410.2794)

- Captures bulk of HH cross-section
- Lower efficiency at high pT
- Combinatorics

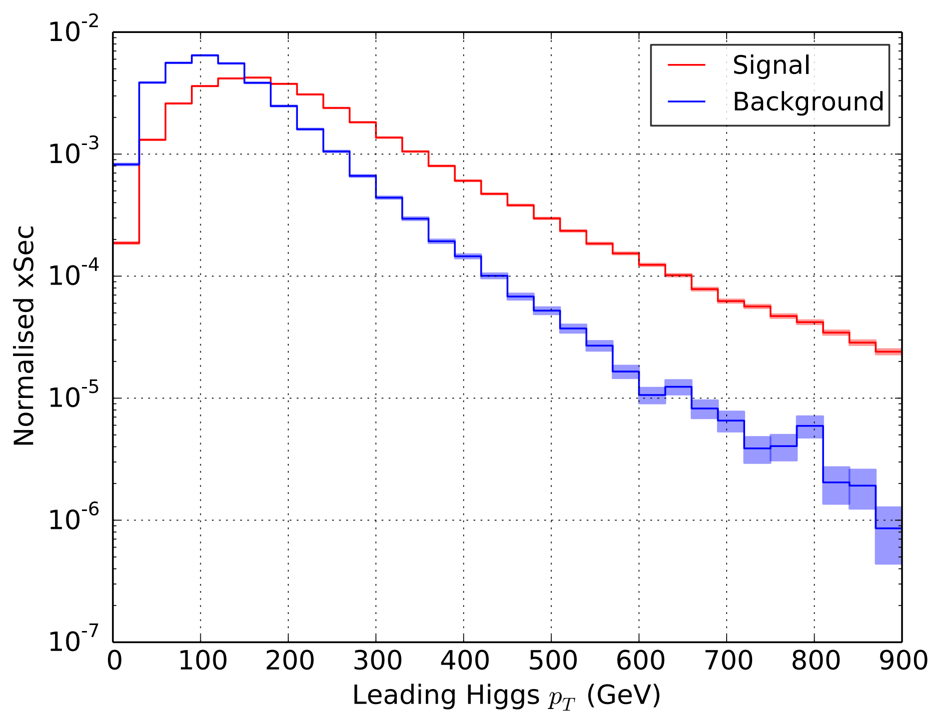
(e.g UCL Study [1410.2794])

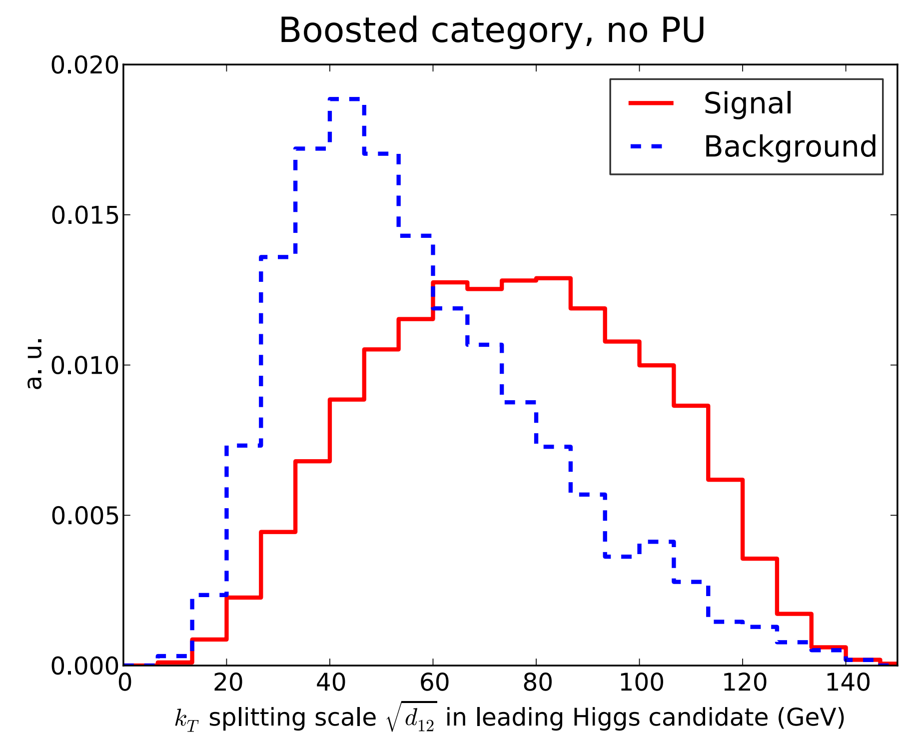
$HH \to (b\bar{b})(b\bar{b})$ ANALYSIS TOPOLOGIES

Boosted Reconstruct Higgs decay products in two (large-R) jets



(Gouzevitch et al 1303.6636)





- Better S/B but lower cross-sections
- Substructure tools available
- Greater sensitivity to pileup

(e.g Durham Study [1404.7139])

OUR APPROACH

- **Consider closely QCD multi-jet background (not just 4b)**

  Assuming relatively optimistic b-tagging parameters

  $$\epsilon_b = 0.8, \quad \epsilon_c = 0.1, \quad \epsilon_j = 0.01$$

- **Handle boosted and resolved analysis topologies**

  Ensure good coverage of final state phase space

- **Use plenty of information on jet substructure**

  MD tags, Splitting scales, N-subjettiness, Energy correlations

- **Investigate how multivariate analysis can boost significances**

  Keep cut-based analysis loose, try to to make as much use as possible of MVA

- **Assess robustness of results under the addition of pile-up (PU)**

  How feasible are these methods in a more realistic environment?

RECONSTRUCTION DETAILS

'Small-R jet'

b-tagged aKT R=0.4 jet

$p_T > 40\,\mathrm{GeV},\ \ |\eta| < 2.5$

'Large-R jet'

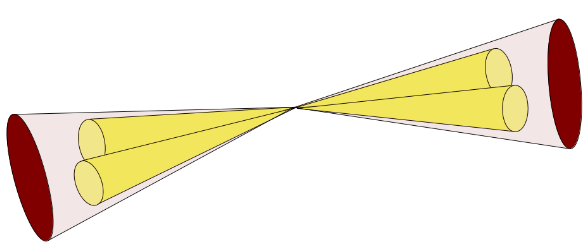
double b-tagged aKT R=1.0 jet

$p_T > 200\,\mathrm{GeV},\ \ |\eta| < 2.5$

Require BDRS mass-drop tag

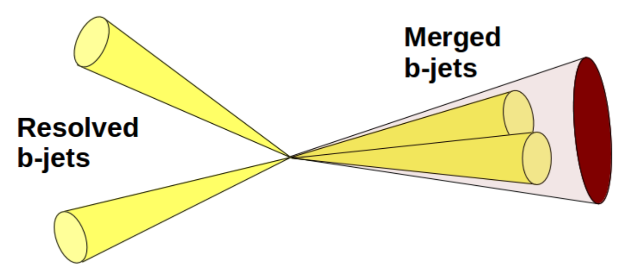
Boosted

Two Large-R jets



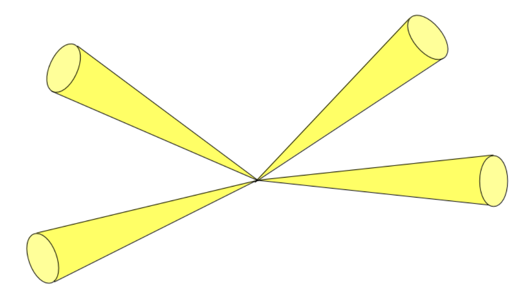
Intermediate

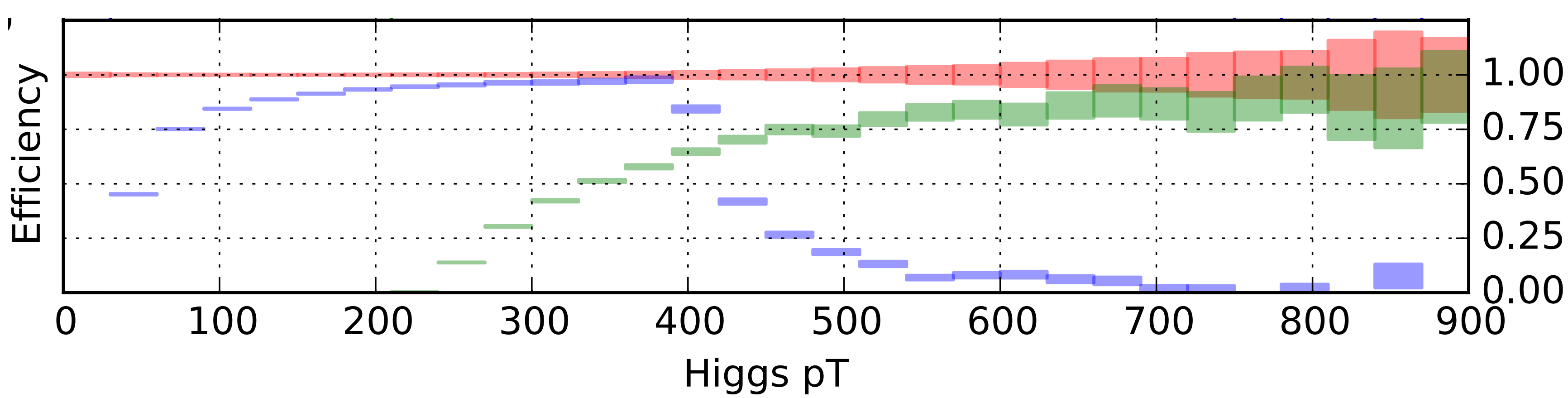
One Large-R jet, two Small-R jets



Resolved

Four Small-R jets





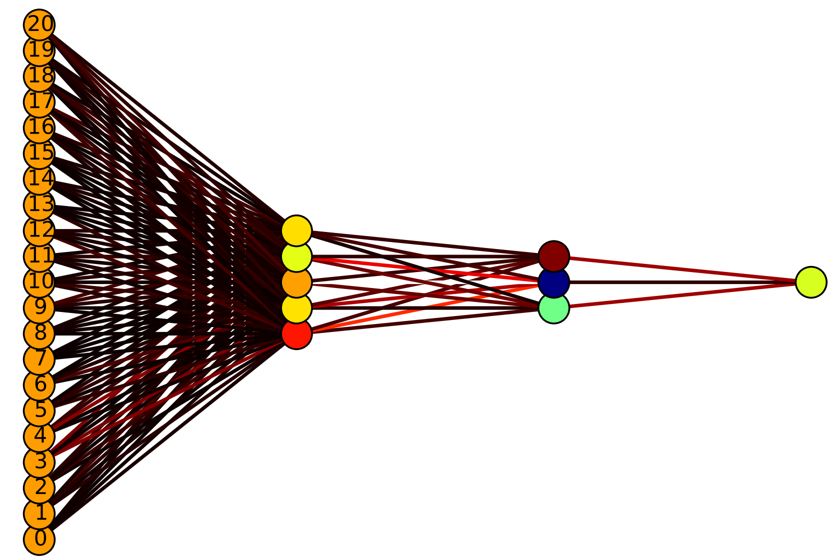
Blue: Resolved efficiency

Green: Boosted efficiency

- Higgs mass window cut $|m - 125| < 40$ GeV
- Prioritise selection Boosted-Intermediate-Resolved

MULTIVARIATE ANALYSIS

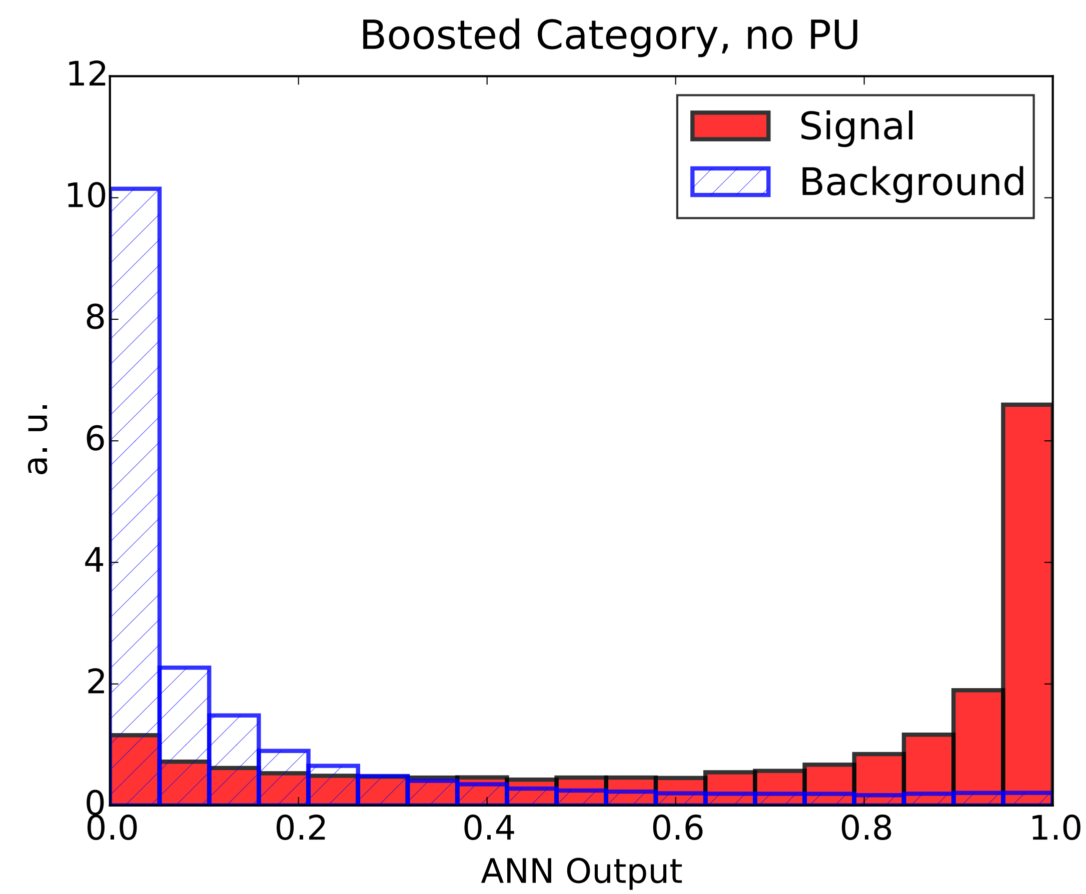
Classify events as signal/background with ANN
(Similar architecture to NNPDF)



MVA input includes standard event kinematics, along with for large-R jets:

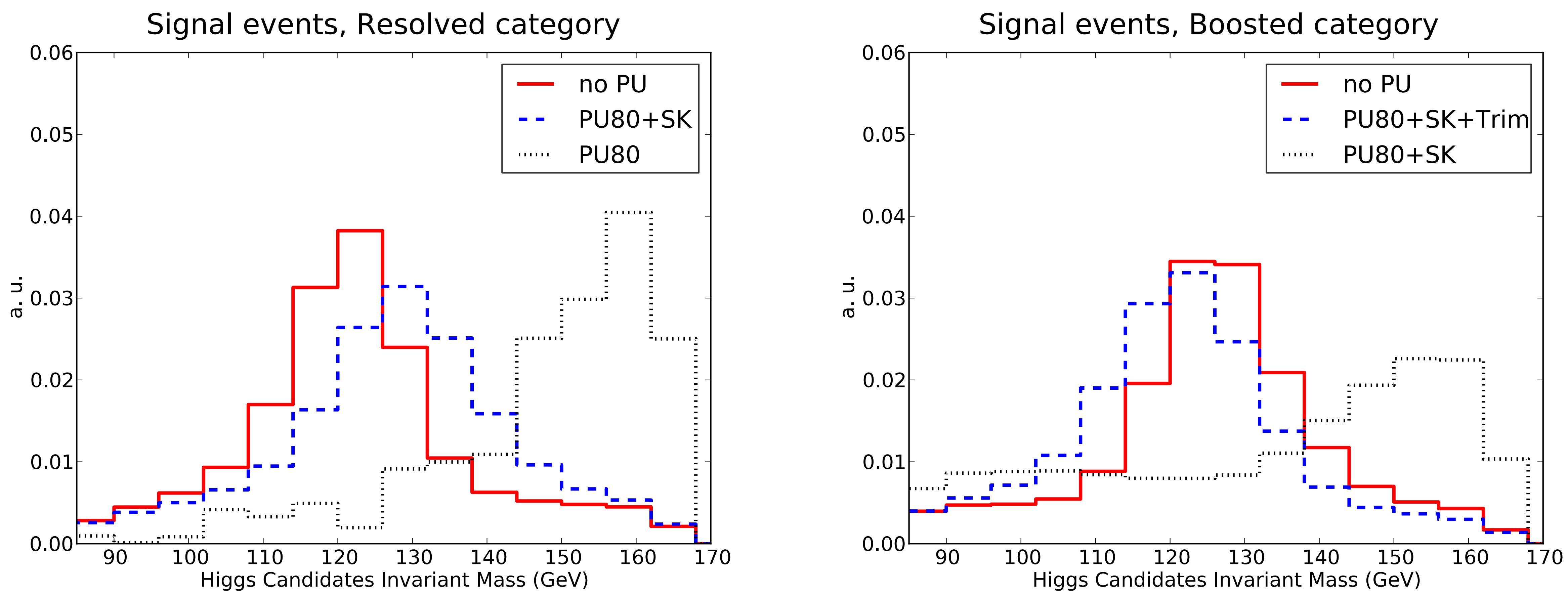
- k_T *-splitting scales*
- *Ratio of 2-to-1 subjettiness*
- *E.C.F double-ratios* C_2, D_2

Channel	#inputs
Boost	21
Inter.	17
Resol.	13



PILEUP SIMULATION

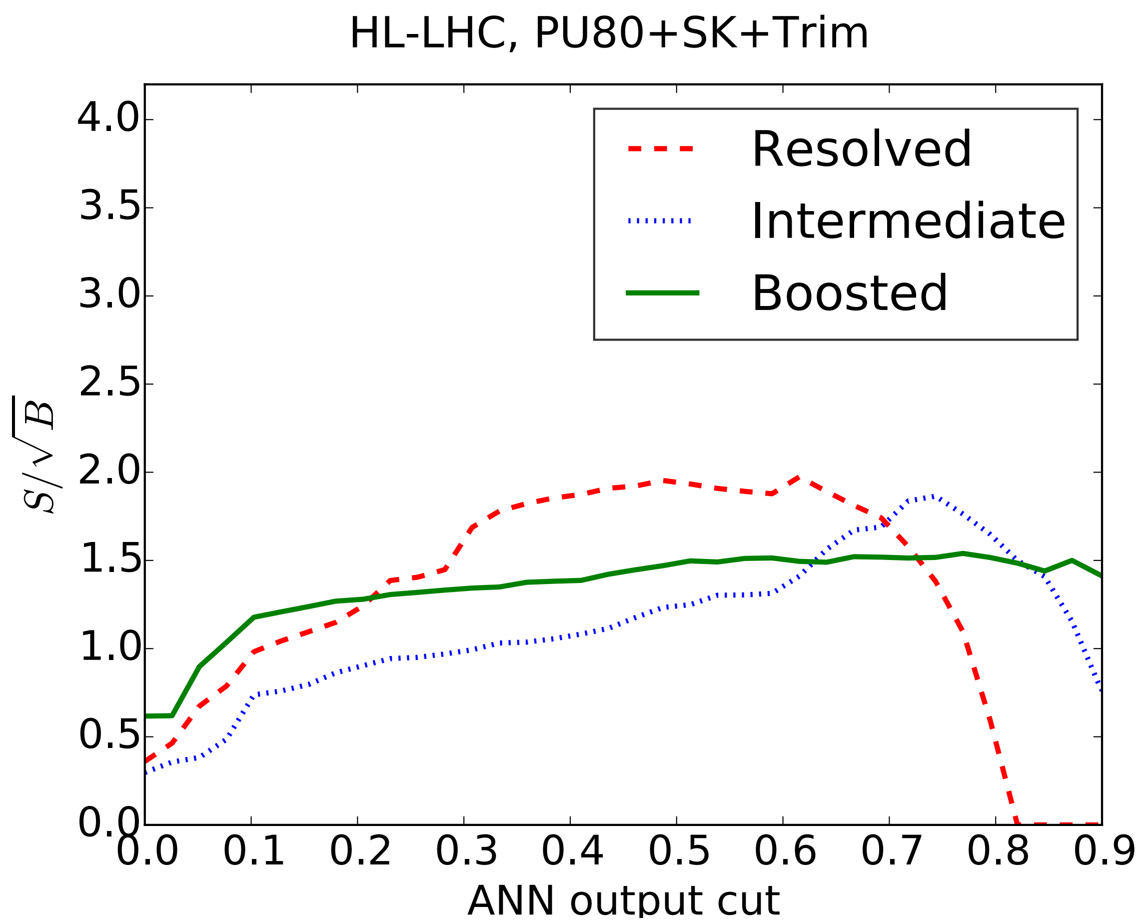
Introduce an additional 80 PU vertices per hard event.

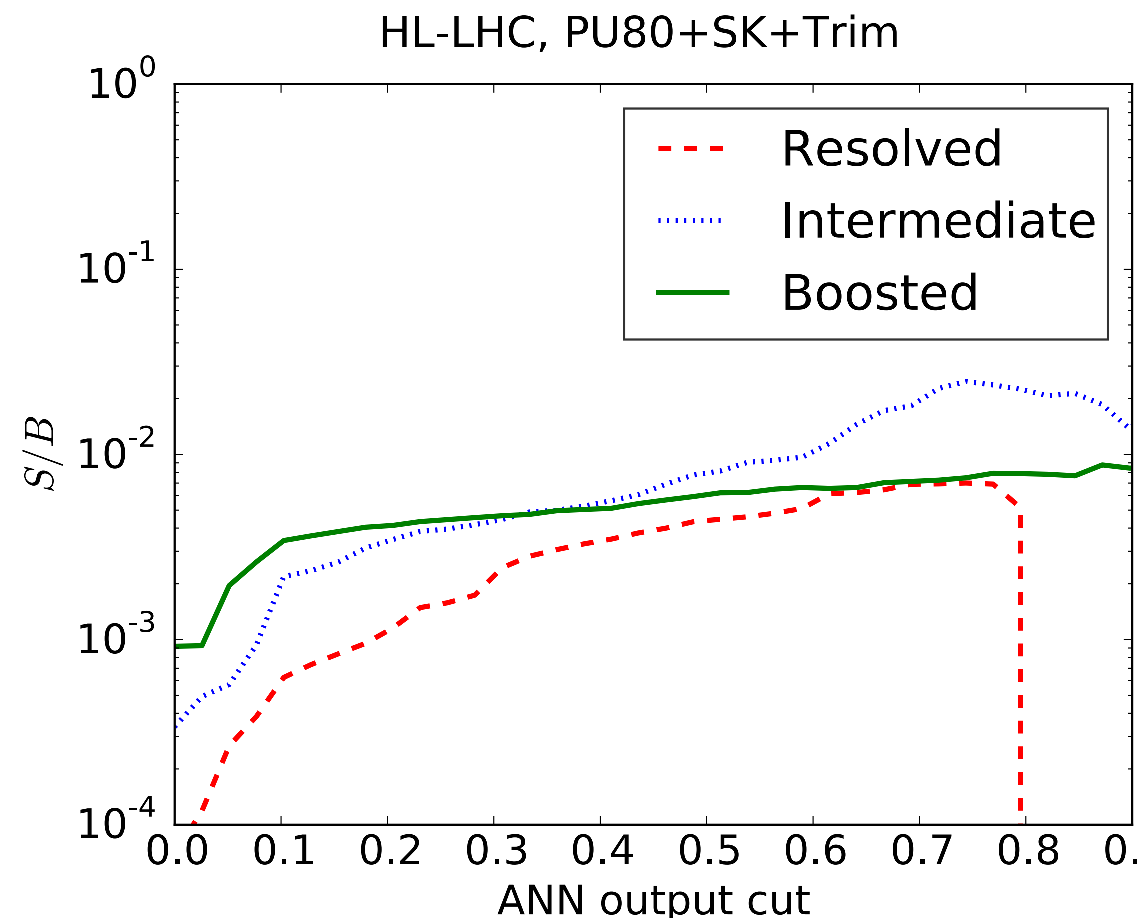


PU is managed by a combination of

- **SoftKiller** subtraction at the event level [Cacciari et al 1407.0408]
- Large-R jets **trimmed** [Krohn et al 0912.1342]

RESULTS





Category		signal	background		$S/\sqrt{B_{\rm tot}}$	$S/\sqrt{B_{\rm 4b}}$	$S/B_{\rm tot}$	$S/B_{\rm 4b}$
		$N_{\rm ev}$	$N_{\rm ev}^{\rm tot}$	$N_{\rm ev}^{\rm 4b}$				
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		

CONCLUSIONS

$HH \to (b\bar{b})(b\bar{b})$ *is a **tough** process to measure, however*

With a combined arms strategy:

- ***Multiple topologies***
- ***Substructure***
- ***Multivariate-analysis***

An observation of HH production in the bbbb channel is feasible at the HL-LHC

Combined $S/\sqrt{B} \sim 3.1$ *after MVA*

Directions for potential improvements:

- ***b-tagging efficiency/purity***
- ***Jet mass resolution***

Next step - *what bounds can we obtain upon* λ?

Thank you for listening!

ANALYSIS BREAKDOWN – NO PILEUP

HL-LHC, Resolved category, no PU										
	Cross-section [fb]						S/B		$S/\sqrt{B}$	
	$hh4b$	total bkg	$4b$	$2b2j$	$4j$	$t\bar{t}$	tot	$4b$	tot	$4b$
C1a	9	$2.2 \cdot 10^8$	$6.9 \cdot 10^4$	$1.5 \cdot 10^7$	$2.0 \cdot 10^8$	$2.1 \cdot 10^5$	$4.0 \cdot 10^{-8}$	$1.3 \cdot 10^{-4}$	0.03	1.9
C1b	9	$2.2 \cdot 10^8$	$6.9 \cdot 10^4$	$1.5 \cdot 10^7$	$2.0 \cdot 10^8$	$2.1 \cdot 10^5$	$4.0 \cdot 10^{-8}$	$1.3 \cdot 10^{-4}$	0.03	1.9
C1c	2.6	$4.4 \cdot 10^7$	$1.6 \cdot 10^4$	$3.2 \cdot 10^6$	$4.1 \cdot 10^7$	$8.8 \cdot 10^4$	$6.1 \cdot 10^{-8}$	$1.6 \cdot 10^{-4}$	0.02	1.1
C2	0.5	$4.9 \cdot 10^3$	$1.7 \cdot 10^3$	$2.9 \cdot 10^3$	$2.1 \cdot 10^2$	47	$1.1 \cdot 10^{-4}$	$2.9 \cdot 10^{-4}$	0.4	0.6

HL-LHC, Intermediate category, no PU										
	Cross-section [fb]						S/B		$S/\sqrt{B}$	
	$hh4b$	total bkg	$4b$	$2b2j$	$4j$	$t\bar{t}$	tot	$4b$	tot	$4b$
C1a	2.8	$8.4 \cdot 10^7$	$2.1 \cdot 10^4$	$5.3 \cdot 10^6$	$7.9 \cdot 10^7$	$3.3 \cdot 10^4$	$3.4 \cdot 10^{-8}$	$1.3 \cdot 10^{-4}$	0.02	1.1
C1b	2.6	$5.8 \cdot 10^7$	$1.4 \cdot 10^4$	$3.6 \cdot 10^6$	$5.5 \cdot 10^7$	$3.0 \cdot 10^4$	$4.5 \cdot 10^{-8}$	$1.9 \cdot 10^{-4}$	0.02	1.2
C1c	0.5	$3.5 \cdot 10^6$	$8.7 \cdot 10^2$	$2.1 \cdot 10^5$	$4.3 \cdot 10^7$	$8.8 \cdot 10^3$	$1.6 \cdot 10^{-7}$	$6.1 \cdot 10^{-4}$	0.02	1.0
C2	0.09	$1.8 \cdot 10^2$	56	96	22	3.1	$5.3 \cdot 10^{-4}$	$1.6 \cdot 10^{-3}$	0.4	0.6

HL-LHC, Boosted category, no PU										
	Cross-section [fb]						S/B		$S/\sqrt{B}$	
	$hh4b$	total bkg	$4b$	$2b2j$	$4j$	$t\bar{t}$	tot	$4b$	tot	$4b$
C1a	3.9	$4.6 \cdot 10^7$	$1.1 \cdot 10^4$	$2.9 \cdot 10^6$	$4.3 \cdot 10^7$	$2.4 \cdot 10^4$	$8.2 \cdot 10^{-8}$	$3.4 \cdot 10^{-4}$	0.03	2.0
C1b	2.7	$3.7 \cdot 10^7$	$7.5 \cdot 10^3$	$2.1 \cdot 10^6$	$3.5 \cdot 10^7$	$2.2 \cdot 10^4$	$7.4 \cdot 10^{-8}$	$3.7 \cdot 10^{-4}$	0.03	1.7
C1c	1.0	$3.9 \cdot 10^6$	$8.0 \cdot 10^2$	$2.3 \cdot 10^5$	$3.7 \cdot 10^6$	$7.1 \cdot 10^3$	$2.6 \cdot 10^{-7}$	$1.3 \cdot 10^{-3}$	0.03	2.0
C2	0.16	$2.5 \cdot 10^2$	53	$1.9 \cdot 10^2$	13	1.6	$5.7 \cdot 10^{-4}$	$2.7 \cdot 10^{-3}$	0.5	1.1

ANALYSIS BREAKDOWN – INCLUDING PILEUP

HL-LHC, Resolved category, PU+SK with $n_{\mathrm{PU}} = 80$										
	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

HL-LHC, Intermediate category, PU+SK+Trim with $n_{\mathrm{PU}} = 80$										
	Cross-section [fb]						S/B		$S/\sqrt{B}$	
	$hh4b$	total bkg	$4b$	$2b2j$	$4j$	$t\bar{t}$	tot	$4b$	tot	$4b$
C1b	2.7	$8.1 \cdot 10^7$	$2.1 \cdot 10^4$	$5.2 \cdot 10^6$	$7.6 \cdot 10^7$	$3.0 \cdot 10^4$	$3.4 \cdot 10^{-8}$	$1.3 \cdot 10^{-4}$	0.02	1.0
C1c	2.6	$6.2 \cdot 10^7$	$1.5 \cdot 10^4$	$3.9 \cdot 10^6$	$5.8 \cdot 10^7$	$2.8 \cdot 10^4$	$4.1 \cdot 10^{-8}$	$1.7 \cdot 10^{-4}$	0.02	1.1
C1d	0.5	$2.8 \cdot 10^6$	$7.9 \cdot 10^2$	$1.9 \cdot 10^5$	$2.7 \cdot 10^6$	$6.5 \cdot 10^3$	$1.8 \cdot 10^{-7}$	$6.2 \cdot 10^{-4}$	0.02	1.0
C2	0.09	$2.6 \cdot 10^2$	47	$1.8 \cdot 10^2$	30	2.2	$3.4 \cdot 10^{-4}$	$1.8 \cdot 10^{-3}$	0.3	0.7

HL-LHC, Boosted category, PU+SK+Trim with $n_{\mathrm{PU}} = 80$										
	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