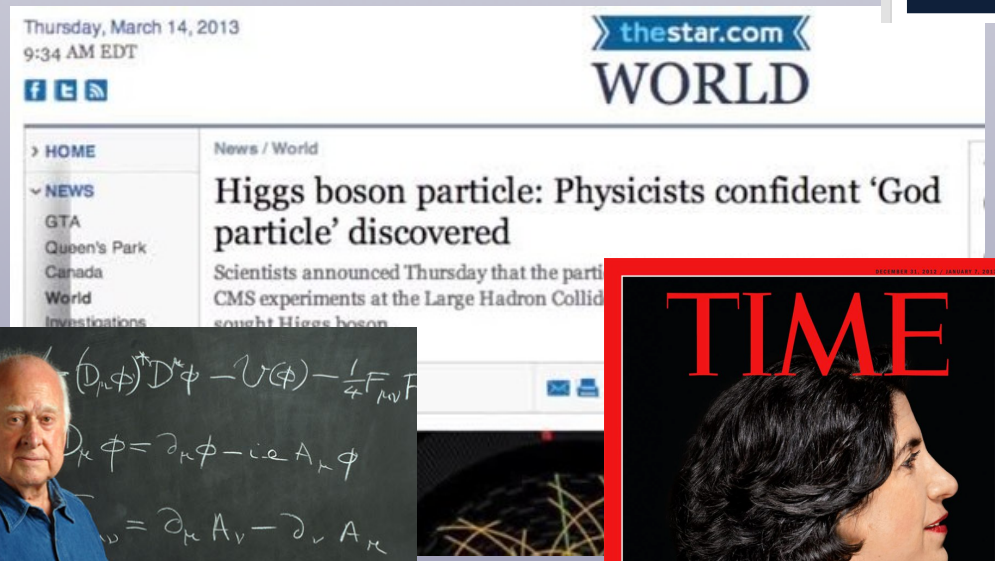
Probing electroweak symmetry breaking with Higgs pair production at the LHC

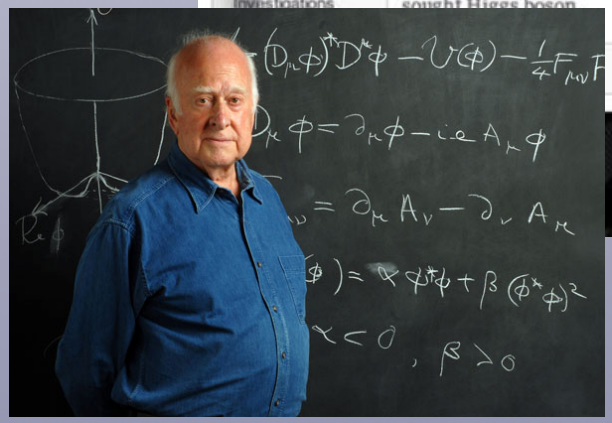
Juan Rojo
STFC Rutherford Fellow
Rudolf Peierls Center for Theoretical Physics
University of Oxford

High Energy Physics seminar
Southampton, 18/02/2016

Particle Physics in the headlines

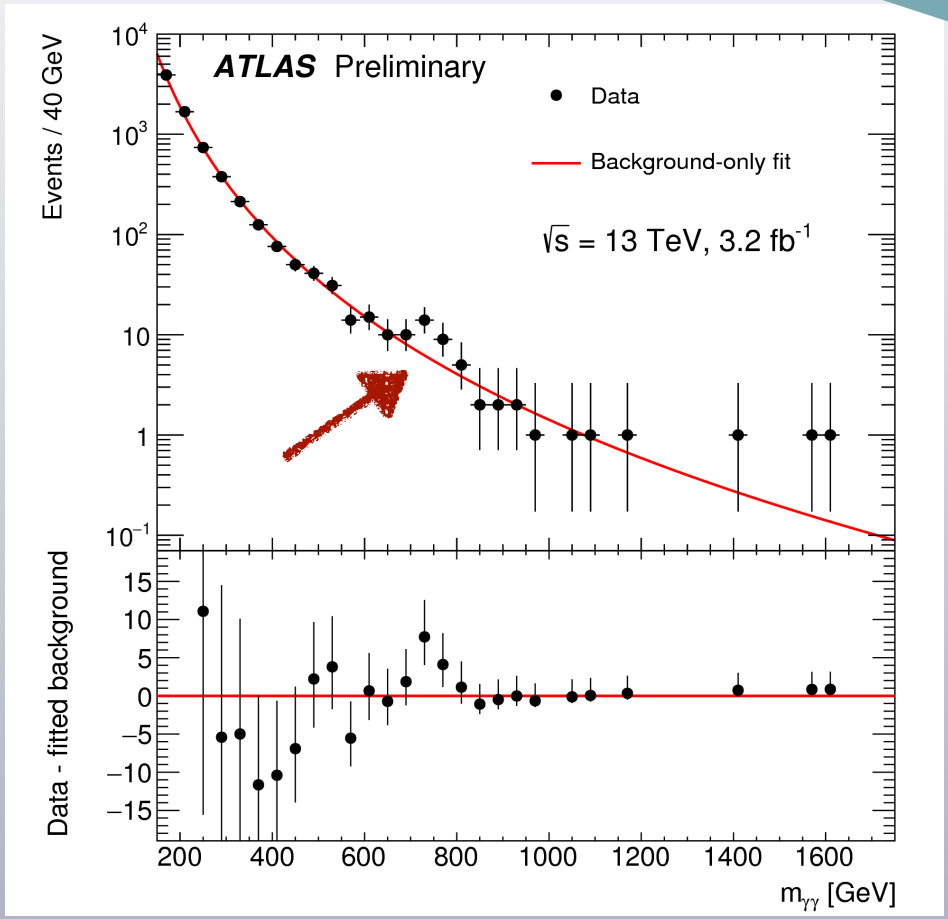
- The **Higgs Boson** is the **most important discovery in particle physics** in 25 years
- The Higgs completes the **extremely successful Standard Model of particle physics**, but at the same time **opens a number of crucial questions** for the field that we need to address
- The **LHC** will play a crucial role in **exploring the energy frontier in the next 20 years**

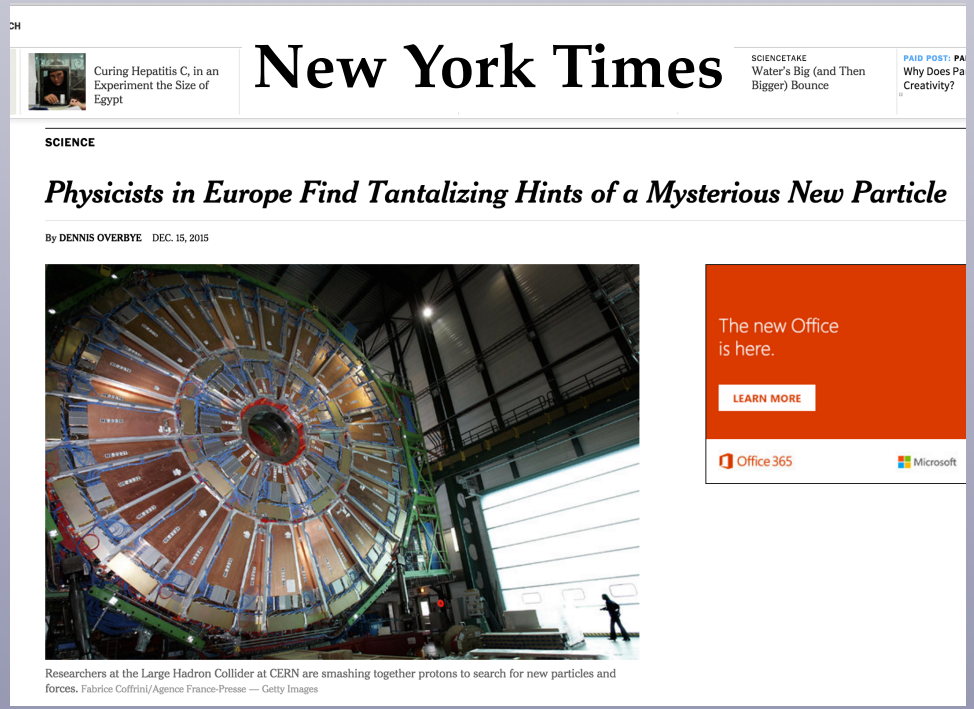


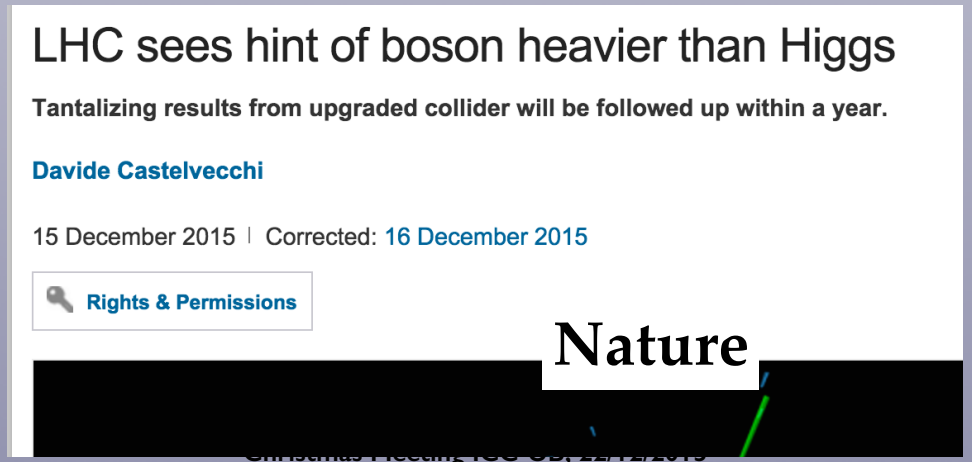


Particle Physics in the headlines AGAIN!

- ☑ **First results from 13 TeV from ATLAS and CMS** announced last december
- ☑ An **excess in the diphoton spectra at 750 GeV** seen by the two experiments, with a **combined significance around 3.0 sigma**
- ☑ Already **O(200)** possible explanations in arXiv ….



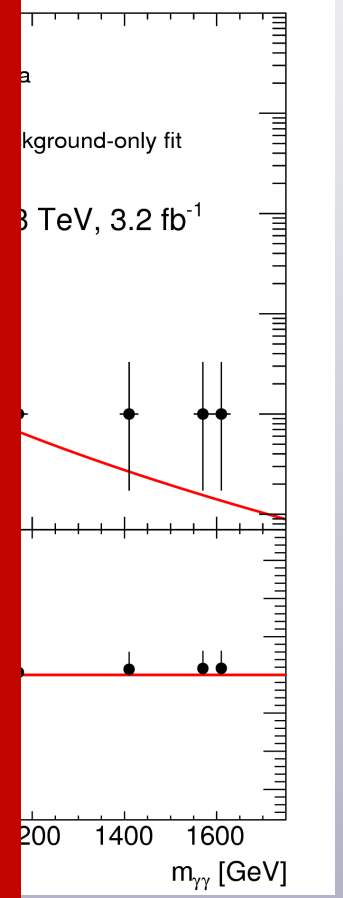




Pa… …s

AGAIN!

- First results fro… announced last w…
- An excess in the… by the two experi… significance arou…
- Already O(100) …



…er than Higgs

…owed up within a year.

…ture

Outstanding Questions in Particle Physics *circa* 2014

… there has never been a better time to be a particle physicist!

Higgs boson and EWSB

- ❑ Is m_H natural or fine-tuned ?
- → if natural: what new physics/symmetry ?
- ❑ does it regularize the divergent $W_L W_L$ cross-section at high $M(W_L W_L)$? Or is there a new dynamics ?
- ❑ elementary or composite Higgs ?
- ❑ is it alone or are there other Higgs bosons ?
- ❑ origin of couplings to fermions
- ❑ coupling to dark matter ?
- ❑ does it violate *CP* ?
- ❑ cosmological EW phase transition

Quarks and leptons:

- ❑ why 3 families ?
- ❑ masses and mixing
- ❑ *CP* violation in the lepton sector
- ❑ matter and antimatter asymmetry
- ❑ baryon and charged lepton number violation

Physics at the highest E-scales:

- ❑ how is gravity connected with the other forces ?
- ❑ do forces unify at high energy ?

Dark matter:

- ❑ composition: WIMP, sterile neutrinos, axions, other hidden sector particles, ..
- ❑ one type or more ?
- ❑ only gravitational or other interactions ?

Neutrinos:

- ❑ *ν* masses and and their origin
- ❑ what is the role of *H*(125) ?
- ❑ Majorana or Dirac ?
- ❑ *CP* violation
- ❑ additional species → sterile *ν* ?

The two epochs of Universe's accelerated expansion:

- ❑ primordial: is inflation correct ? which (scalar) fields? role of quantum gravity?
- ❑ today: dark energy (why is Λ so small?) or modification of gravity theory ?

Many of these crucial questions will be addressed at the Large Hadron Collider

Outstanding Questions in Particle Physics *circa* 2014

… there has never been a better time to be a particle physicist!

Higgs boson and EWSB

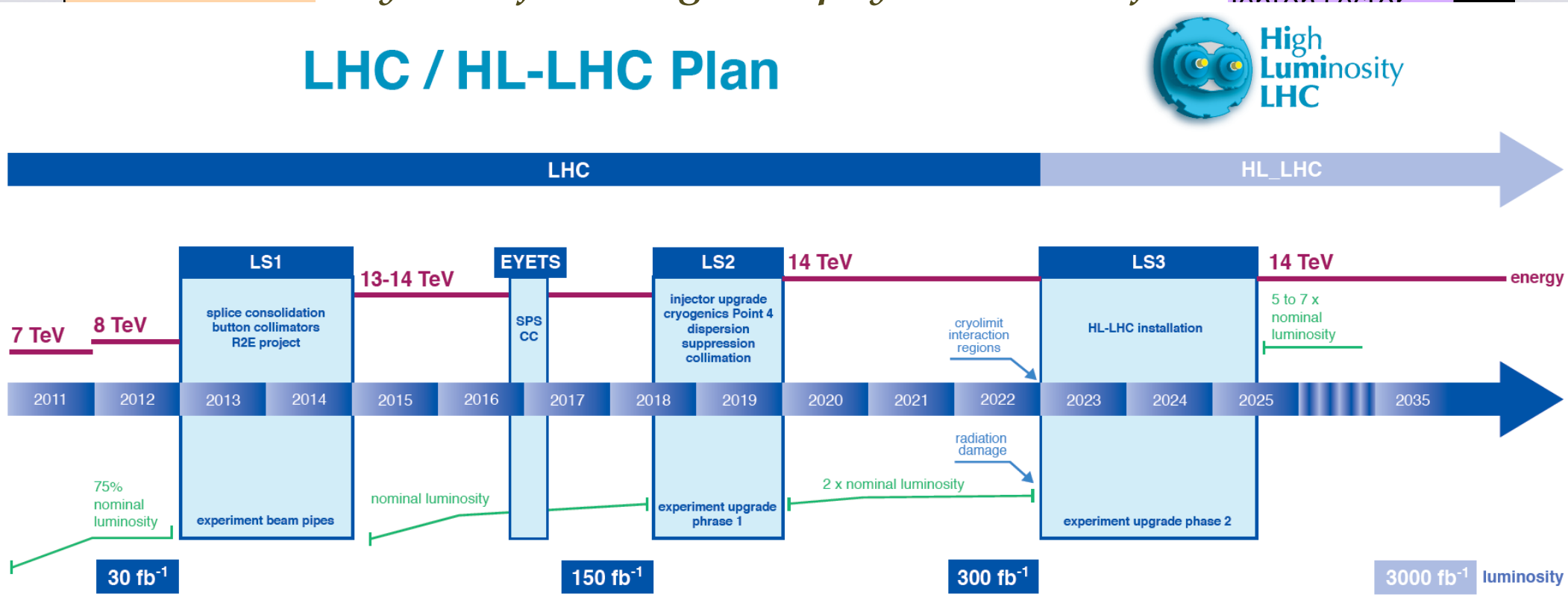
- ❑ Is m_H natural or fine-tuned ?
- → if natural: what new physics/symmetry ?
- ❑ does it regular

Quarks and leptons:

- ❑ why 3 families ?

lepton sector

20 years of exciting LHC physics ahead of us!

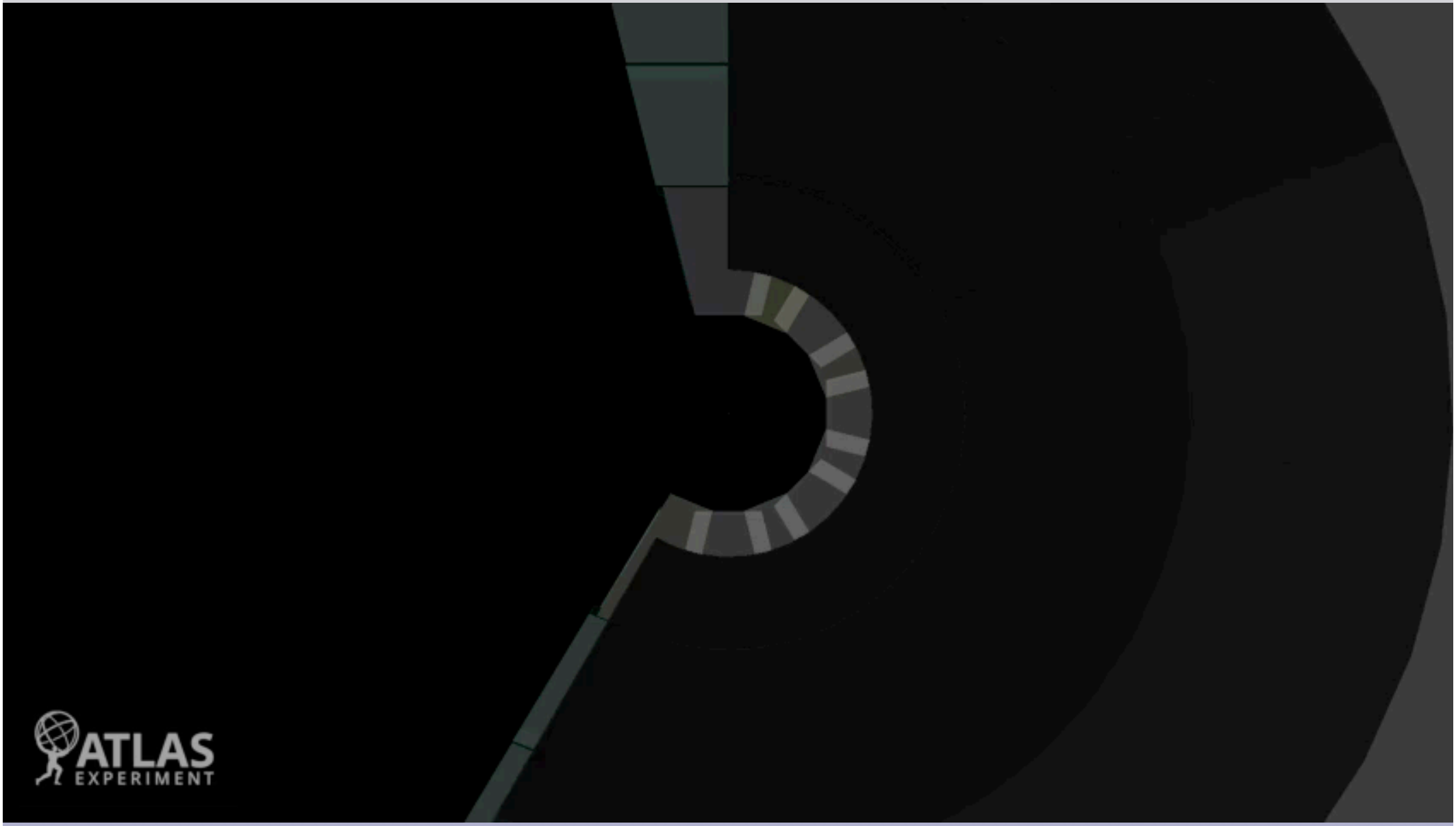


- ❑ primordial: is inflation correct ?
  which (scalar) fields? role of quantum gravity?
- ❑ today: dark energy (why is Λ so small?) or modification of gravity theory ?

- ❑ additional species → sterile *ν* ?

Many of these crucial questions will be addressed at the Large Hadron Collider

A dijet event at 13 TeV



Higgs Pair Production at the LHC

Why Higgs couplings?

Higgs couplings may indicate new physics: a few percent precision is a good target

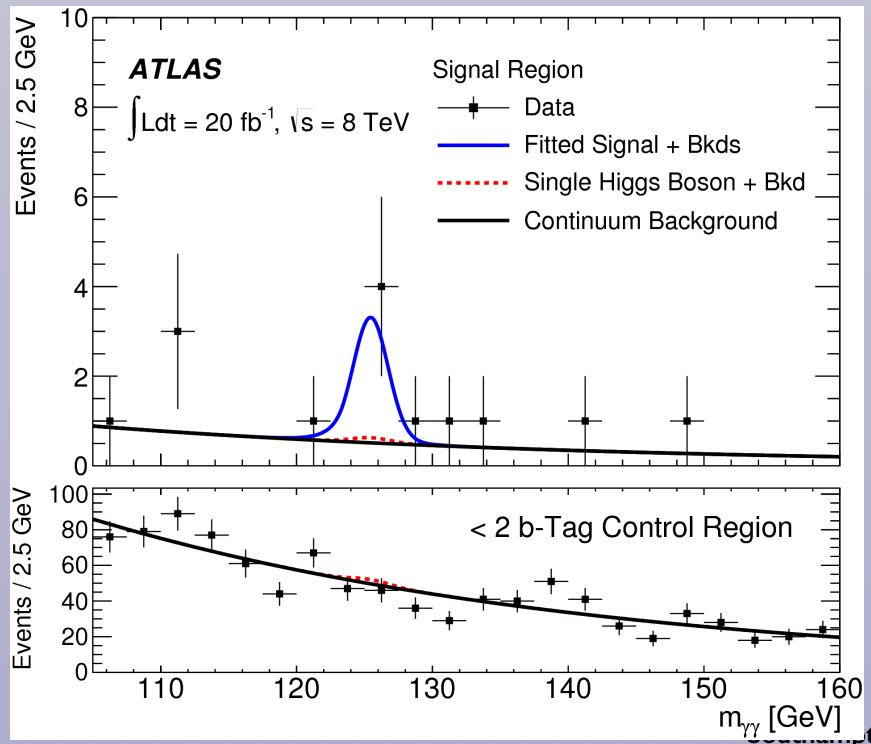
Higgs Snowmass report (arXiv:1310.8361)
Deviation from SM due to particles with M=1 TeV

Model	κ_V	κ_b	κ_γ
Singlet Mixing	$\sim 6\%$	$\sim 6\%$	$\sim 6\%$
2HDM	$\sim 1\%$	$\sim 10\%$	$\sim 1\%$
Decoupling MSSM	$\sim -0.0013\%$	$\sim 1.6\%$	$\sim -.4\%$
Composite	$\sim -3\%$	$\sim -(3-9)\%$	$\sim -9\%$
Top Partner	$\sim -2\%$	$\sim -2\%$	$\sim +1\%$

Future LHC data: measure H couplings at 2-8% level (cf 20-50% today), and to access rare decays such as H→μμ

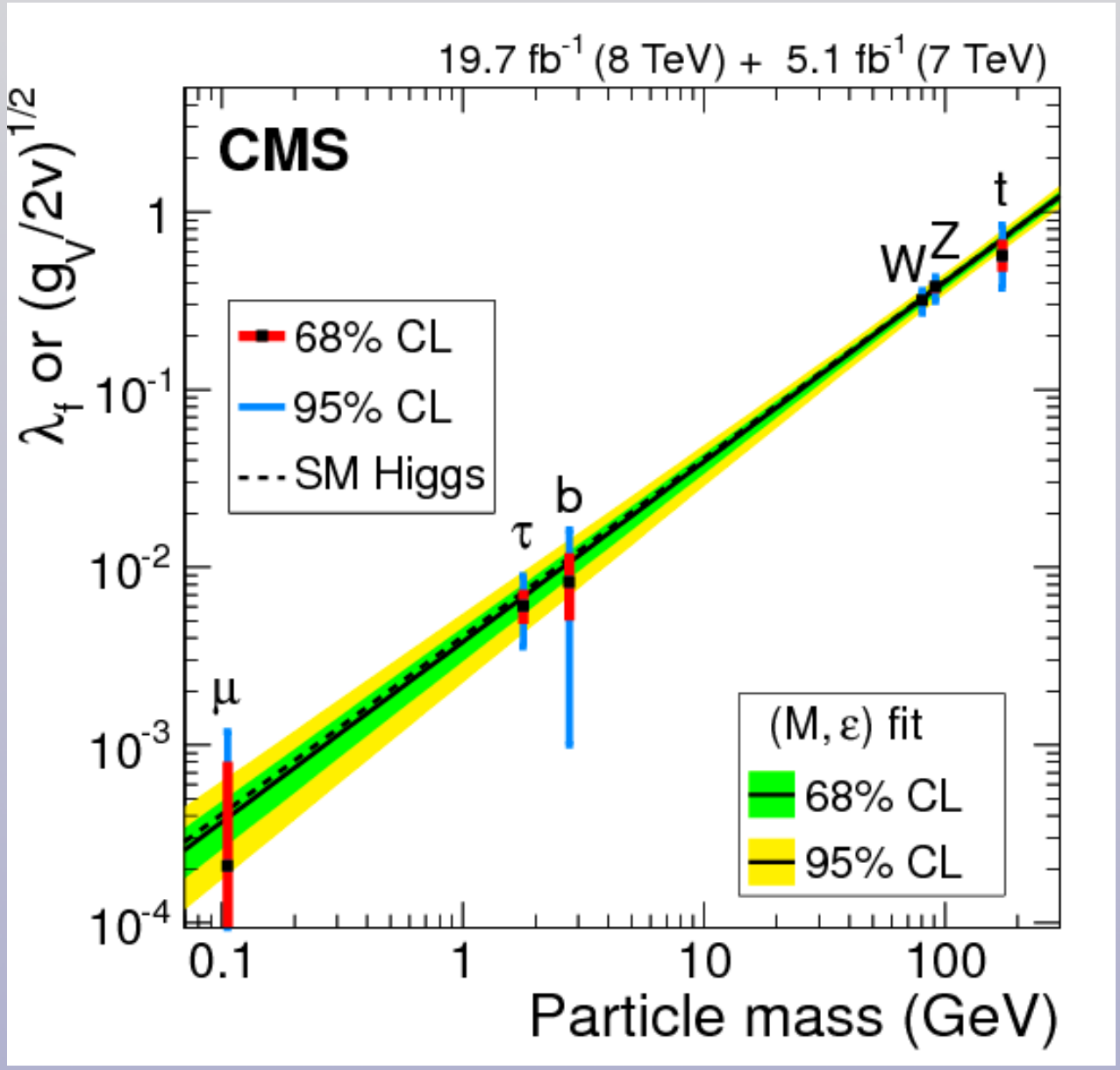
Higgs Pair Production at the LHC

- **Double Higgs production** allows accessing crucial components of the Higgs sector:
  - ☑ Reconstruct the **full electroweak symmetry breaking potential**
  - ☑ Probe the **Higgs self-interaction**
  - ☑ Probe the **doublet nature** of the Higgs by means of the **hhVV coupling**
- In the 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.
- Rates for double Higgs production **generically enhanced in BSM scenarios**, and **LHC searches** in various final states have already started at **Run I**



ATLAS search for resonant hh -> bbγγ production

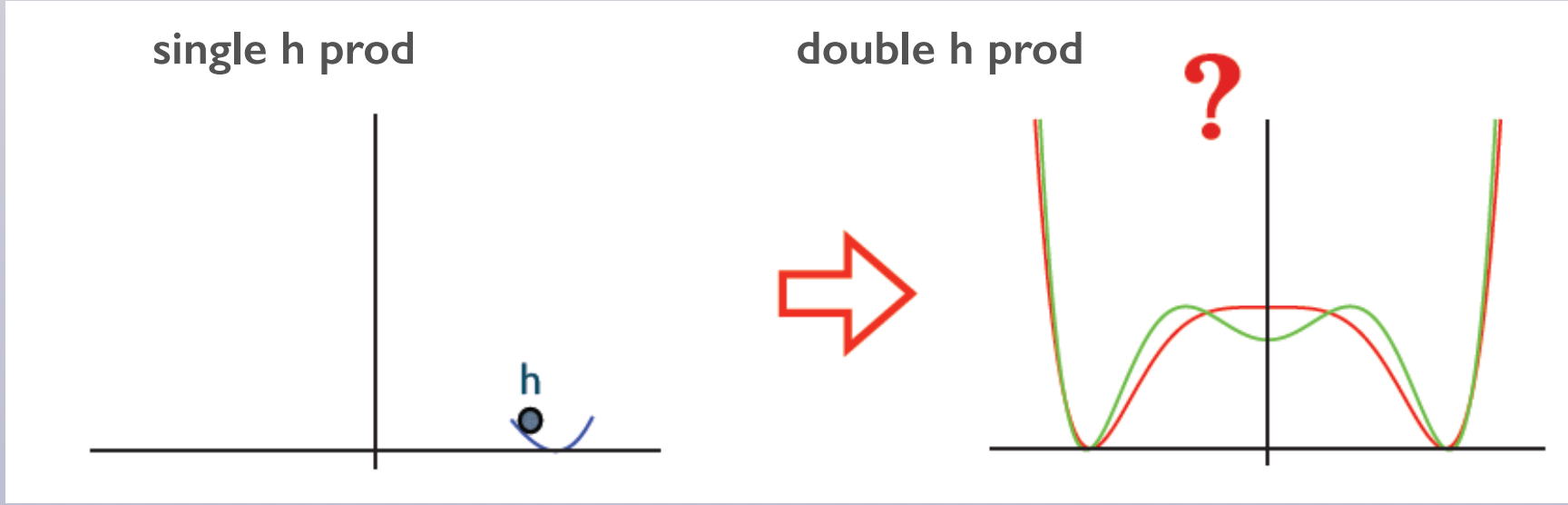
Probing Electroweak Symmetry breaking



- **Yukawa/Couplings** between Higgs and SM particles **proportional to mass**
- The Higgs boson is **responsible to break EW symmetry and give particles mass**
- However, we still lack understanding of **why and how** EWS is broken

Probing Electroweak Symmetry breaking

- **Current measurements** (couplings in single Higgs production) probe **Higgs potential close to minimum**
- Double Higgs production essential to **reconstruct the full Higgs potential** and clarify EWSB mechanism
- The Higgs potential is *ad-hoc*: **many other EWSB mechanisms conceivable**



Higgs mechanism

$$V(h) = m_h^2 h^\dagger h + \frac{1}{2}\lambda (h^\dagger h)^2$$

Coleman-Weinberg mechanism

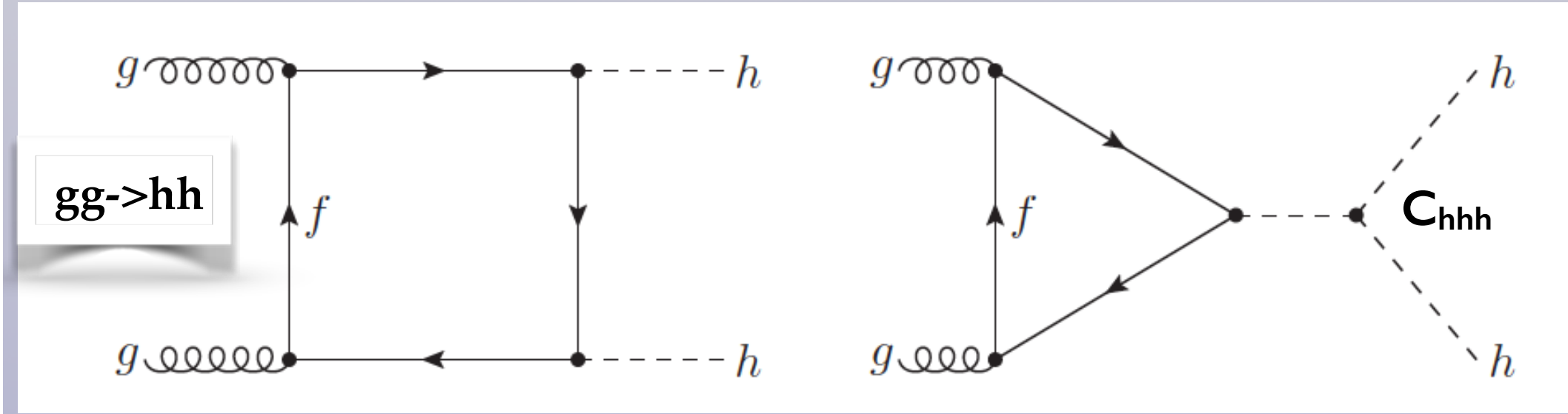
$$V(h) \to \frac{1}{2}\lambda (h^\dagger h)^2 \log\left[\frac{(h^\dagger h)}{m^2}\right]$$

Each possibility associated to **completely different EWSB mechanism**, with crucial implications for the **hierarchy problem**, the structure of quantum field theory, and **New Physics at the EW scale**

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

Exploiting the 4b final state

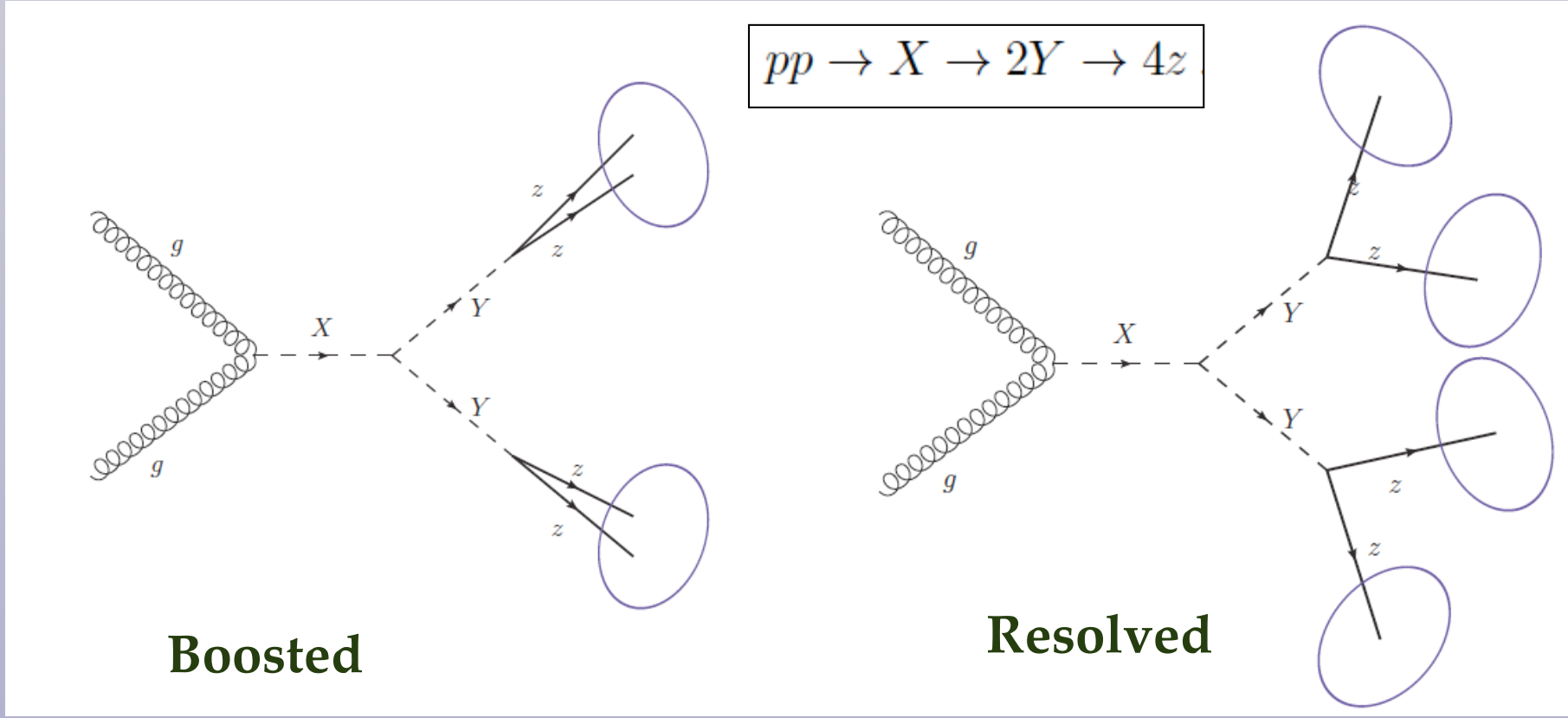
- Focus on the **hh->4b final state:** largest rates, but **overwhelming QCD multijet background**
- Made competitive requiring the **di-Higgs system to be boosted** and exploiting kinematic differences between signal and QCD background with **jet substructure**
- Mandatory to optimize the **boosted b-tagging techniques,** impressive recent progress by ATLAS and CMS



Interference between box and triangle complicates the extraction of the self-coupling

Exploiting the 4b final state

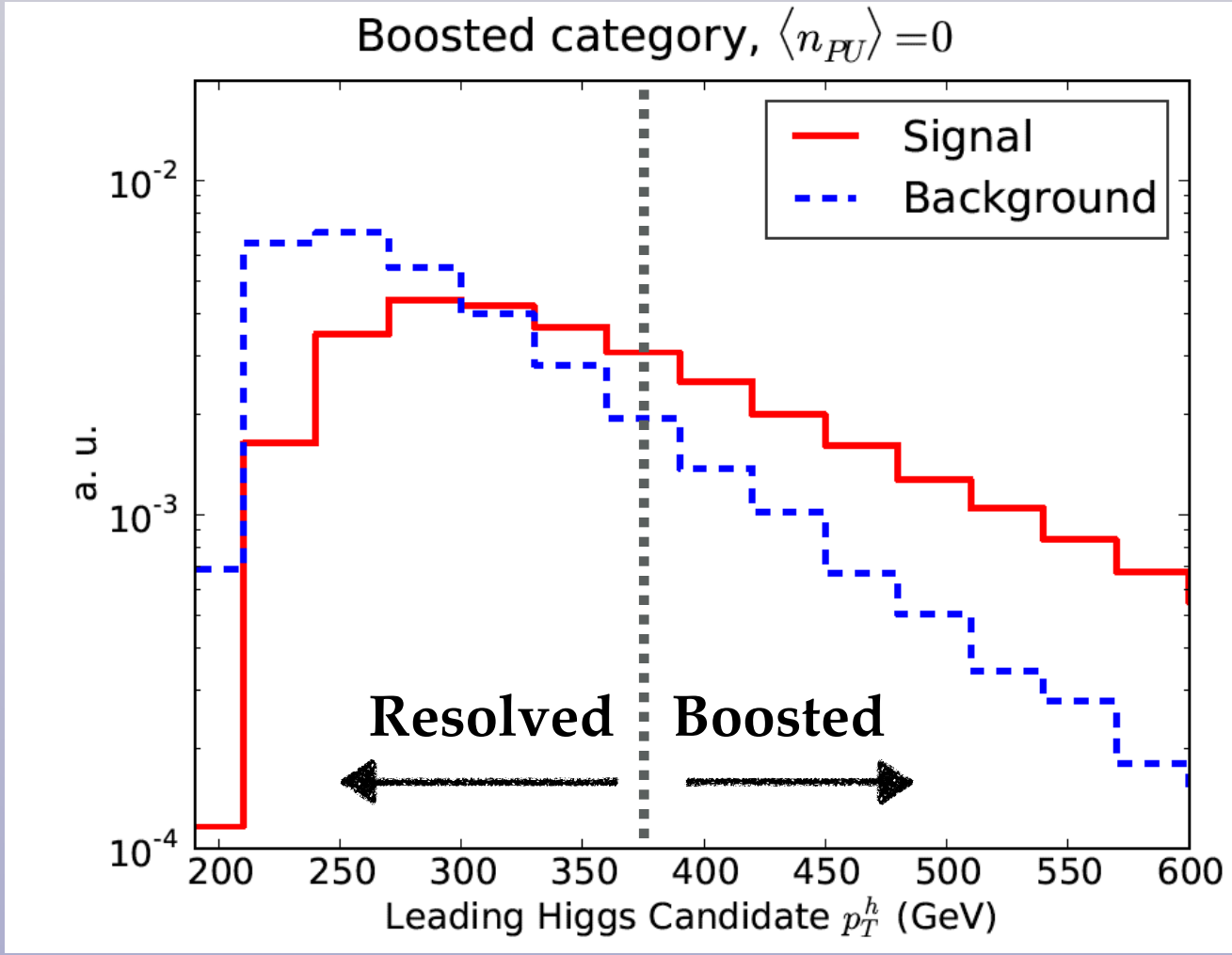
- Focus on the **hh->4b final state:** largest rates, but **overwhelming QCD multijet background**
- Made competitive requiring the **di-Higgs system to be boosted** and exploiting kinematic differences between signal and QCD background with **jet substructure**
- Mandatory to optimize the **boosted b-tagging techniques,** impressive recent progress by ATLAS and CMS



Scale-invariant resonance tagging
Gouzevitch, Oliveira, Rojo, Rosenfeld, Salam and Sanz, arxiv:1303.6636

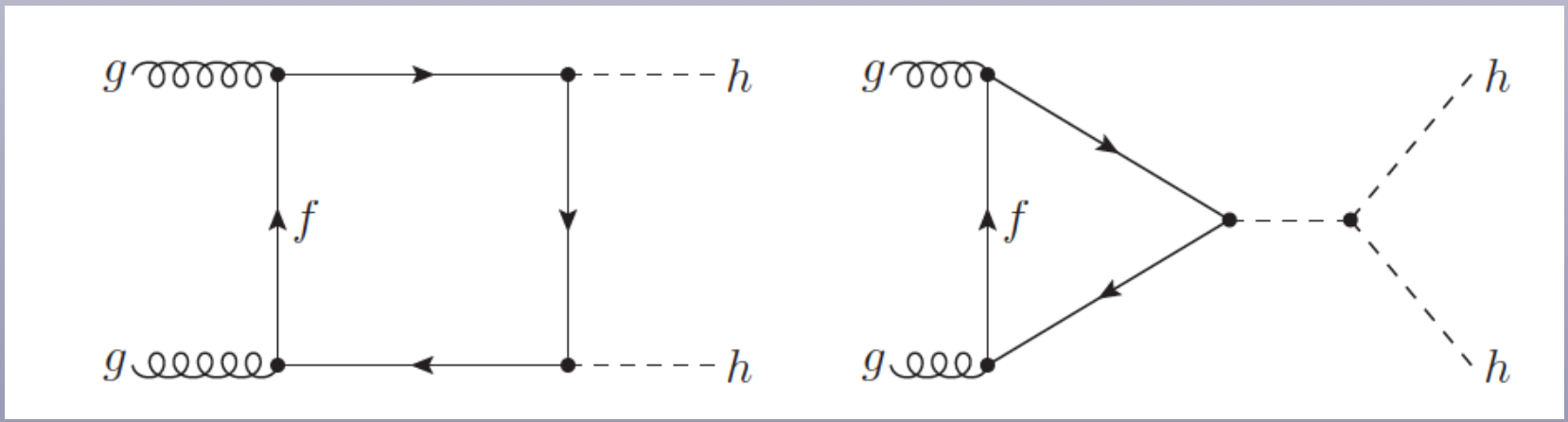
Exploiting the 4b final state

- Focus on the **hh->4b final state:** largest rates, but **overwhelming QCD multijet background**
- Made competitive requiring the **di-Higgs system to be boosted** and exploiting kinematic differences between signal and QCD background with **jet substructure**
- Mandatory to optimize the **boosted b-tagging techniques,** impressive recent progress by ATLAS and CMS



Boosting Higgs Pair Production in the 4b final state with Multivariate Techniques

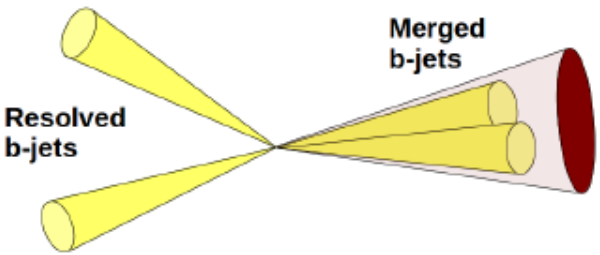
Behr, Bortoletto, Frost, Issever, Hartland, JR, ***arxiv:1512.08928***

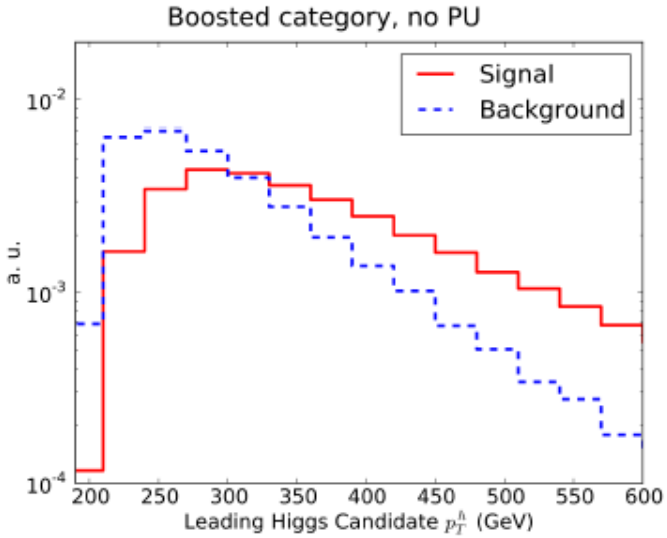


Selection Strategy

Separation between b-jets from $h \to b\bar{b}$ decays shrinks like: $\Delta R(b, \bar{b}) \sim \frac{2m^h}{p_T^h}$

$\Rightarrow$ Single large-R jet for $p_T^h \sim 250$ GeV





Resolved Higgs

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

b-tagging efficiencies

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

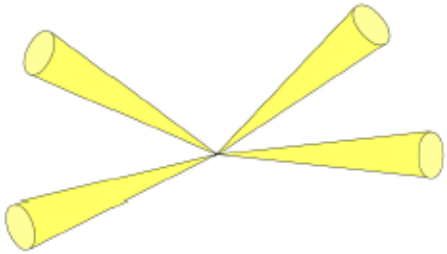
Merged Higgs

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

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

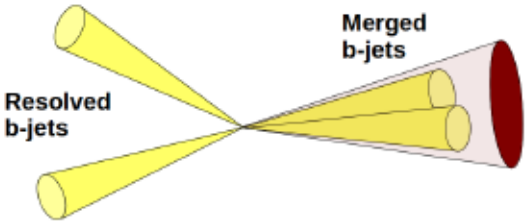
Selection Strategy

Resolved



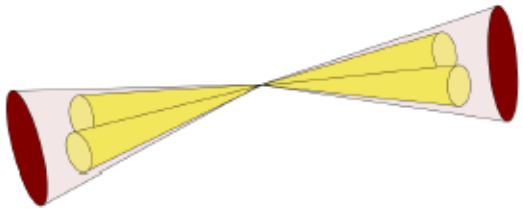
- ≥ 4 b-tagged small-R jets
- Higgs reconstruction from leading 4 jets
- Choice that minimises mass difference between dijet systems

Intermediate



- $= 1$ large-R jet (Higgs-tagged + b-tagged) (leading Higgs)
- ≥ 2 b-tagged small-R jets
- $\Delta R > 1.2$ w.r.t. large-R jet
- Higgs reconstruction from leading 2 small-R jets
- Choice that minimises mass difference of dijet system and large-R jet

Boosted

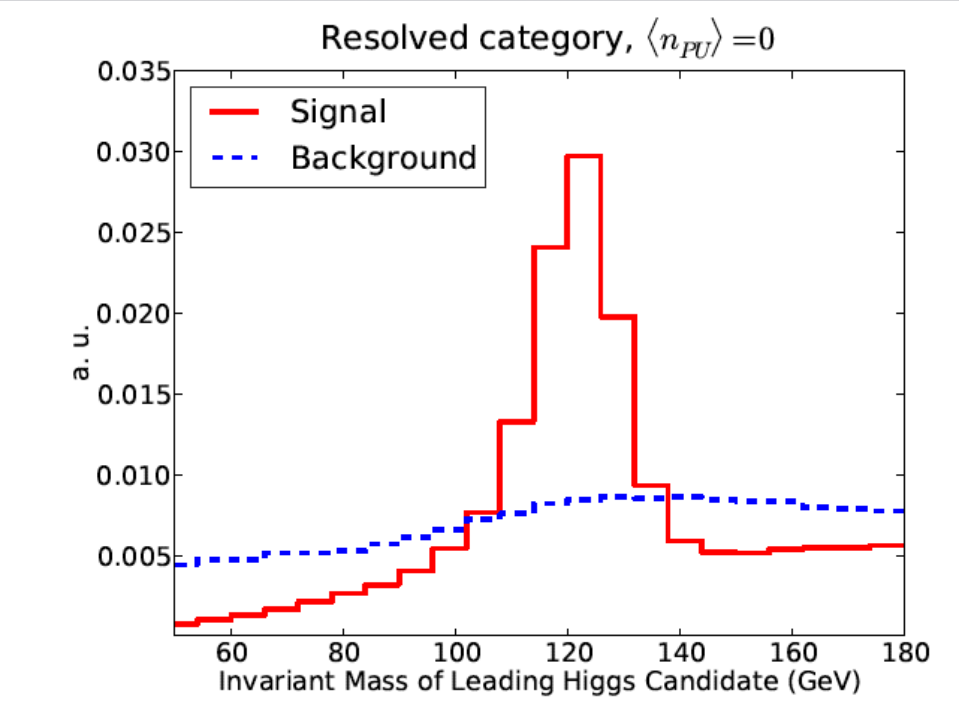


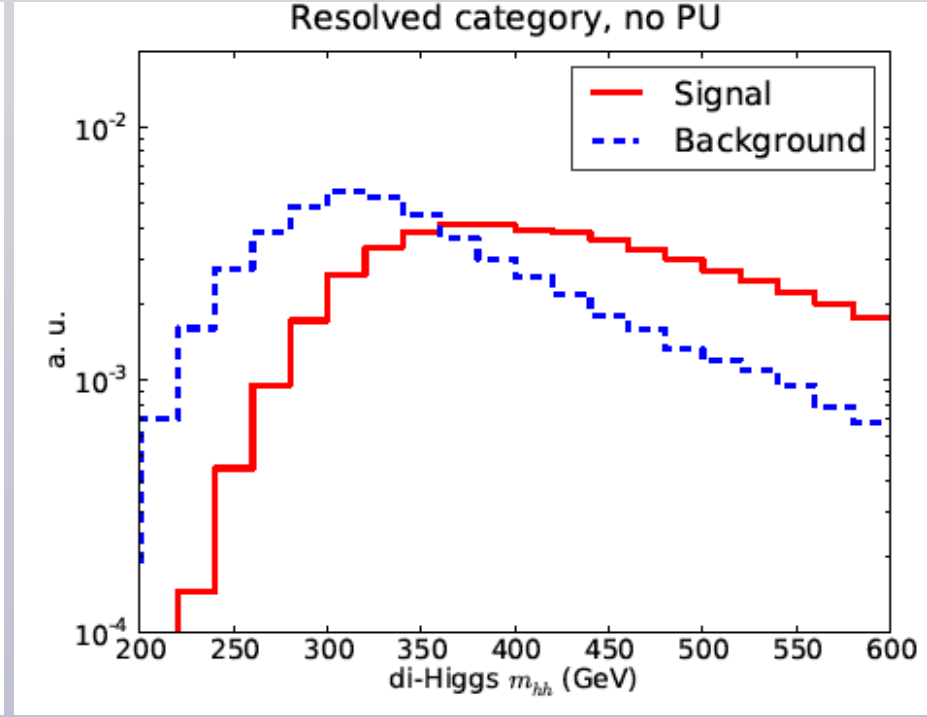
- ≥ 2 large-R jets (Higgs-tagged + b-tagged)
- Leading two jets taken as Higgs candidates

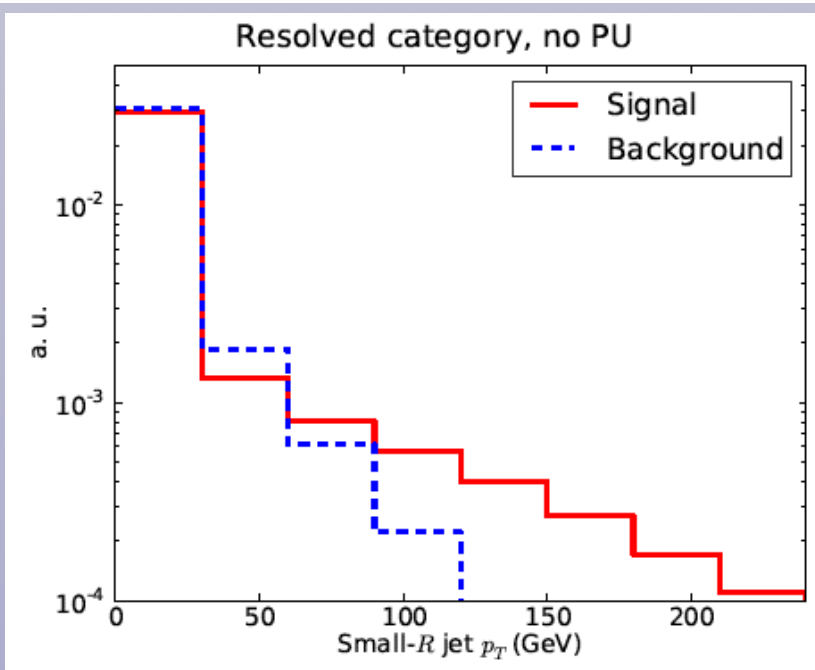
\+ **Loose Higgs mass window cut**: $|m_{h,j} - 125\ \mathrm{GeV}| < 40\ \mathrm{GeV}$, $j = 1, 2$

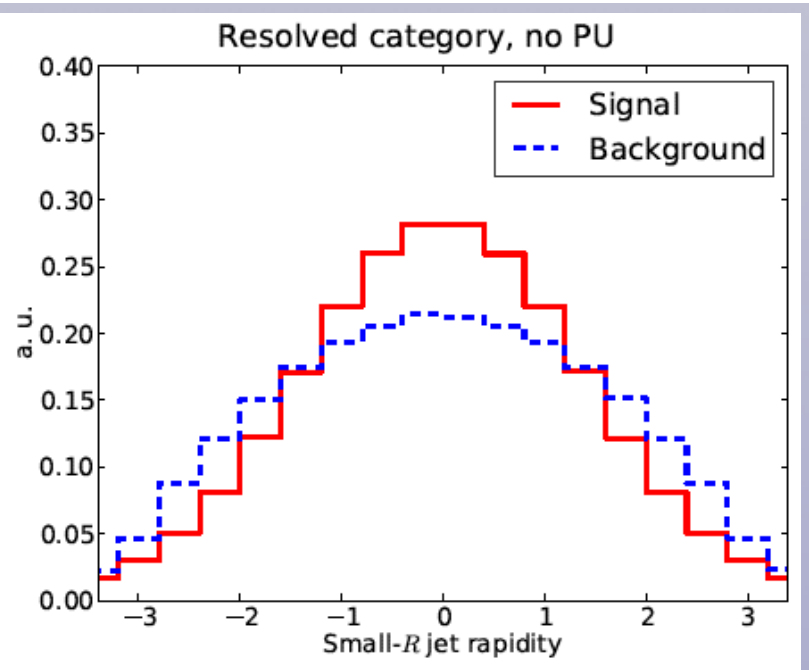
\+ **Rank categories** by $S/\sqrt{B}$ to make them **exclusive**: boosted > intermediate > resolved

pre-MVA results - resolved category

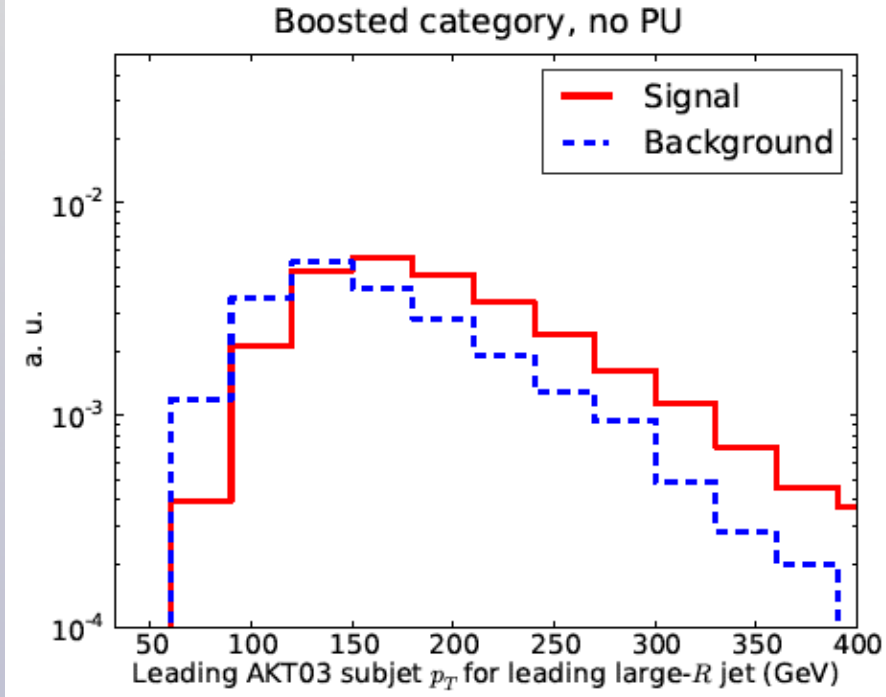


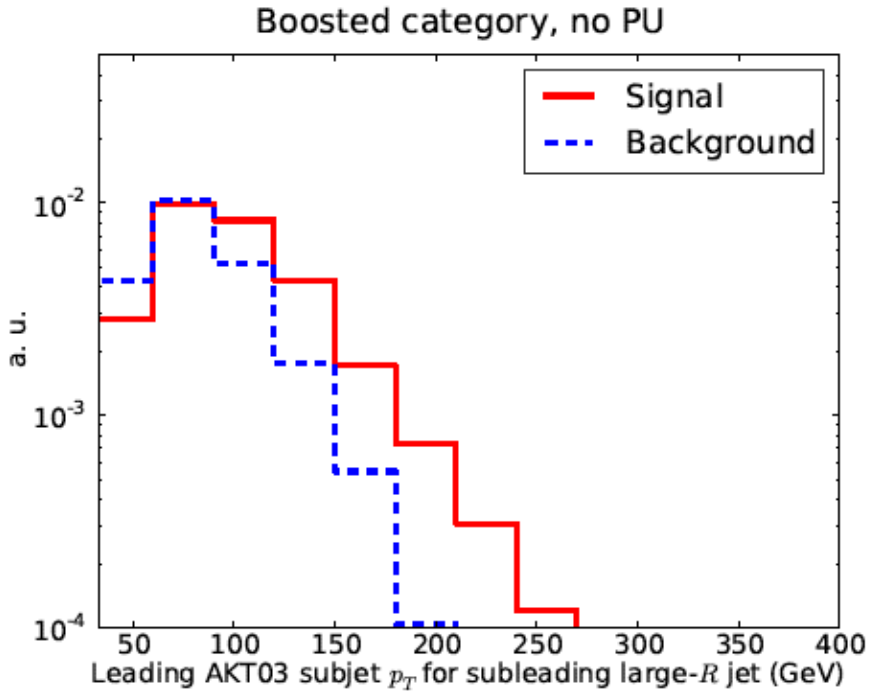


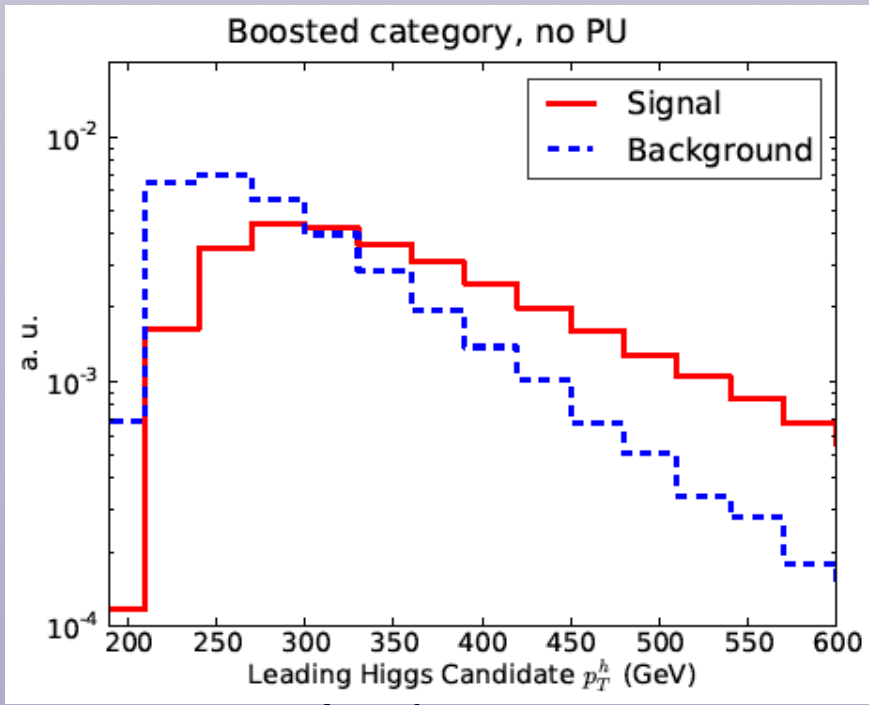


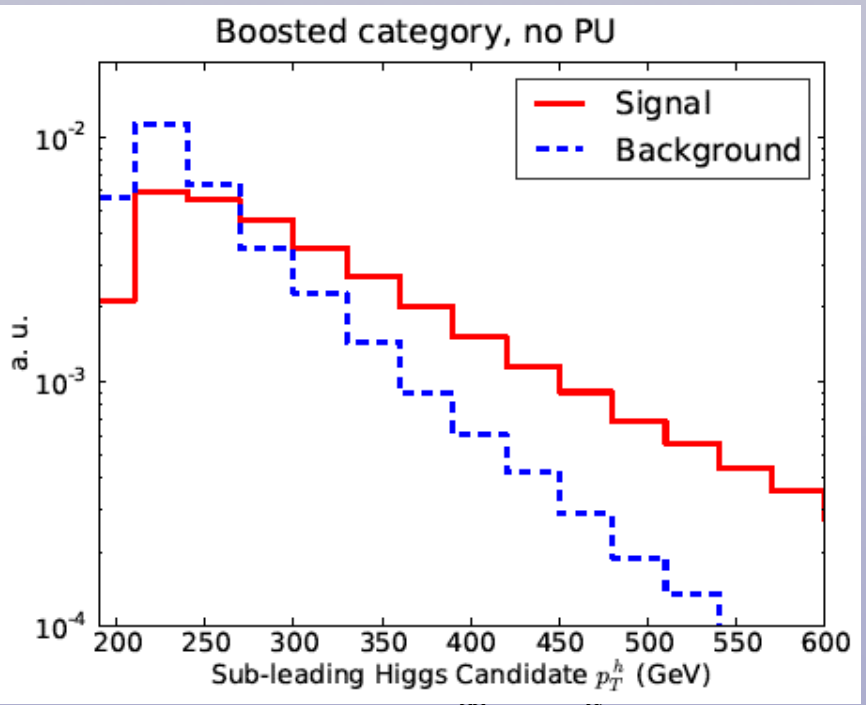


pre-MVA results - boosted category









Jet substructure variables

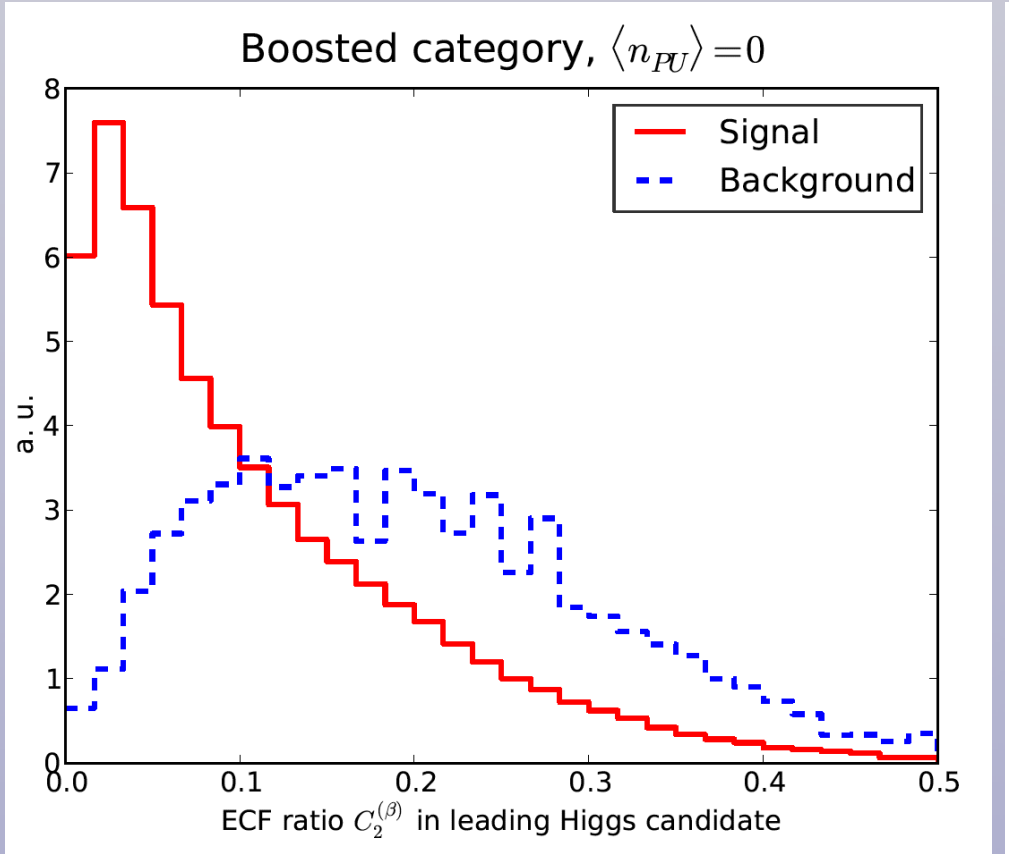
- **Substructure variables** quantify differences in internal structure between **QCD jets** and **jets from the decay of heavy resonances**
- QCD radiation tends to be **soft** and **collinear,** while decay products of resonances **share momentum evenly**

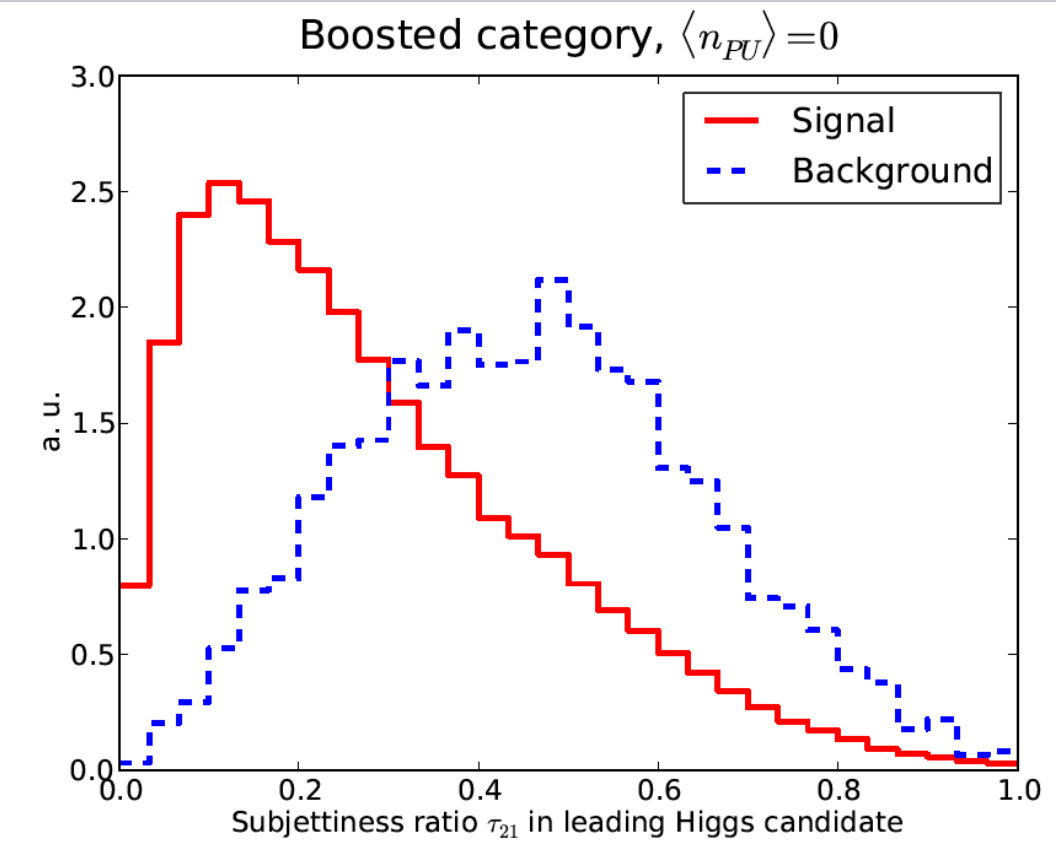
$$C_2^{(\beta)} \equiv \frac{\mathrm{ECF}(3,\beta)\mathrm{ECF}(1,\beta)}{[\mathrm{ECF}(2,\beta)]^2}$$

Energy Correlation Functions ratio

$$\tau_N \equiv \frac{1}{d_0}\sum_k p_{T,k} \cdot \min\left(\delta R_{1k}, \ldots, \delta R_{Nk}\right)$$

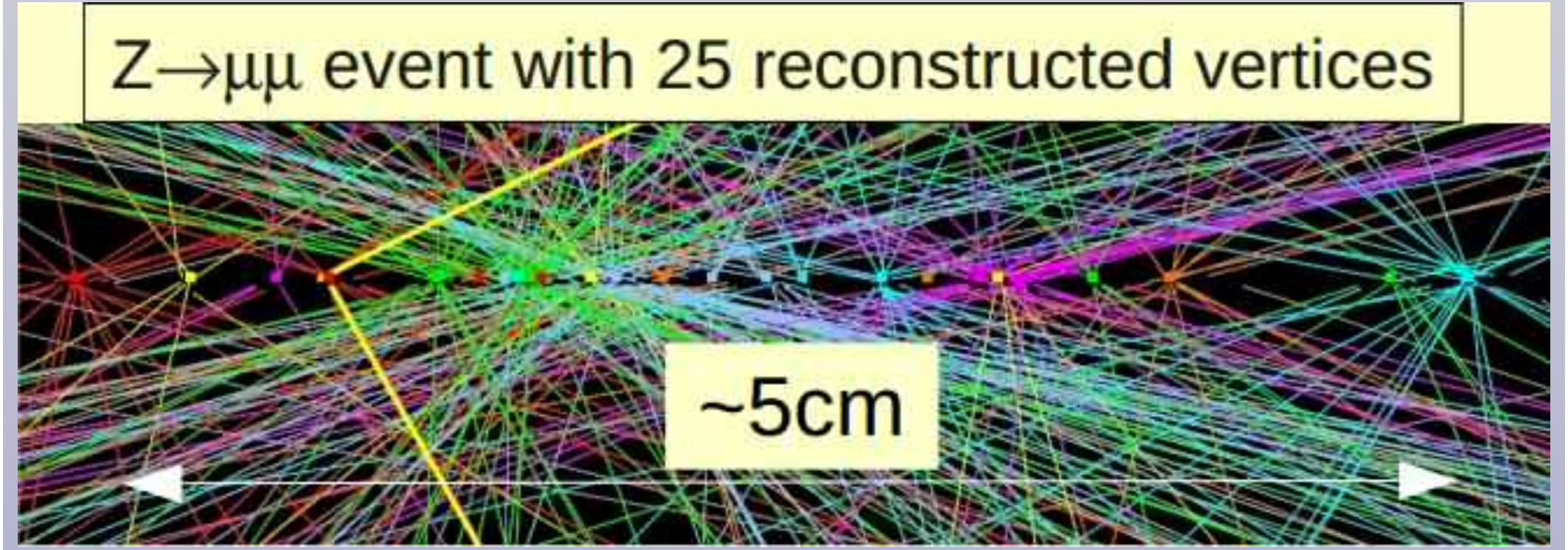
τ_{21}: 2-to-1 Subjettiness ratio





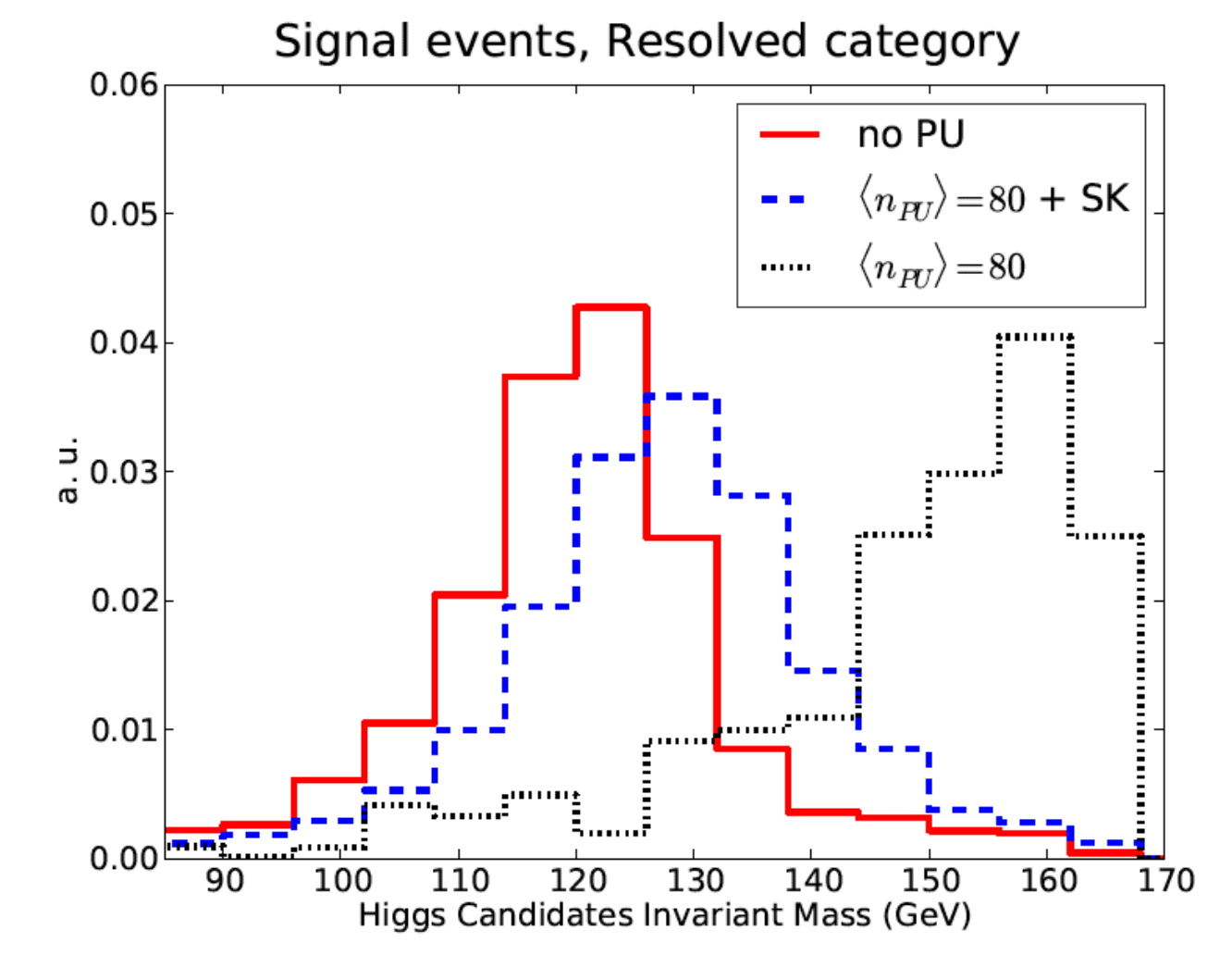
Pile-up at the HL-LHC

- **Pile-up (PU)**: multiple interactions between different pairs of protons in the same **bunch crossing**
- At the **high luminosities** of the HL-LHC, PU will be the **major contamination** in most analysis
- **$\langle n_{PU} \rangle$=150** corresponds to **embedding each hard event into 150 minimum bias events**



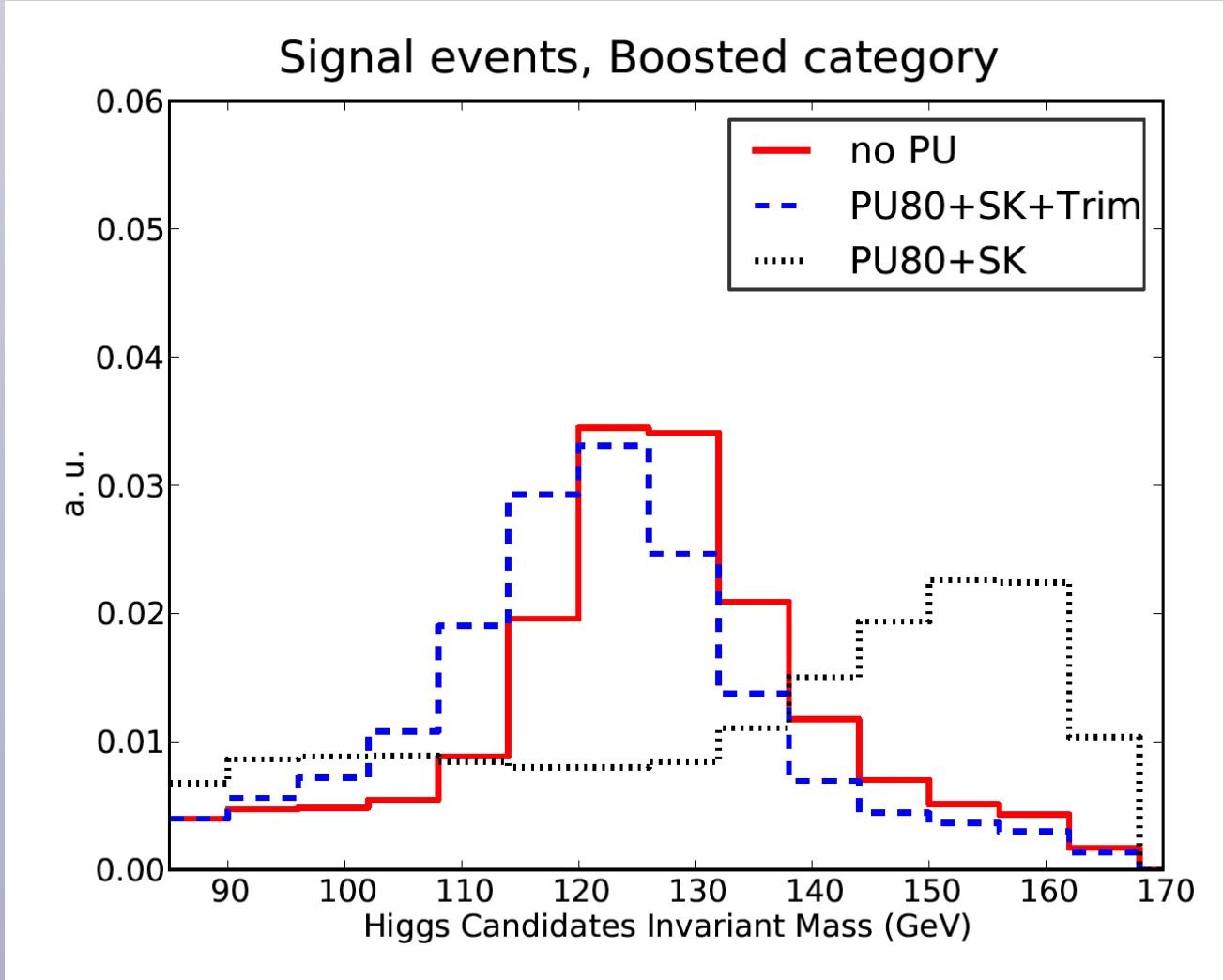
Pile-up at the HL-LHC

- Recent development of **PU subtraction methods** allows to overcome these limitations
- Use **SoftKiller (SK)** to **subtract PU contamination** and maintain most of the signal/background discrimination power of the no-PU scenario SoftKiller: Cacciari, Salam and Soyez, arxiv:1407.0408
- In addition, **jet trimming** is also performed on the **large-*R* jets**



Pile-up at the HL-LHC

- Recent development of **PU subtraction methods** allows to overcome these limitations
- Use **SoftKiller (SK)** to **subtract PU contamination** and maintain most of the signal/background discrimination power of the no-PU scenario SoftKiller: Cacciari, Salam and Soyez, arxiv:1407.0408
- In addition, **jet trimming** is also performed on the **large-*R* jets**



Multivariate techniques

Large number of kinematic variables to disentangle signal and background, **how to combine them?**

Multivariate techniques: Identify automatically kinematical variables with most discrimination power

Higgs pT

Higgs m

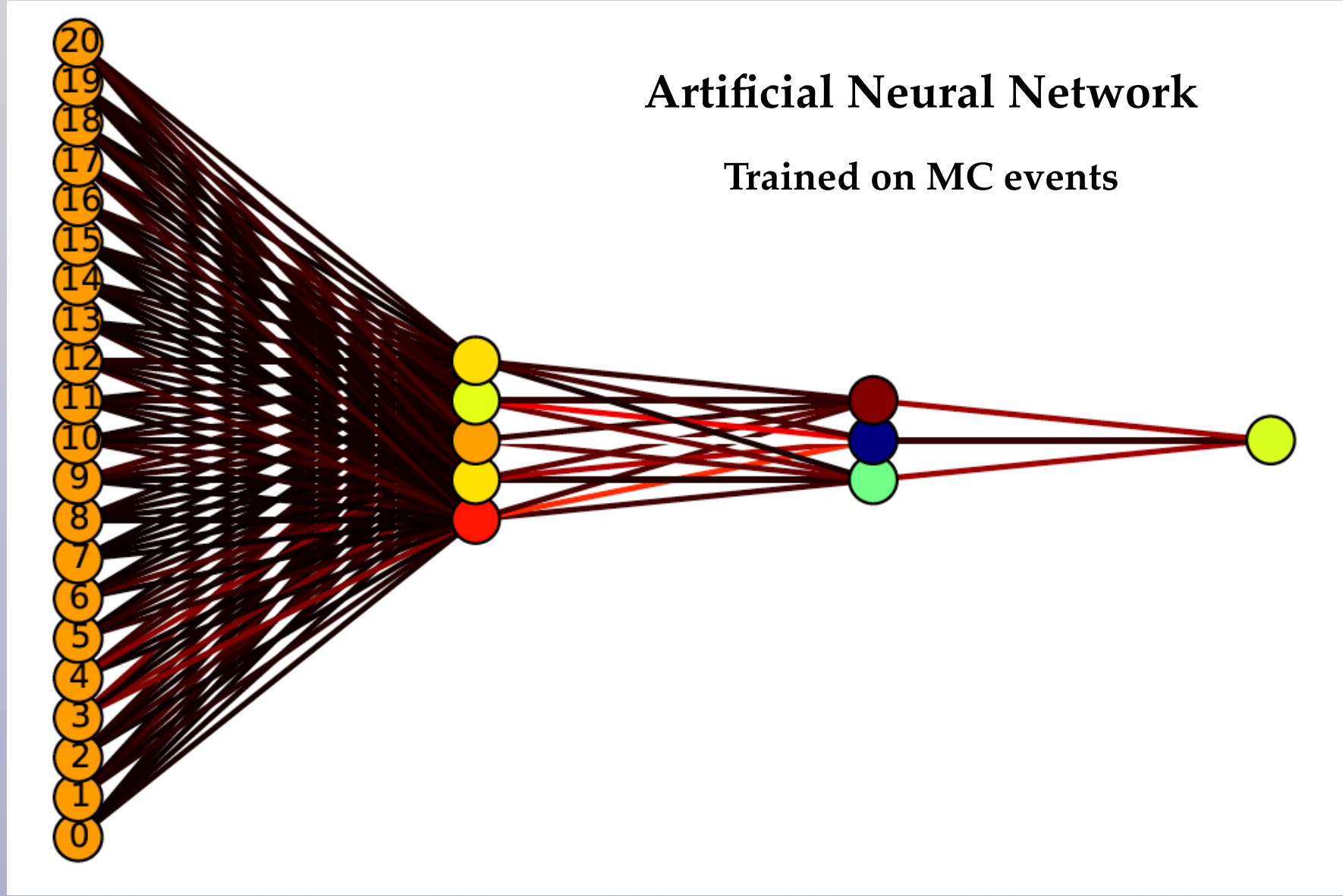
di-Higgs m

ECF

τ_{12}

Subjet pT

.....



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

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

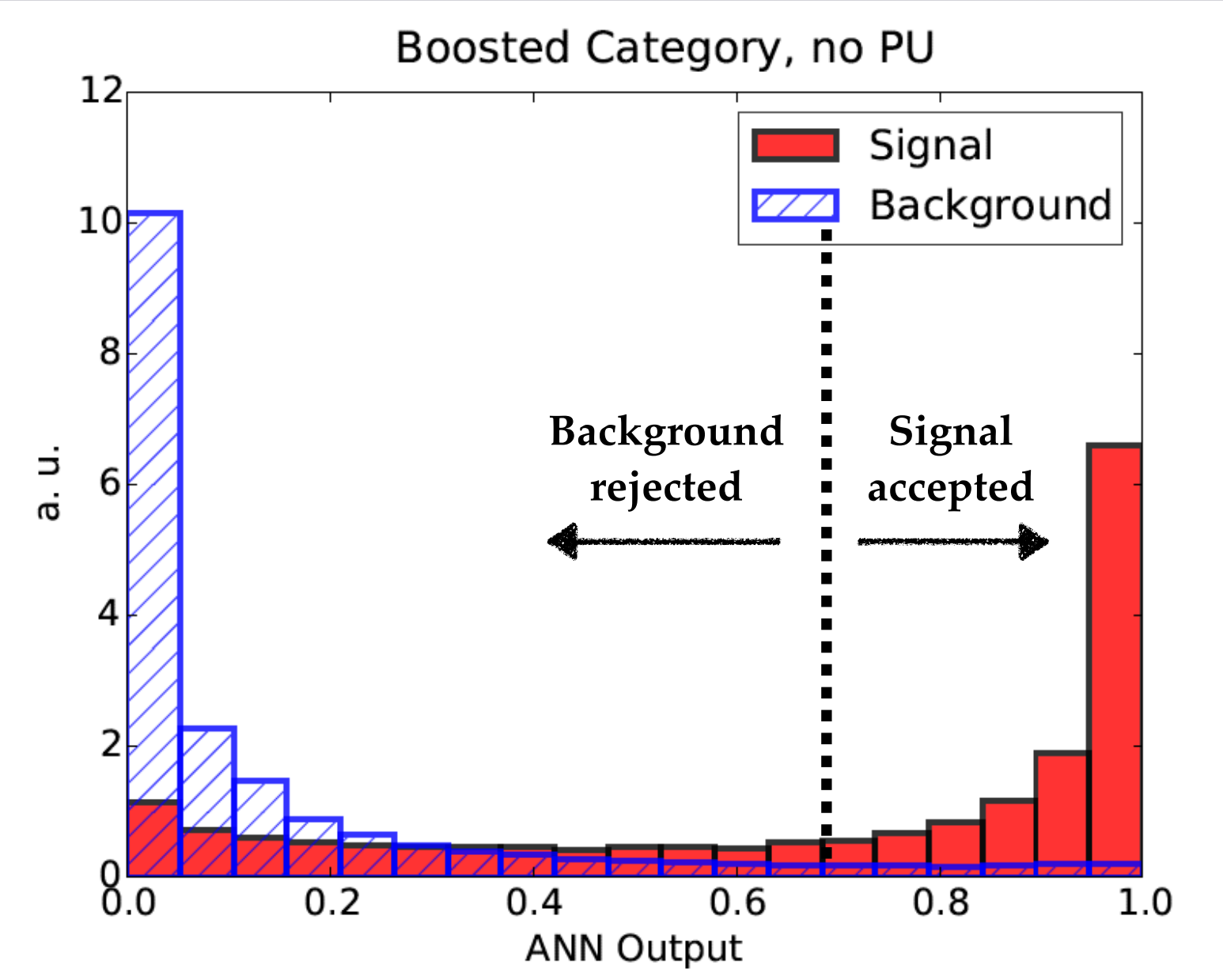
Thus the **error function to be minimised during the training** is the **cross-entropy:**

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

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

Multivariate techniques

Combining information from all kinematic variables in MVA: **excellent signal/background discrimination**

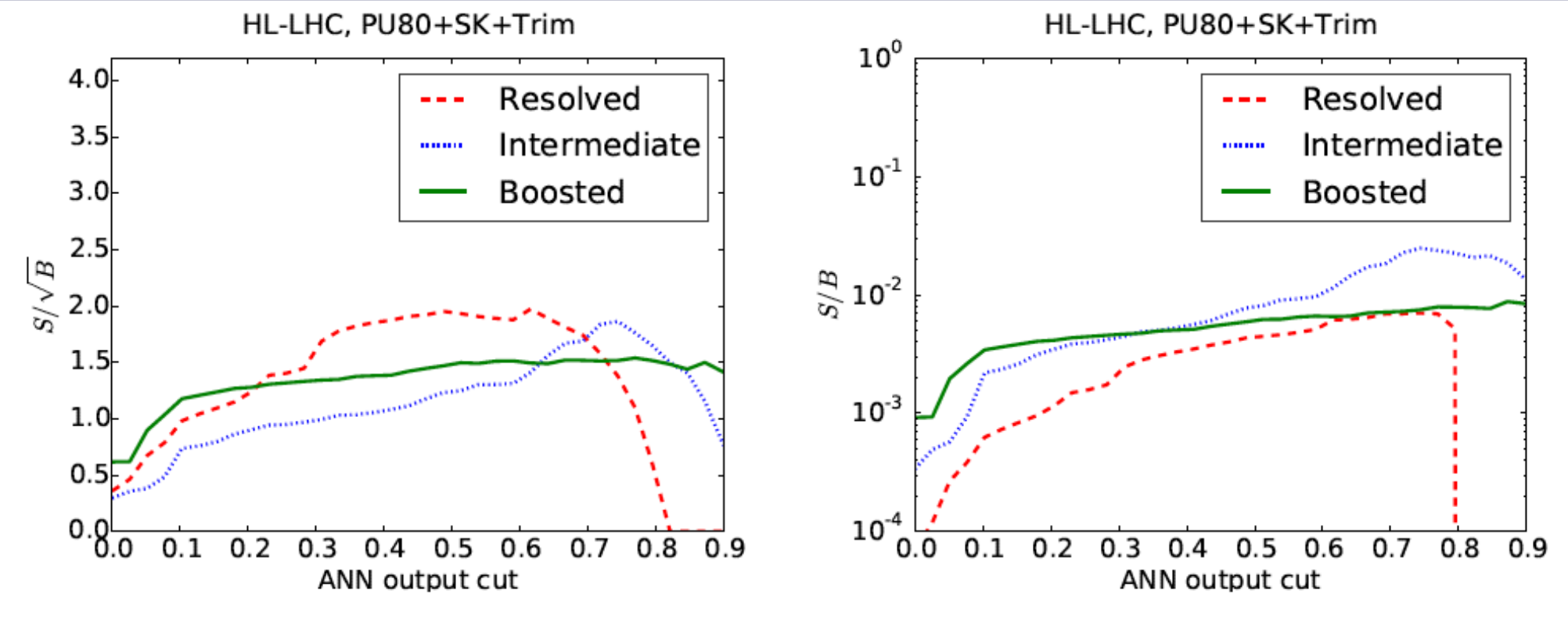


Signal significance

- Use of **multivariate techniques** allows to **substantially improve the signal significance** for this process as compared to a **traditional cut -based analysis**
- The total combined significance is enough to **observe Higgs pair production in the 4b final state** at the HL-LHC. Substantial improvement if reducible backgrounds (fakes) can be eliminated

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

$$\left(\frac{S}{\sqrt{B_{4\text{b}}}}\right)_{\text{tot}} \simeq 4.7\ (1.5)\,, \quad \mathcal{L} = 3000\ (300)\ \text{fb}^{-1}$$



Signal significance

- Use of **multivariate techniques** allows to **substantially improve the signal significance** for this process as compared to a **traditional cut -based analysis**
- The total combined significance is enough to **observe Higgs pair production in the 4b final state** at the HL-LHC.

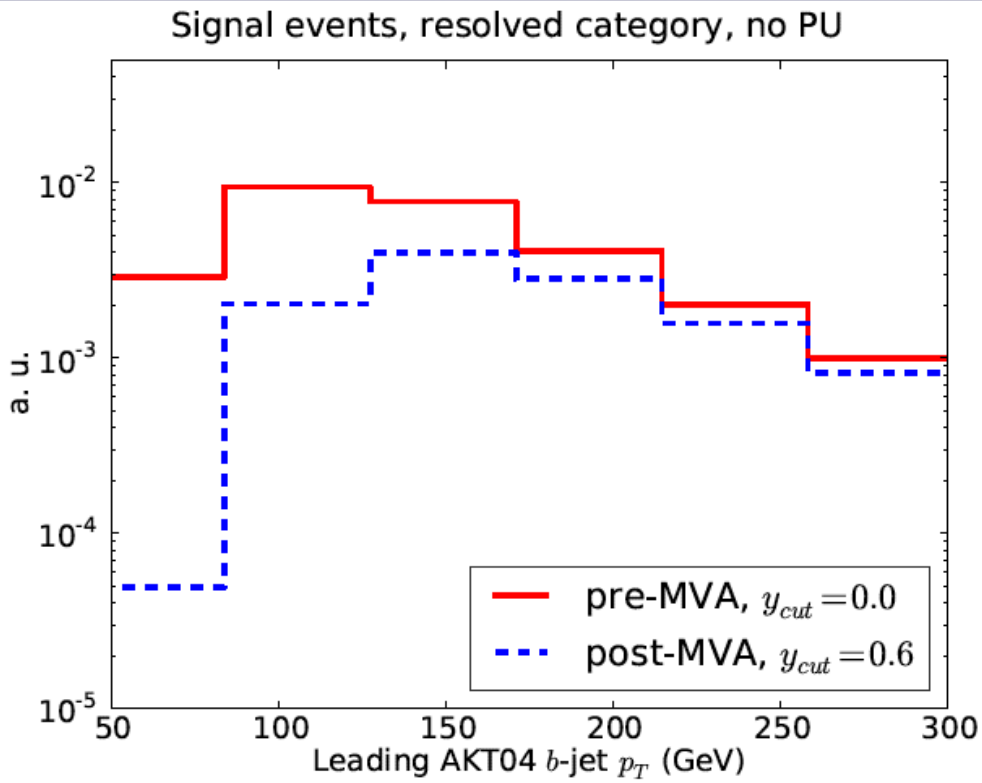
HL-LHC, PU80+SK+Trim					
Category		N_{ev} signal	N_{ev} back	$S/\sqrt{B}$	S/B
Boosted	$y_{\text{cut}} = 0$	410	$4.5 \cdot 10^5$	0.6	10^{-3}
	$y_{\text{cut}} = 0.8$	290	$3.7 \cdot 10^4$	1.5	0.01
Intermediate	$y_{\text{cut}} = 0$	260	$7.7 \cdot 10^5$	0.3	$3 \cdot 10^{-4}$
	$y_{\text{cut}} = 0.75$	140	$5.6 \cdot 10^3$	1.9	0.03
Resolved	$y_{\text{cut}} = 0$	1800	$2.7 \cdot 10^7$	0.4	$7 \cdot 10^{-5}$
	$y_{\text{cut}} = 0.60$	640	$1.0 \cdot 10^5$	2.0	0.01

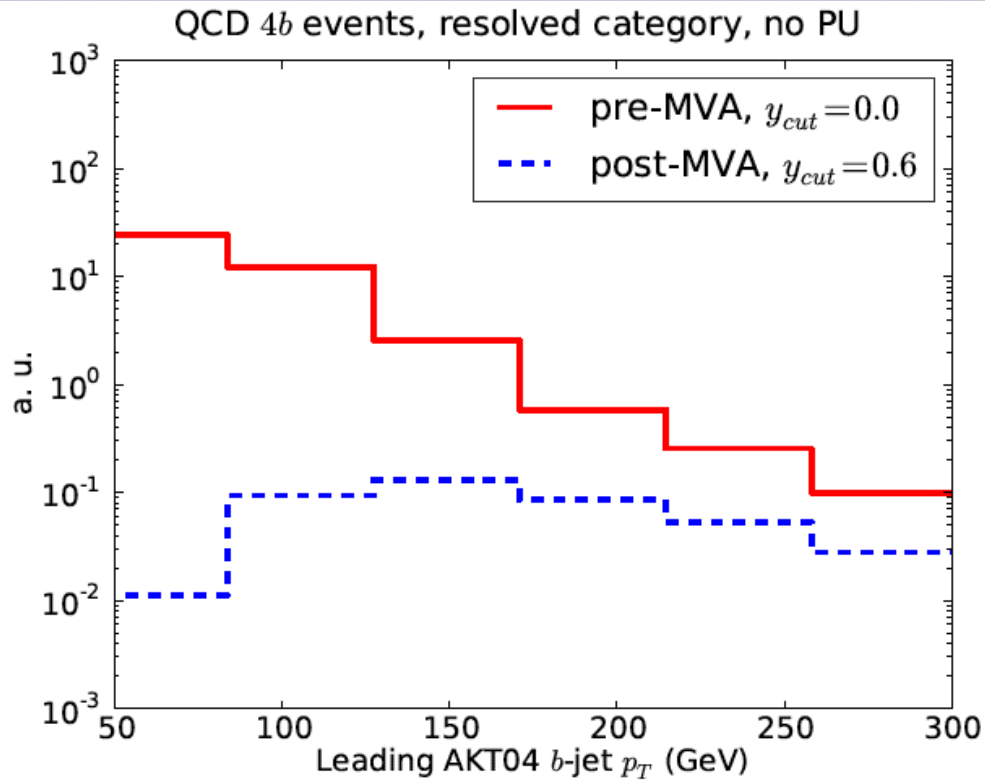
The MVA achieves a **substantial background reduction** with **high signal efficiency**

Caveat: low signal over background ratio, need high control on systematic uncertainties

Opening the Black Box

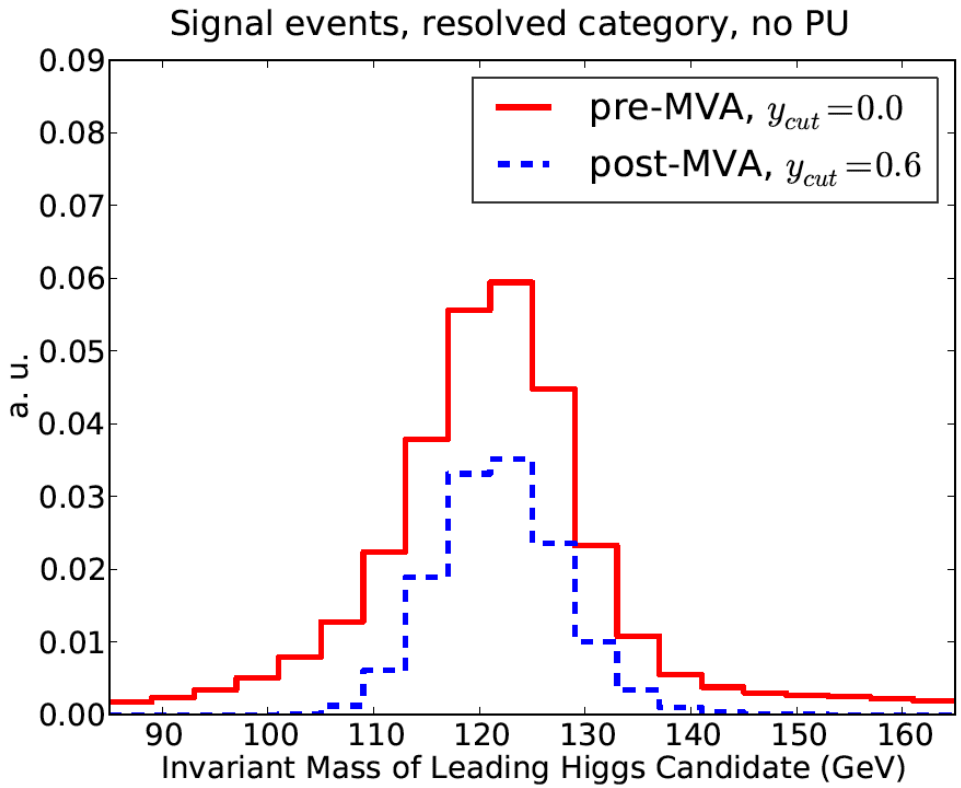
- ANNs are sometimes criticised by acting as **black boxes**, with little control/understanding of what is happening inside them
- But ANNs are simply a **set of combined kinematical cuts**, nothing mysterious in them
- To verify this, plot kin distributions **after and before the ANN cut**: we can then determine the **effective kinematic cuts** are being optimised by the MVA
- This info should be enough to perform a cut-based analysis and achieve similar signal significance

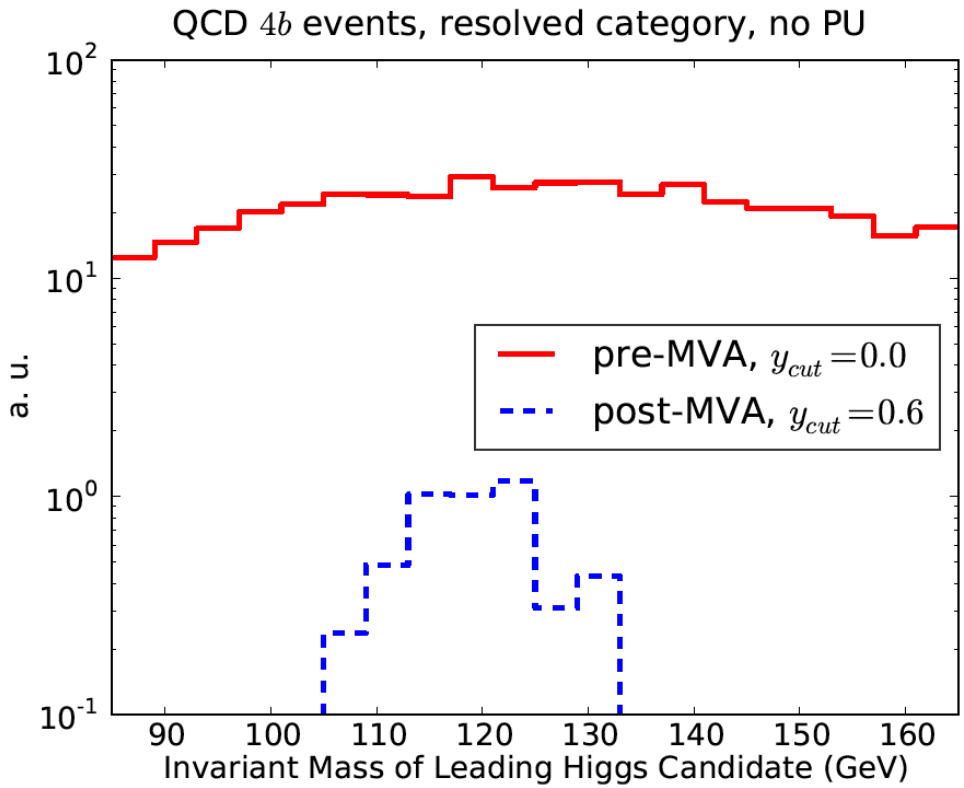




Opening the Black Box

- ANNs are sometimes criticised by acting as **black boxes**, with little control/understanding of what is happening inside them
- But ANNs are simply a **set of combined kinematical cuts**, nothing mysterious in them
- To verify this, plot kin distributions **after and before the ANN cut**: we can then determine the **effective kinematic cuts** are being optimised by the MVA
- This info should be enough to perform a cut-based analysis and achieve similar signal significance

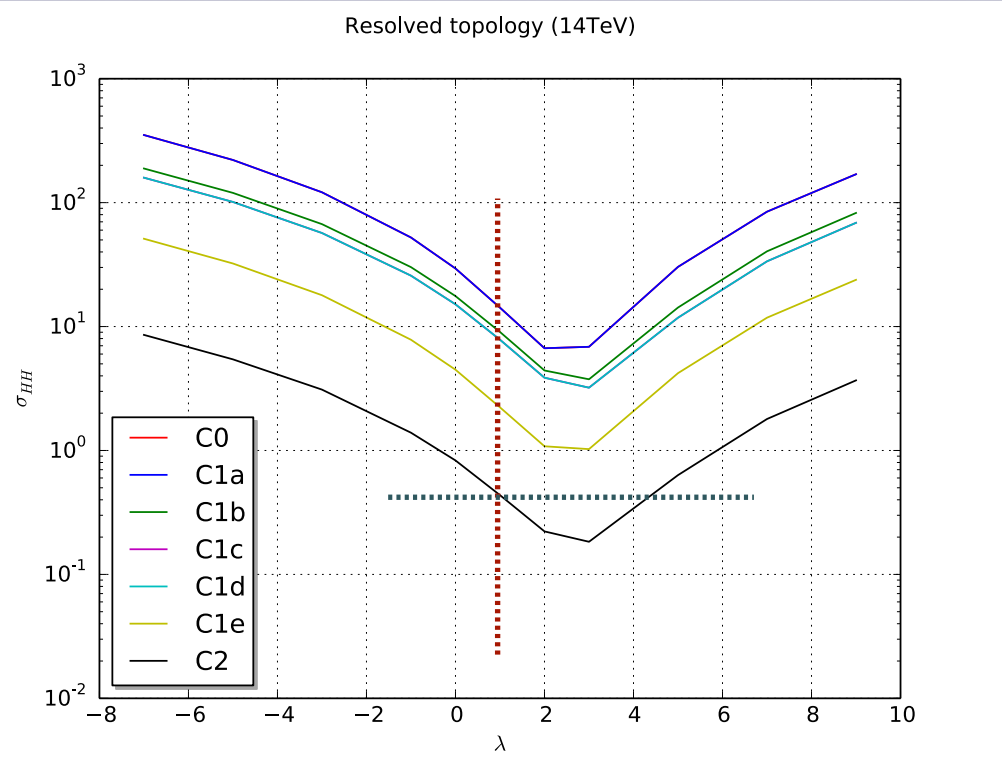


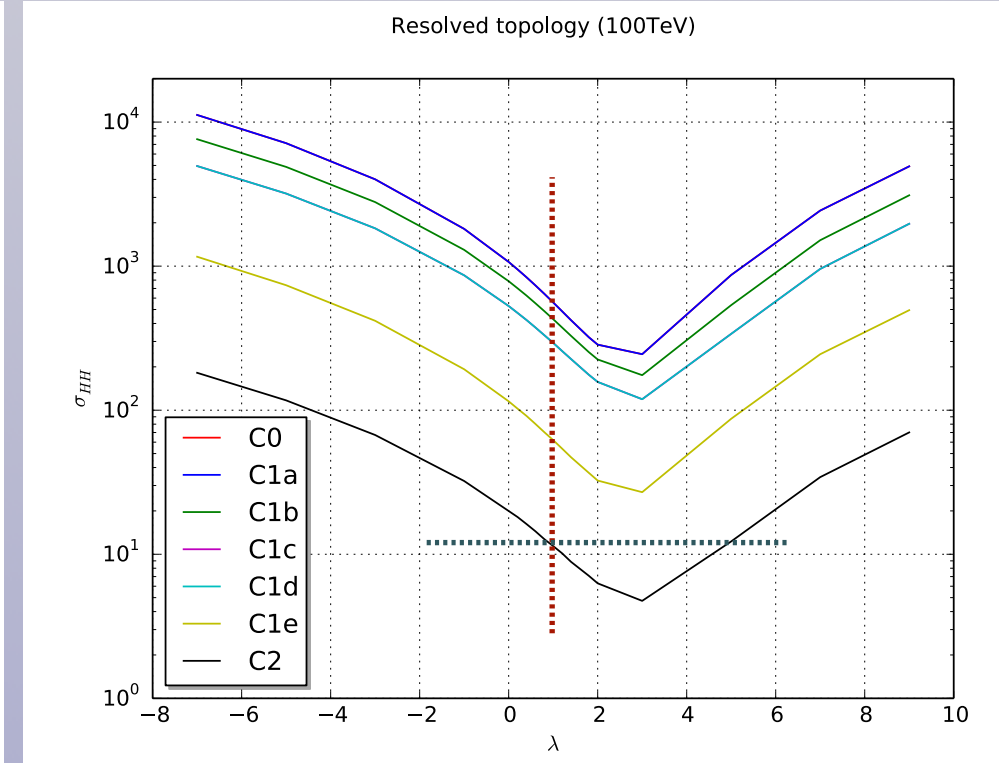


The background "Higgs mass peak" now mimics the signal one

Towards a self-coupling determination

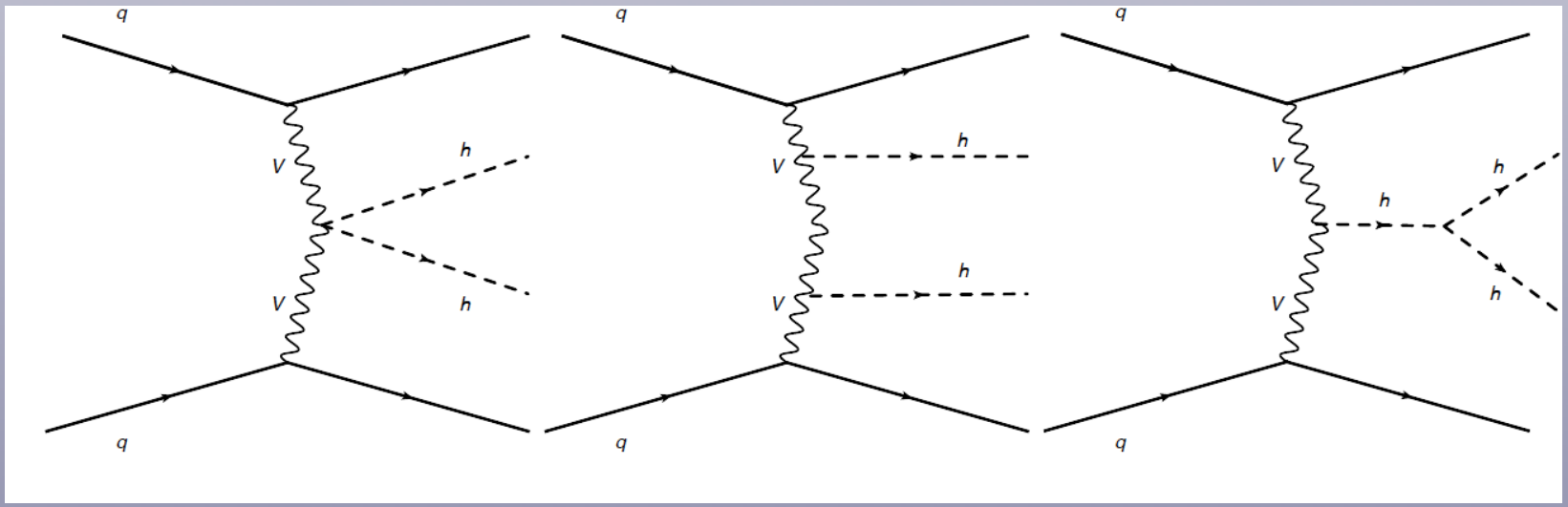
- Now working on estimating the accuracy on the **extraction of the Higgs self-coupling** that can be achieved at the **LHC Run II**, the **HL-LHC** and at a **100 TeV FCC**
- **The optimization of the extraction of the Higgs self-coupling might require the use of an additional MVA** (different learning task than signal/background classification)
- Crucial the estimate of the impact of **experimental systematic uncertainties as** real measurement
- We train with MC samples, a measurement the **MVA should be trained on LHC data** in control regions





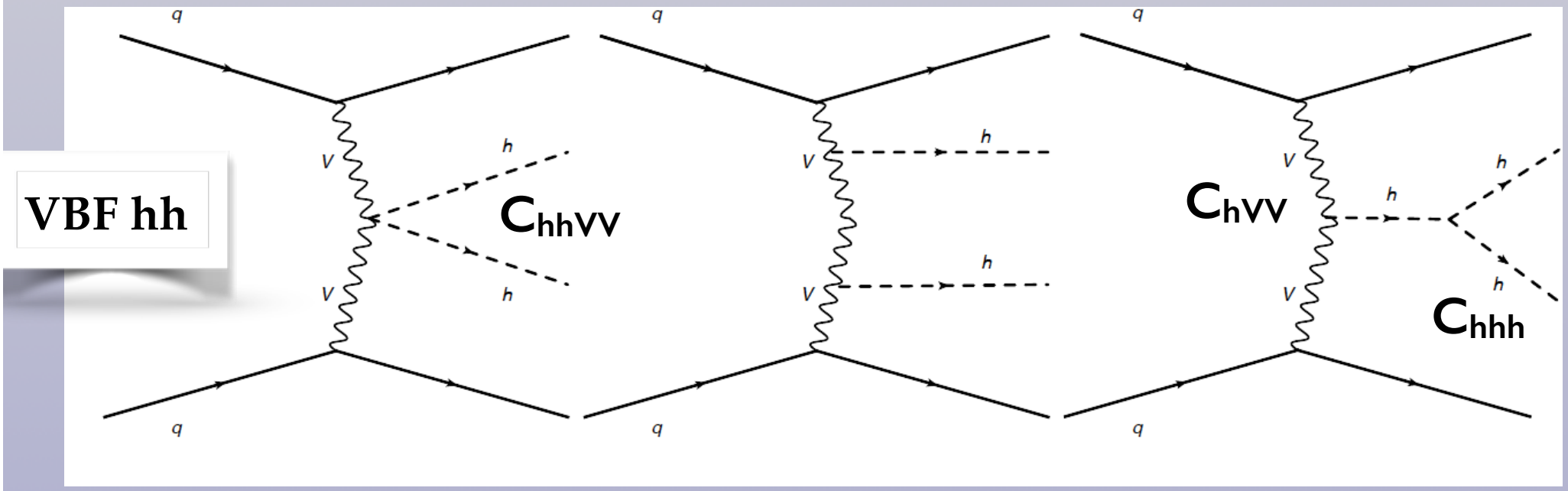
Strong Higgs Pair Production in the Vector-Boson Fusion Channel

Bishara, Contino, JR, in preparation



Higgs Pair Production in Vector Boson Fusion

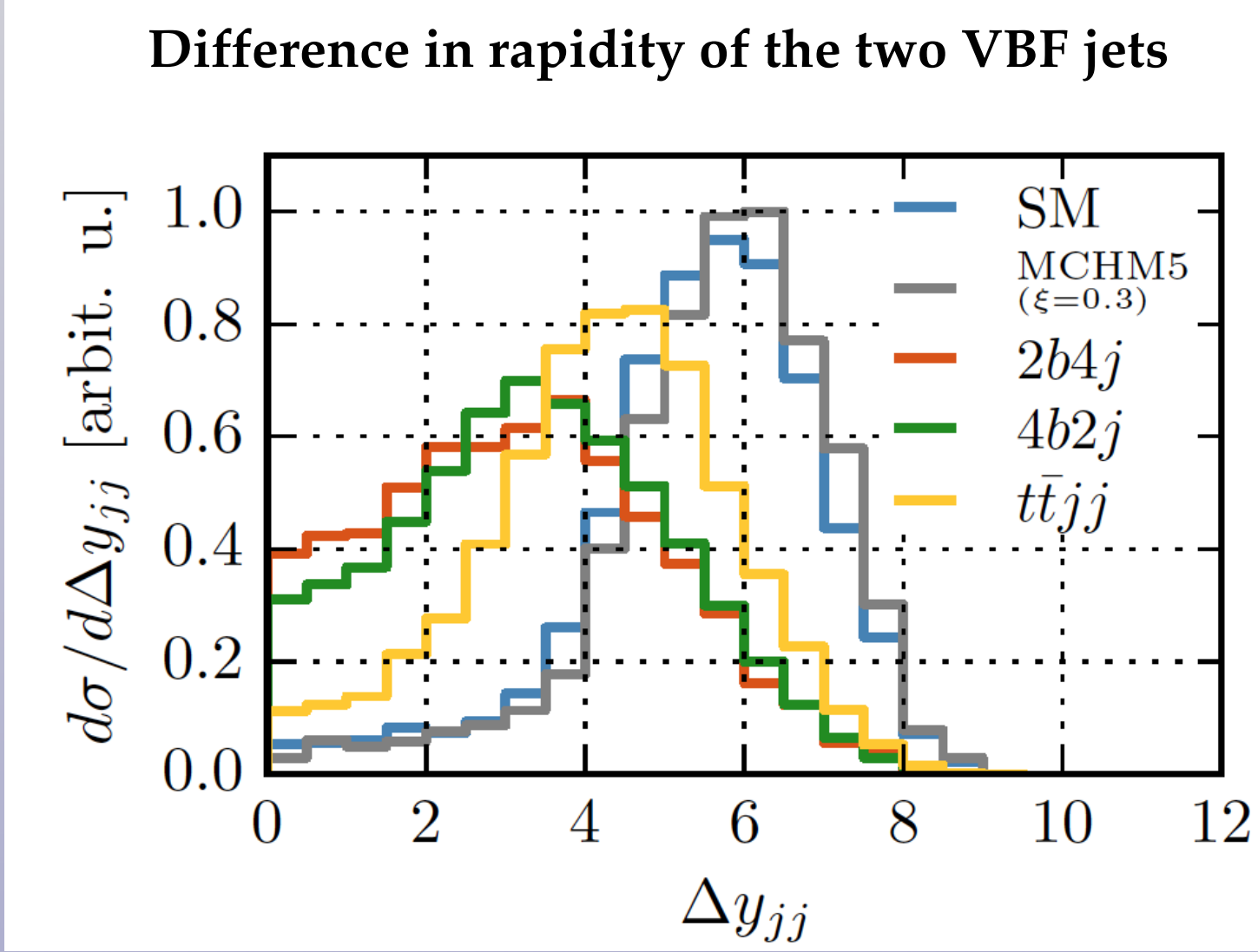
- Higgs pair production in VBF allows direct access of the **hhVV coupling:** direct test of the **doublet nature** of the Higgs boson
- Cannot be accessed in single Higgs production
- Small SM production rates: at 14 TeV cross-section is only **2 fb.** But signal yields can be **enhanced by orders of magnitude** even for **small deviations wrt SM couplings**



Higgs Pair Production in Vector Boson Fusion

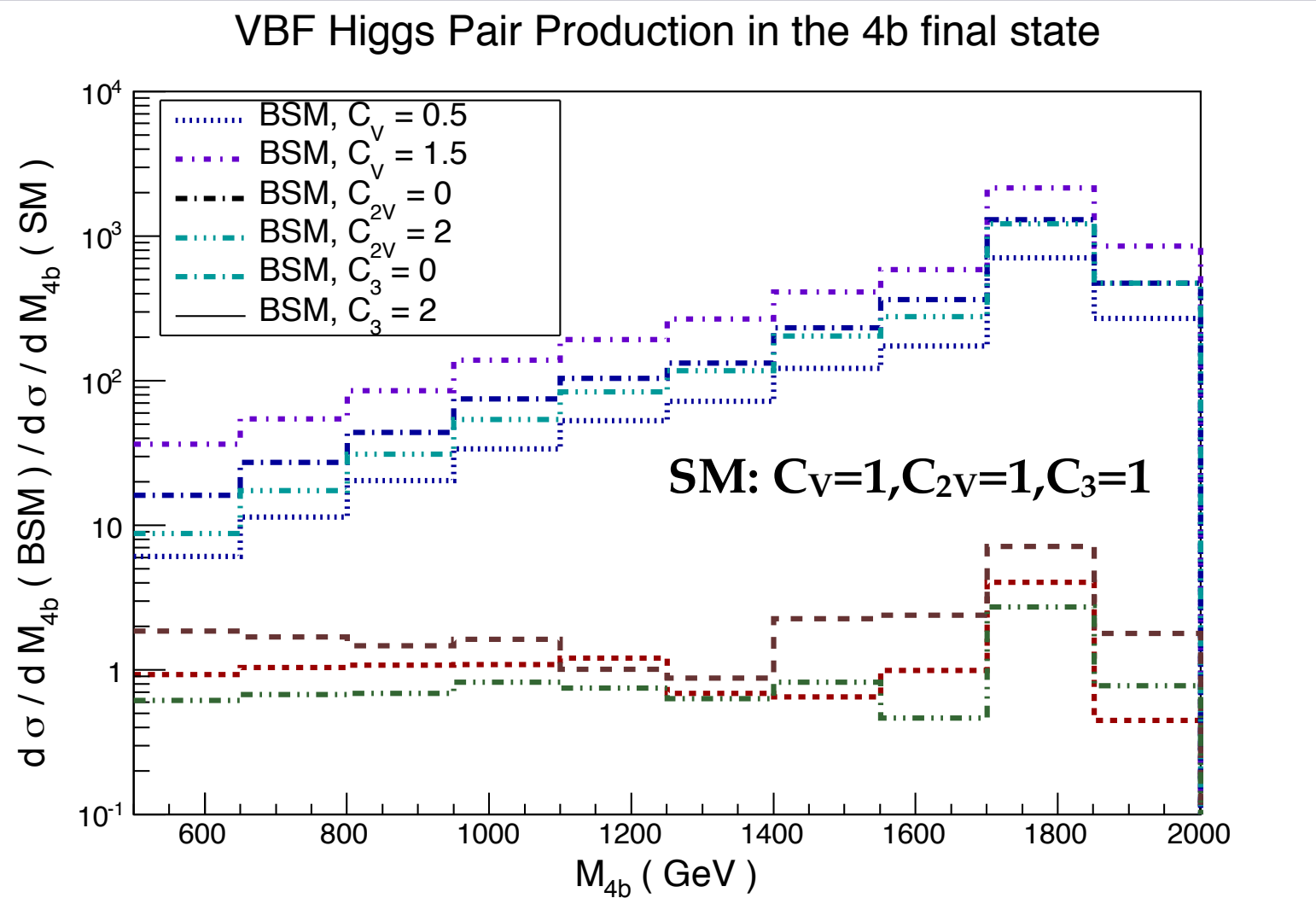
Also in this case overwhelming QCD background, can be tamed with **selection cuts** and **jet substructure**

Exploit the **vector-boson-fusion topology**: two light jets, well separated in rapidity, and a central gap with reduced hadronic activity



Higgs Pair Production in Vector Boson Fusion

- Deviations from the Standard Model in the **hVV and hhVV couplings** lead to **dramatic growth with energy of the di-Higgs invariant mass** distribution
- Smoking gun for New Physics!



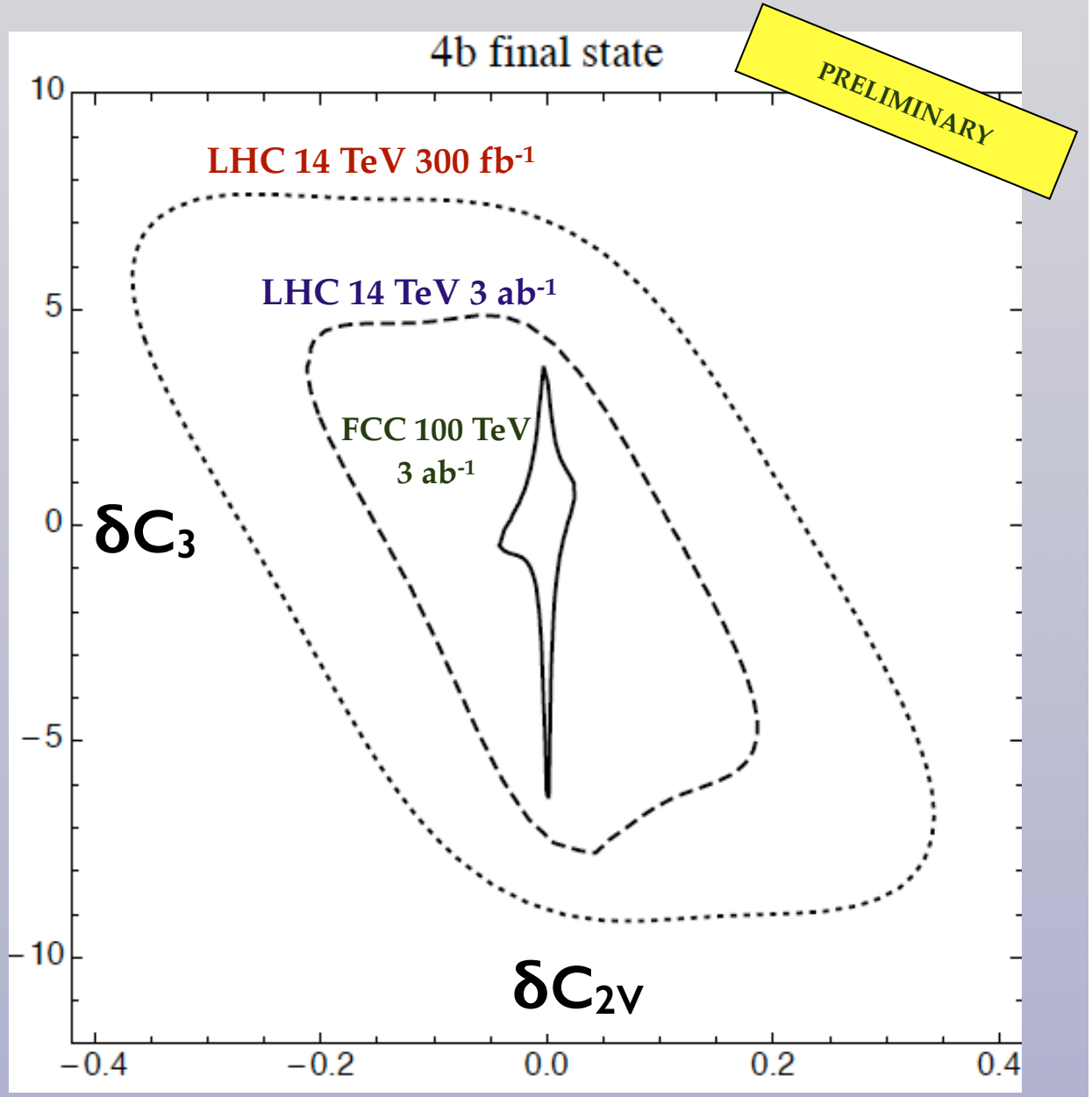
Higgs Pair Production in Vector Boson Fusion

- In the **4b final state,** 14 TeV with 300 fb^{-1} (3000 fb^{-1}) the **hhVV** coupling can be measured with **good precision: ~25-30% (10-15%)**

- At a 100 TeV **FCC,** the **hhVV** coupling can be pinned down with **very high, few percent precision**

- Encouraging to begin to explore **Higgs pair production in VBF** already at the **LHC Run II**

- Unique probe of the Higgs mechanism to **unitarize electroweak symmetry breaking**



Summary and outlook

- Higgs pair production is a **cornerstone of the LHC program** for the next years
- Double Higgs production would provide evidence of **Higgs self-interactions**, allow to reconstruct the **EWSB potential** and test the **doublet nature** of the Higgs boson
- The **4b final state** offers the highest yields, but requires taming an overwhelming QCD background
- At the **HL-LHC**, **"observation" signal significances** will be achieved in this final state. A first measurement might even be possible **by the end of Run II**, but extremely challenging.