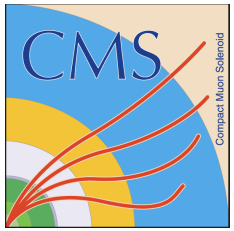
Jets in Higgs searches with the CMS Experiment

Arnaud Pin[1]
On behalf of the CMS Collaboration

1 : Manchester University.
CMS membership with the center for Cosmology, Particle Physics and Phenomenology,
Université Catholique de Louvain, Belgium

arnaud.pin@cern.ch



17th July– Jet vetoes and multiplicity observables 2014 – IPPP Durham, UK

Jet in Higgs searches at CMS

Example of Higgs searches using jets at CMS
Not a complete review of Higgs analysis.

Higgs decays to pair of b-quarks

- Vector boson: **VH(bb)** [PhysRevD.89.012003]
  - Sensitivity to jet energy resolution: Mbb
  - → How to improve it ?
- Quarks: **VBF H to bb** [CMS-PAS-HIG-13-011]
  - Need to discriminate between quark and gluon jets
  - → Quark / Gluon jet tagger

- H→WW: [CMS-PAS-HIG-13-023]
  - Categorization in jet multiplicity

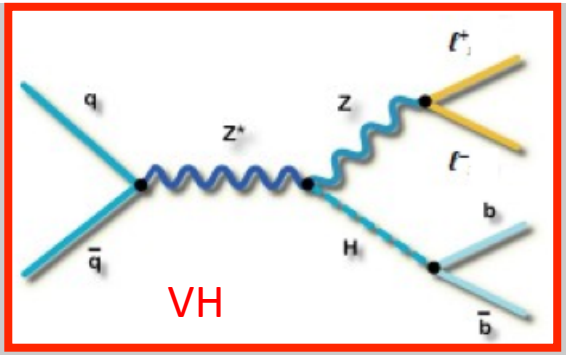
Searches and Measurement of the SM Higgs boson:

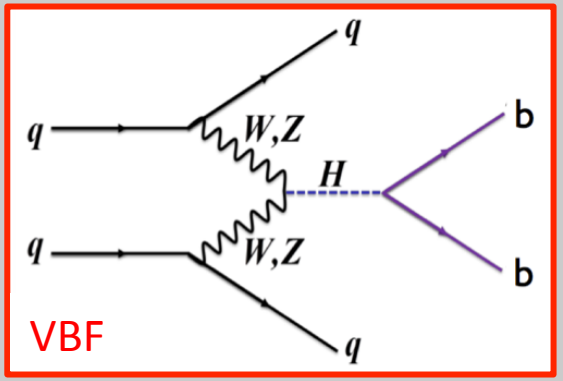
- Search for H→bb: test for fermionic final state.
- Large QCD background → Inclusive H(bb) production too challenging

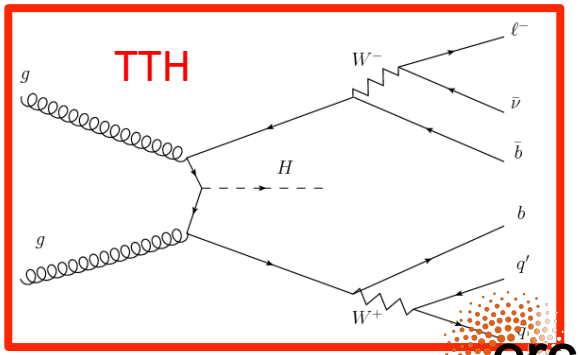
3 Search Analysis:

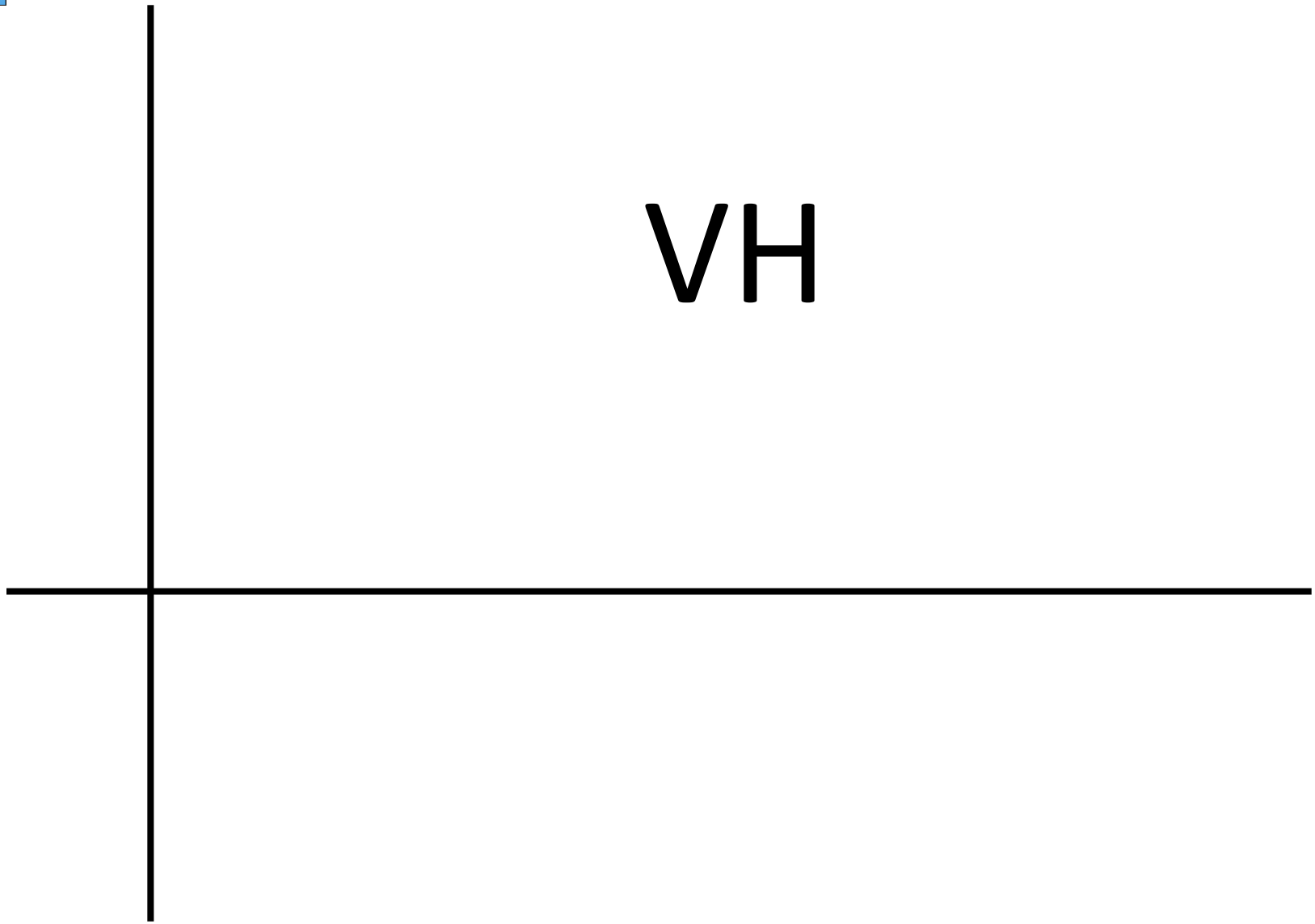
- Need to search for H→bb in associated production:
  - Less background.
  - Lower Cross-section.
  - Associated production with:
    - Vector boson: **VH(bb)** [PhysRevD.89.012003]
    - Quarks: **VBF H to bb** [CMS-PAS-HIG-13-011]
    - Top quarks: **ttH H to bb** [CMS-PAS-HIG-13-019]

[Not discussed in this talk]








VH

Analysis strategy:

1) Event Selection and Categorization

- Select boosted events on P_t V.
- Vector boson decaying leptonically:
  - **6 channels**: 2x **W(lν)**, 2x **Z(ll)**, **Z(νν)**, **W(τν)**

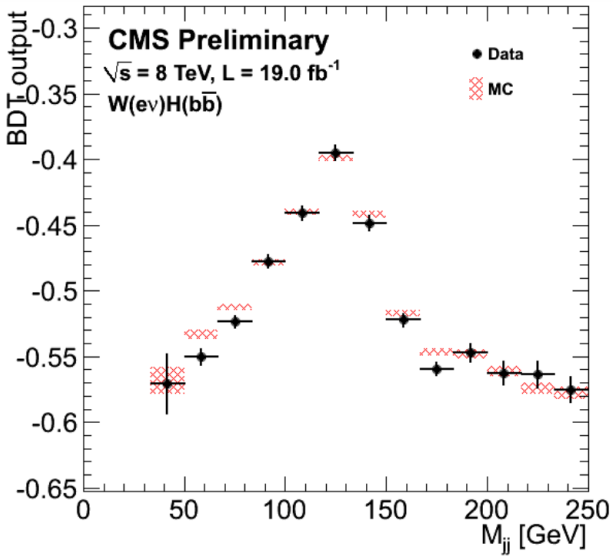
- Categorization depends on **P_t V**:

Pt V [GeV]	W(lν)	W(τν)	Z(ll)	Z(νν)
low	100 – 130	> 120	50 – 100	100 – 130
intermediate	130 – 180			130 – 170
high	> 180		> 100	> 170

- Large azimuthal opening angle, $\Delta\varphi$(V,H)
- Total of 14 Categories (e/μ).

2) Multivariate analysis: BDT

Improved bb Mass resolution



Why improving Mbb:

Highly correlated with the BDT discriminant output

Mass resolution of the two b-jets from Higgs decay: ~10% (depending on Pt) +Shift value

Principle:

Specialized BDT regression to provide corrected jet energy.

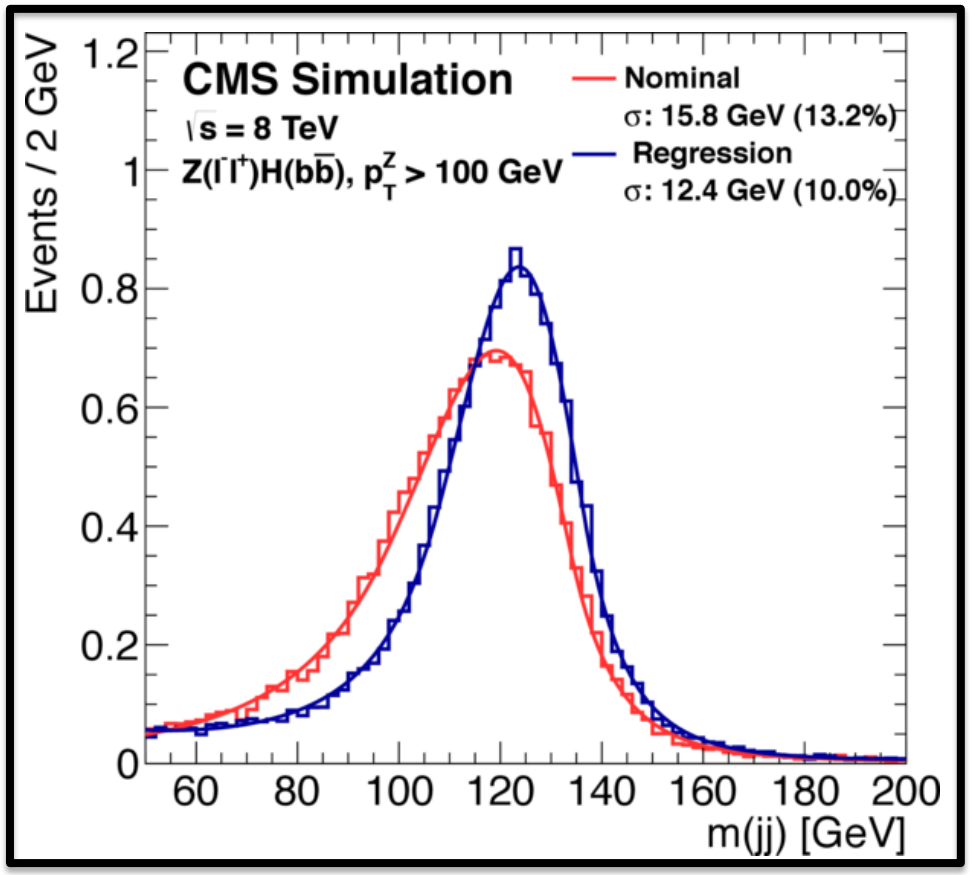
Based on detailed jet structure informations including:

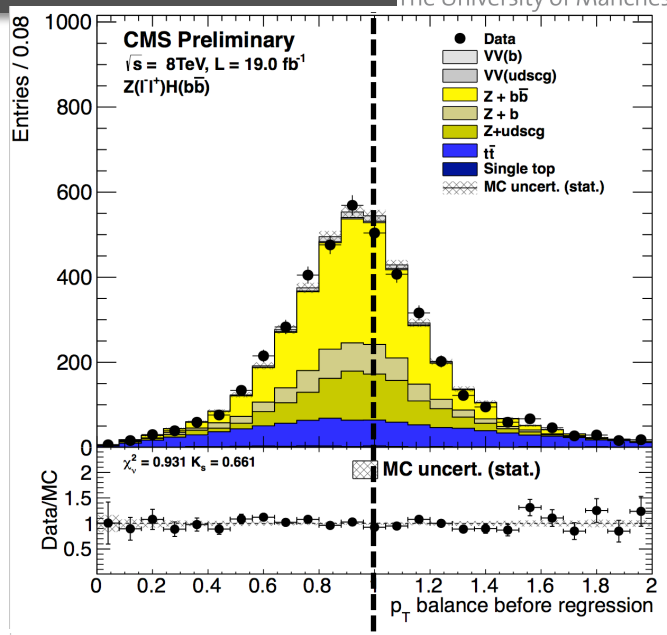
- secondary vertex properties (when reconstructed)
- track and jet constituent informations (PF candidate), transverse mass
- energy of charge particle in the jet
- b-jet contains in average more lepton and larger fraction of Etmiss
  - use of the soft lepton in jet properties.
- For WH: $\Delta R(l,j)$, P_T of lepton
- For ZH: E_{Tmiss} and its azimuthal angle with each jet.

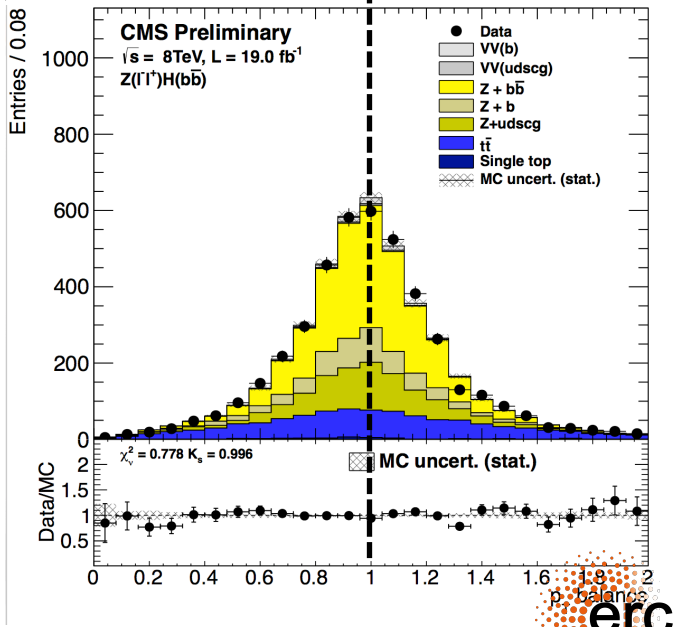
arXiv: 1107.3026v1

Improved bb Mass resolution

M_{bb} Improvement of ~15%
→ Analysis sentitivity increase between 10-20%







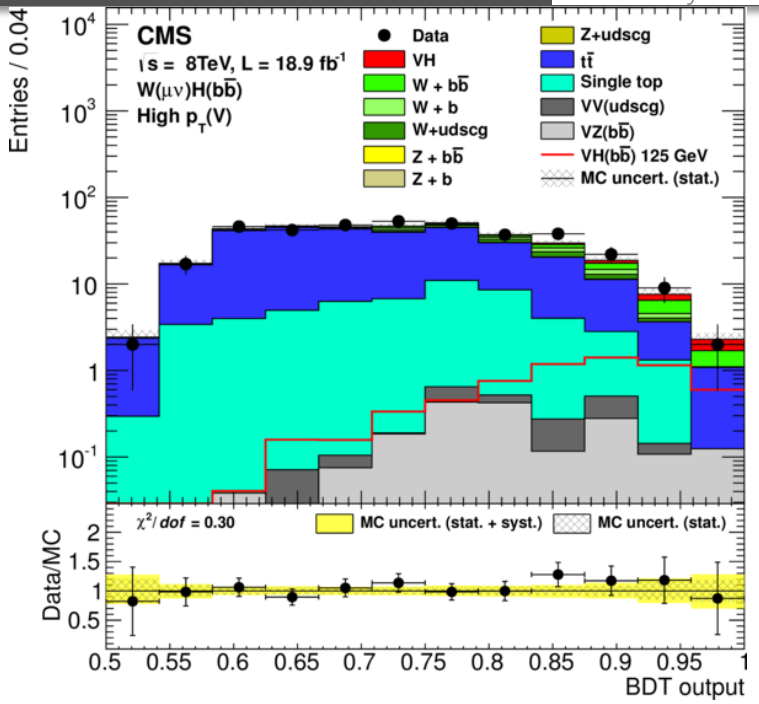
BDT approach

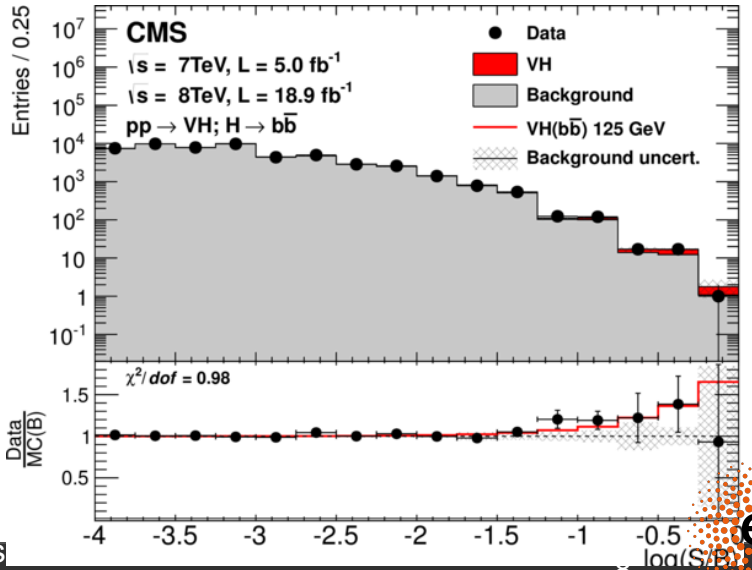
- Cascade BDT.
- Samples divided in **four** subsets.
- Each enriched in: tt, V+jets, dibosons, and VH
- **14 BDT distributions are considered**
  - each P_T V categories
  - Electron/muon for Z(ll) & W(lv)

Background:

- Control regions identified in data.
- Used to validate the simulation modeling of the distributions used as input to the BDT discriminants

- To obtain scale factors used to adjust the simulation event yield estimates for the most important background processes





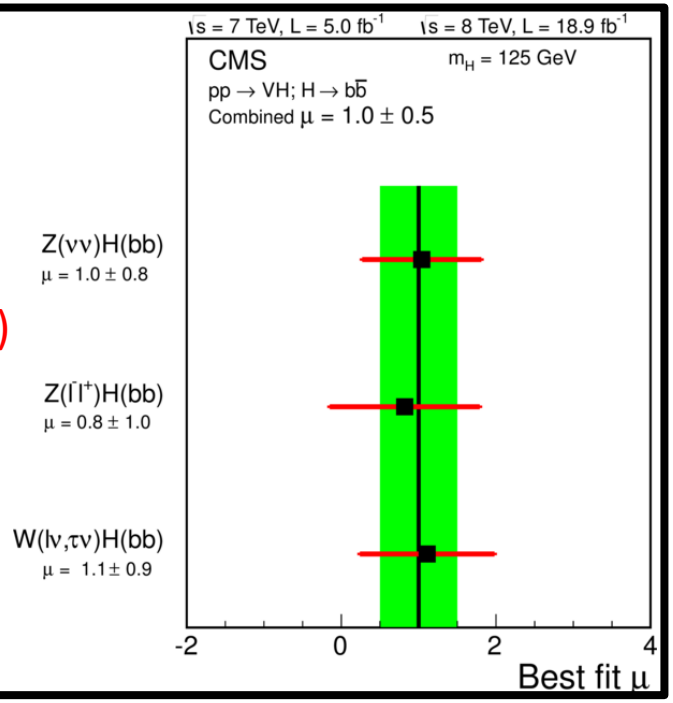
Z/W + H(bb): **Results**

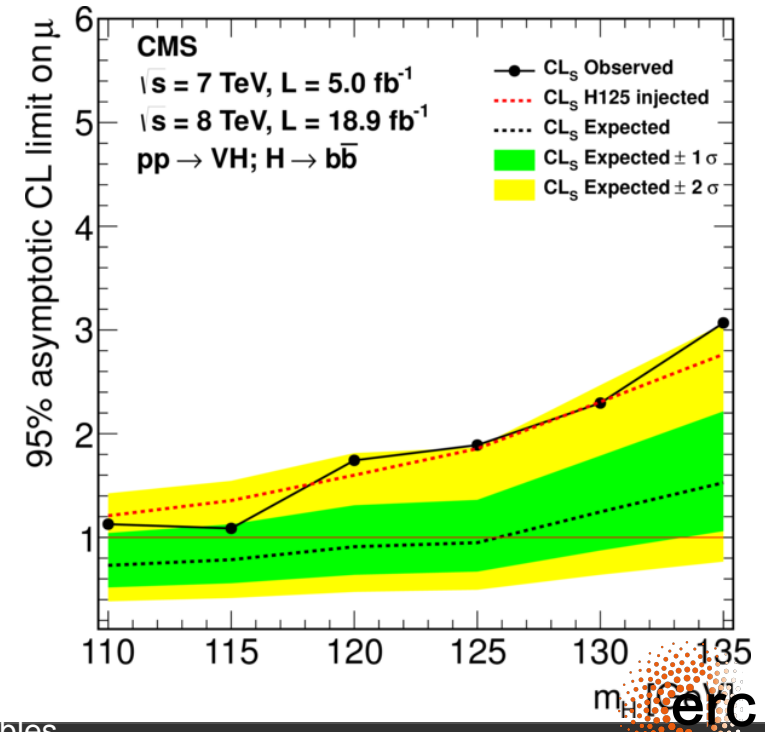
$m_H = 125\,GeV$	σ/σ_{SM} (95% CL) median expected	σ/σ_{SM} (95% CL) observed	Significance expected	Significance observed
$W(\ell\nu,\tau\nu)H$	1.6	2.3	1.3	1.4
$Z(\ell\ell)H$	1.9	2.8	1.1	0.8
$Z(\nu\nu)H$	1.6	2.6	1.3	1.3
All channels	0.95	1.89	2.1	2.1

- **Excess** of events is **observed** above the expected background.
- Observation **agrees** with Standard Model expectation for Higgs boson @ 125 GeV.

Local Significance:

- **Significance:**
  2.1σ obs. (2.1σ exp.)
- **Best fit:**
  μ=1.0±0.5

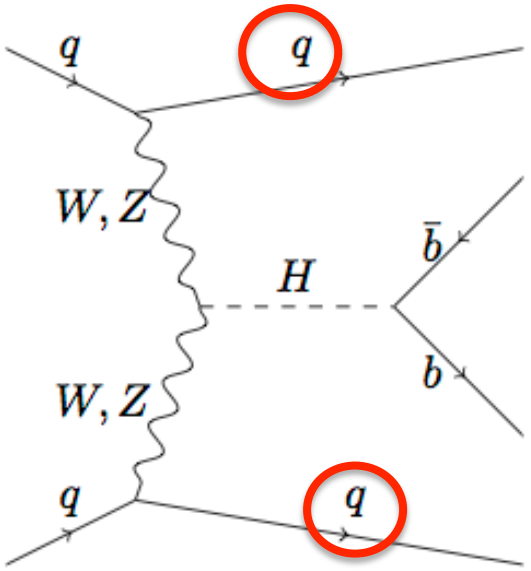




Vector boson Fusion

Fully hadronic final state:

- **VBF selection**
  - 4 jets
  - Dedicated trigger:
    - CMS L1: 3 jets.
    - Only one of the two leading jets allowed to be forward
  - HLT: four jets (one b-tagged), one pair with large m_{jj} $\Delta\eta_{jj}$
- Use of Quark/gluon-jet tagger
  - Distinguish quarks/gluon jet up to $|\eta|<4.7$
  - Reconstruct VBF jet pair "qq"
    1. Pair of least b-tagged jets
    2. Largest pseudorapidity separation.
- **Discriminate** VBF against QCD
  - ANN: MVA considering distributions and correlation of various discriminating variables.



Quark /Gluon Jet tagger

Measurement have shown that jet initiated by gluons exhibit differences with respect to jet from light flavor quark. (OPAL & ALEPH)*

- Charge multiplicity higher in gluon jet
- difference of fragmentation function
- gluon jet are less colimated.

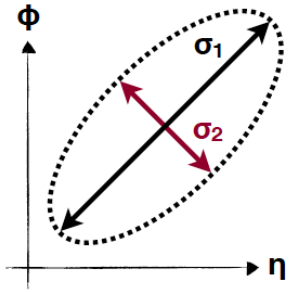
→ **Can be distinguish.**

From particle flow:

Charged multiplicity
Neutral multiplicity
Total multiplicity

Jet shape information:

RMS of PF candidate η-φ spread
Major axis in η-φ: σ_2
Minor axis in η-φ: σ_1



and $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$

Energy variables:

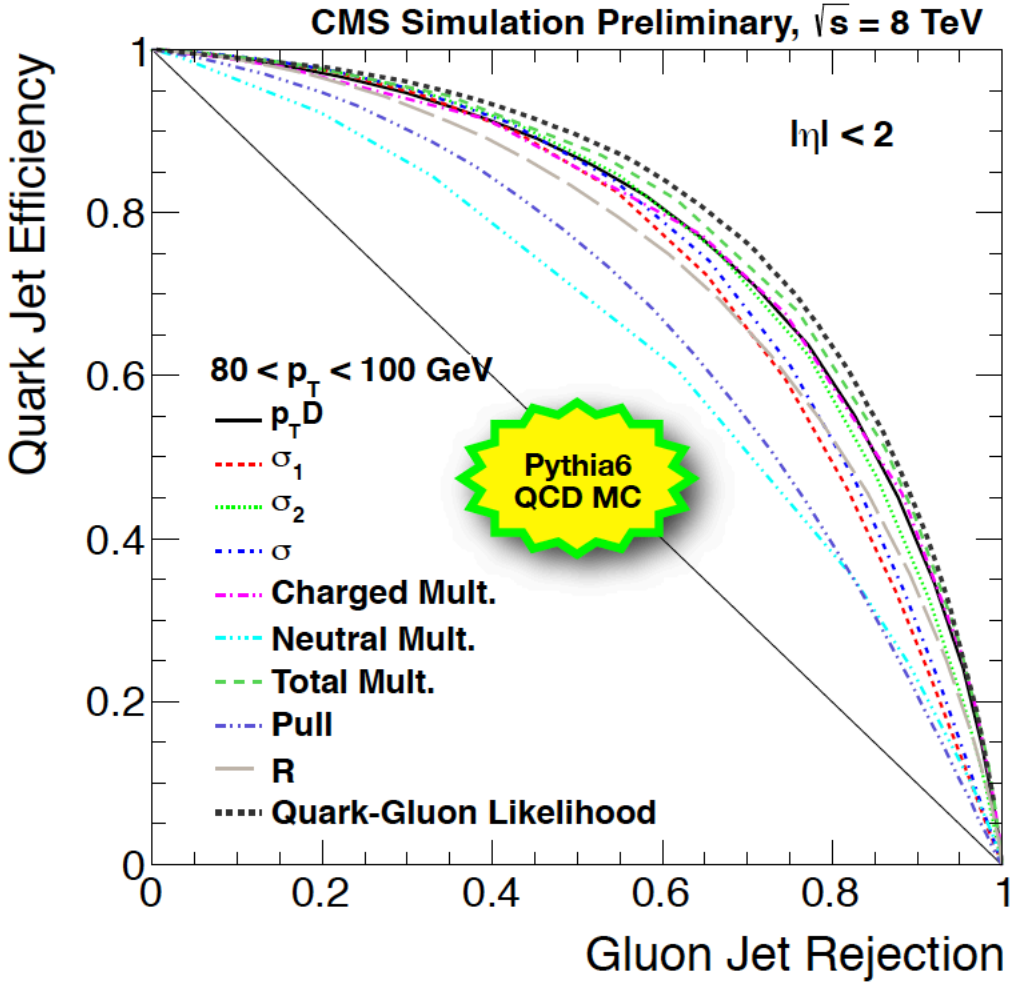
Pull:
R: Energy fraction carried by the leading constituent.
Fragmentation function related variable, $P_T D$:

$$p_T D = \frac{\sqrt{\sum_i p_{T,i}^2}}{\sum_i p_{T,i}}$$

Quark /Gluon Jet tagger

Discrimination power

CMS PAS JME-13-002

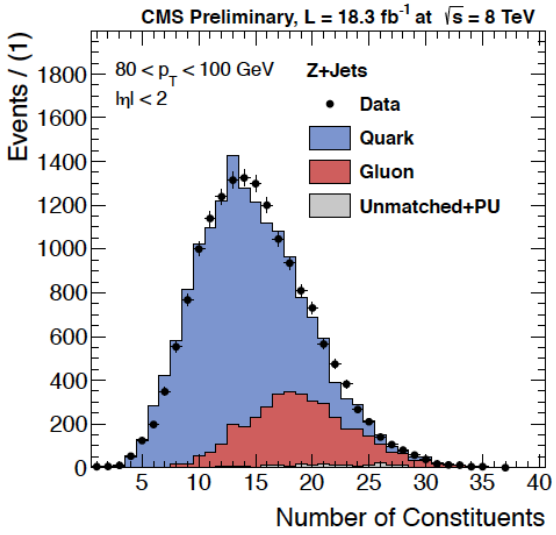


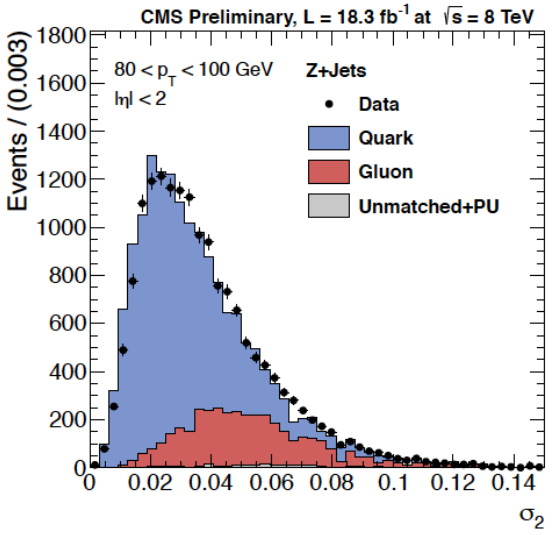
Choice of three variables

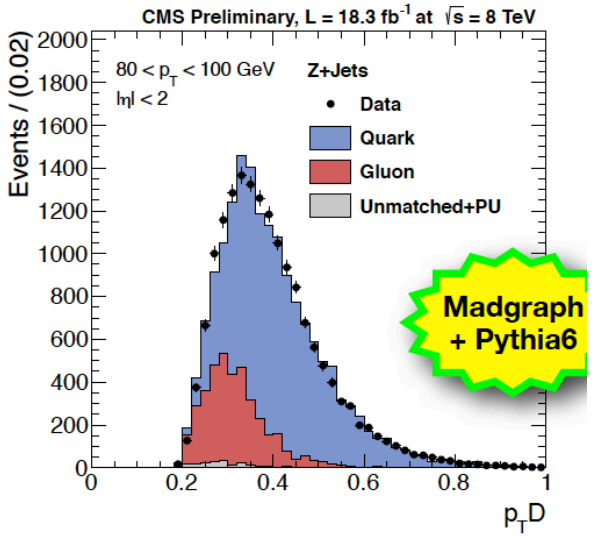
- Total Jet constituent (PF) multiplicity.
- The Minor axis σ2.
- Fragmentation function $P_T D$

Quark /Gluon Jet tagger

3 inputs

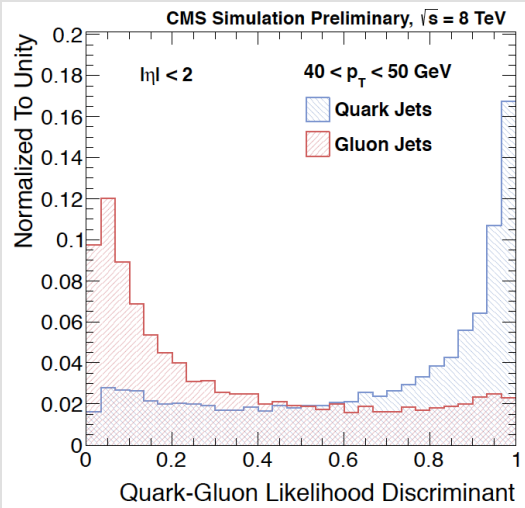


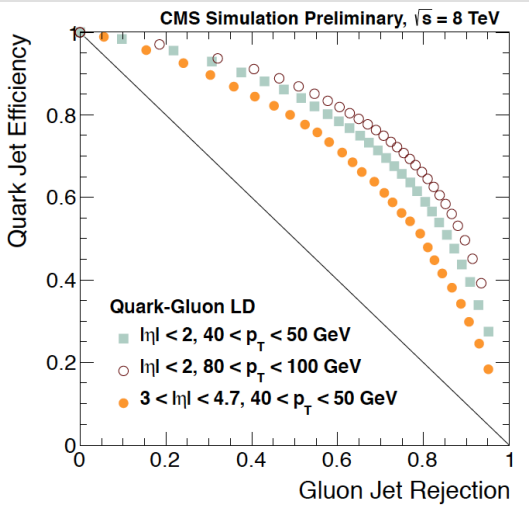




Likelihood discriminant

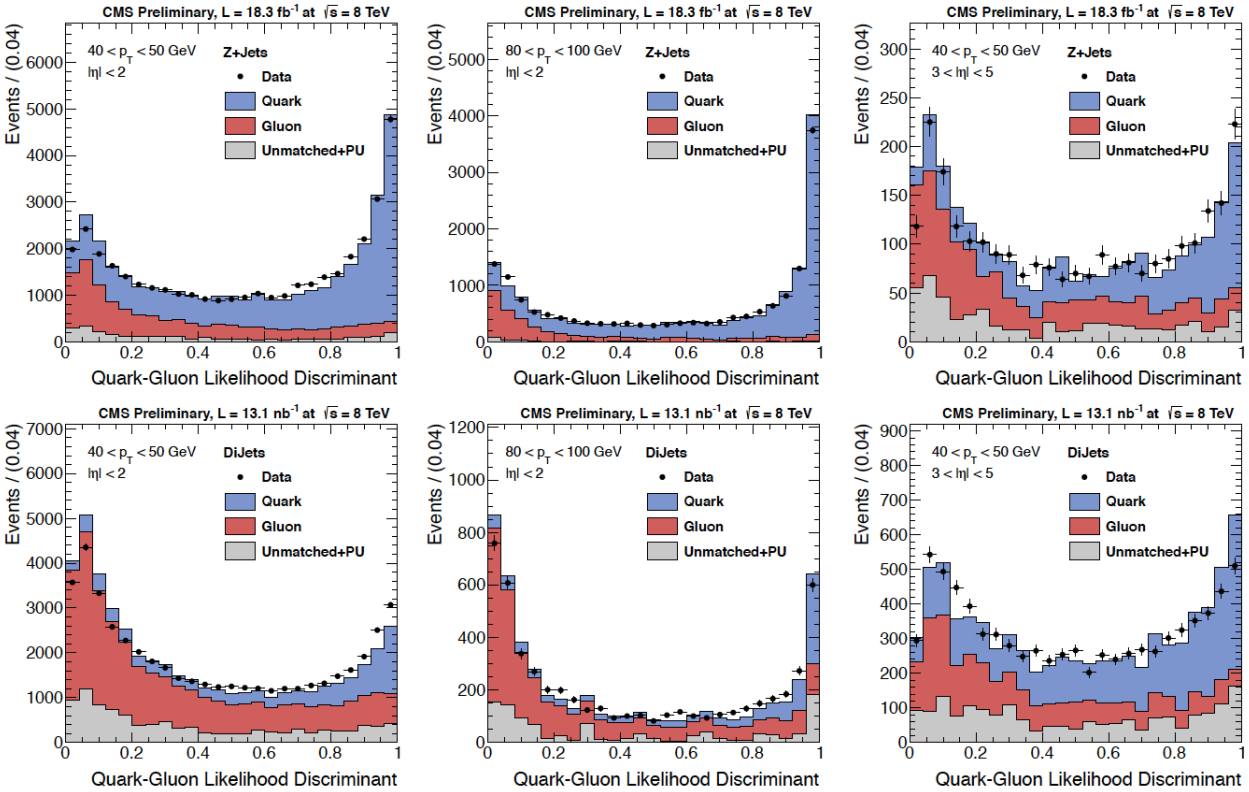
Discrimination up to $|\eta| = 5$





Quark /Gluon Jet tagger

Validation



Z+jets sample

- Mainly quark jets
- ~70% hard and central.

Di-jets sample

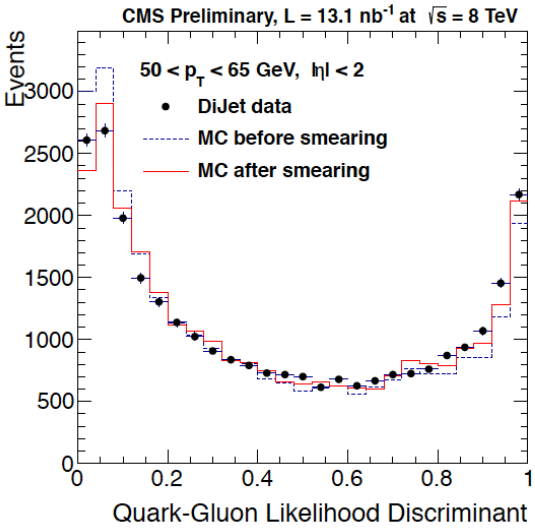
→ Discrepancy in gluon fraction observed in the di-jets sample:

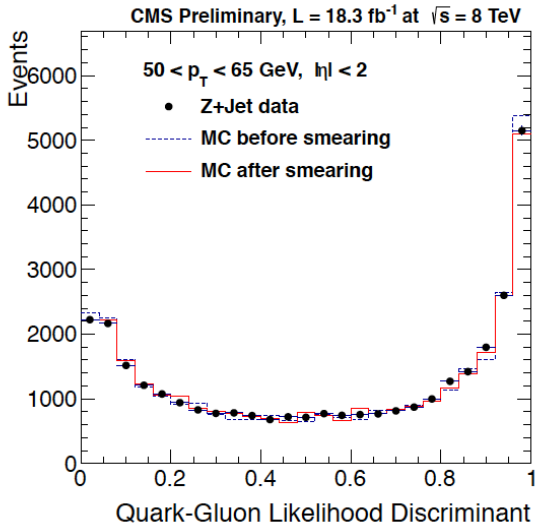
- Smear MC to take into account the discrepancy.
- NOT a reweighting but a jet per jet correction

Quark /Gluon Jet tagger

Validation

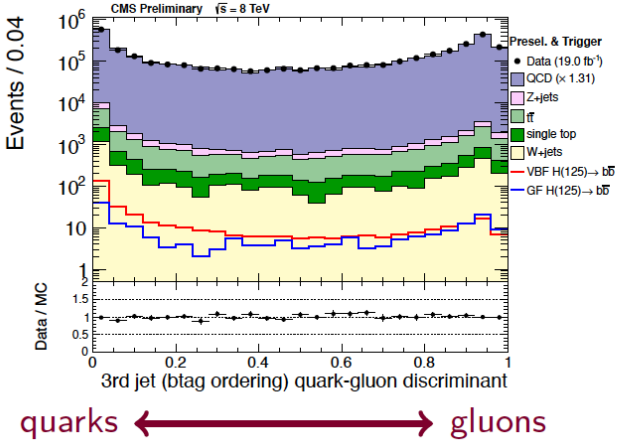
- Extracted on di-jet sample.
- Applied also on the Z+jet sample for validation

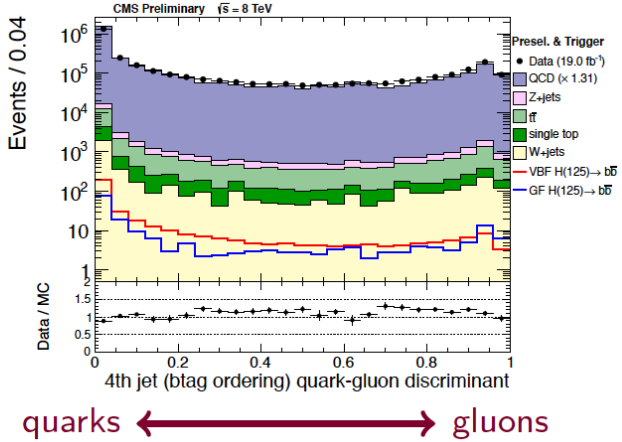




In the VBF analysis:

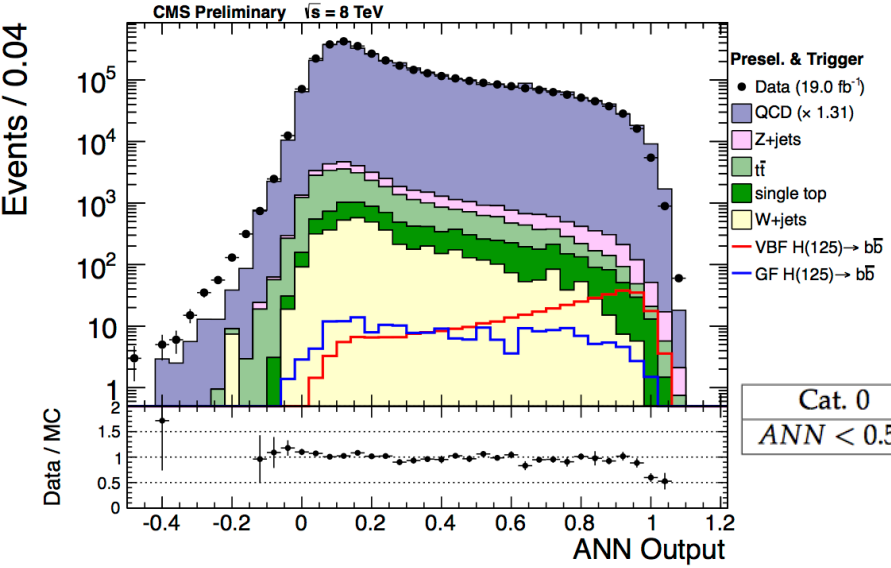
Good agreement between data and MC.





VBF H(bb)

Measurement Strategy:



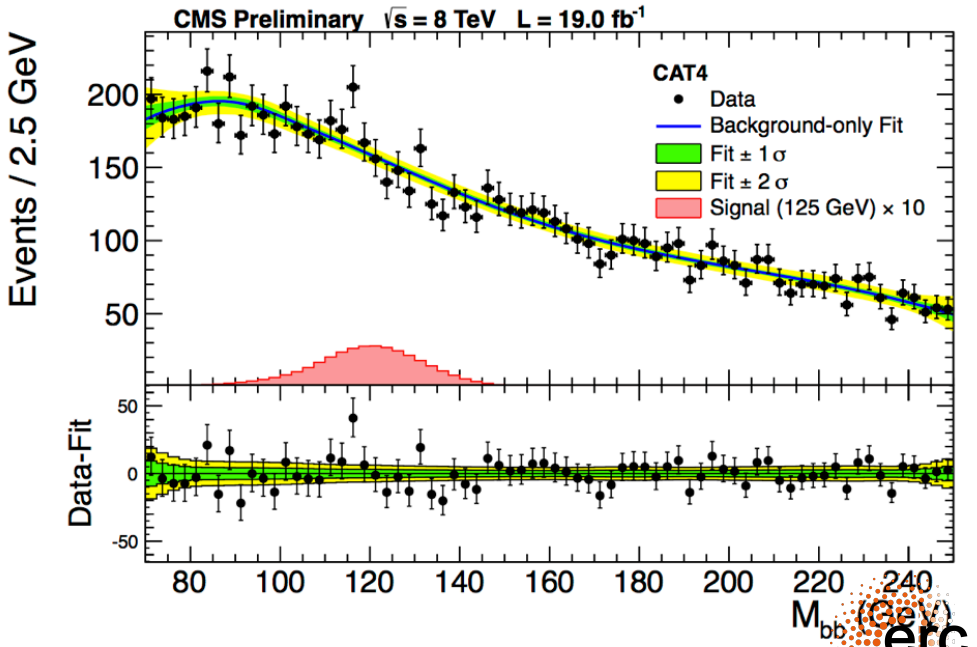
Categorization according to neural net output:

- No kinematic information of b-tagged jets
- **Minimal correlation** with M(bb)

Cat. 0	Cat. 1	Cat. 2	Cat. 3	Cat. 4
$ANN < 0.52$	$0.52 \leq ANN < 0.76$	$0.76 \leq ANN < 0.90$	$0.90 \leq ANN < 0.96$	$ANN \geq 0.96$

→ **Use M(bb) to extract signal:**

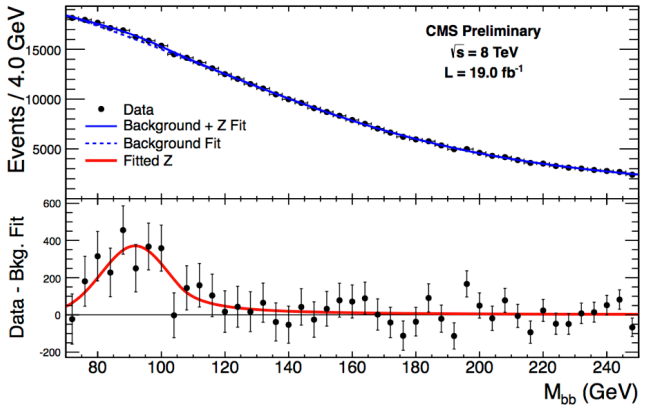
- In the 4 categories
- B-jet energy regression to improve resolution. (same as VH analysis)



VBF H(bb): Cross-check & **Results**

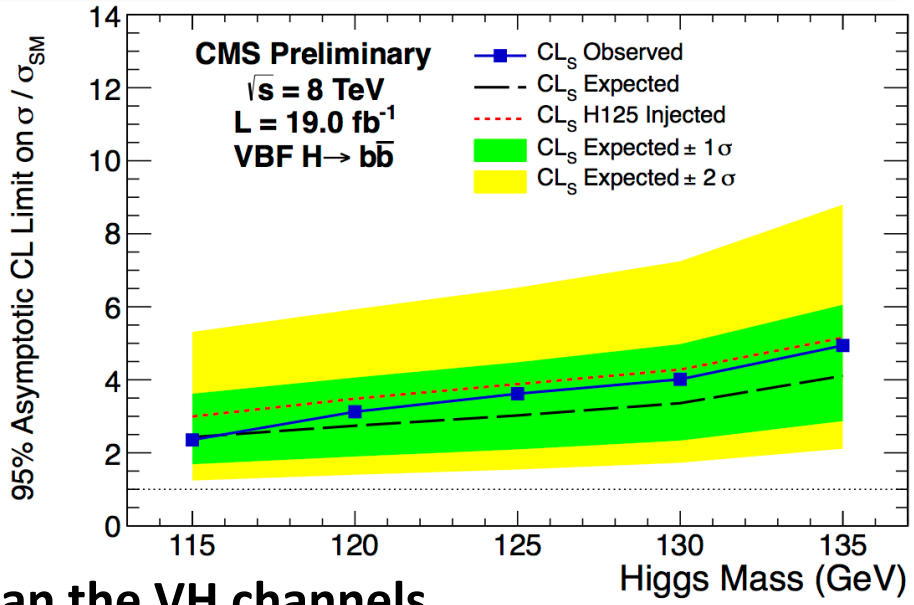
Cross-check : Search for **Z→bb** peak in the m(**bb**) spectrum.

- Same event pre-selection.
- Same fitting procedure.
- Result in agreement with the expectations:
  - observed (expected) significance: 8.0 (6.8)

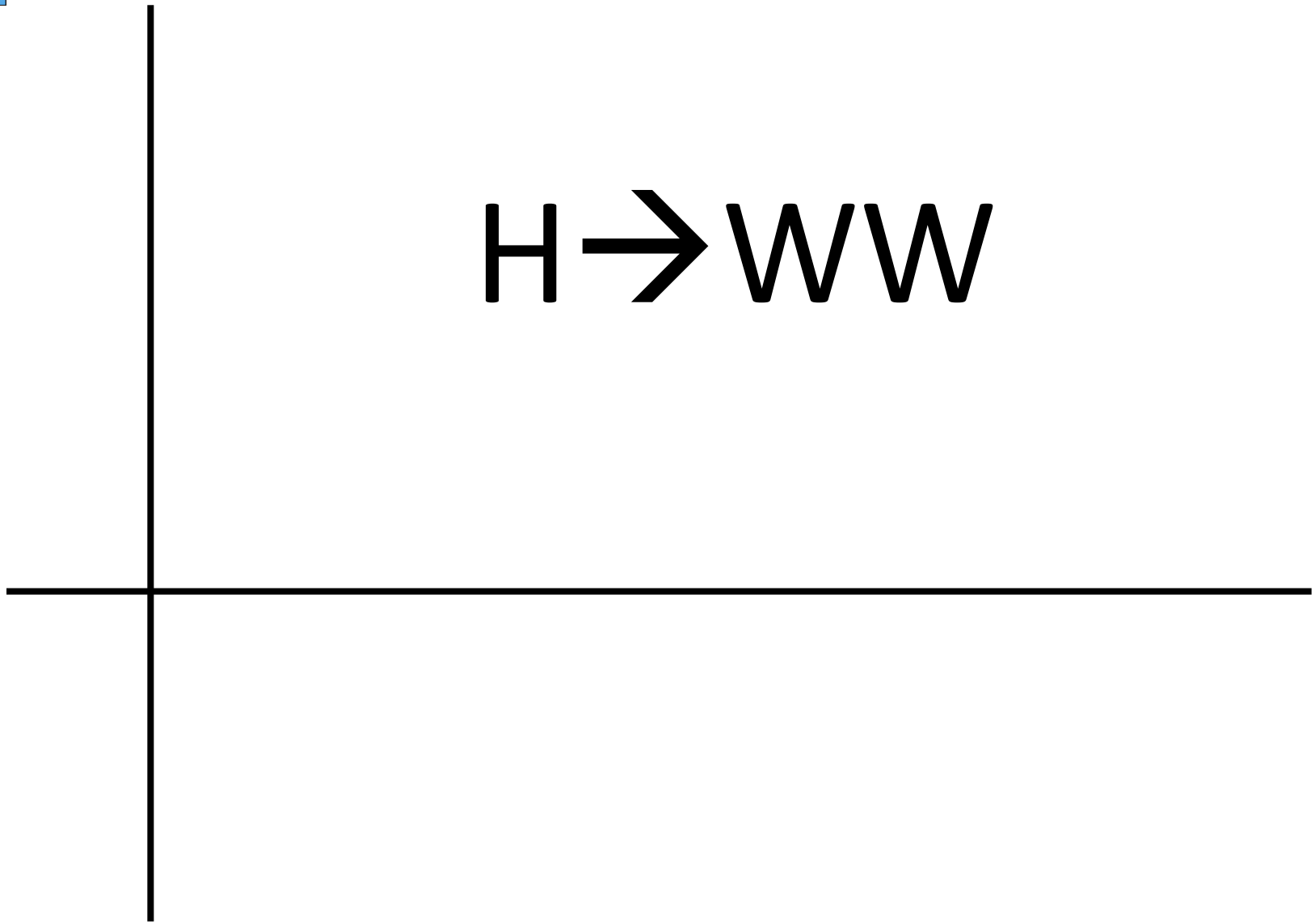


Results:

- **Limit at 125 GeV:**
  - Observed: **3.6** x SM
  - Expected **3.0** x SM
- **Best fit:** μ = **0.7 ± 1.4**



Less sensitivity than the VH channels


H→WW

Jet Multiplicity and H→WW

Search for a final state with two charged leptons: **H→WW→2l2ν.**

To increase sensitivity: categorization in jet multiplicity
- 0 jet, 1 jet and two or more jets.

→ Signal and background differ by categories.

	Zero-jet and one-jet ggH tag	Two-jet VBF tag	Two-jet VH tag
Number of jets	$= 0/1$	≥ 2	≥ 2
Default analysis	binned shape-based ($e\mu$) counting (ee, $\mu\mu$)	binned shape-based ($e\mu$) counting (ee, $\mu\mu$)	counting
Alternative analyses	parametric shape-based counting	counting	binned shape-based
VBF tagging	—	applied	vetoed
Main backgrounds	WW, top-quark, W + jets, $W\gamma^{(*)}$	WW, top-quark	WW, top-quark

Category	ggH (%)	VBF (%)	VH (%)	Total H → WW yield $\sqrt{s} = 7$ TeV	$\sqrt{s} = 8$ TeV
Two-lepton analyses					
0-jet different-flavor (shape-based)	95.7	1.2	3.1	52.6	245
0-jet same-flavor (counting)	98.1	0.9	1.0	10.4	58.5
1-jet different-flavor (shape-based)	81.6	10.3	8.1	19.8	111
1-jet same-flavor (counting)	83.6	11.2	5.2	3.1	19.6
2-jet VBF tag different-flavor (shape-based)	22.3	77.7	0.0	1.3	6.4
2-jet VBF tag same-flavor (counting)	14.2	85.8	0.0	0.3	2.3
2-jet VH tag different-flavor (counting)	55.5	4.7	39.8	0.8	4.3
2-jet VH tag same-flavor (counting)	65.1	4.1	30.8	0.2	2.8

Jet characteristic to VBF process.

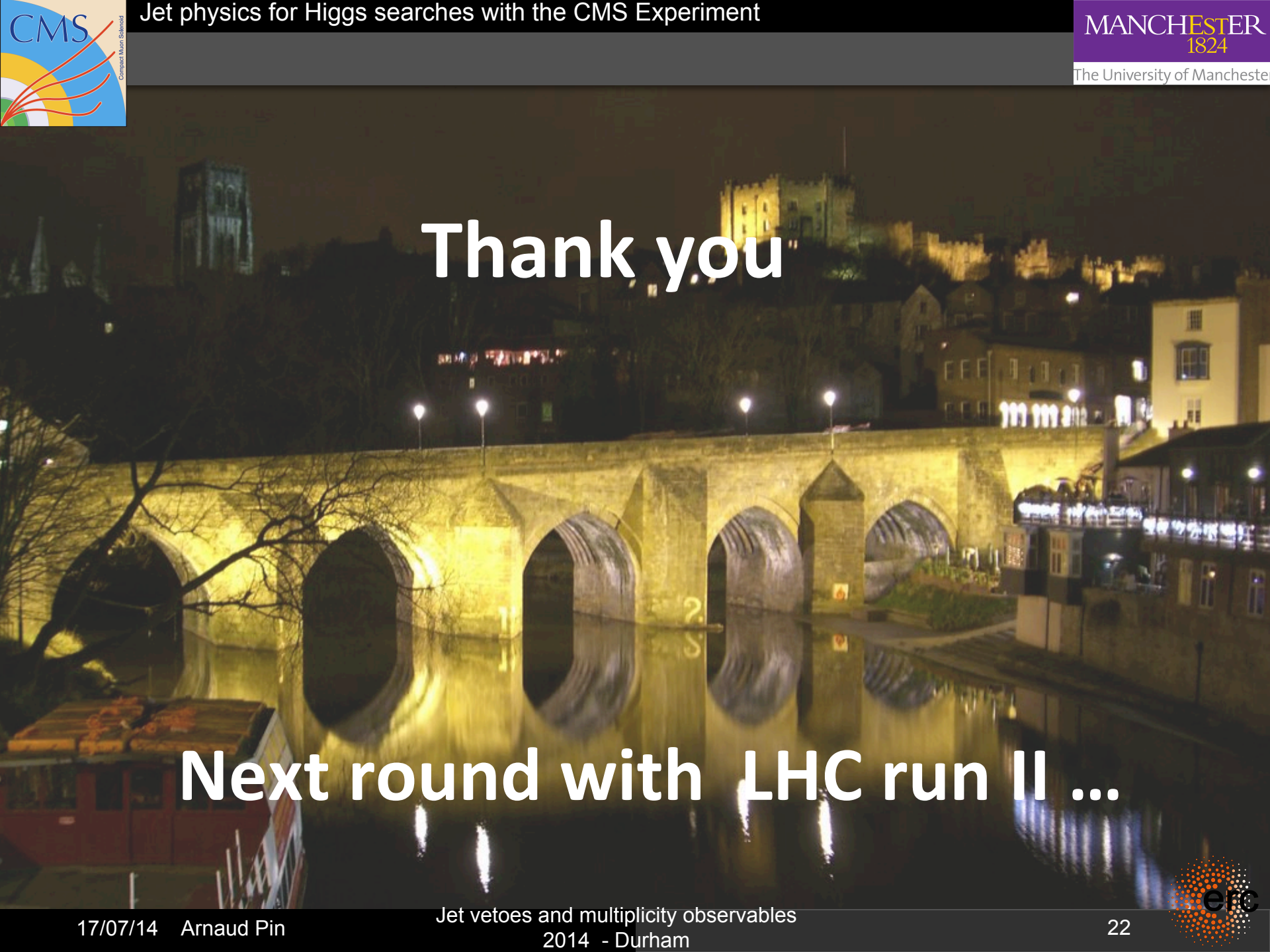
Centrally produced jets from decay of W/Z

Jet counting categorization → systematic uncertainty: - of 7-20% on the signal yields
- of ~ 5% on VV yields

➢[CMS-PAS-HIG-13-023]

Summary

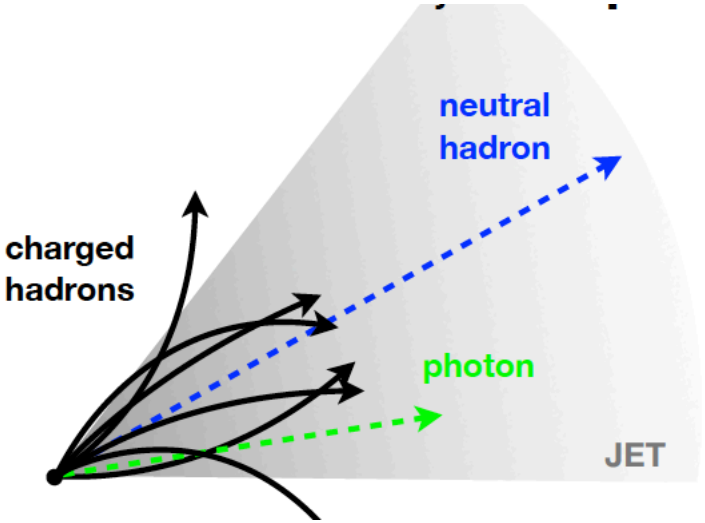
- Three production modes investigated for H→bb:
  - Associate production with vector boson (WH / ZH).
    - → **Most sensitive analysis** [PhysRevD.89.012003] (5.0 + 19.5 fb^{-1})
    - -Sensitity improved due to b-jet energy correction.
  - Higgs production by vector boson fusion [CMS-PAS-HIG-13-011] (19.5 fb^{-1})
    - -Successfully employed quark-gluon separation
      - Multiplicity, width, energy sharing
      - Simple likelihood with three inputs, up to $|\eta| = 5$
      - Quarks well modeled by MC, gluons not so well
- Search for H→WW:
  - Categorization in term of number of jet.
  - Analysis optimized for each category → to reach the best sensitivity.

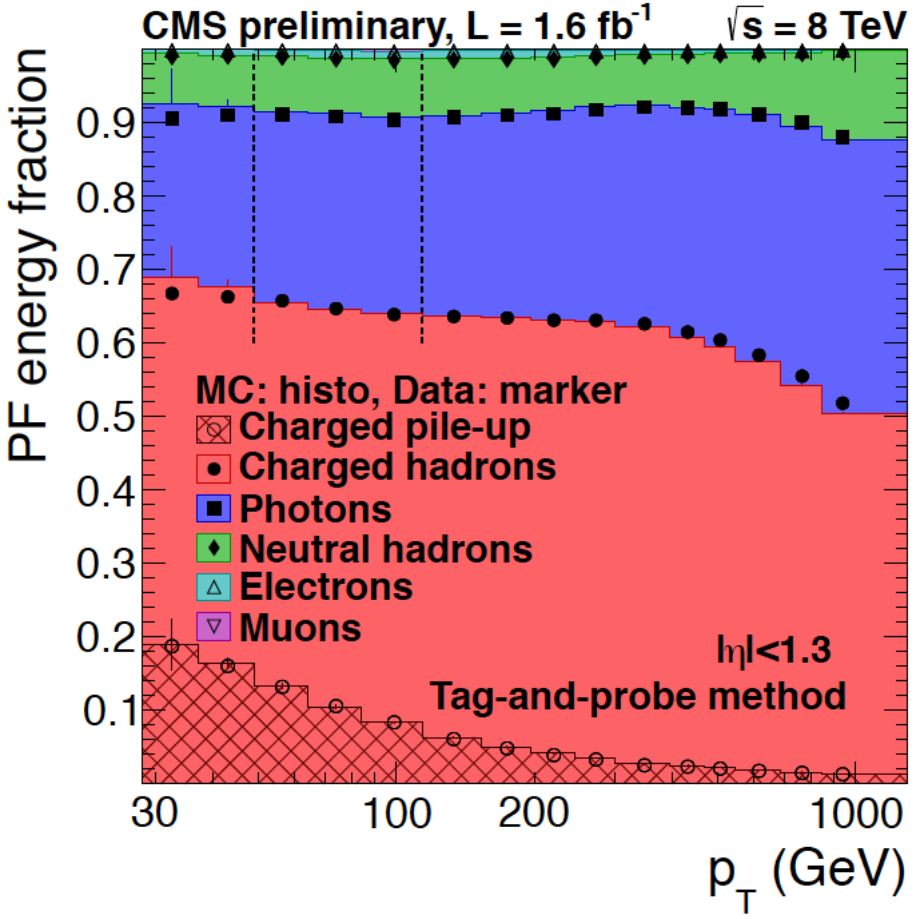

Thank you
Next round with LHC run II …

Backup

Particle flow jet.

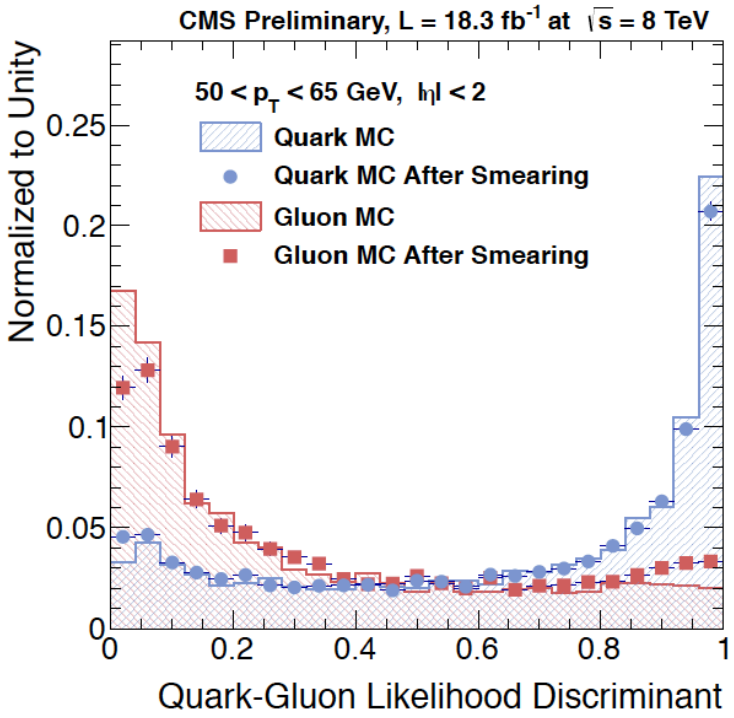
- **Particle Flow:**
  - Reconstructing all stable particles in event
  - Using all detectors in unison
- Powerful tool for jet structure
  - Particle-level information
  - Full access to jet shape





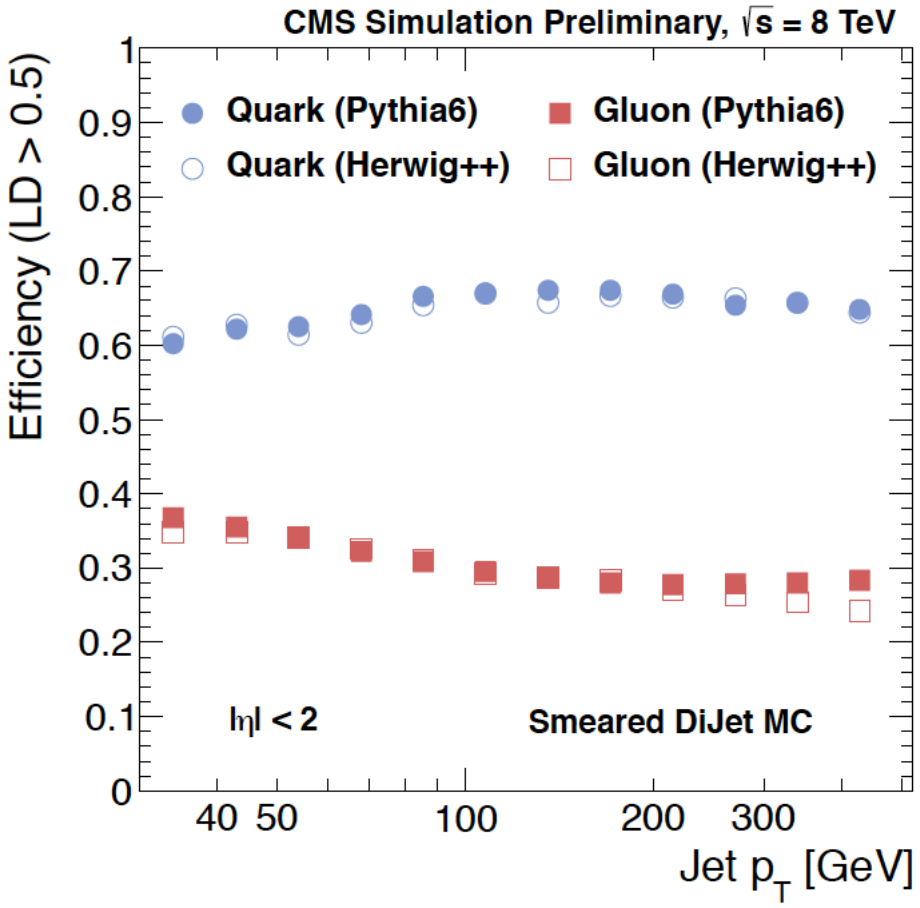
Gluon quark tagger smearing

- Take likelihood (LD) distributions for quarks and gluons separately
- Define smearing function: $g(x,a,b) = \tanh\left(a \operatorname{arctanh}(2x-1)+b\right)/2 + \frac{1}{2}$
  - Changes value of LD on jet-per-jet basis
  - Not a reweighting
  - Reduces LD discrimination
- Smear until data and MC agree
  - χ^2 minimization



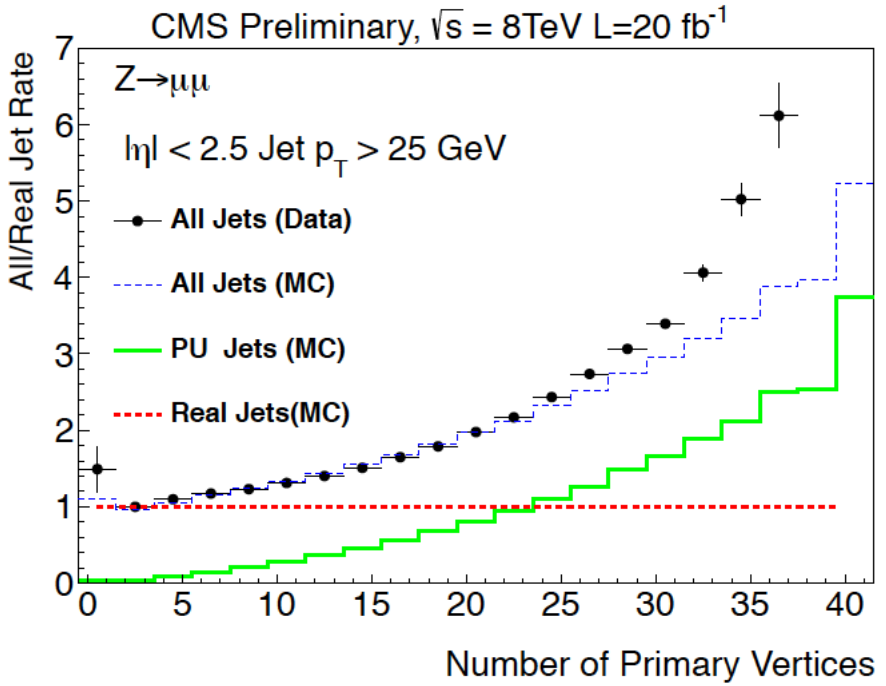
Gluon qaurk tagger smearing

Pythia – Herwig agreement:



Pile-Up

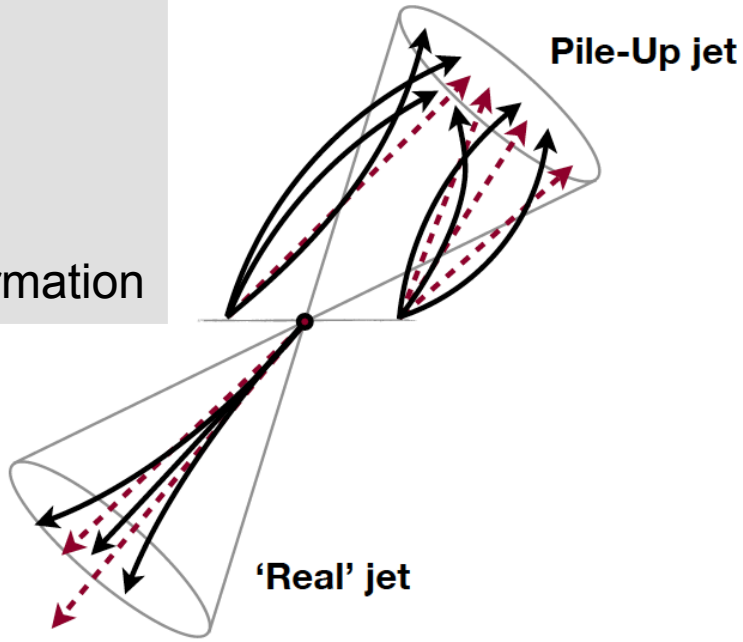
- Current LHC running: high-pileup conditions
  - Average of 23 extra interactions (and up to 40!)
- Additional collisions produce soft jets
  - But can overlap (combinatorics!)
  - Resulting jets can have **p**T > 25 GeV
- Pile-up jet ID
  - Crucial for analyses with low-**p**T jets



Pile-Up ID

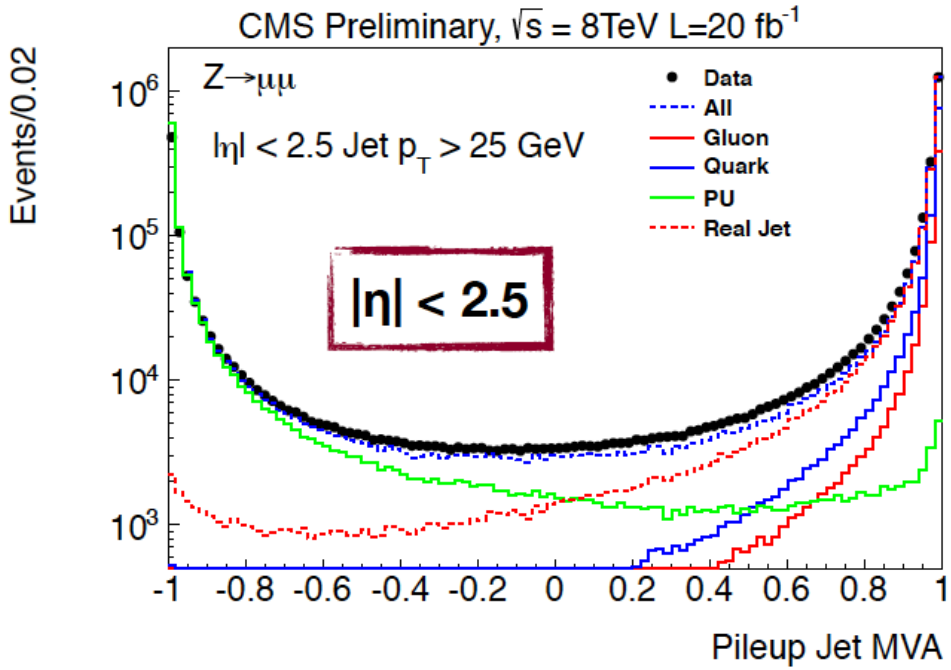
Pile up jets mainly overlap of soft jets from extra interactions

- Two main characteristics:
  - Tracks incompatible with primary vertex
  - Clustered particles more diffuse
- Selected 12 variables (scanned >80)
  - 4 track variables ($|\eta| < 2.5$)
  - 8 shape variables ($|\eta| < 5$)
- Again making use of powerful Particle Flow information



Pile-Up ID

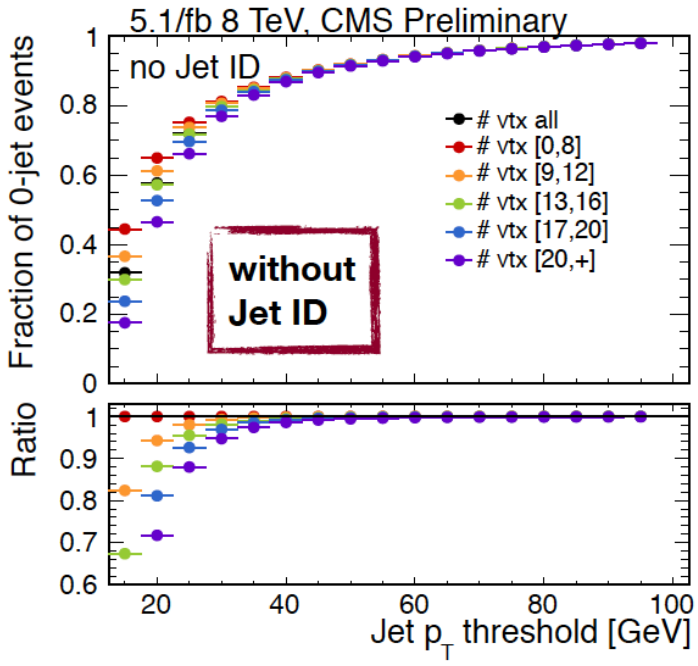
- Twelve variables fed to an MVA (BDT)
  - Trained separately in **four |η| bins**
  - Tested on Z→μμ data
- Best discrimination in **central region**

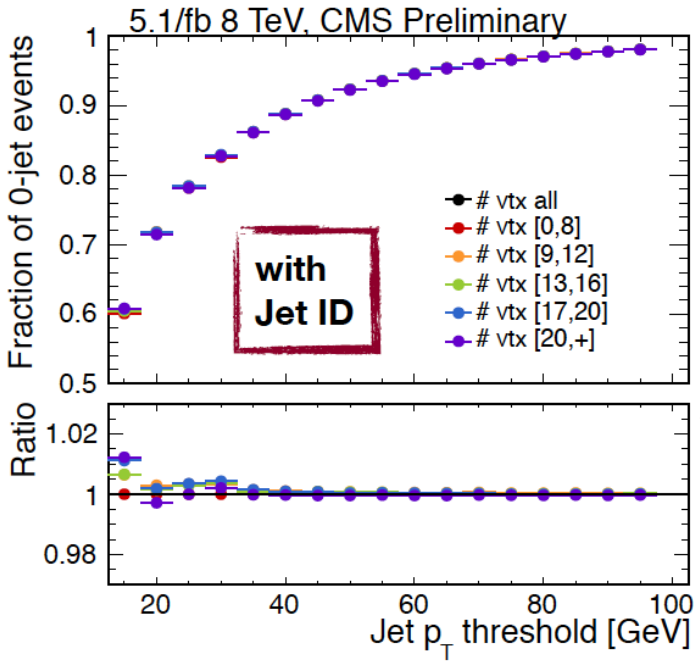


Application of Pile-up ID

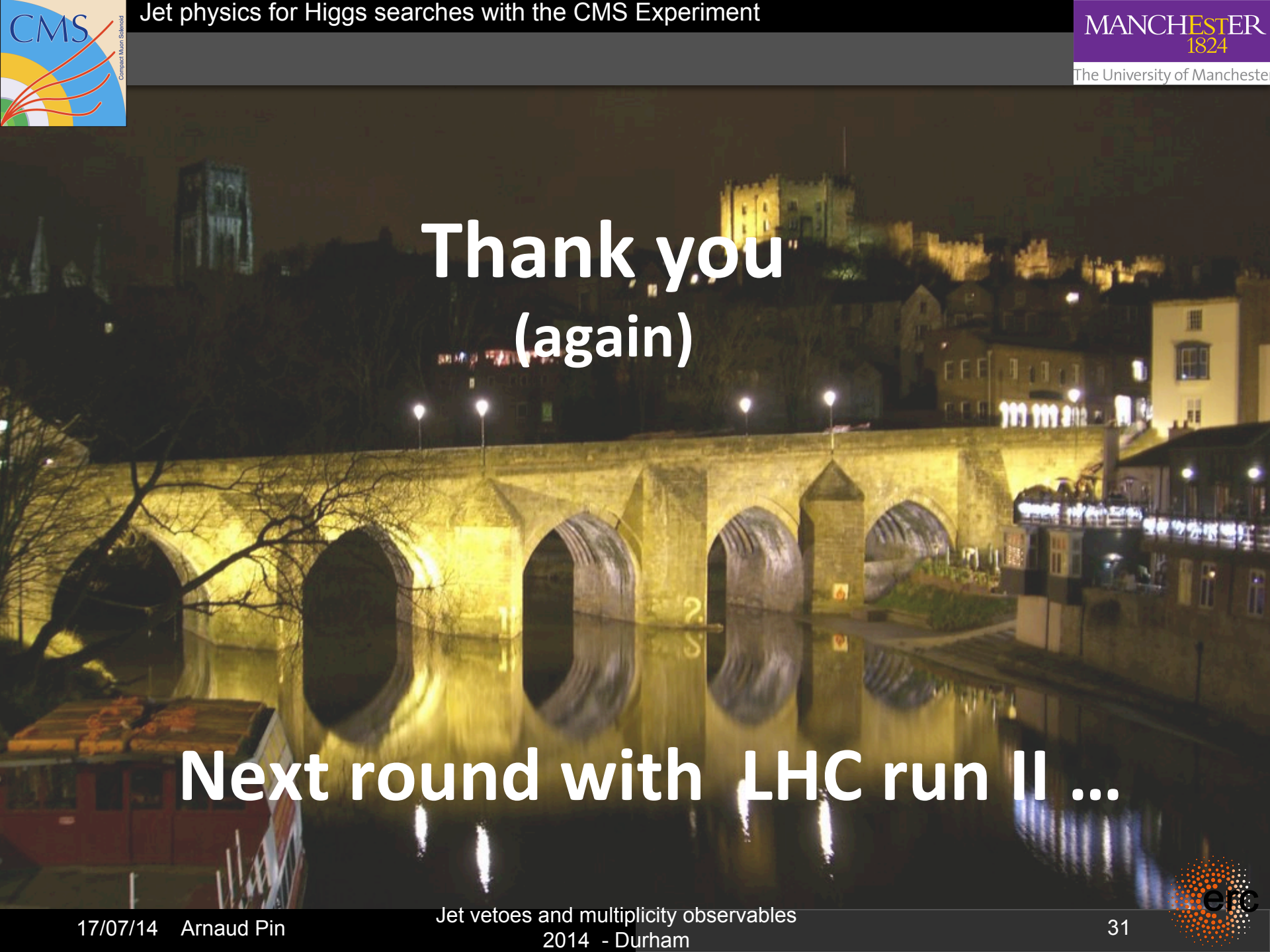
Jet Veto

- An example: **jet veto** in Z→μμ analysis
  - Without pile up ID: large dependence on pile up for $p_T < 40$ GeV
  - With pile up ID: **no dependence**





different axis scale!



Thank you
(again)

Next round with LHC run II …