Invisible Higgs & trigger challenges on ATLAS

A discussion on Dark sector,
Higgs boson,
Trigger, and
ML on FPGA

Tae Min Hong

University of Pittsburgh

Université de Genève
September 28, 2021

Outline

Motivation (10 min)

- Dark sector
- Higgs boson

Analyses (20 min)

- $H_{125} \rightarrow$ Dark matter pair — ATLAS Collab, ATLAS-CONF-2020-008 (2020)
- $H_{125} \rightarrow$ Dark photon $+ \gamma$ — ATLAS Collab, ATLAS-CONF-2021-004 (2021)

Trigger (30 min)

- E_T^{miss}, VBF — ATLAS, JHEP 08 (2020) 080
- ML on FPGA — Hong et al., JINST 16 (2021) P08016

Motivation

Dark sector
Higgs boson

Many evidence of dark matter

Here: colliding galaxy clusters

3 MLy (MACSJ0025, 2008)

Inferred distribution of matter by lensing

– X-rays from known matter

= Dark matter (Known unknown)

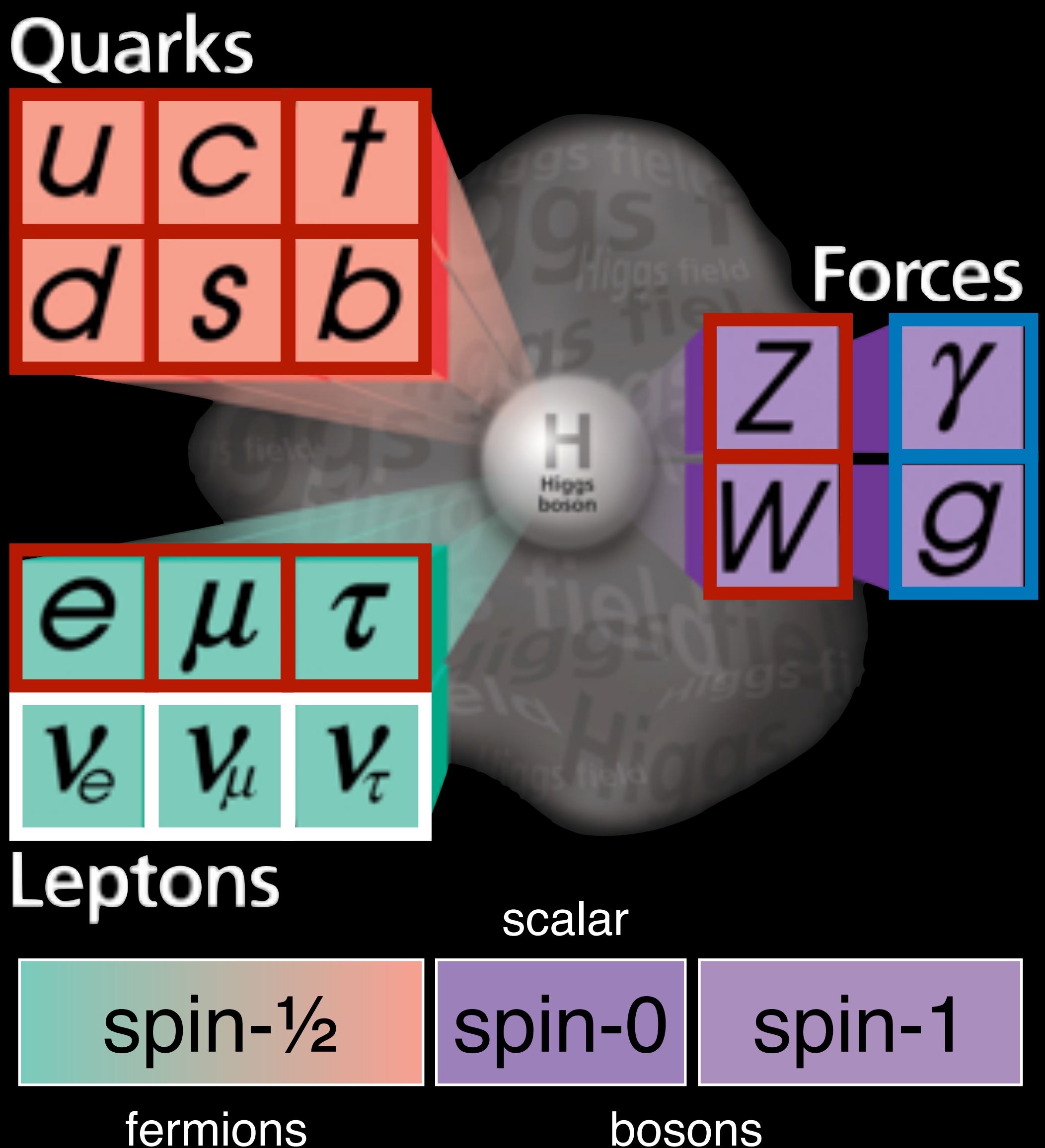
Source: https://apod.nasa.gov/apod/ap080917.html

Can we create dark matter? Related to Higgs?

Higgs boson couplings

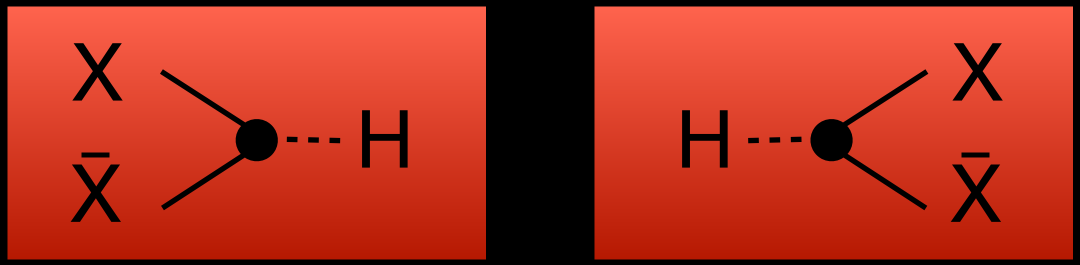
Higgs couples to everything

Table of elementary particles

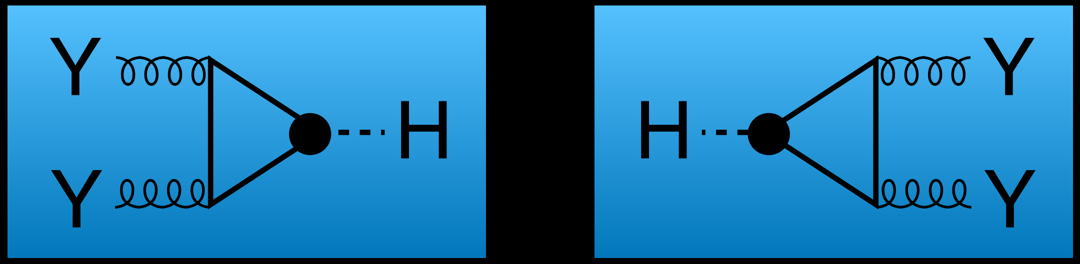


Feynman diagrams

Tree relation for massive X



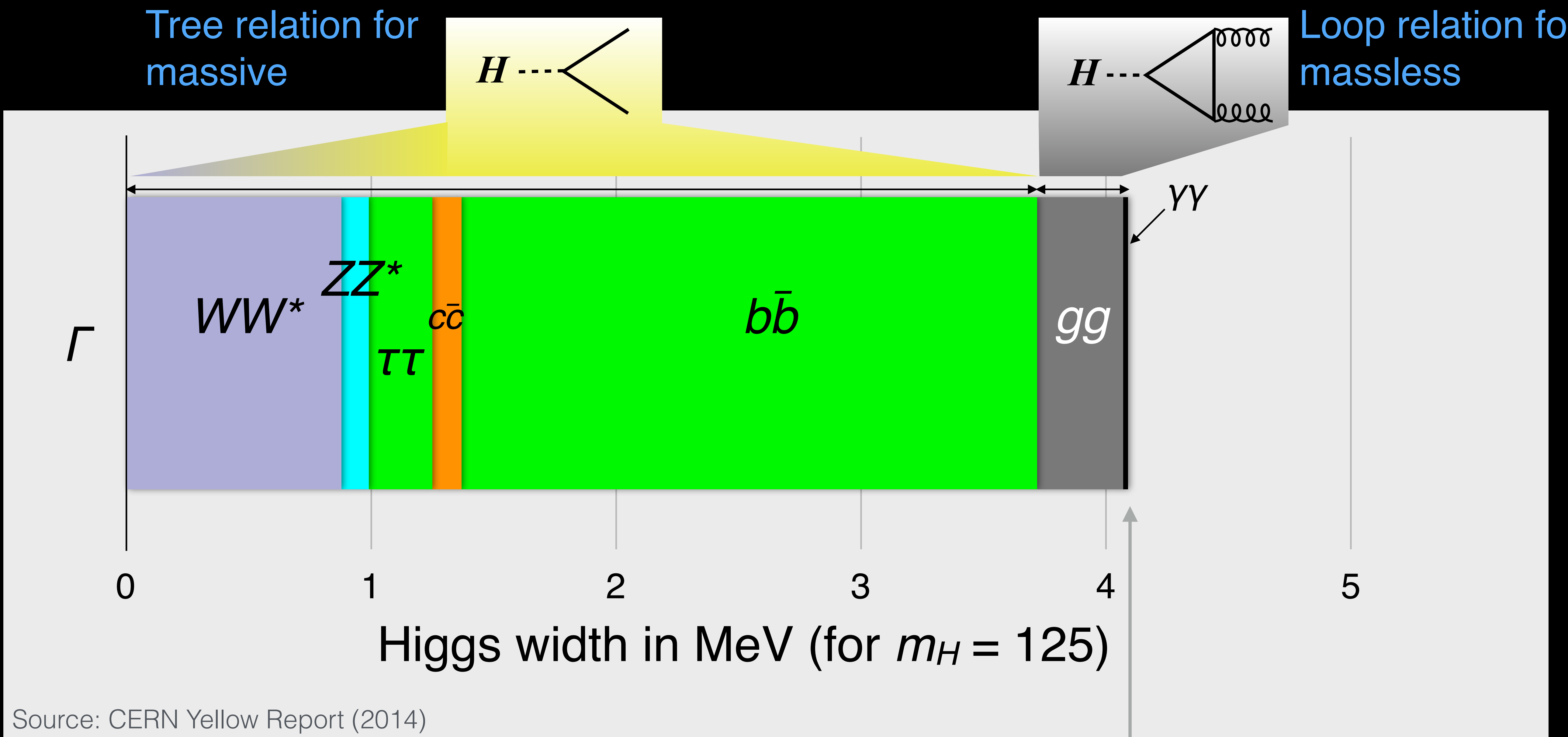
Loop relation for massless Y



Source: Fermilab

Higgs coupling to each particle is determined given m_H

Higgs is so narrow because 125 GeV



Massive bosons
p.s. suppressed
($m_H \ll 2 \cdot m_{W,Z}$)

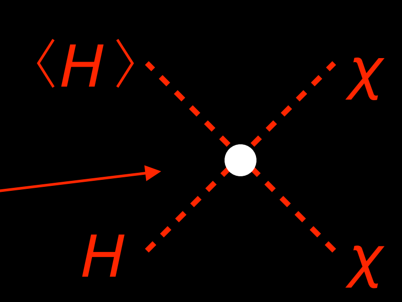
Fermion m_F
means tiny Yukawa couplings
($t\bar{t}$ large, but $m_H \ll 2 \cdot m_{top}$)

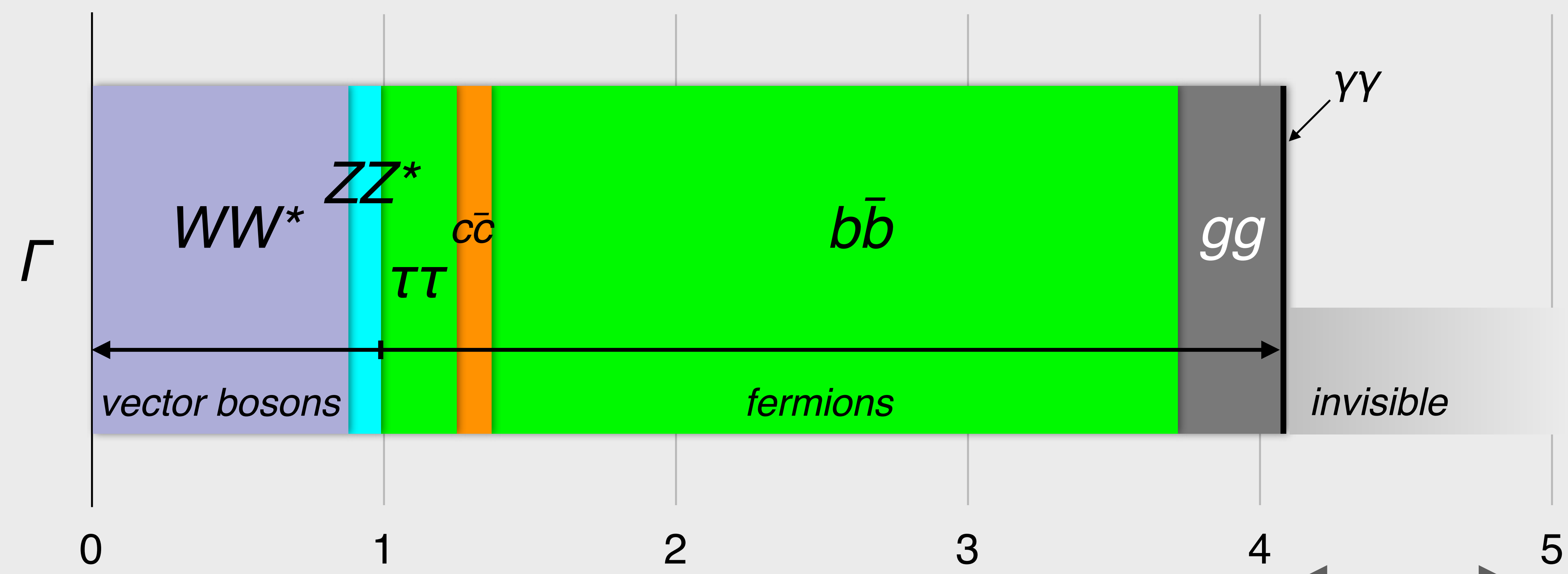
Massless bosons
loop suppressed

Everything is suppressed

O(MeV) is ~~not un~~reasonable

Portals to NP can look like





Measure these branching ratios

to **indirectly** limit invisible

Search

Source: CERN Yellow Report (2014)

Theory

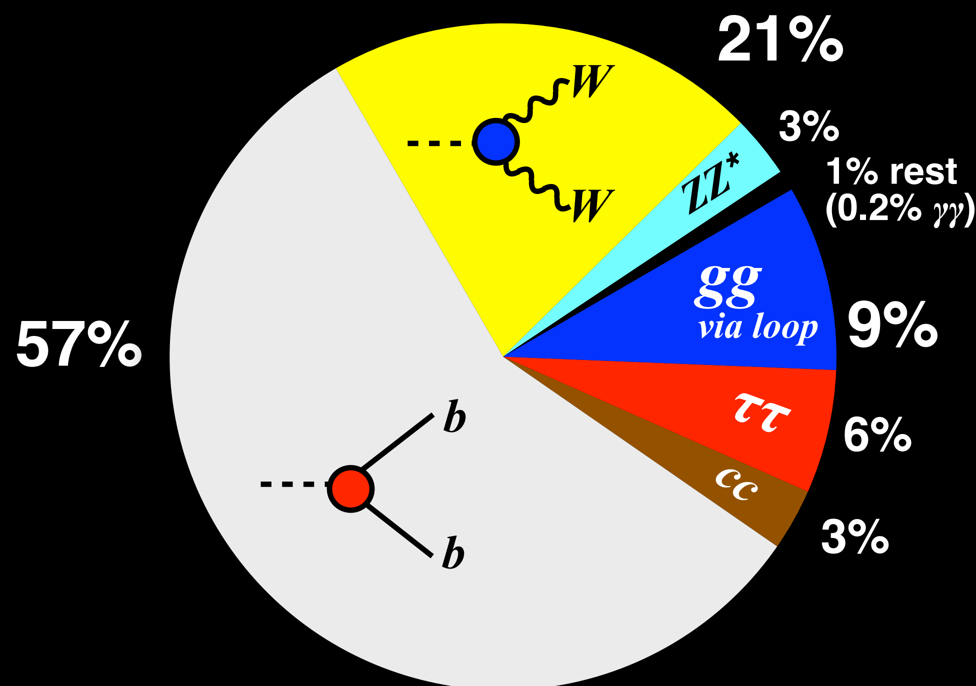
- SM singlet ~ g $H^2 \chi^2$
- Fully renormalizable

Goncalves, Han, Mukhopadhyay, PRL 120, 11801 (2018)
Curtin, Essig, Zhong, 1412.4779 (2015)
Curtin +12 others, 1312.4992 (2014)
Chang +3 others, 0801.4554 (2008)
Silveira & Zee, PL B161,136 (1985)
and many many more papers.

Coupling at 0.01 → MeV-level modification → Large rate

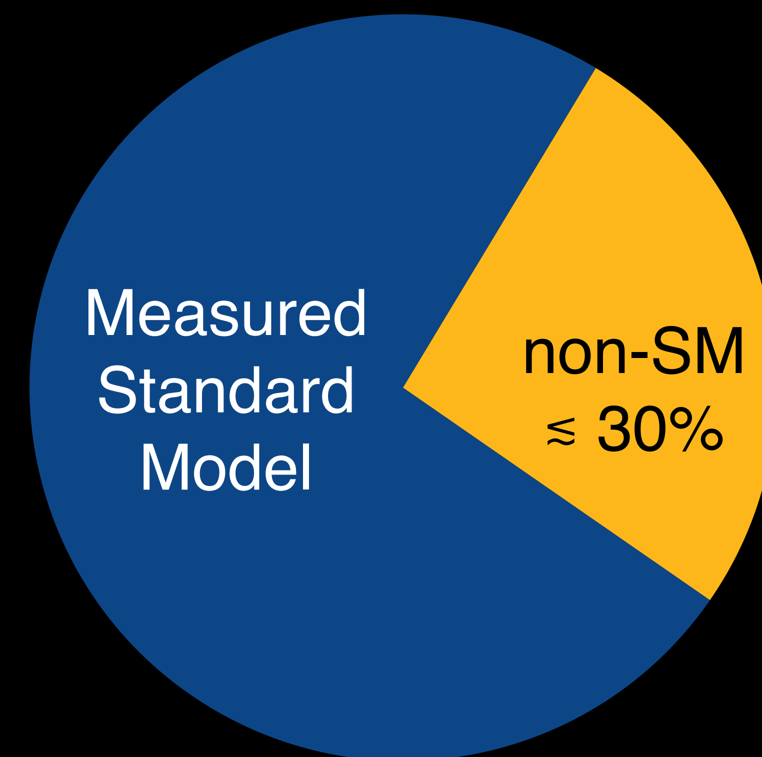
What's allowed by individual measurements?

Standard Model



Measure each predicted slice

Allowed



Constrains non-Standard Model
(with caveats)

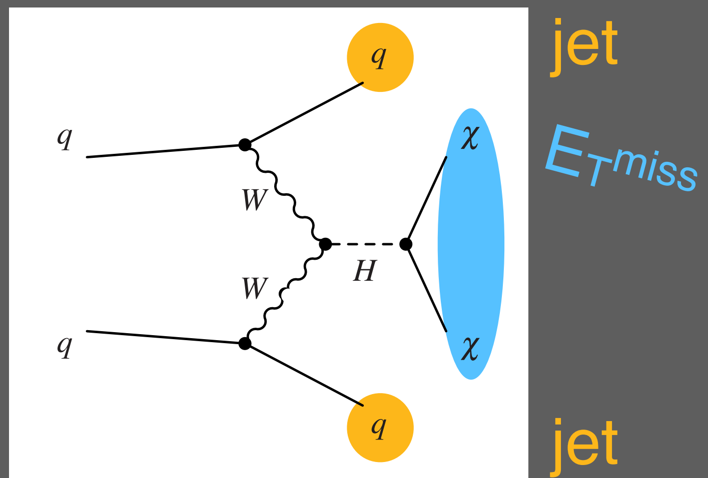
Source: http://cdsweb.cern.ch/record/2629412

A fraction of Higgs decays *could* be related to our hypothesis

Dark matter (m>0)

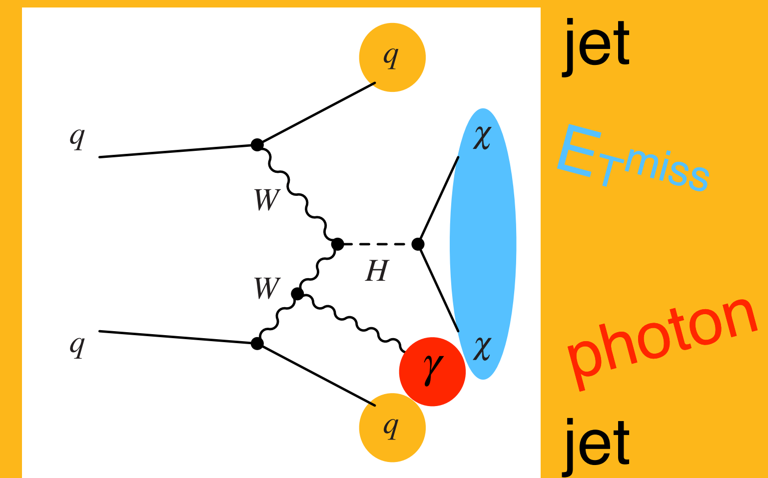
Related production channels

Dark matter



→

Dark matter + γ



Higgs production

smaller σ

- ggF — No observable!
- ggF + 1 jet — Overwhelming strong bkg'd
- VBF — Depends on trigger threshold →
- ZH — Suppressed by σ • B
- WH — " and neutrino / hadronic W

Comparison to without γ

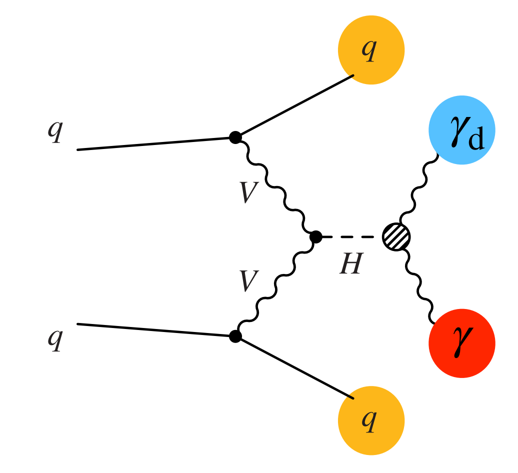
- Smaller signal size, but clean
- Adding γ reduces strong background
- Depends on trigger threshold
- Added bonus in interpretation (next slide)

Trigger threshold is critical aspect of the study

Dark matter (m>0) → Dark photon (m=0)

Maybe there is a Dark sector

Dark photon



jet

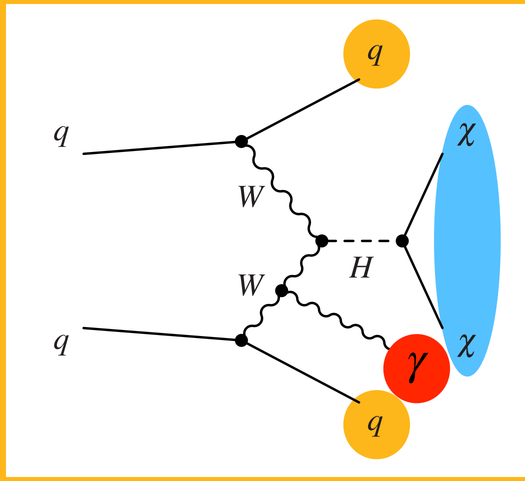
E_T^{miss}

photon

jet

←

Dark matter + γ



jet

E_T^{miss}

photon

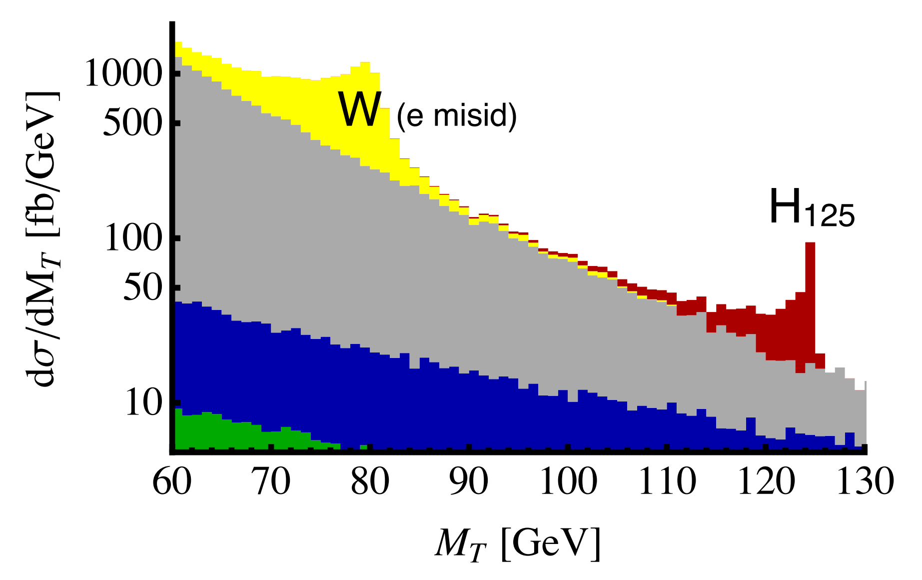
jet

same final state

Theory

- Unbroken $U(1)_{dark}$ with enhanced $H \to \gamma\, \gamma_{dark}$
- Signal peaks in the m_T of (E_T^{miss},γ) system ⟶

Gabrielli, Heikinheimo, Mele, Raidal, PRD 90, 055032 (2014)



Expand the scope of the search with alternate signal models

Analyses

$H_{125} \rightarrow$ Dark matter pair
$H_{125} \rightarrow$ Dark photon + γ

Physics of VBF $H \rightarrow$ *invisible*

VBF production of the Higgs is established

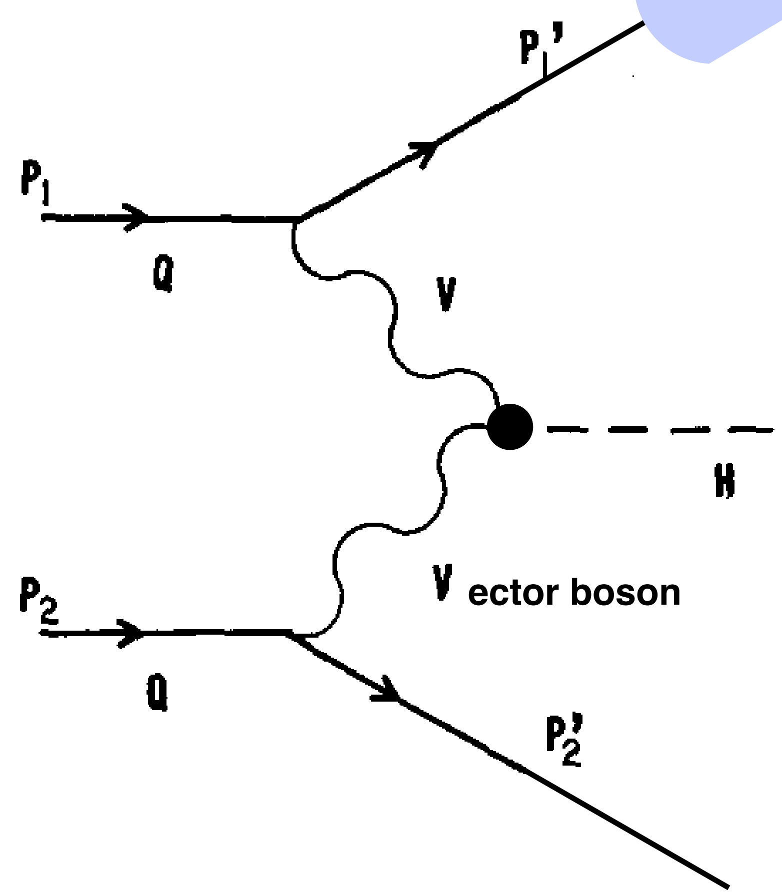
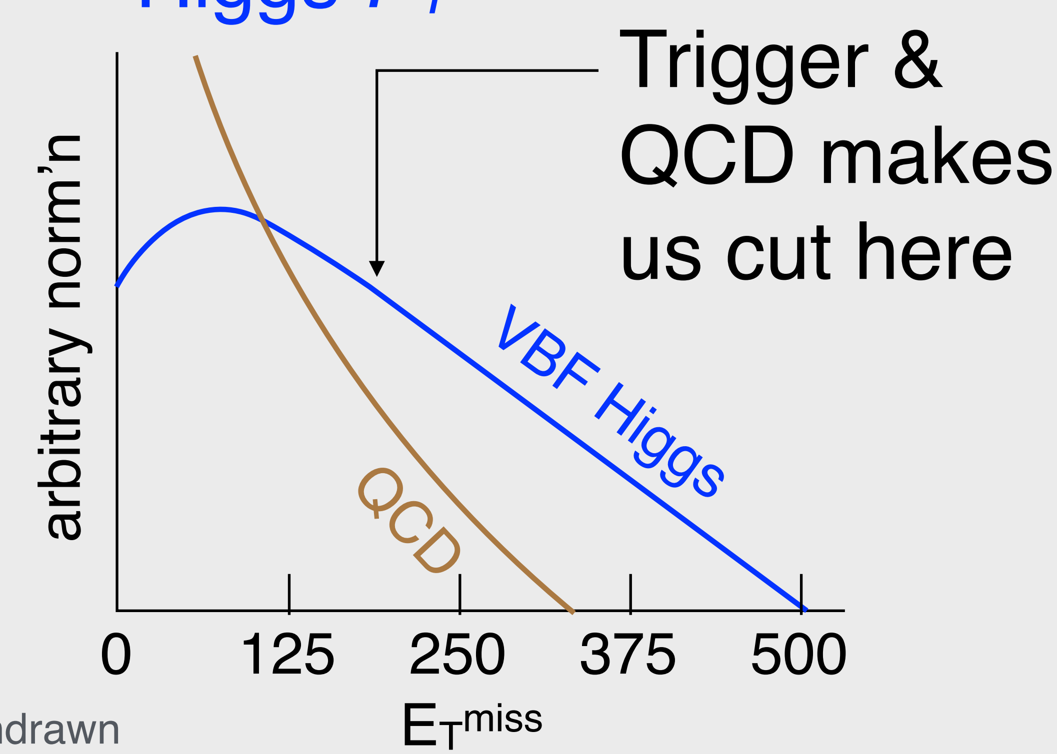


Fig. 1. Higgs boson production from virtual vector boson pairs (V = W or Z).

Source: Cahn, Dawson, PLB 136 (1984) 196

- Energetic jets with large η gap
- No hadronic activity
- m_{jj}, $\Delta\eta_{jj}$, $N_{\text{central jets}}$

- E_T^{miss} ~ Higgs P_T

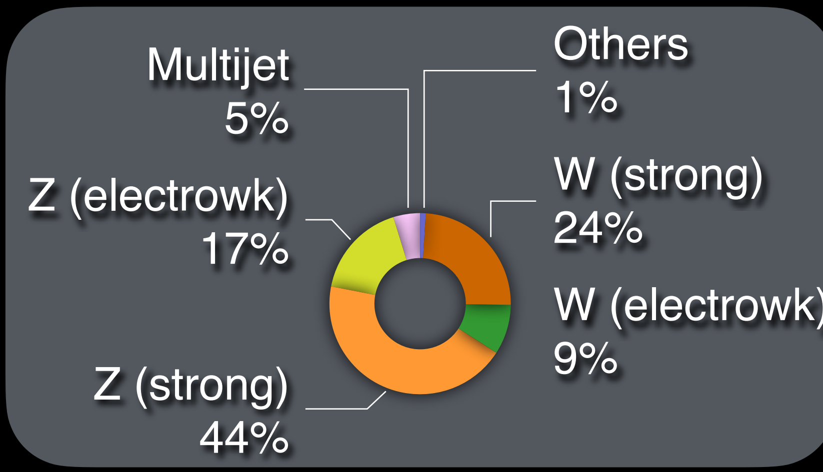


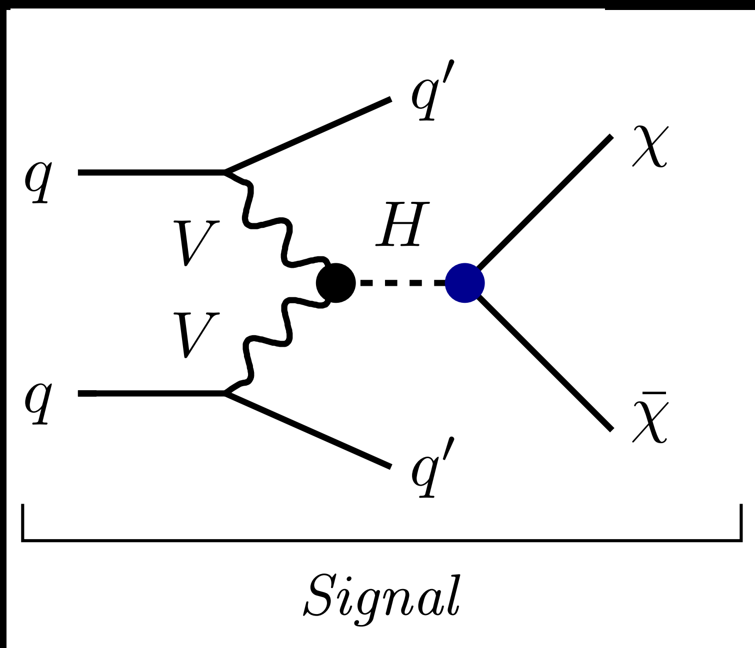
Source: handrawn

Not many handles, background estimation is crucial

Signal & background

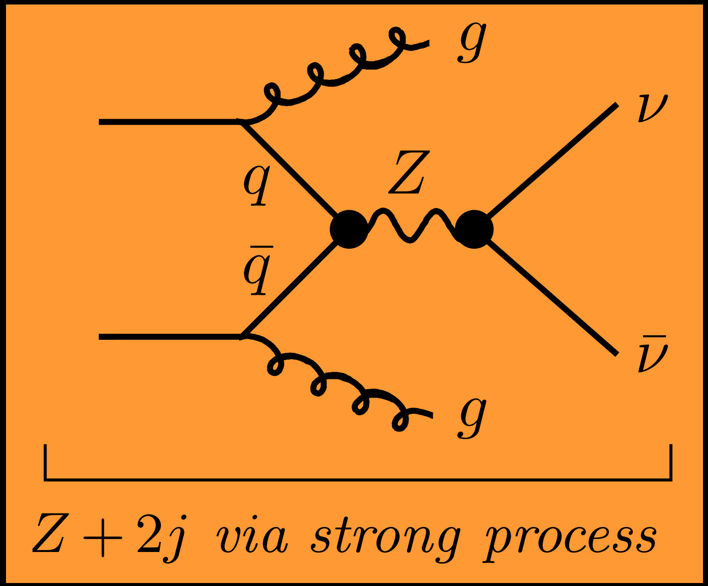
Z & W are largest

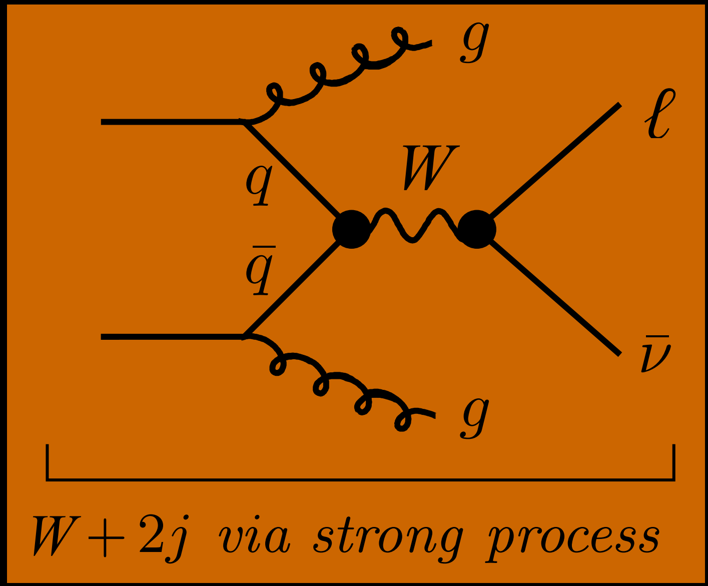


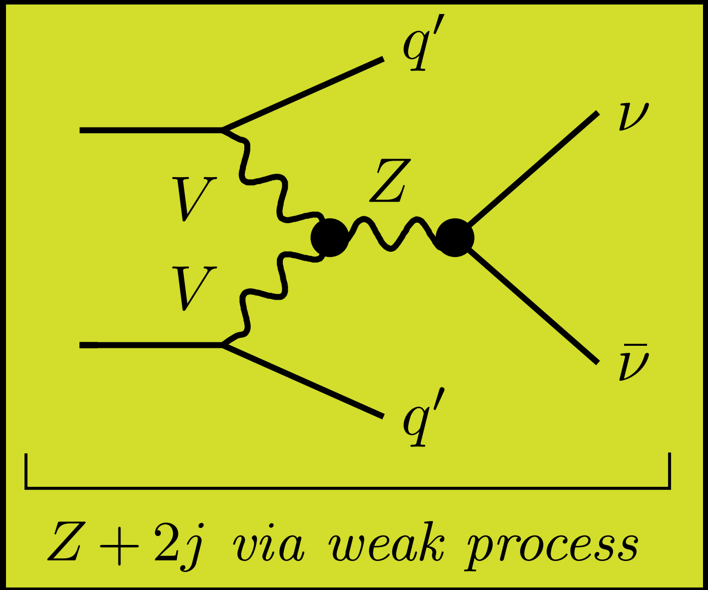


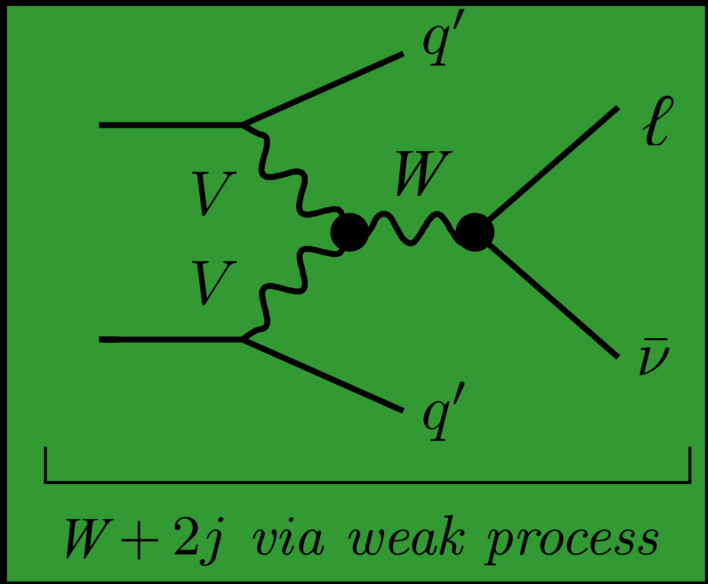
"Irreducible"

Lost lepton





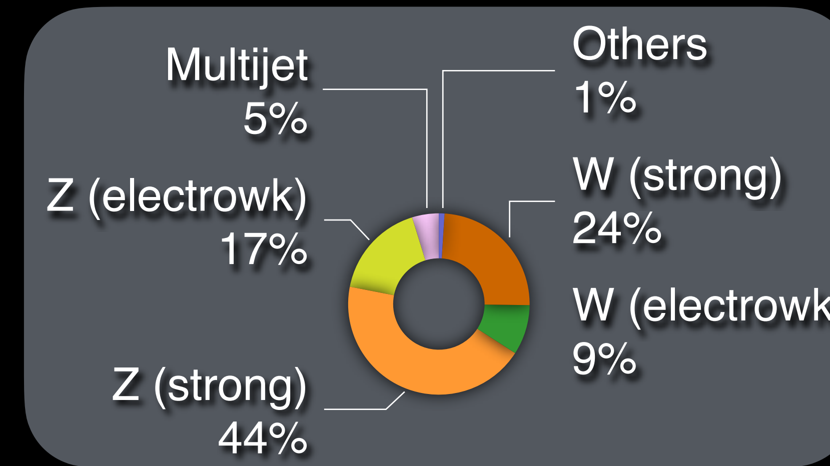


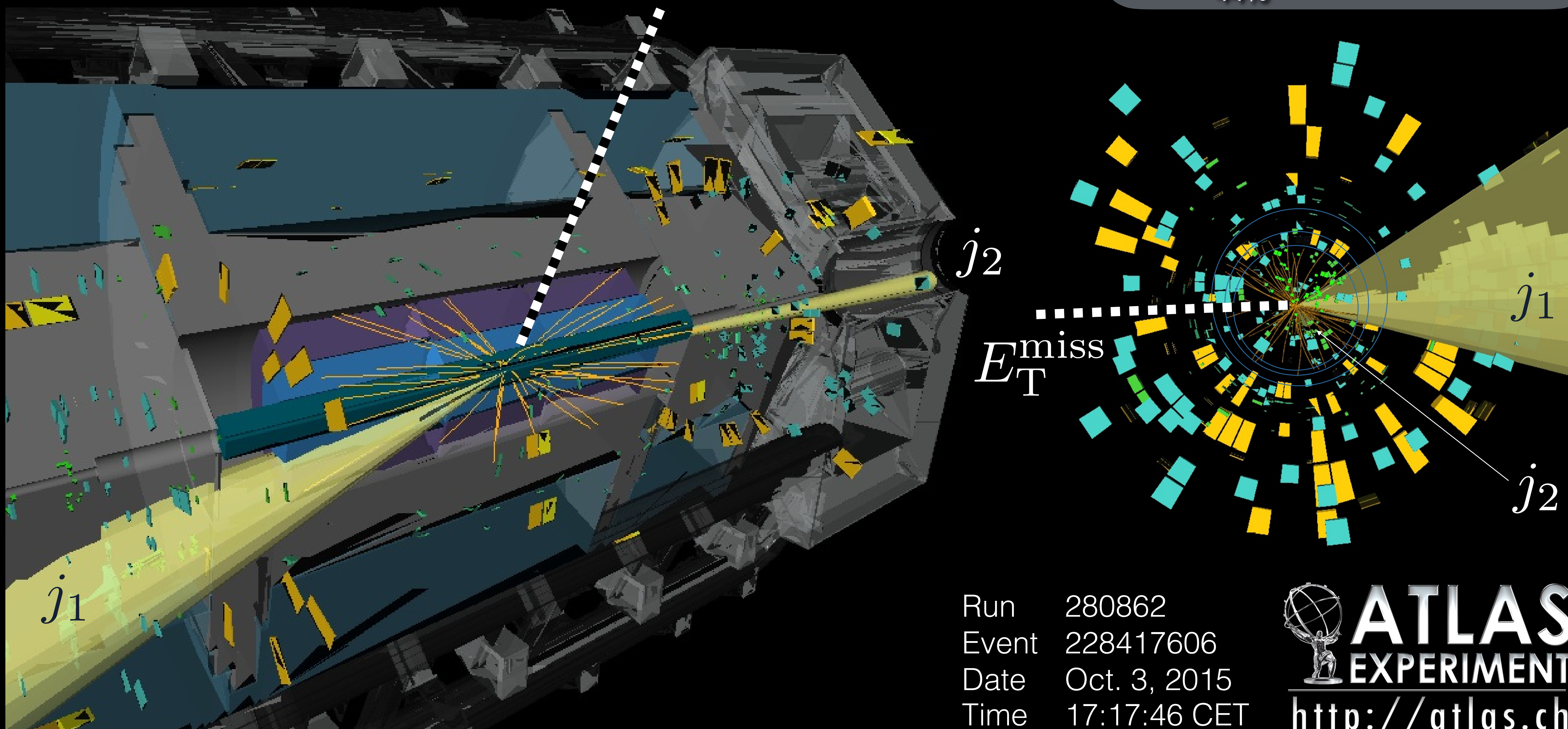


Strong process is dominant, but weak process imp't at higher m_{jj}

Event display

High E_T^{miss} (564 GeV) + High m_{jj} (3.6 TeV)

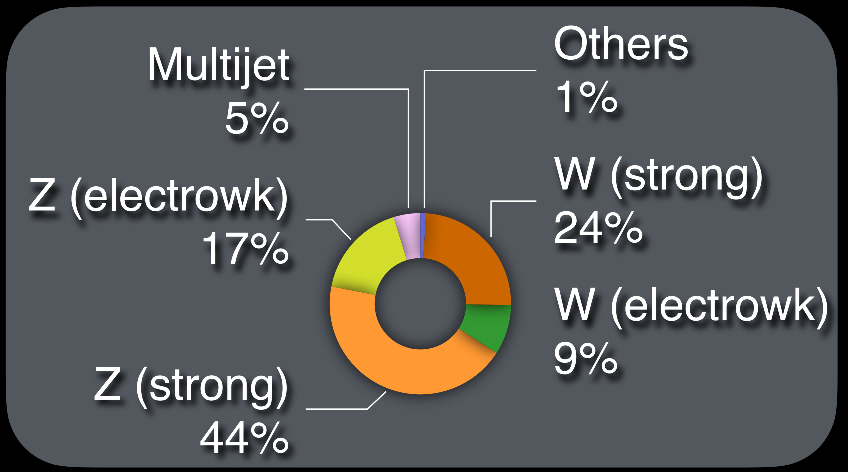




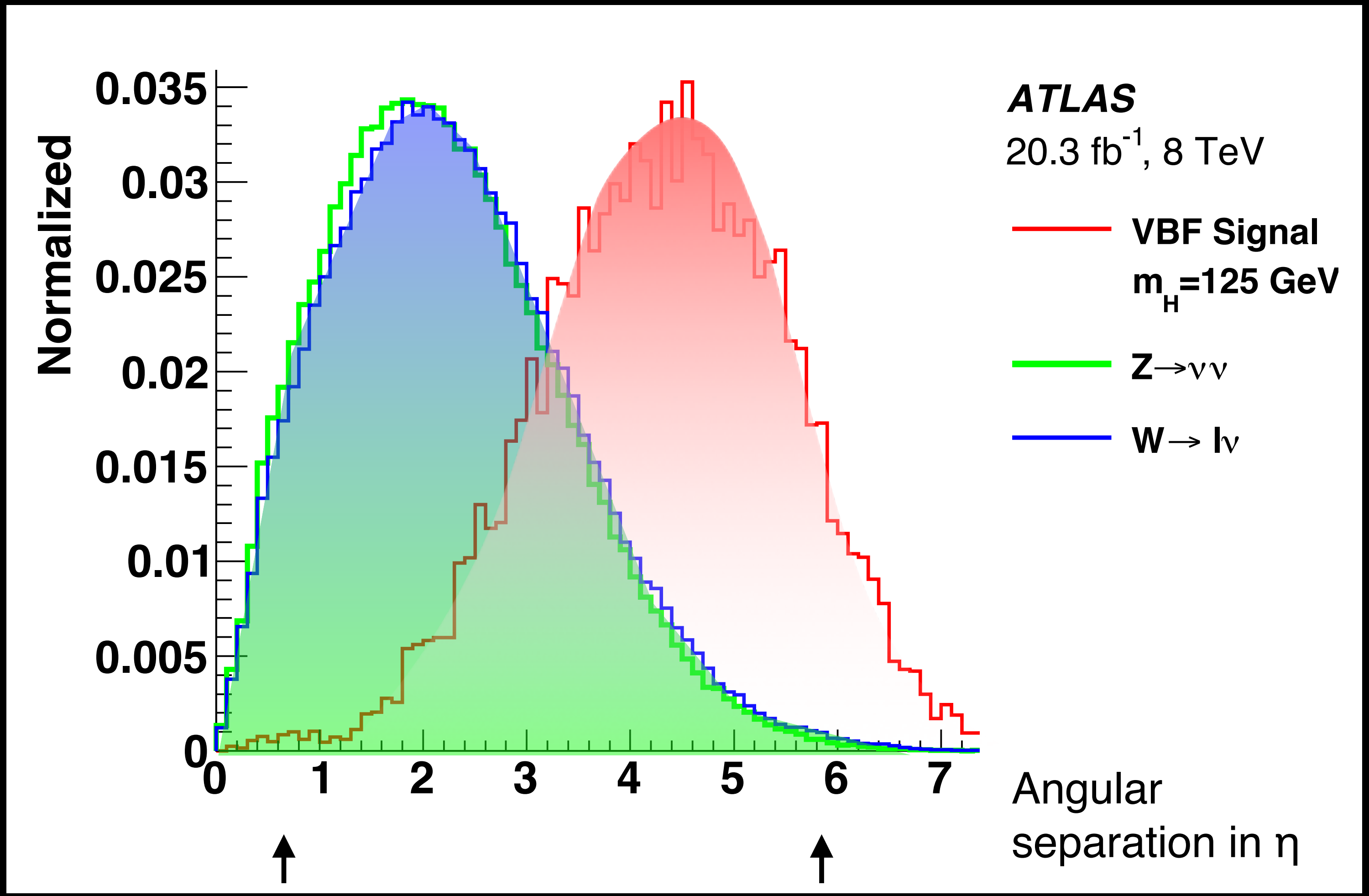
High E_T^{miss} event + two energetic jets with large $\Delta\eta$

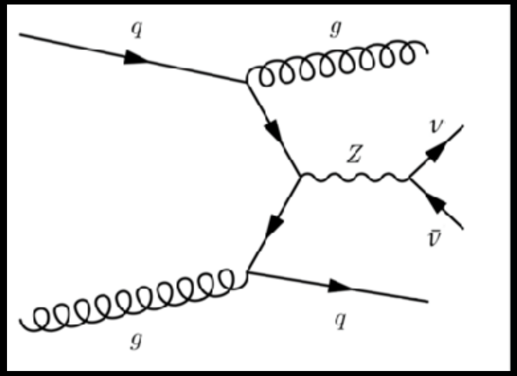
Angular characteristics

Data distribution of separation in η



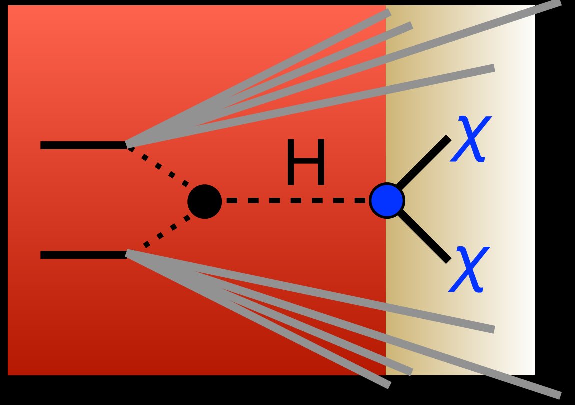
ATLAS, J. High Energy Phys. 01 (2016) 172





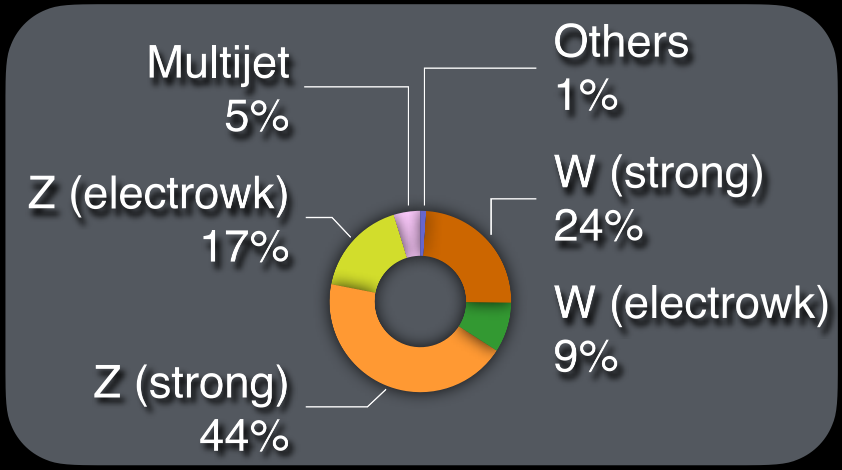
Known processes

Higgs signal process



Selection

Signal region, control regions

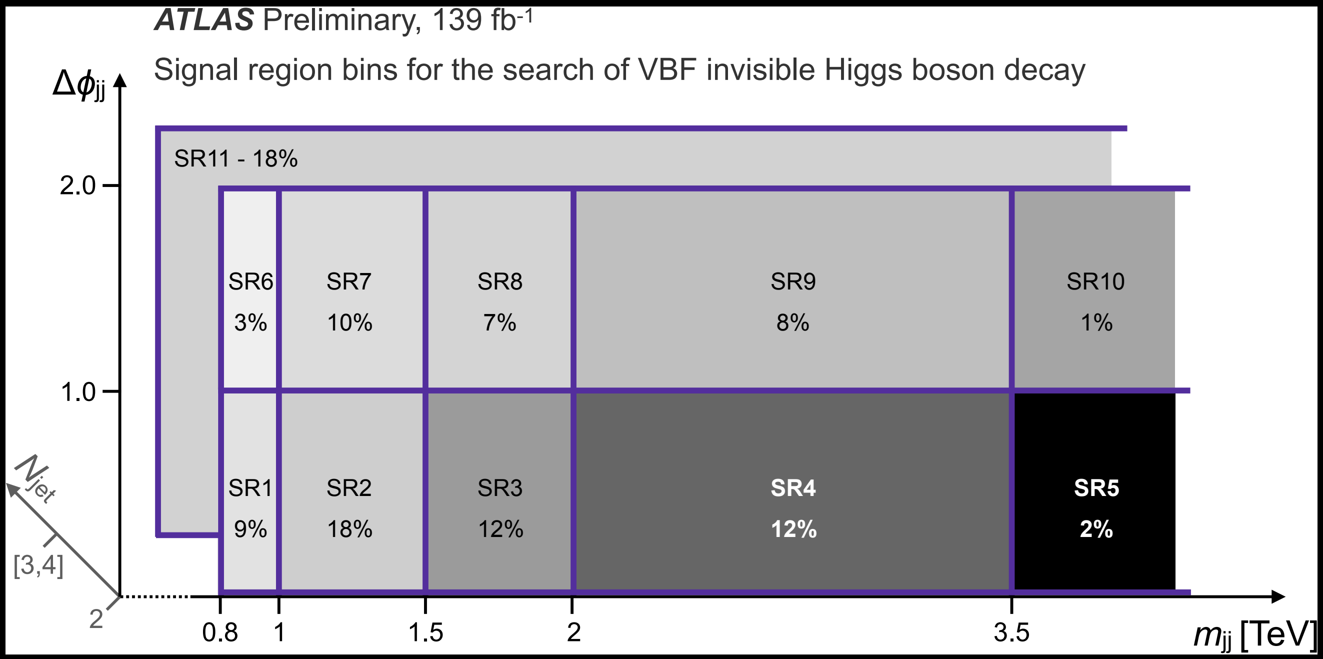


Cuts

- Two jets > 80 GeV, 50 GeV
- Centrality of additional jets
- E_T^{miss} > 200 GeV

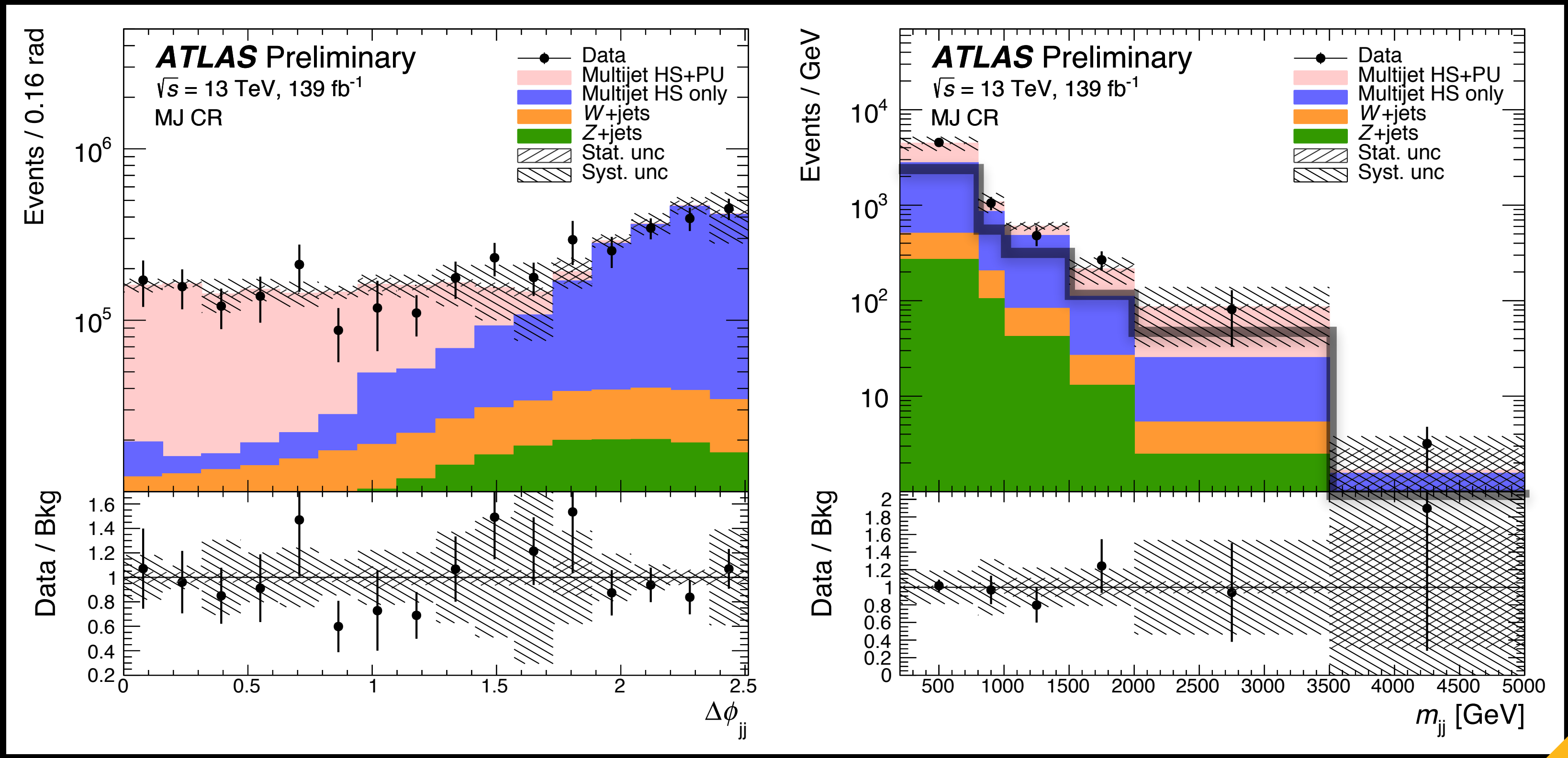
Signal region →

- Bin in $m_{jj} \otimes N_{jet} \otimes \Delta\phi_{jj} = 5 \otimes 2 \otimes 2$



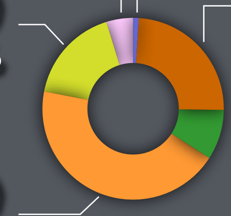
Control region

- W → ℓν
- Z → ℓℓ
- Multijet by rebalance & smear →

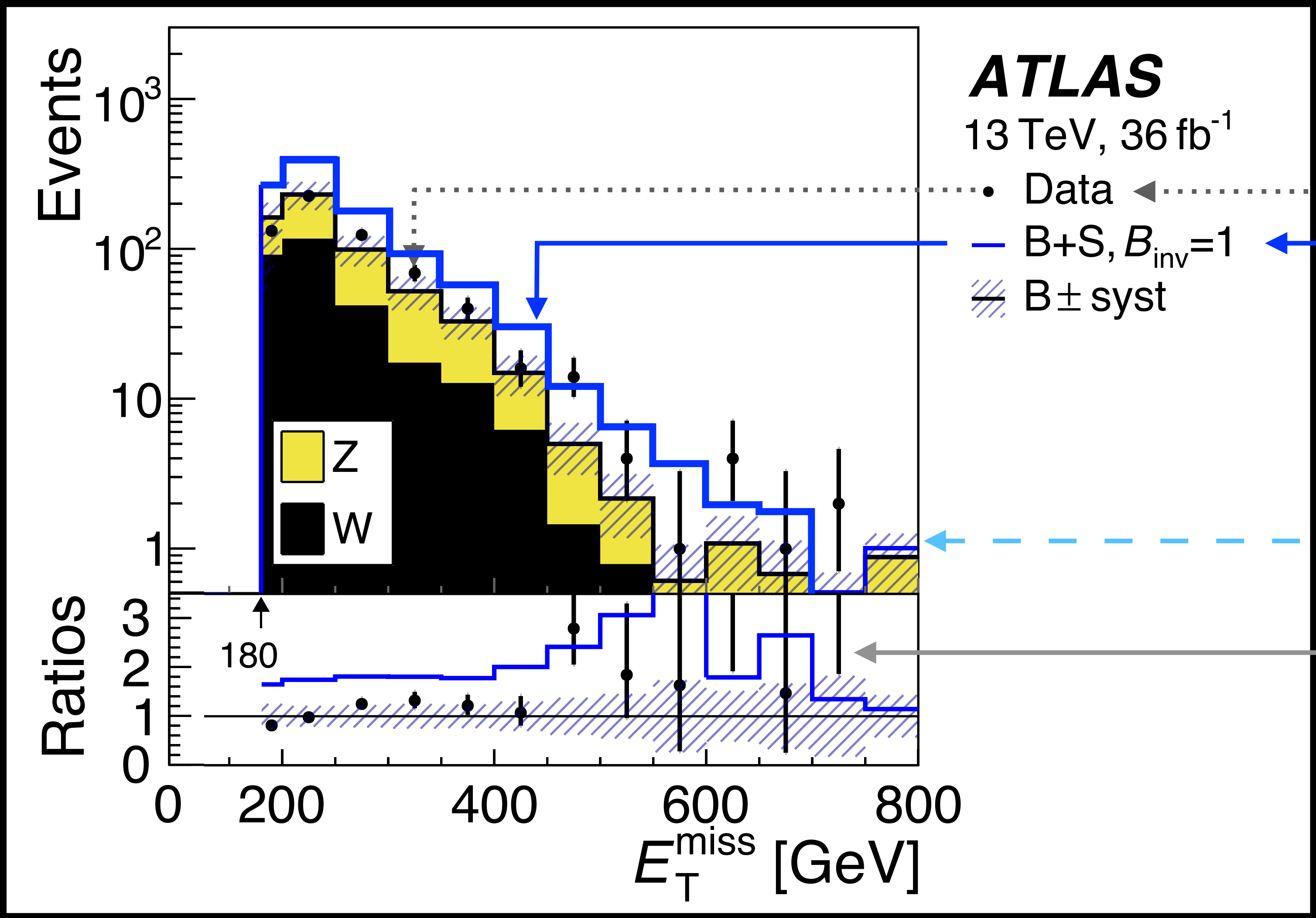


Results (2019)

Limits, systematics



ATLAS, Phys. Lett. B 793, 499 (2019)



Implied recoil

Data is consistent with no signal

If 100% of the Higgs decayed invisibly

Systematic errors

- simulation samples
- hadronic jet energy

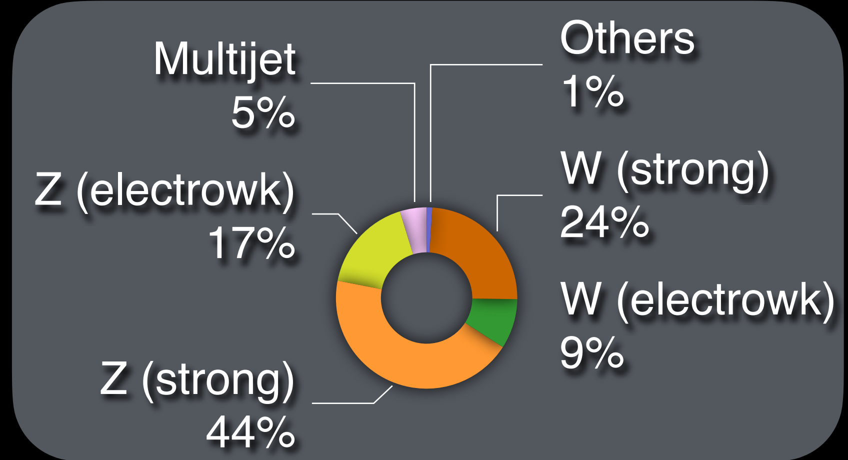
Statistical errors

- large for sensitive bins
- need more data

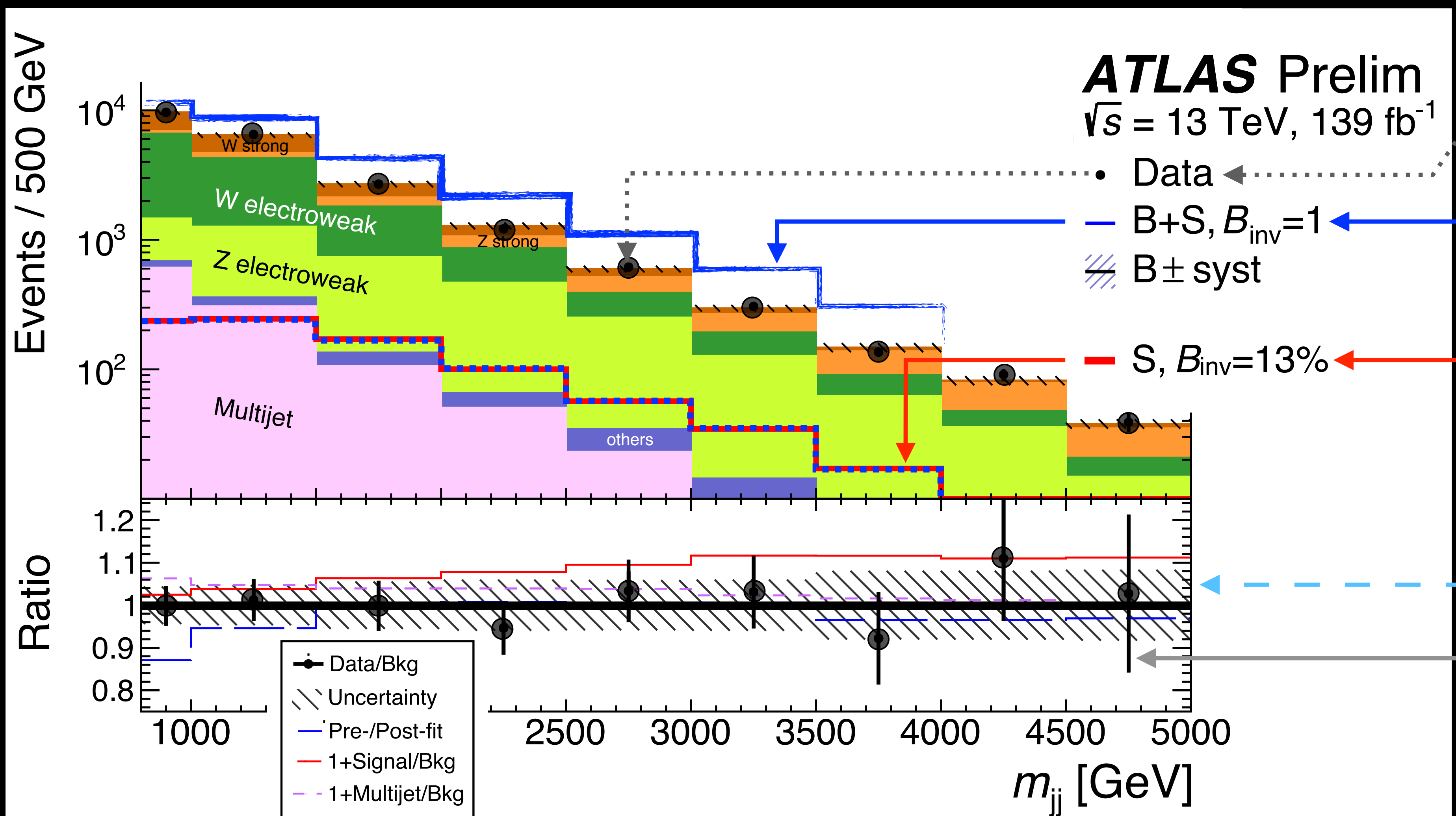
We can say that Higgs decays less than 37% to invisible final states

Update (2020)

Add 4x more data & improve methods



ATLAS, ATLAS-CONF-2020-008 (2020)



Two-jet invariant mass

Data is consistent with no signal

If 100% of the Higgs decayed invisibly

If 13% of the Higgs decayed invisibly

Systematic errors

- simulation samples 8%
- multijet estimation 7%
- jet energy 6%

Statistical errors

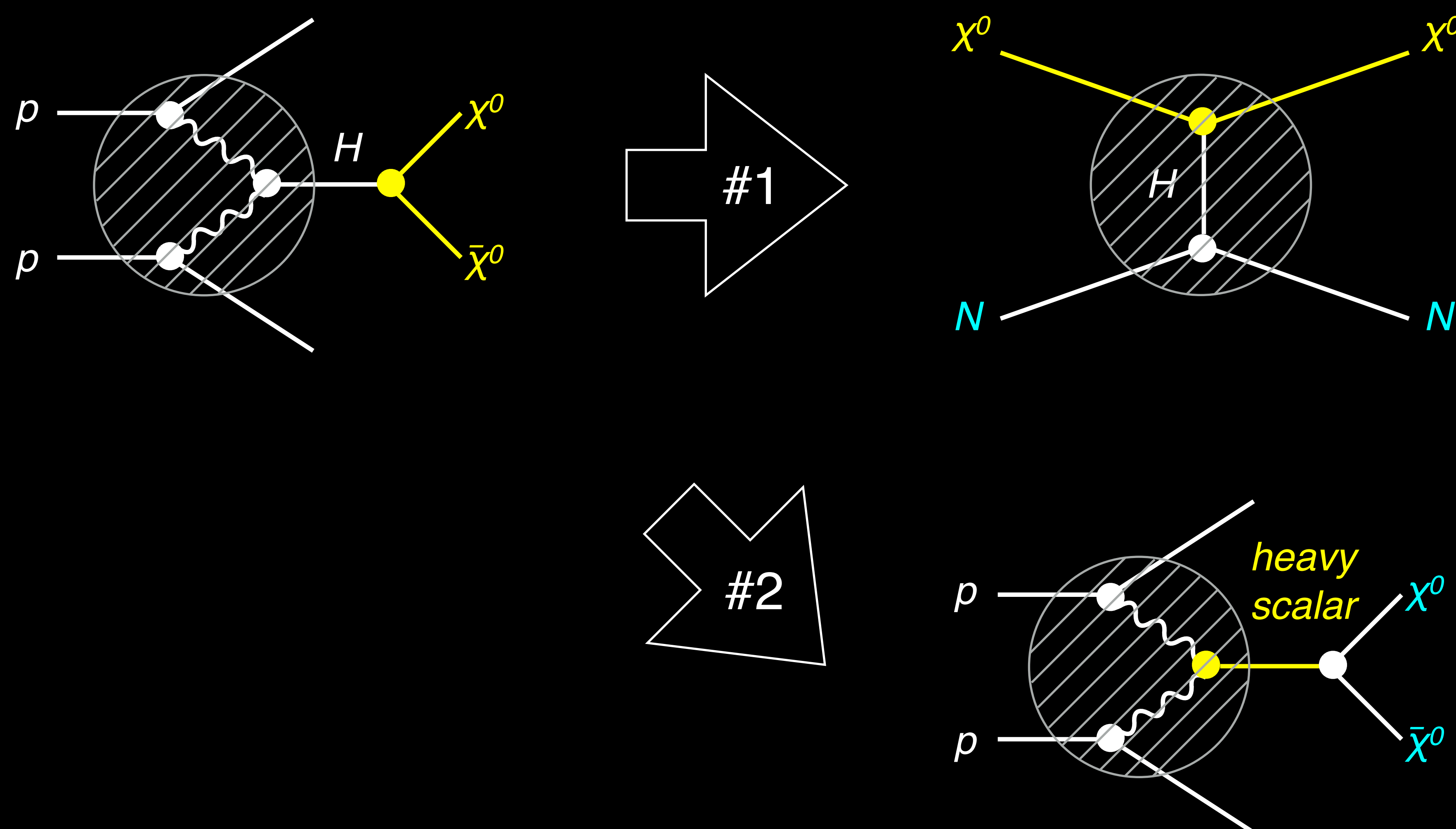
- large for sensitive bins
- need more data 17%

We can say that Higgs decays less than 13% to invisible final states

Interpretations

ATLAS result interpreted in #1, #2

ATLAS result | Interpretation | Quantity

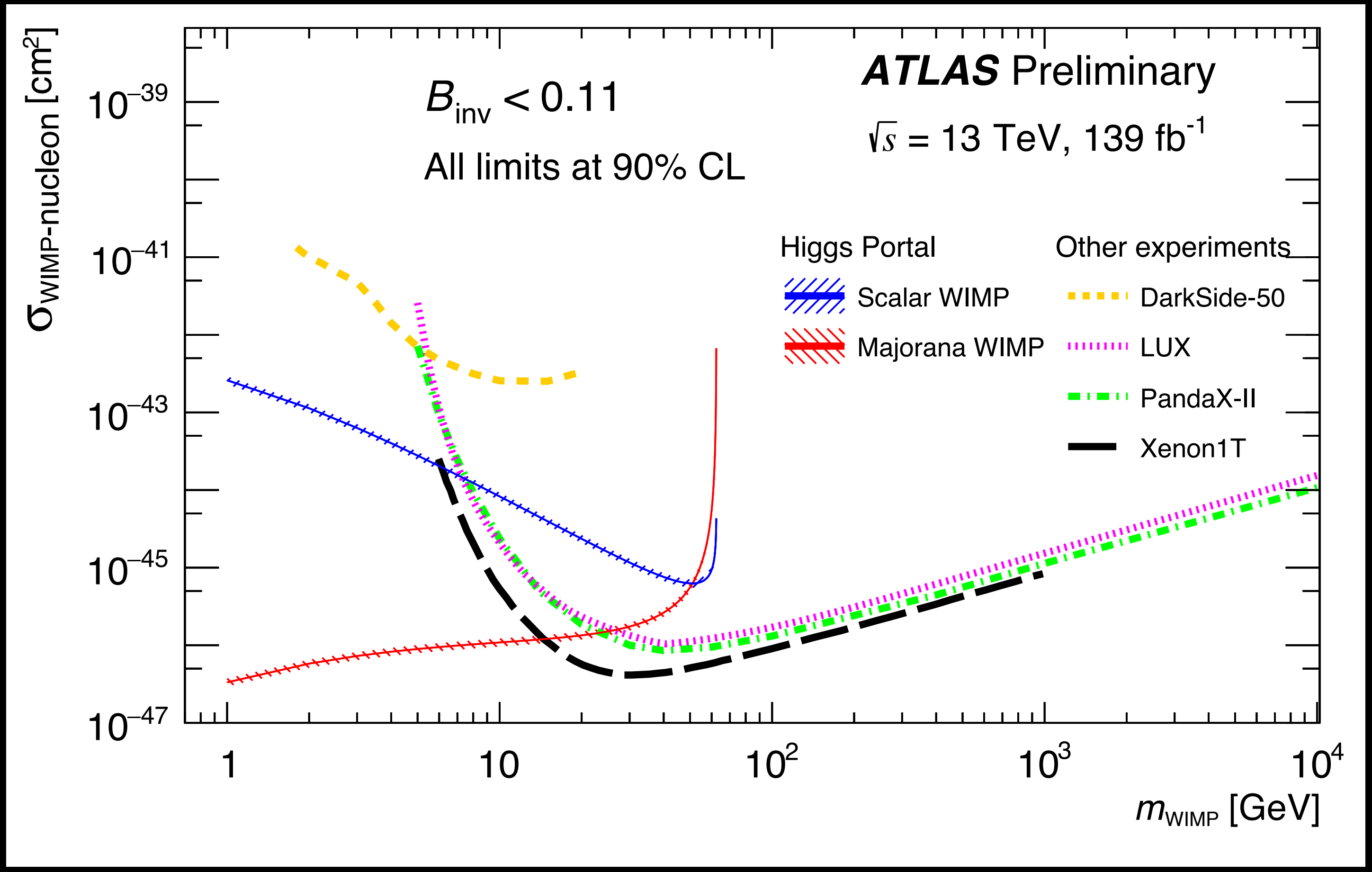


$= \sigma_{\mathrm{WIMP}}$

$= \sigma_{\mathrm{scalar}} \cdot B_{\mathrm{inv}}$

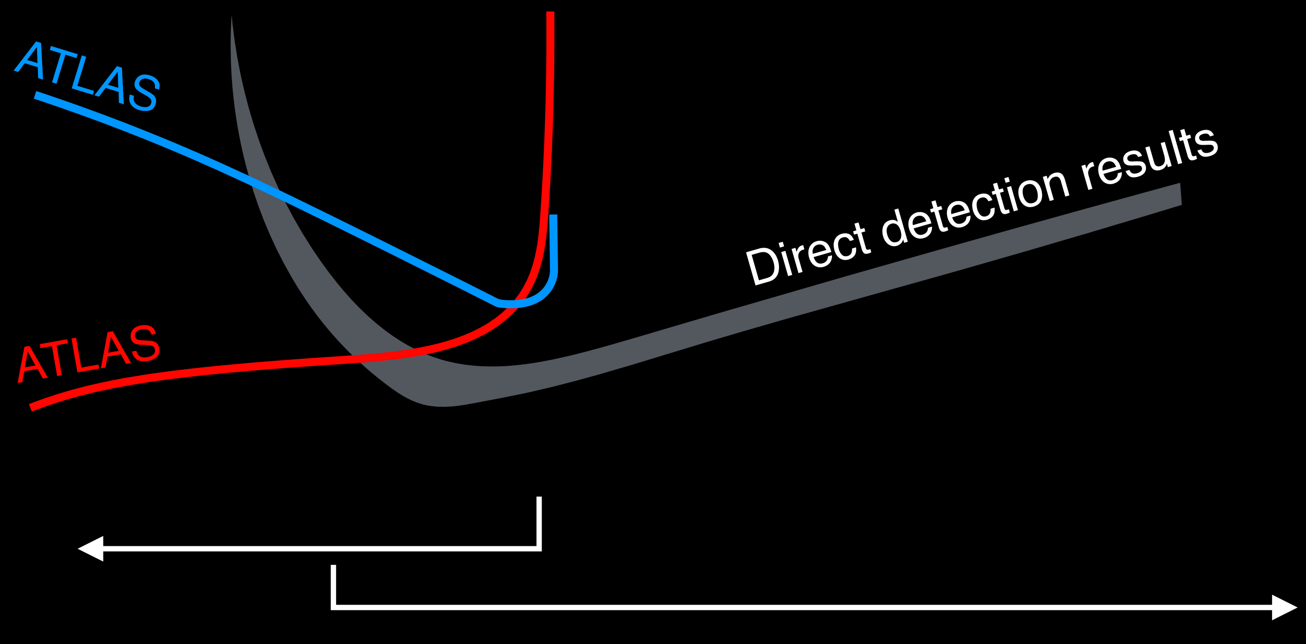
Interpretation #1

ATLAS result + Higgs portal



scalar: $\Gamma_{inv} \cdot (m_\chi)^{-2} \sim \sigma_{WIMP}$

fermion: $\Gamma_{inv} \cdot \text{const} \sim \sigma_{WIMP}$

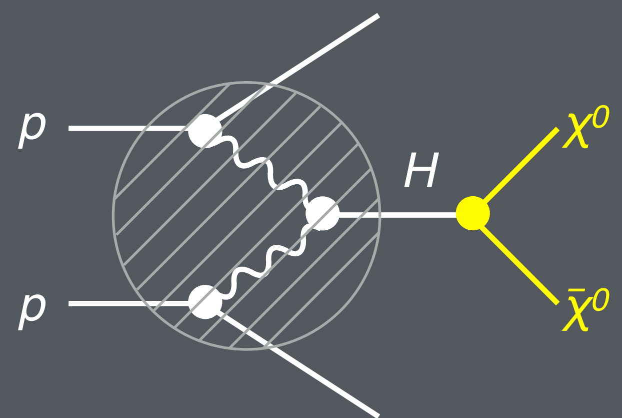


DM interpretation is complementary to direct detection

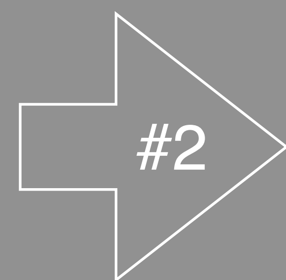
Interpretation #2

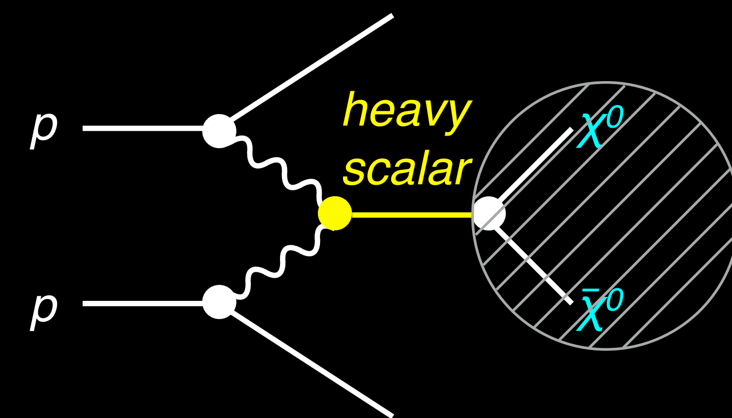
ATLAS result + "No model"

$$\sigma_H \cdot \frac{\Gamma_{inv}}{\Gamma_H}$$



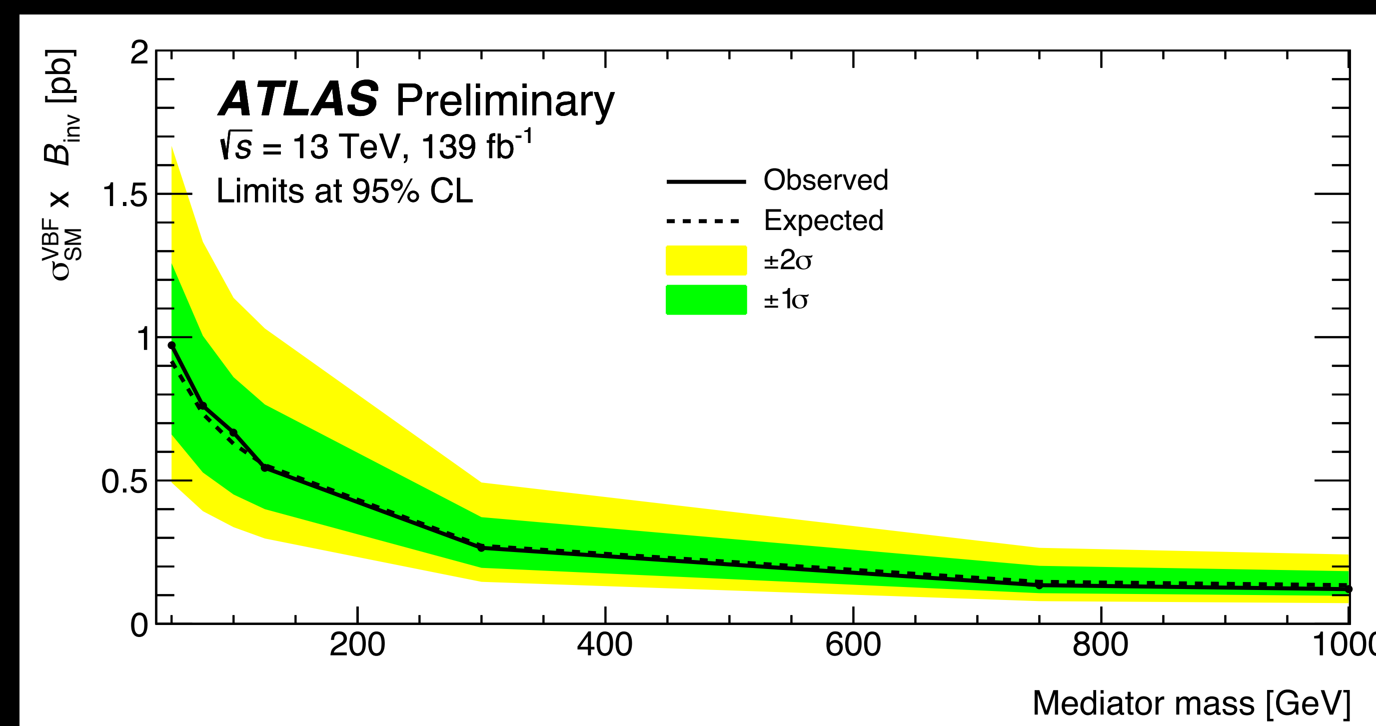
$$\cdot \frac{\varepsilon_{125}}{\varepsilon_{scalar}}$$





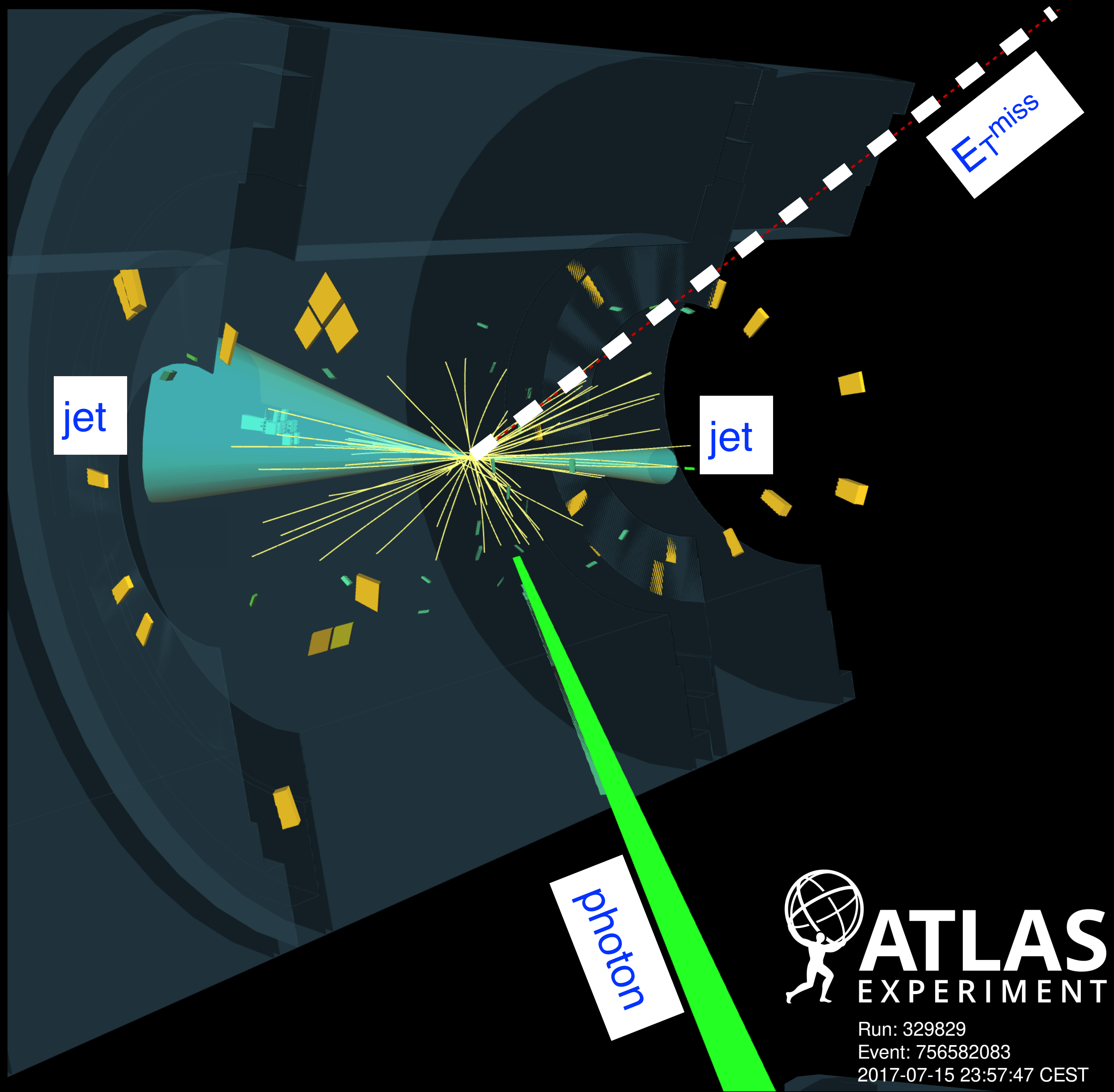
$$= \sigma_{scalar} \cdot \frac{\Gamma_{inv}}{\Gamma_H}$$

$$= \sigma_{scalar} \cdot B_{inv}$$



TeV scale particle limit at ~ ¼ pb

Add a photon to it

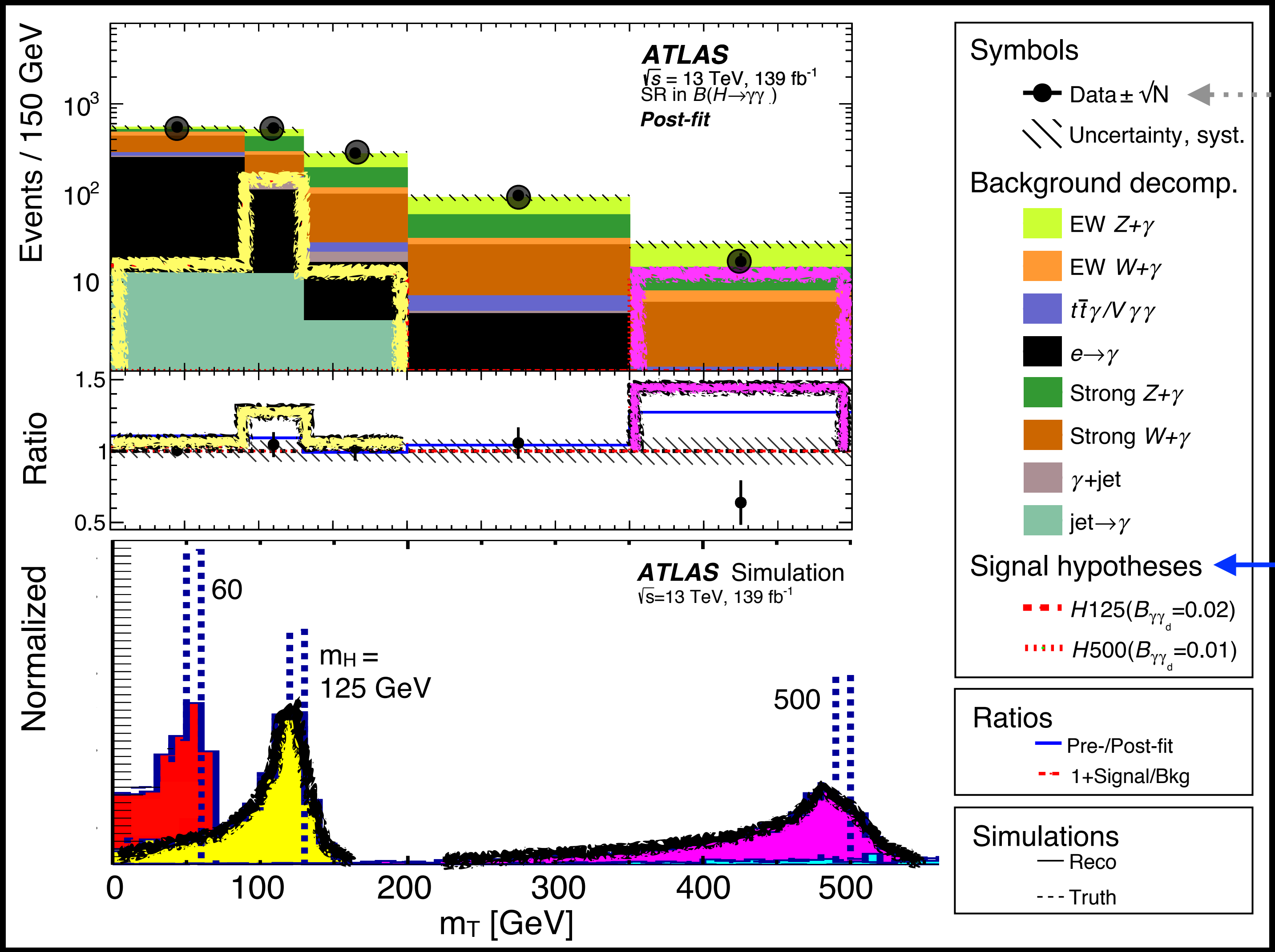


http://cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2021-004/fig_19.pdf

Interpretation

H_{125} to $\gamma\gamma_{dark}$

ATLAS, [arXiv:2109.00925], submitted to Eur. Phys. J. C



Data is consistent with no signal

- need more data

If Higgs decayed to dark photons

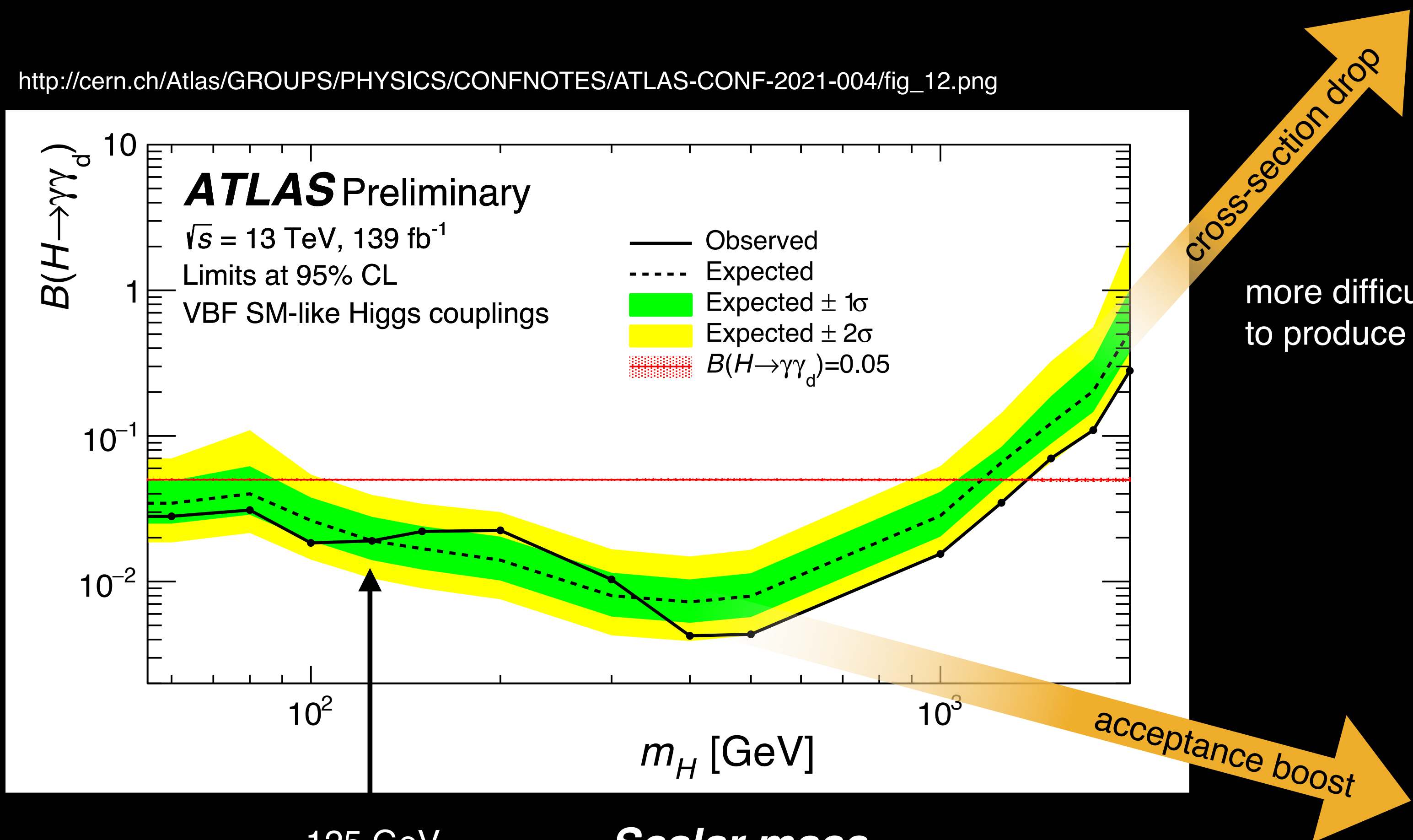
Observed limit B_{dark} of 0.014

Expected limit B_{dark} of 0.017 ± 0.006

Interpretation

Scalar particle to $\gamma\gamma_{dark}$

http://cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2021-004/fig_12.png



more difficult
to produce

harder γ
larger E_T^{miss}

125 GeV

Scalar mass

Exclude at few percent up to a few TeV

Trigger

E_T^{miss}, VBF
ML on FPGA

E_T^{miss} trigger rate vs. ⟨μ⟩

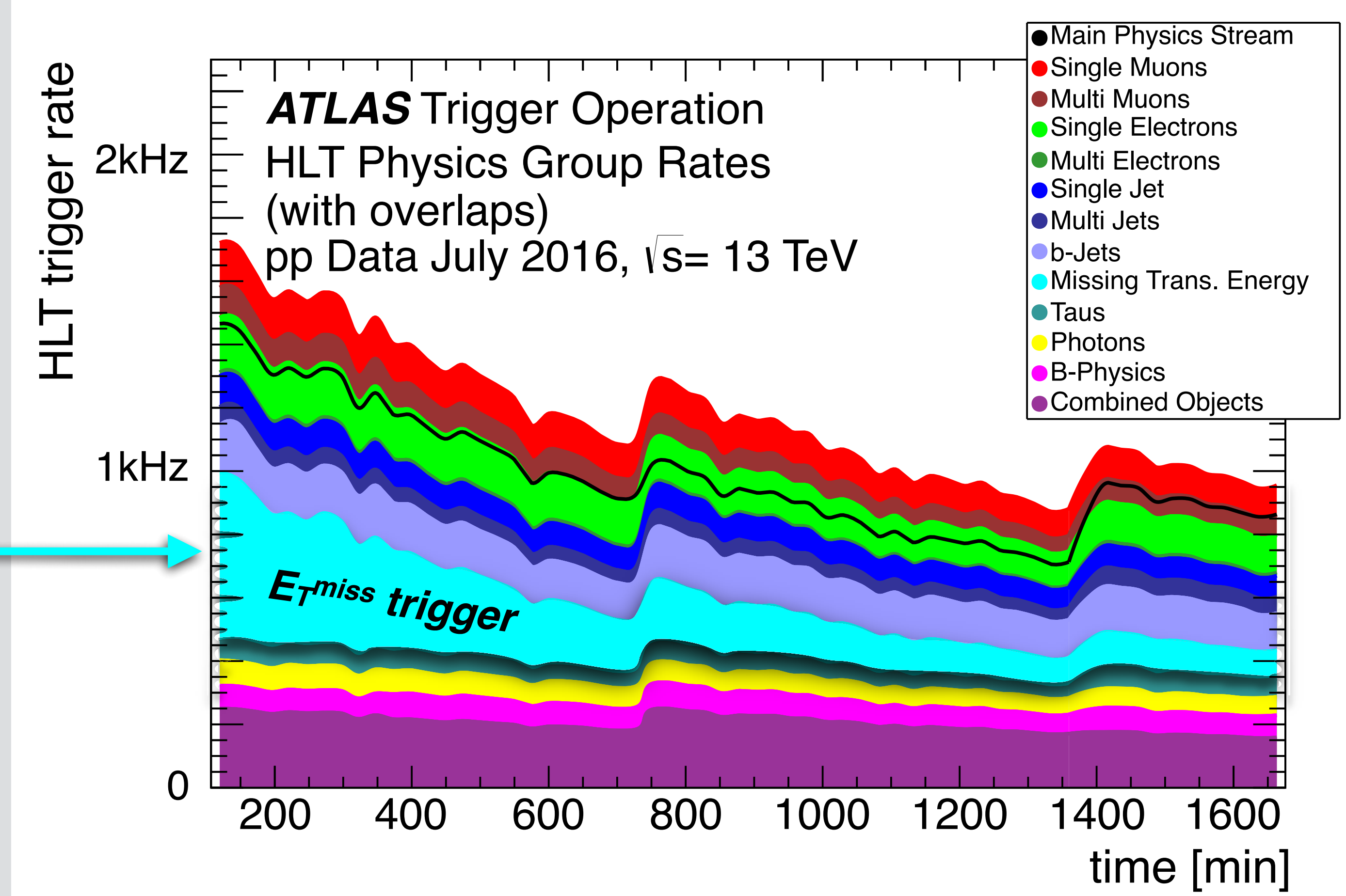
Rate blows-up with ⟨μ⟩, so very large at the beginning of runs

Physics

- Signal: Higgs p_T gives rise to E_T^{miss}
- Background: Combinatorics esp. vs. ⟨μ⟩

Challenge

- Problem: E_T^{miss} trigger is bandwidth limited
- Non-solution: Can't increase threshold to reduce the rate beyond ~150 GeV bec. signal peaks at low values
- Solution: Be smarter about reducing background while maintaining signal (sounds like physics analysis!)



E_T^{miss} is large fraction → E_T^{miss} is smaller fraction

Pileup lower as run goes on

NB. 150 → 180 GeV reduces signal by ~30%

E_T^{miss} trigger rate challenge

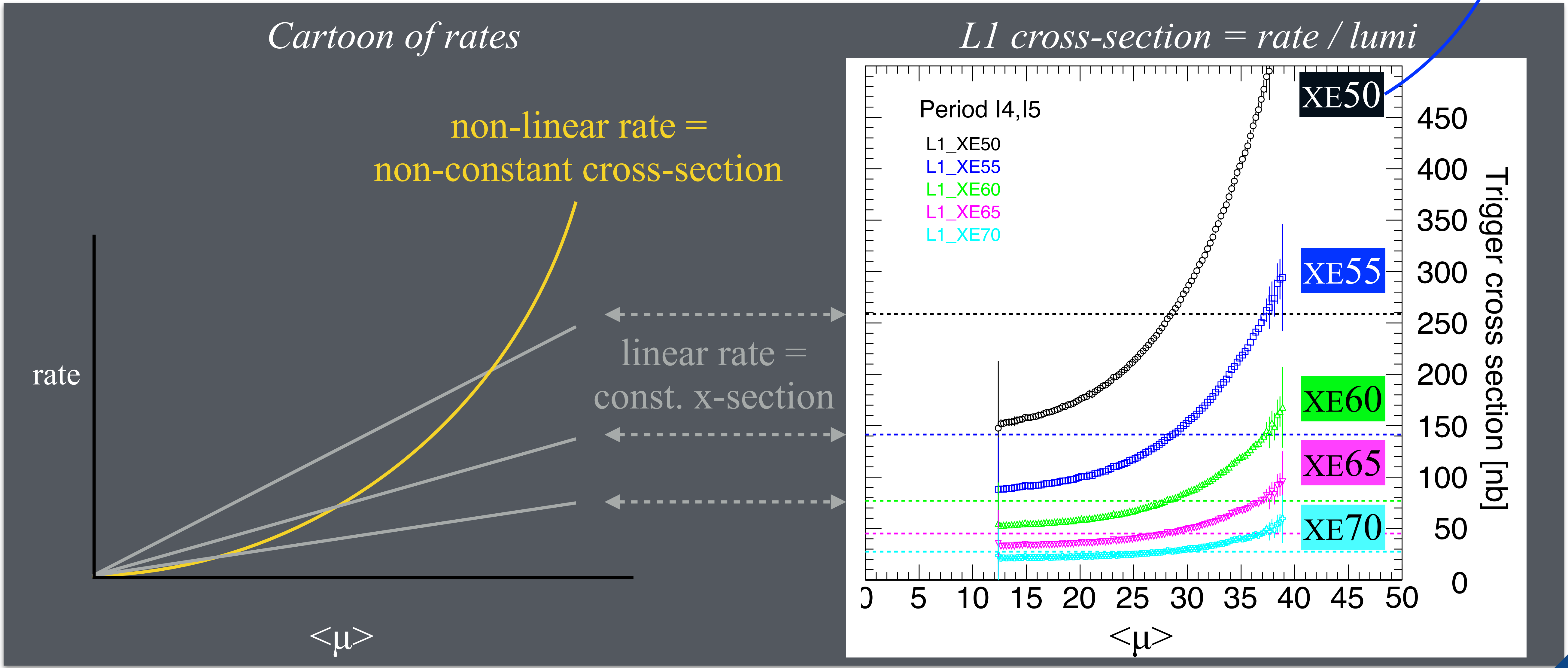
Lots of trouble with pileup driving up MET rates

- Conceptually, linear rate v. <μ> means "no pileup dependence," see left
- Rates show non-linear <μ> dependence, see right plot

Level-1 XE50 blowing up

Despite periods with very high L1 rates, we kept XE50

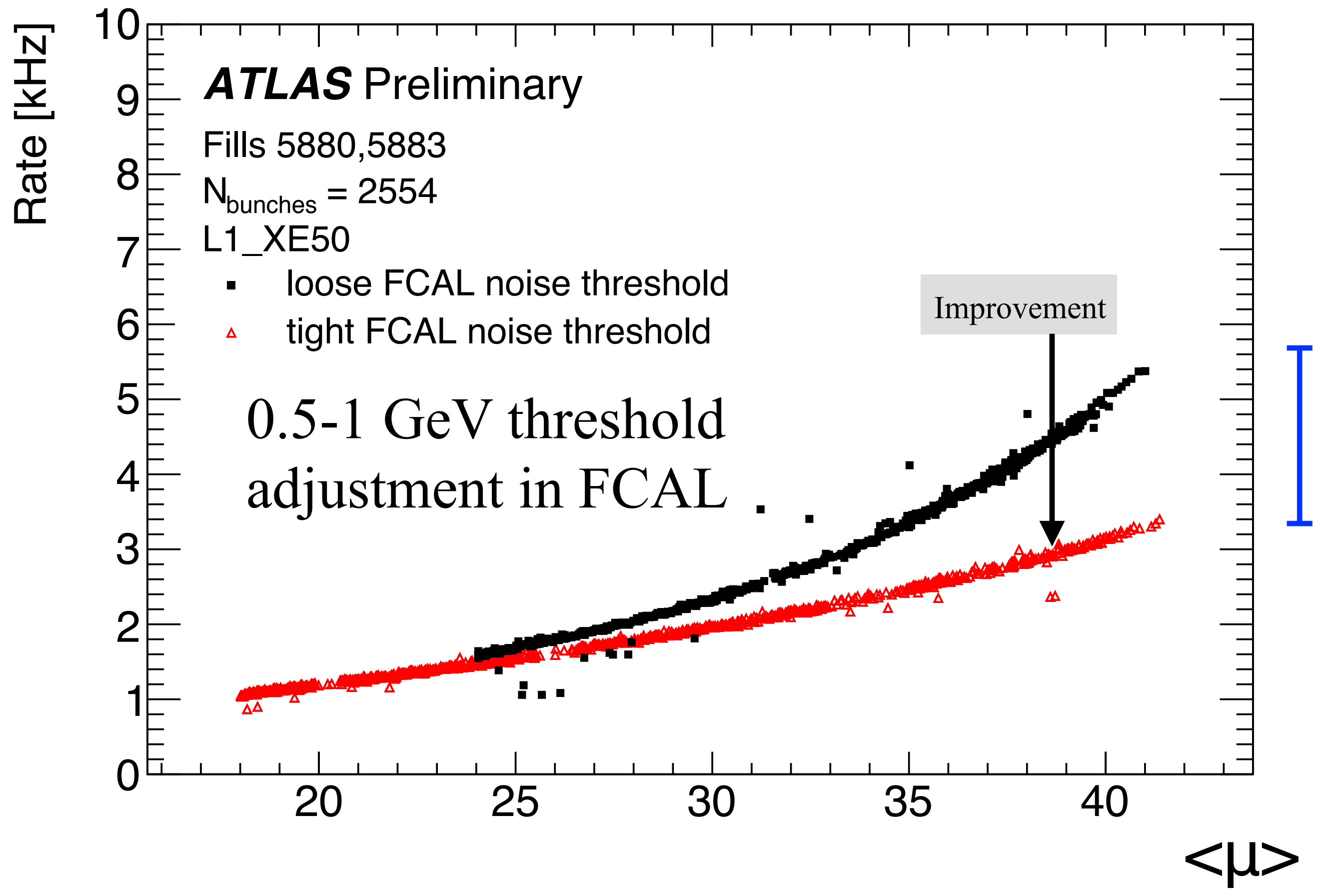
- The total rate did peak to almost 8 kHz a few times, but this was only at the start of



At level-1: Smarter noise cuts

Frequent adjustments to noise cuts control the rate

- Adjusted as the pileup increases or the filling scheme changed (right)
- The noise cuts were adjusted three times in 2017 and once in 2018
- Plot on the left shows the impact of the first change



Noise cuts to reduce L1 rate

Signal efficiency similar (before / after noise update with lower rates)

At HLT: Smarter algorithm (2016)

*mht*110 (default for post-CHEP 2016)

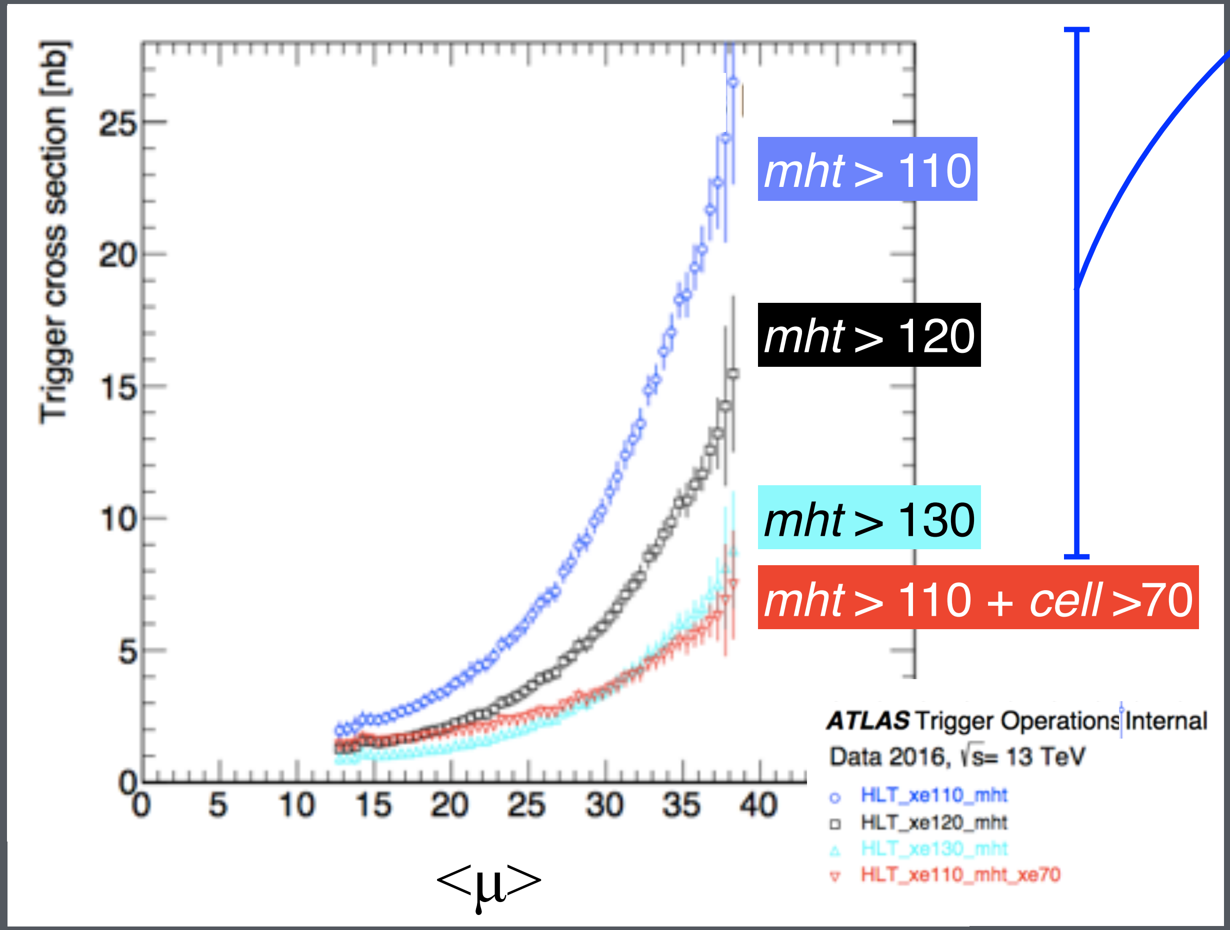
- Rate for μ > 45-50 too high, see left plot

Try combo to reduce rate in 2016

We found backup: *mht*110 + *cell*70*, but kept **mht110**

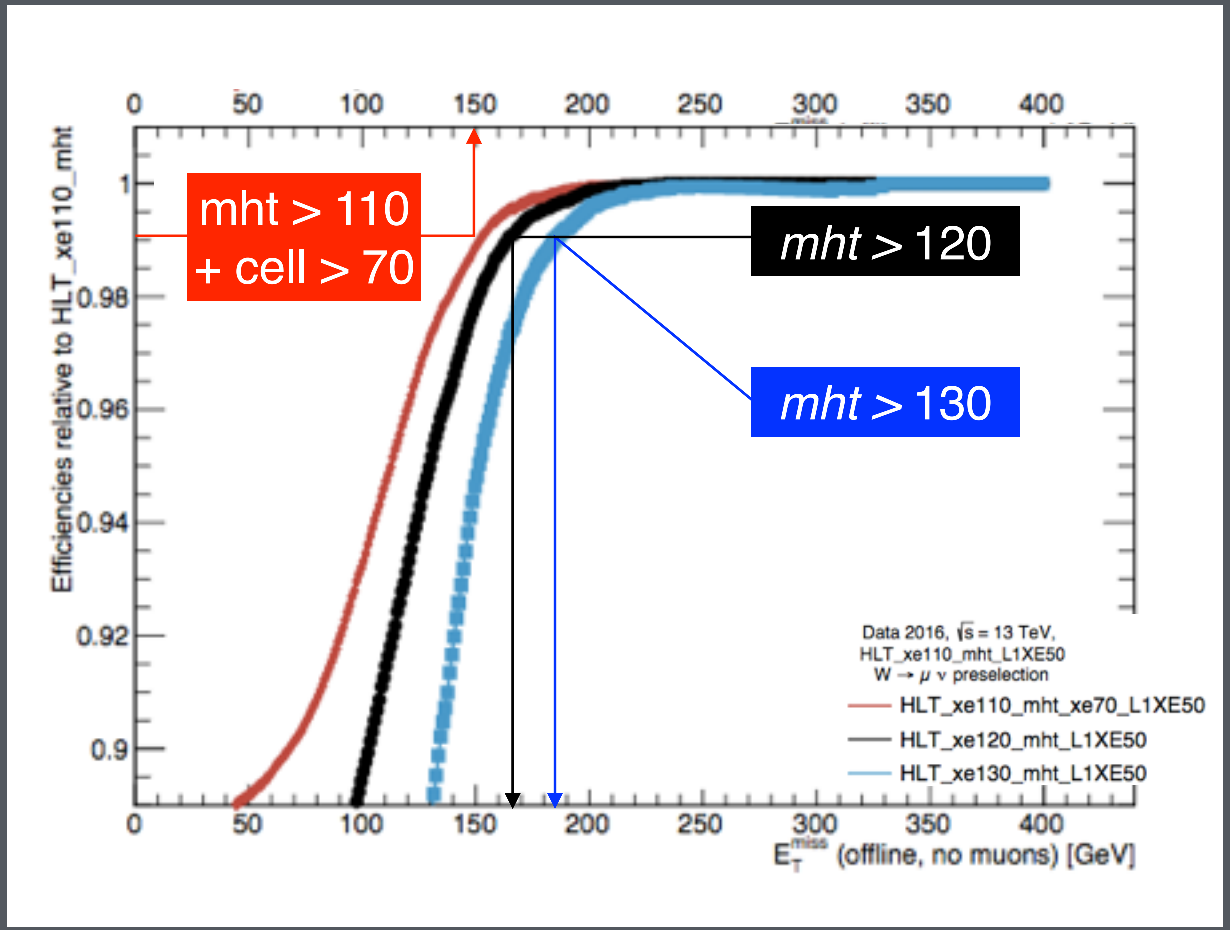
- The performance is much better when mht is combined with cell MET
- Efficiency is better compared to *mht*130, see right plot^

HLT cross-section = rate / lumi



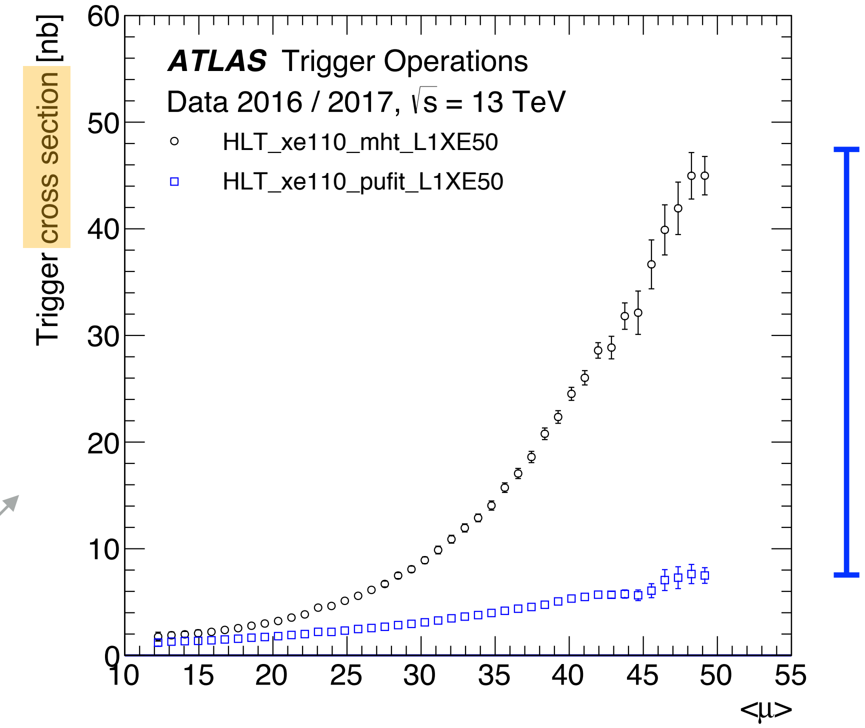
https://cds.cern.ch/record/2215888/files/ATL-COM-DAQ-2016-132.pdf

Efficiency relative to mht110



At HLT: Smarter (2017)

χ^2 based "pileup fit" algorithm

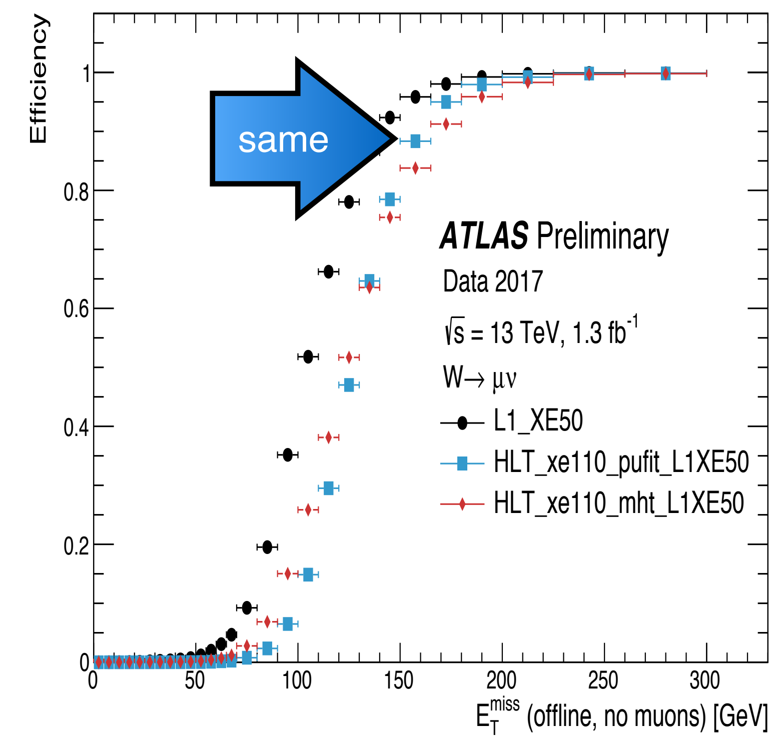


Algorithm

- Divide η-ϕ space in ~0.4^2 grid
- Assume uniform underlying pileup energy in η-ϕ, float magnitude given momentum conservation in xy

Result

- Trigger rate drastically reduced
- Signal efficiency is similar



Algorithm & threshold evolution

- Rapid development

Year	Trigger name	HLT algorithm [GeV]	L1 threshold [GeV]	HLT threshold [GeV]	$\int \mathcal{L}\, dt$ [fb^{-1}]
2015	HLT_xe70_mht_L1XE50	mht	50	70	3.5
2016	HLT_xe90_mht_L1XE50	mht	50	90	12.7
2016	HLT_xe110_mht_L1XE50	mht	50	110	30.0
2017	HLT_xe90_pufit_L1XE50	pufit, cell	50	90, 50	21.8

Rapid algorithm R&D to retain E_T^{miss} threshold

At HLT: Combo (2018)

Use combinations of algorithms

Algorithm

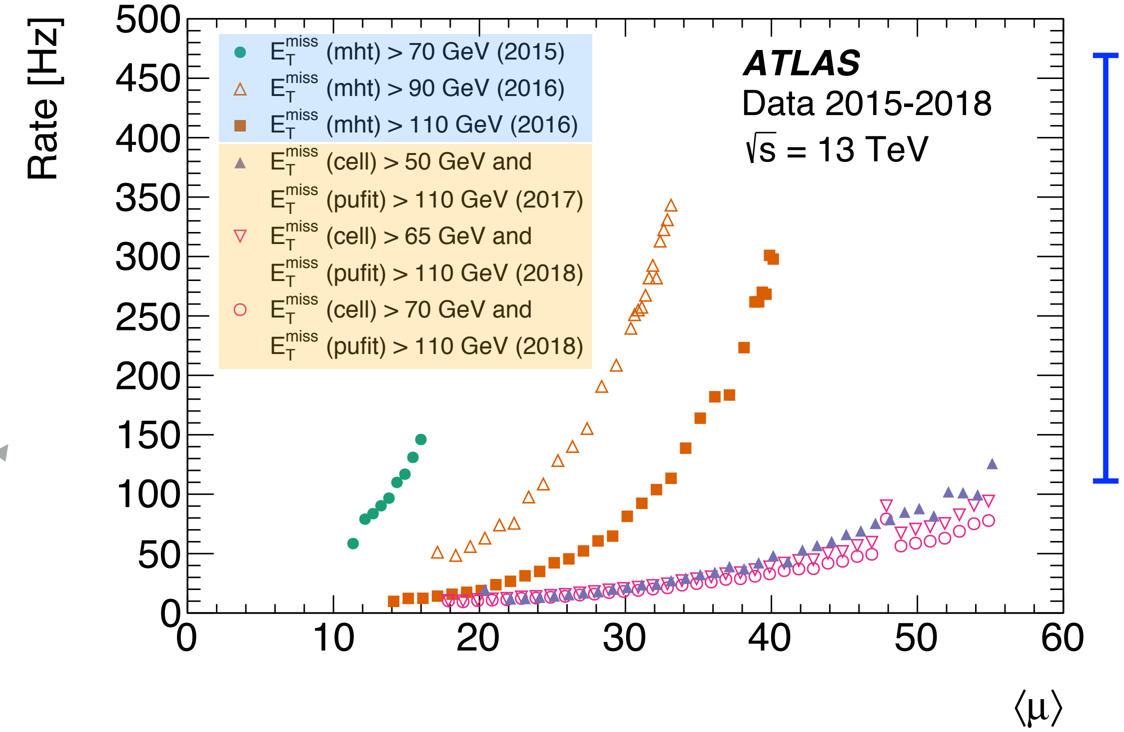
- χ^2 based algorithm from prev. slide
- Cell-based algorithm using ~200k LAr cells
- Use both algorithms!

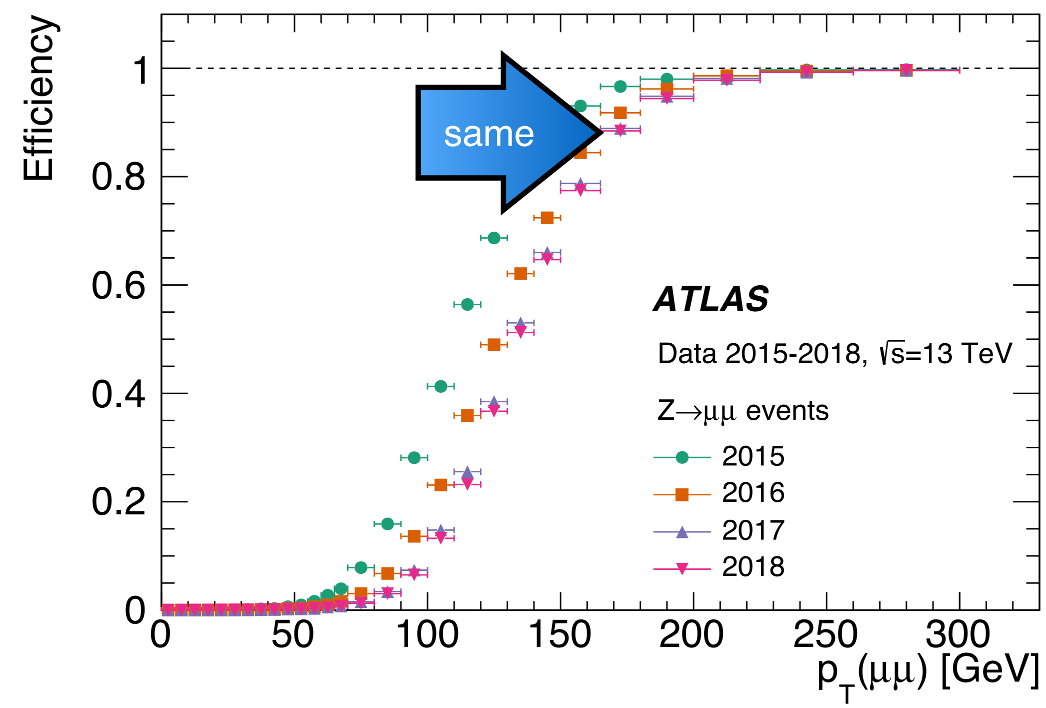
Result

- Trigger rate drastically reduced
- Signal efficiency is similar

Algorithm & threshold evolution

- Rapid development





Year	Trigger name	HLT algorithm [GeV]	L1 threshold [GeV]	HLT threshold [GeV]	$\int \mathcal{L}\, dt$ [fb^{-1}]
2017	HLT_xe100_pufit_L1XE50	pufit, cell	50	100, 50	33.0
2017	HLT_xe110_pufit_L1XE50(55)	pufit, cell	50 (55)	110, 50	47.7
2018	HLT_xe110_pufit_xe65_L1XE50	pufit, cell	50	110, 65	57.0
2018	HLT_xe110_pufit_xe70_L1XE50	pufit, cell	50	110, 70	62.6

Rapid algorithm R&D to retain E_T^{miss} threshold

E_T^{miss} trigger

Summary of the Run 2 history & my outlook on Run 3

Approach

- Start data taking → higher pileup than before → E_T^{miss} rates are too high → threaten to raise thresholds → Develop & deploy clever solutions
- Repeat every few months

Obvious question (& answer)

- Why not pre-develop in advance? Rates are notoriously difficult to simulate

My view for Run 3

- Keep a similar theme of innovating on algorithms, combining algorithms as we did in HLT
- May want to use non-E_T^{miss} triggers for the VBF + E_T^{miss} (+ soft) analyses

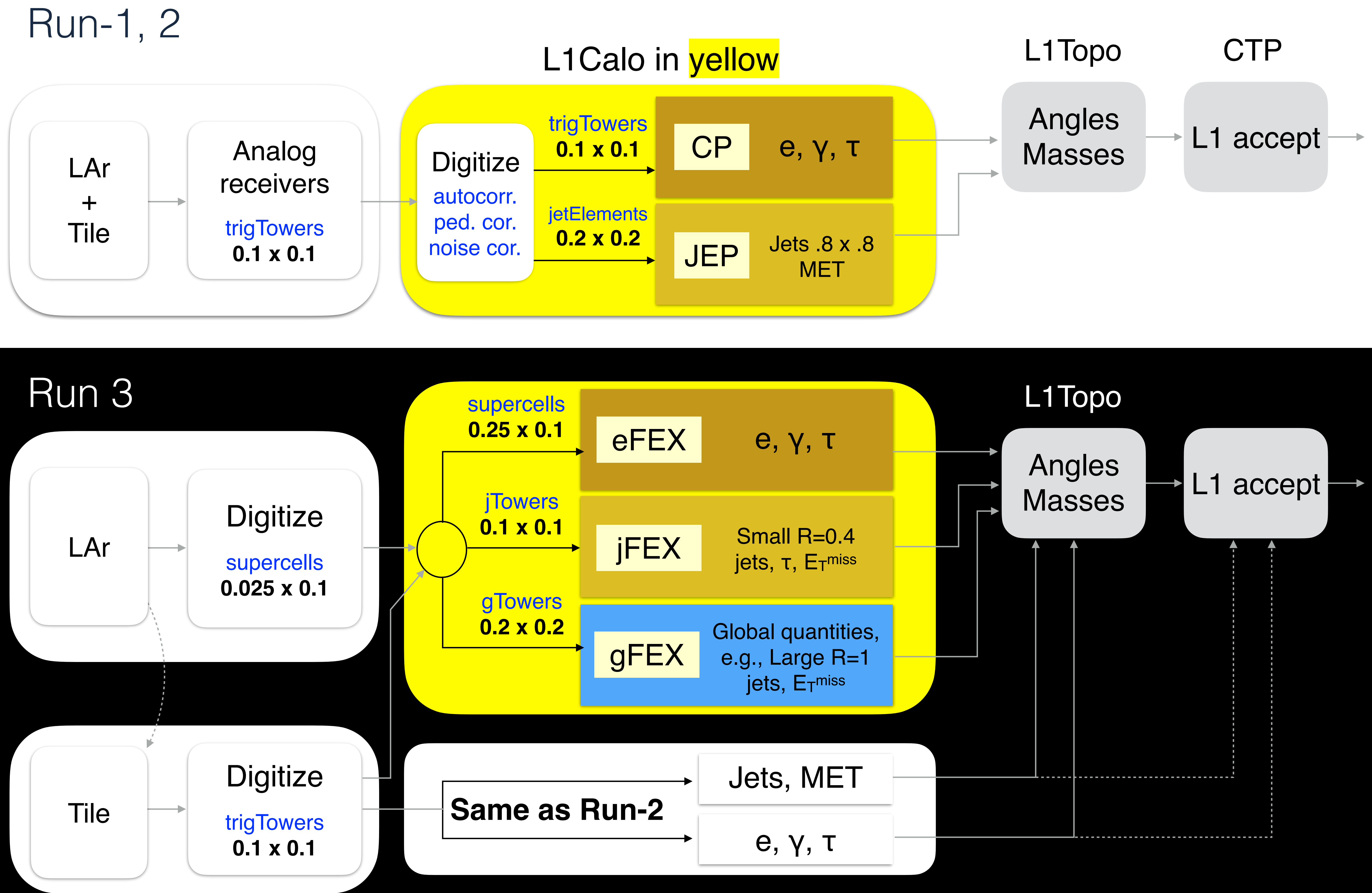
VBF + E_T^{miss} (+ soft) analyses

VBF + E_T^{miss} (+ soft) analyses

VBF + E_T^{miss} (+ soft) analyses

Rapid algorithm R&D to retain E_T^{miss} threshold

Upgrade of the level-1 architecture

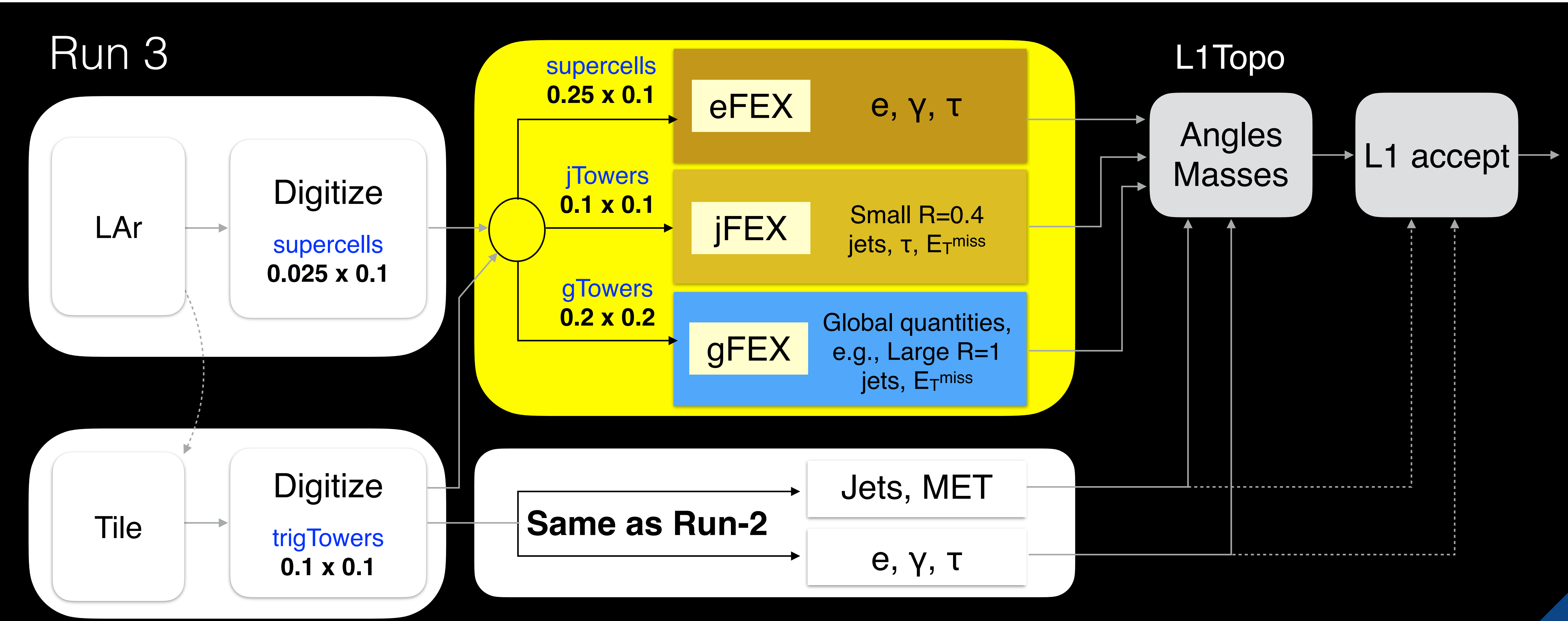


Run 3 architecture

My guess

More speculative

- We'll start with baseline E_T^{miss} algorithms in jFEX-gFEX
- We'll take data and probably realize that we need to do better than baseline
- We'll probably improve & add jFEX-gFEX algorithms (like we did before in HLT)
- We'll combine jFEX-gFEX outputs (like we did before in HLT) → use ML?



Source: L1Calo TDR, https://cds.cern.ch/record/1602235. This diagram courtesy B. Carlson.

ML on FPGA

Boosted decision trees

http://fwx.pitt.edu

JINST

Published by IOP Publishing for Sissa Medialab

Received: *April 9, 2021*
Accepted: *June 29, 2021*
Published: *August 4, 2021*

Nanosecond machine learning event classification with boosted decision trees in FPGA for high energy physics

T.M. Hong,* B.T. Carlson, B.R. Eubanks, S.T. Racz, S.T. Roche, J. Stelzer and D.C. Stumpp

Department of Physics and Astronomy, University of Pittsburgh, 100 Allen Hall, 3941 O'Hara St., Pittsburgh, PA 15260, U.S.A.

E-mail: tmhong@pitt.edu

Abstract: We present a novel implementation of classification using the machine learning/artificial intelligence method called boosted decision trees (BDT) on field programmable gate arrays (FPGA). The firmware implementation of binary classification requiring 100 training trees with a maximum depth of 4 using four input variables gives a latency value of about 10 ns, independent of the clock speed from 100 to 320 MHz in our setup. The low timing values are achieved by restructuring the BDT layout and reconfiguring its parameters. The FPGA resource utilization is also kept low at a range from 0.01% to 0.2% in our setup. A software package called fwXmachina achieves this implementation. Our intended user is an expert in custom electronics-based trigger systems in high energy physics experiments or anyone that needs decisions at the lowest latency values for real-time event classification. Two problems from high energy physics are considered, in the separation of electrons vs. photons and in the selection of vector boson fusion-produced Higgs bosons vs. the rejection of the multijet processes.

Keywords: Digital electronic circuits; Trigger algorithms; Trigger concepts and systems (hardware and software); Data reduction methods

ArXiv ePrint: 2104.03408

*Corresponding author.

https://doi.org/10.1088/1748-0221/16/08/P08016

Machine learning at L1 trigger

TM Hong

FW X Machina

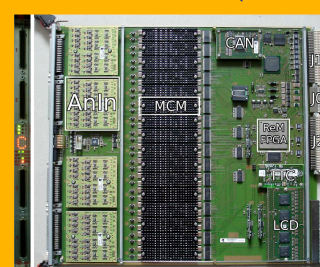
65 TB/s

Partial data

...
99
...
0

Level	*Latency*
L1	**O(1) µs**
HLT	O(1) s

FPGA



Custom electronics

Reduce rate to
100 kHz

yes / no

160 GB/s

50k CPUs

Software

Reduce rate to
1-10 kHz

yes / no

1.5 GB/s

Isn't there already a package? (yes, now there are 2)

TM Hong

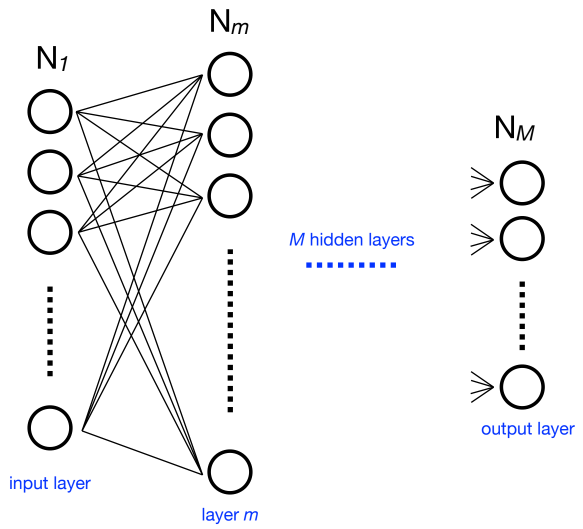
Deep Neural Network

Duarte et al., J. Instrum. 13, P07027 (2018)

- Popular method for signal vs. background
- But can't be very deep for FPGA, so ~3 "deep" in paper
- $y = \Theta(\mathbb{M} \cdot x + b)$

Fancy activation

Multiplication (limited resource on FPGA)



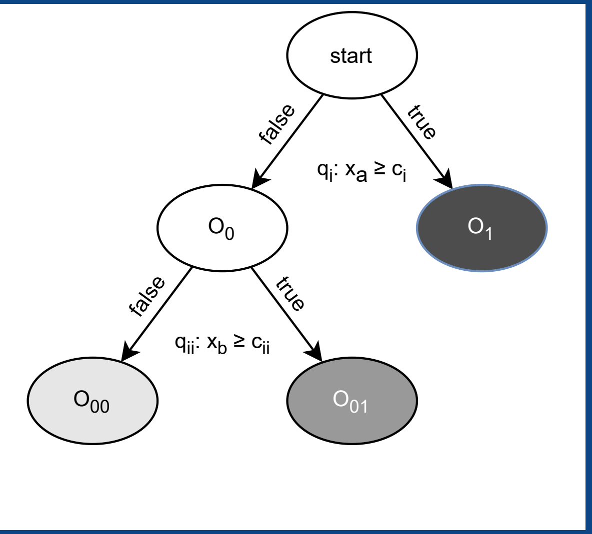
Standard Decision Tree

- Another popular classification
- $y = \Theta(x < \text{threshold})$

Boolean

No multiplication (bin search problem)

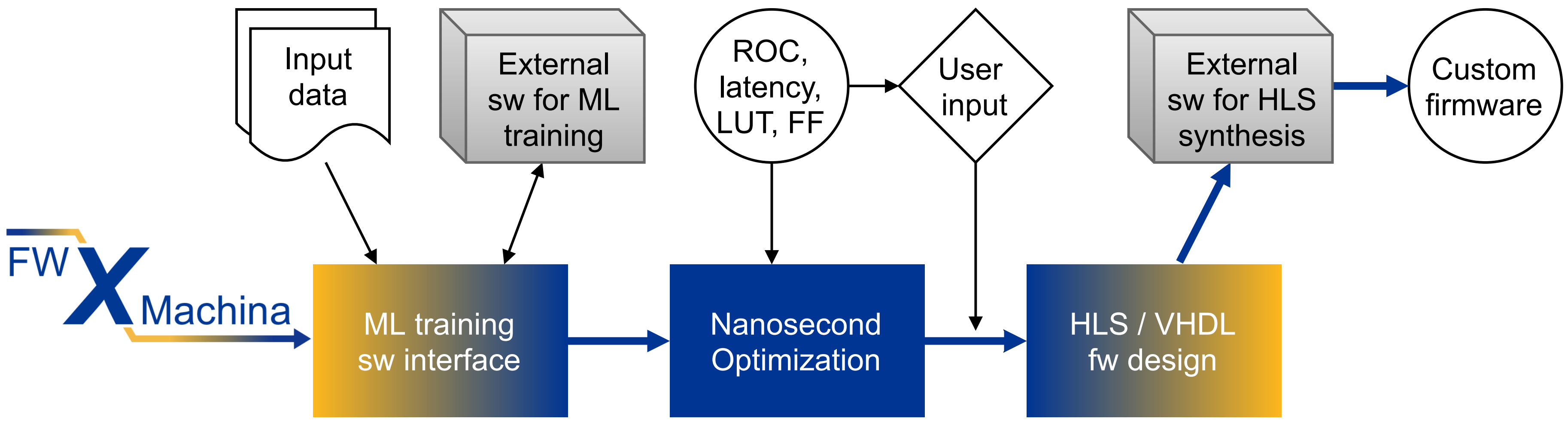
Hong et al., J. Instrum. 16, P08016 (2021)



Decision Tree for FPGA

- Smart bit integer precision, bit shifting
- Flattened (also "deep")
- All variables processed in parallel
- One step algorithm, ns fast, tiny footprint

- **Workflow**

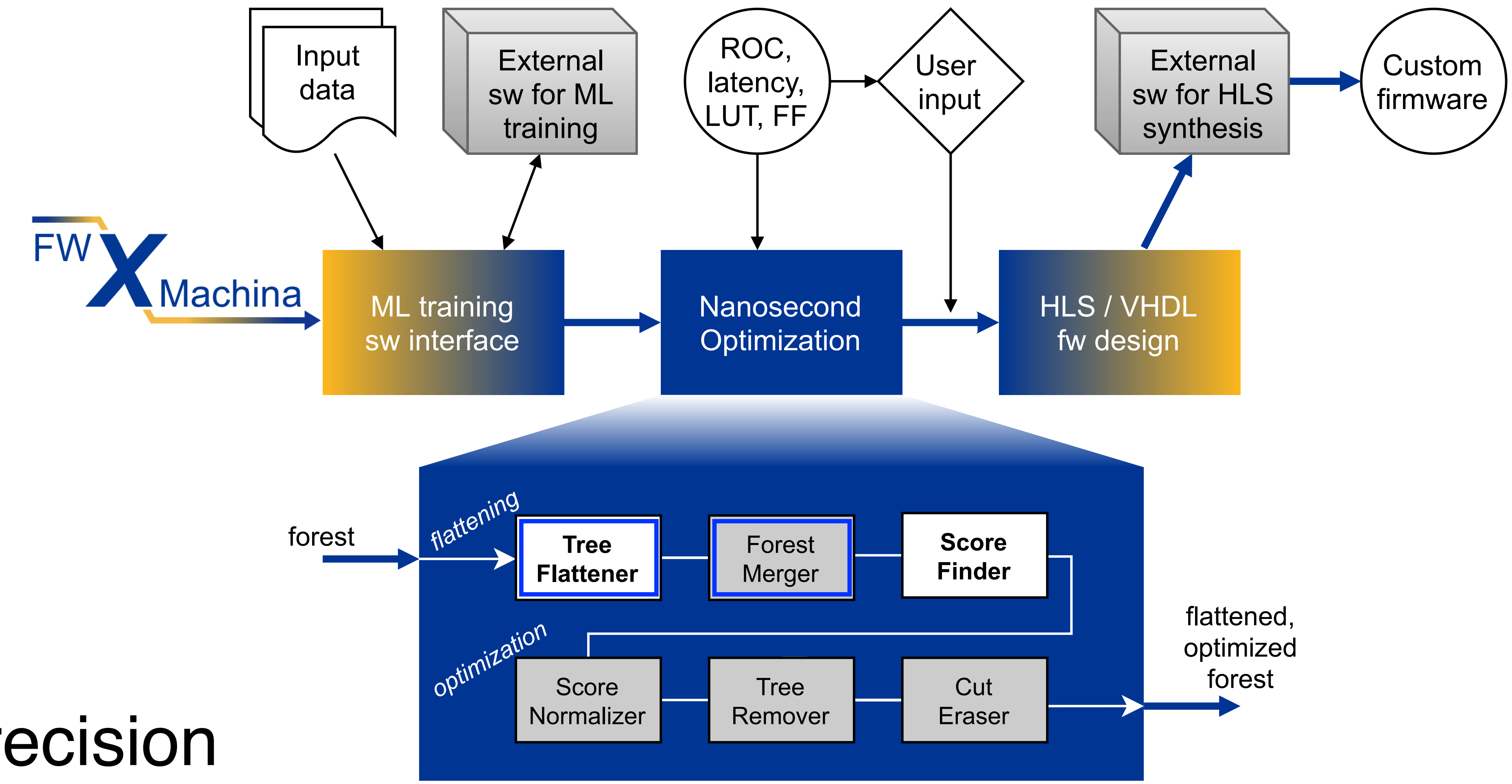


- Optimization
- Use bit integer precision

structure (2)

TM Hong

- Workflow
- **Optimization**
- Use bit integer precision



- Will discuss next:

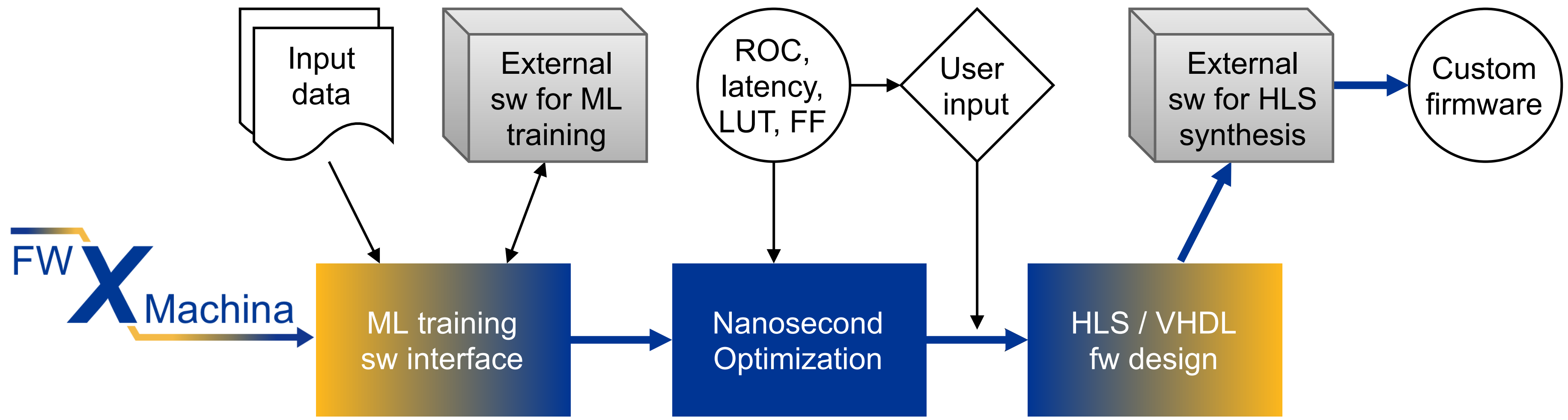
  Tree Flattener

  Forest Merger

structure (3)

TM Hong

- Workflow



- Optimization

- **Use bit integer precision**

  E.g., ap_int⟨8⟩ means the variable is represented by a range from 0 to 255.

- Advantages & subtleties

  Bit integers represents a wide range without sacrificing float precision

Pre-evaluate f ***Firmware only adds***

Transformation $$c_{\mathrm{int}} = f(c_{\mathrm{float}}) = \left\lfloor \frac{c_{\mathrm{float}} - c_{\min}}{c_{\max} - c_{\min}} \cdot \left(2^N - 1\right) \right\rfloor$$

$$f(x_1 + x_2) = f(x_1) + f(x_2)$$

Floor operation

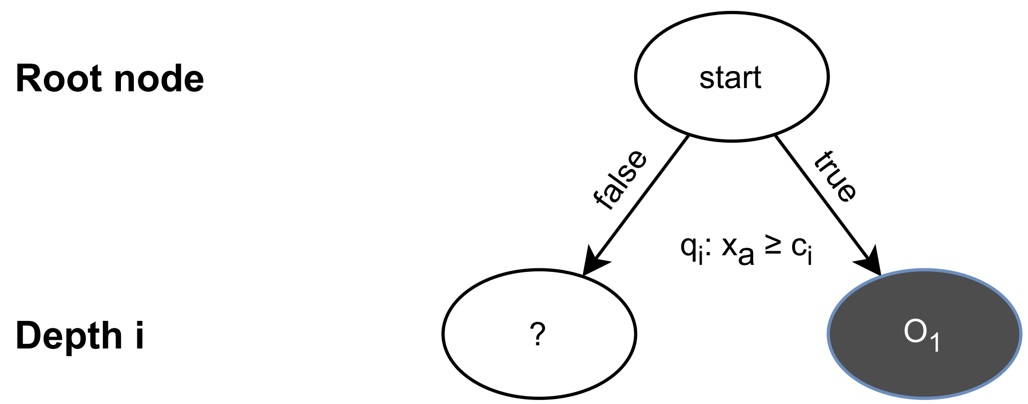
Equal up to one bit because of floor

Decision tree, 2 var example

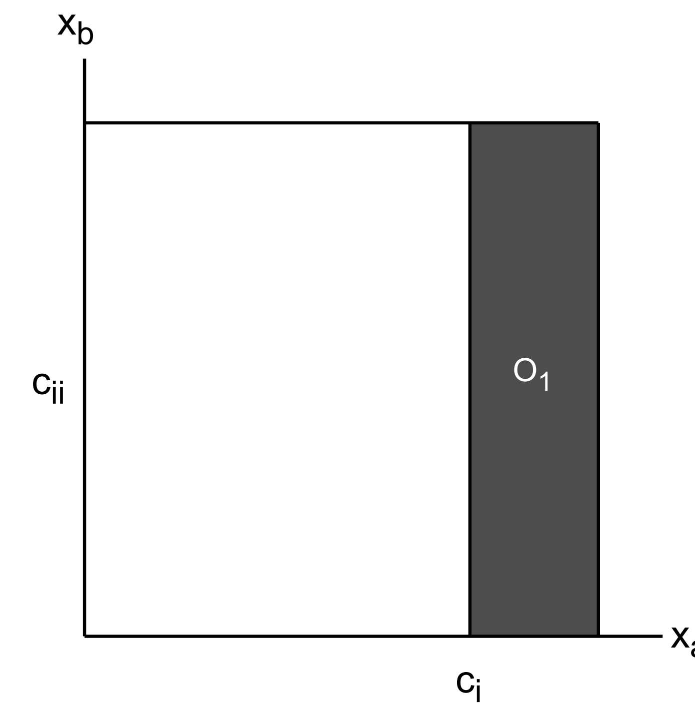
TM Hong

First step

Conventional tree structure



2d plane: x_a vs. x_b



- Advantages & subtleties
  - Cut thresholds & weights determined during training
  - Danger of "memorizing" boundaries (overtraining), so must consider a forest

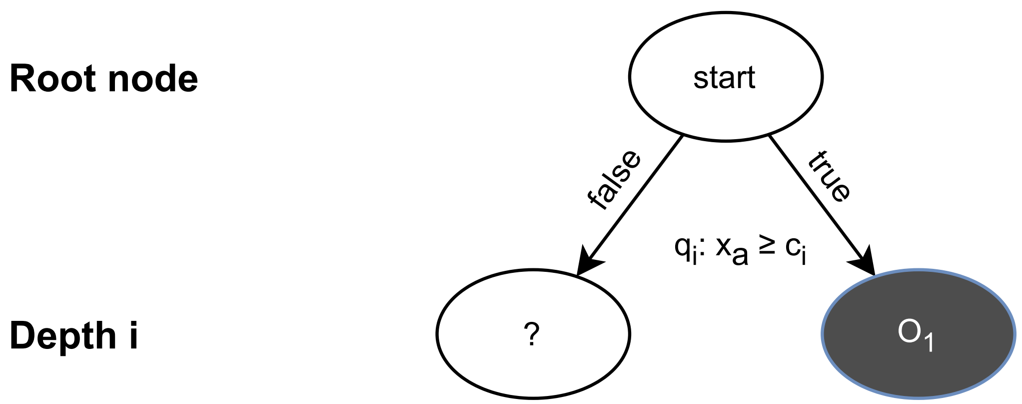
Decision tree, 2 var example (2)

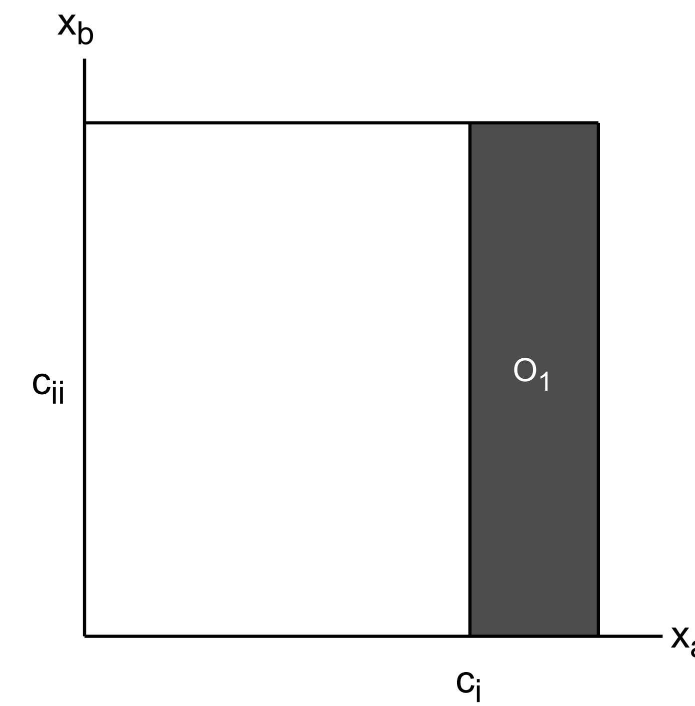
TM Hong

Conventional tree structure

2d plane: x_a vs. x_b

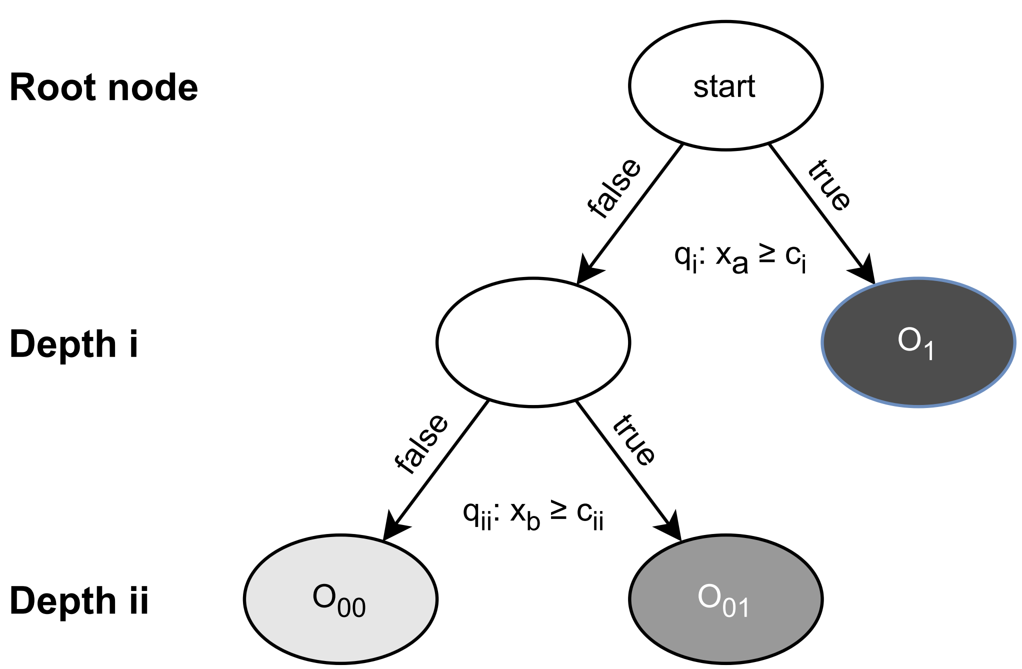
First step

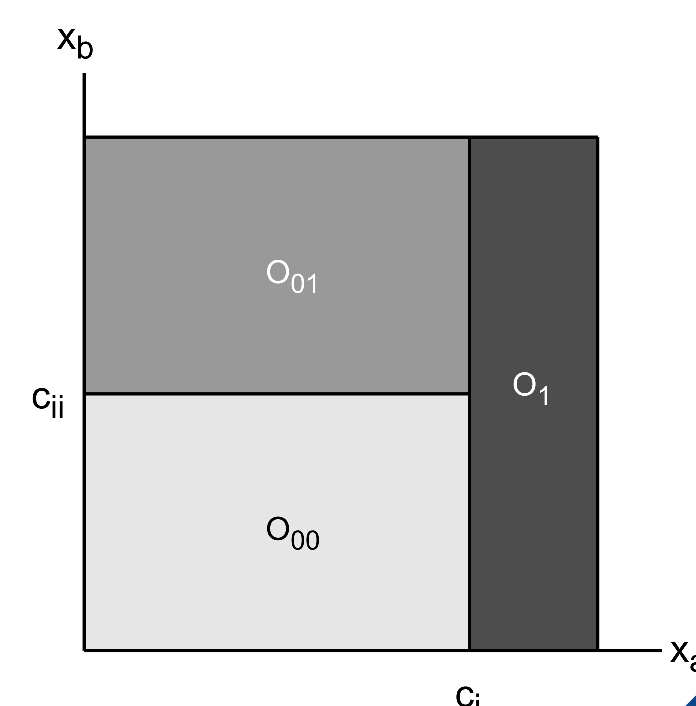




- Advantages & subtleties
  - Deterministic, conventional style
  - Cuts in each axis is not independent of each other, so recursive

Full tree

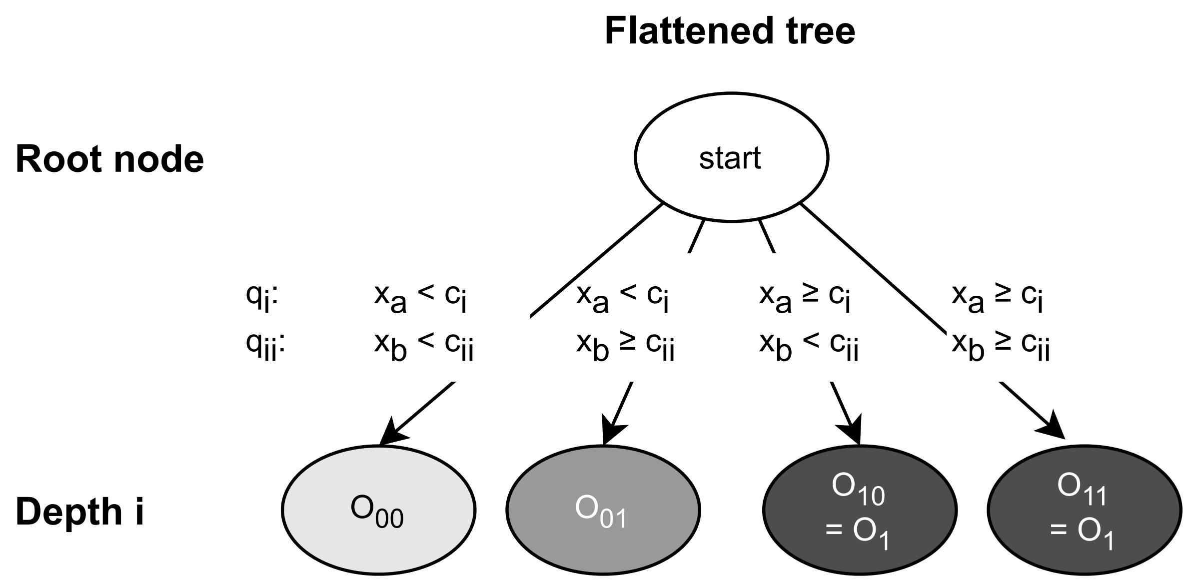


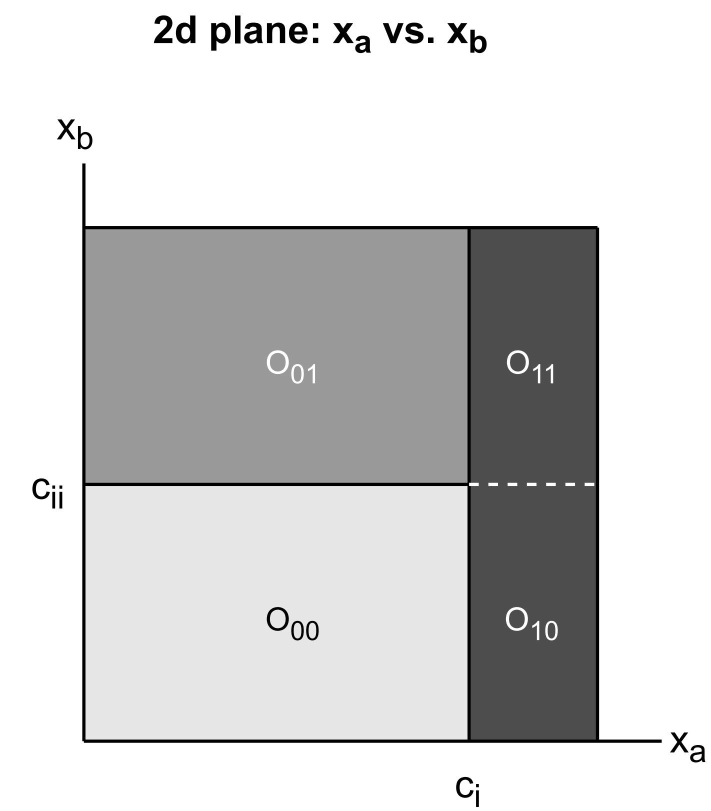


Decision tree, 2 var example (3)

TM Hong

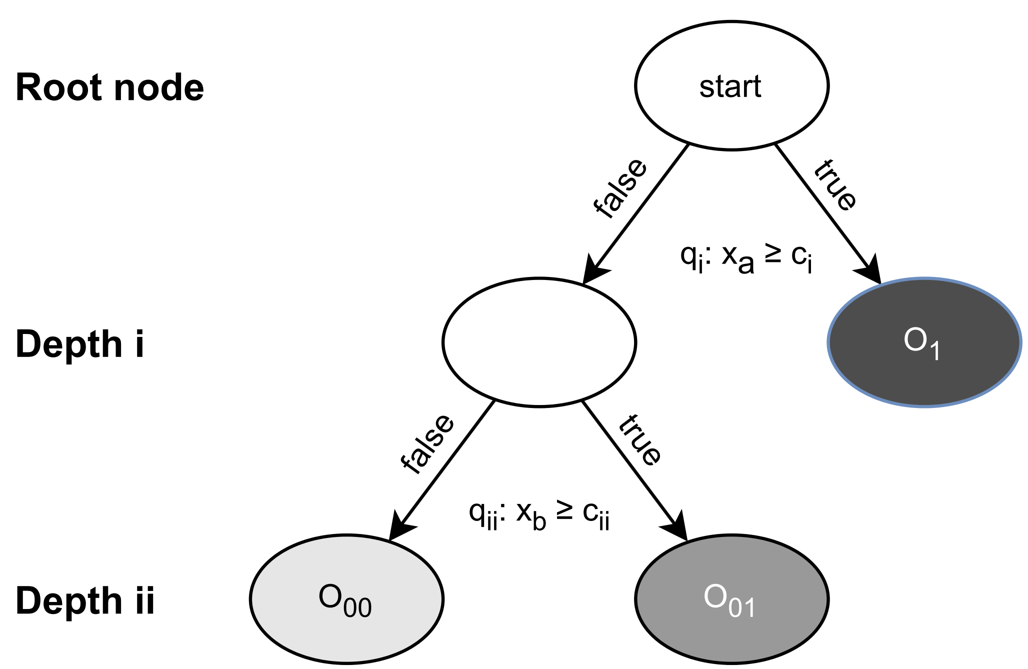
Our approach

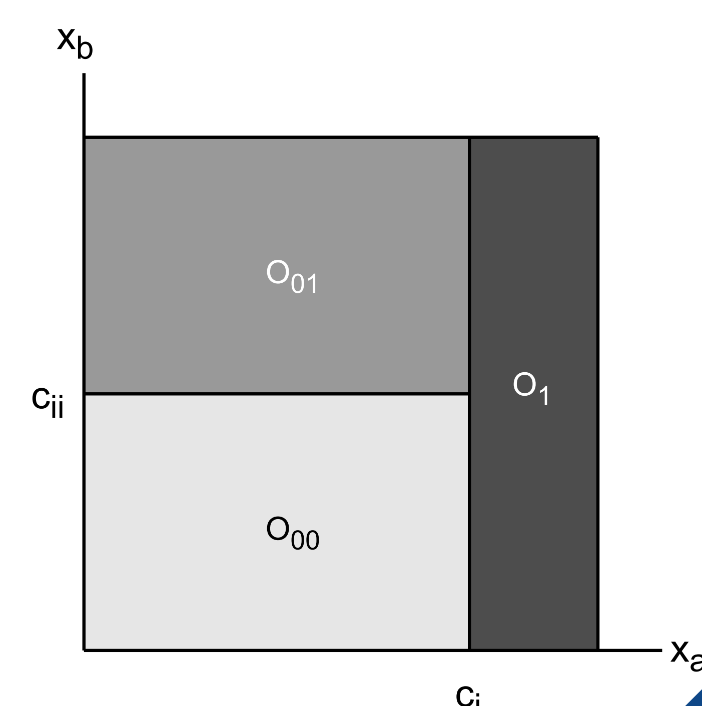




- Advantages & subtleties
  - Each axis is independent of each other → Bin search problem on a grid
  - Does not scale well for very deep trees (but do you really need it at L1?)

Full tree





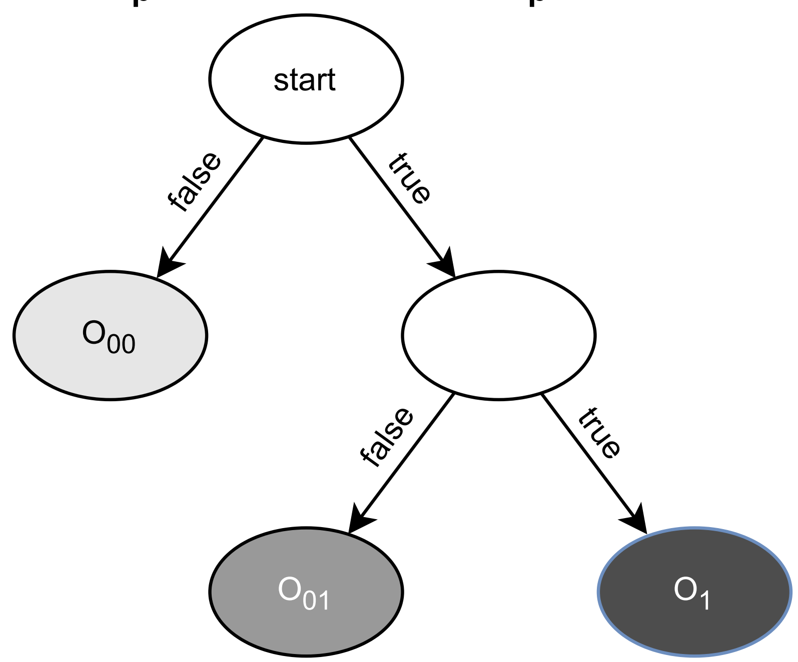
Forest of boosted decision trees

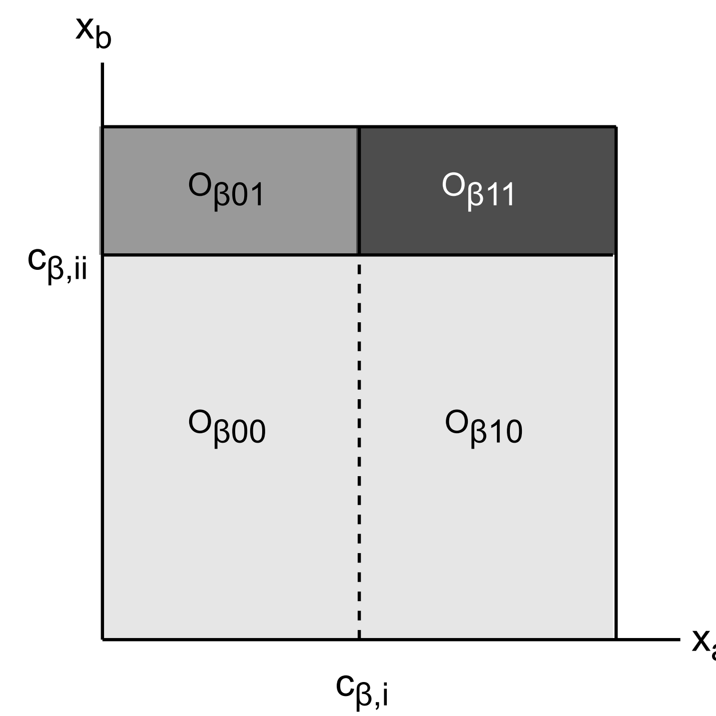
- Advantages & subtleties
  - Use TMVA software to train the BDT (support for other sw coming)

Our approach

2nd tree

Decision tree τ_β with boost weight w_β

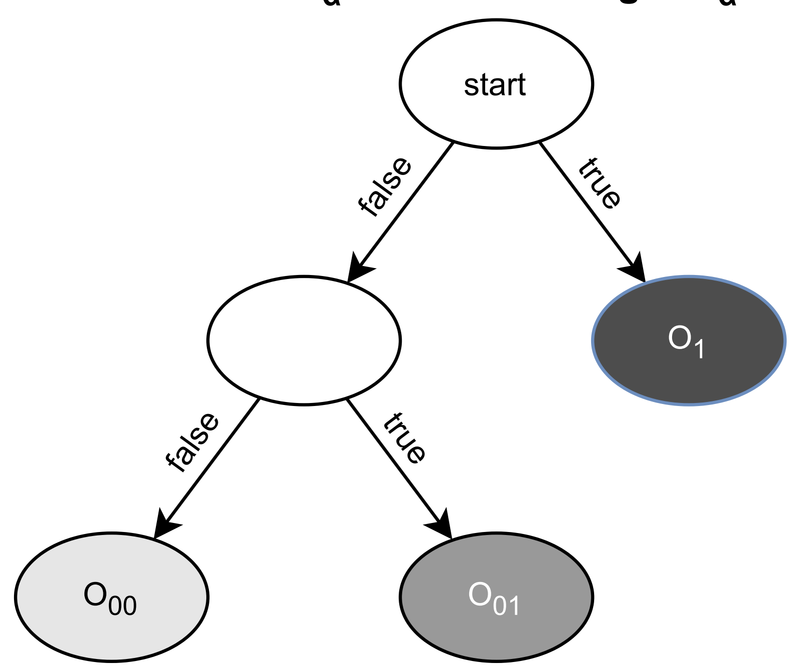


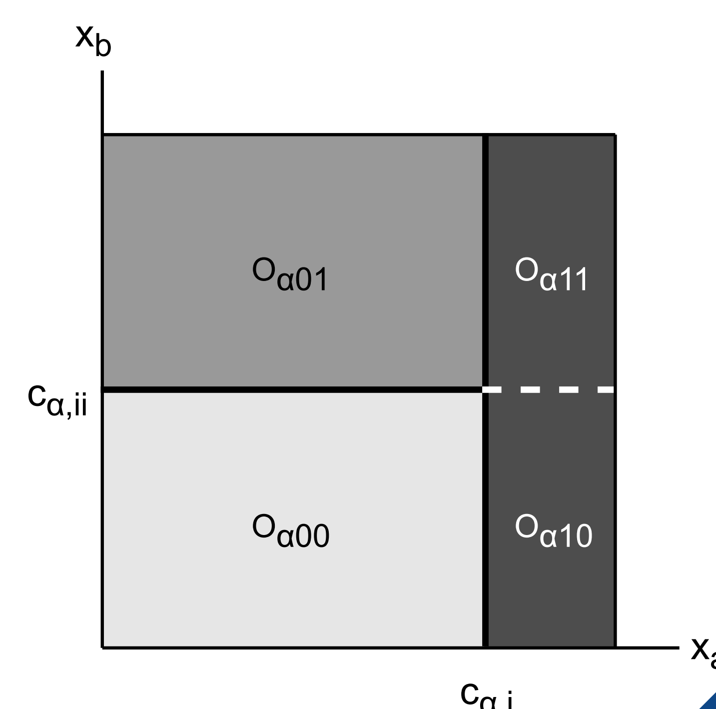


- Can we pre-merge the trees for firmware? Yes, next slide.

Decision tree τ_α with boost weight w_α

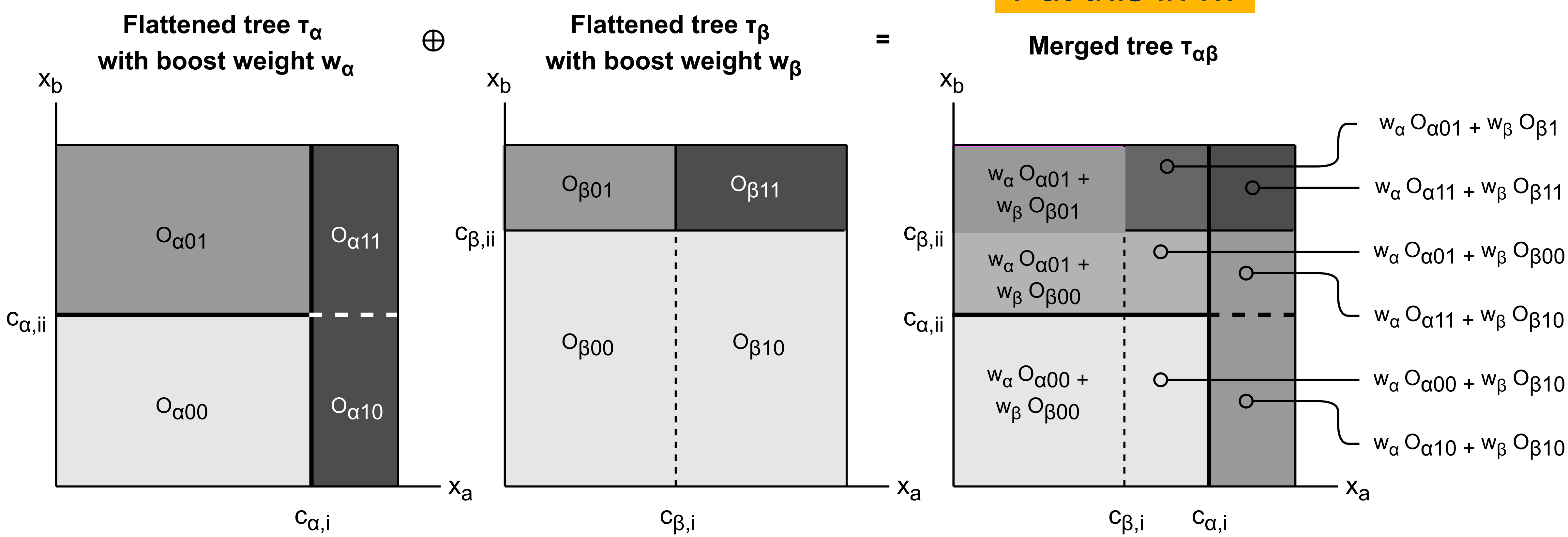
1st tree





Merging of the forest

Put this in fw



- Advantages & subtleties
  - Merging is pre-processed before implementation in firmware
  - This is using adaptive boosting. Gradient boosting cannot pre-merge, but we have approximations for that method to improve performance.

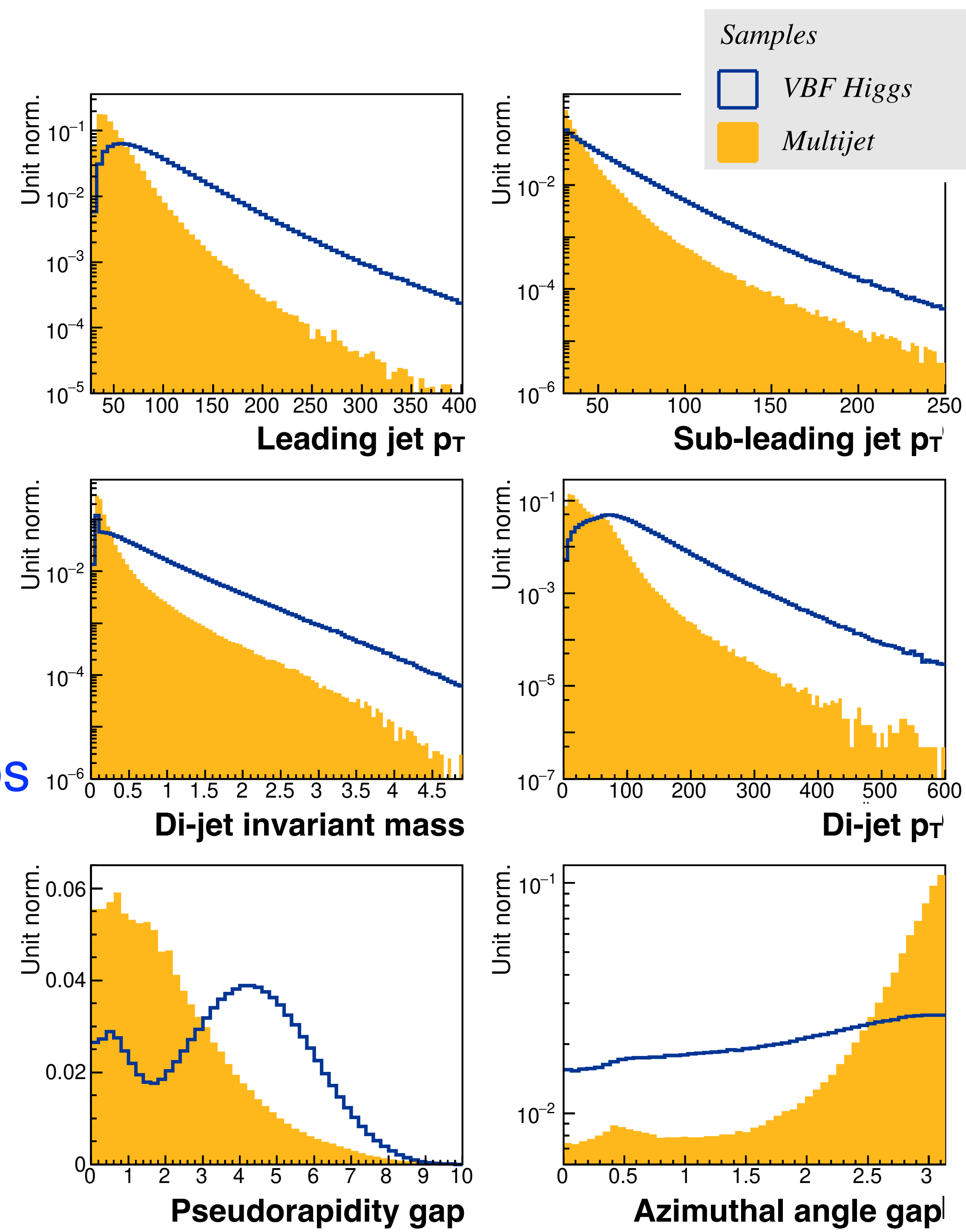
- Physics impact of flattening & merging
  - None, bec. encodes the entirety of conventional approach
  - Firmware is a giant look-up table problem

Physics: VBF Higgs vs. multijet

TM Hong

VBF Higgs vs. Multijet background

- σ_{Higgs} = 4 pb, two widely separated high-p_T jets
- σ_{pp} = 80 mb, dominant process at LHC
- Distributions given on the right



We consider two decays of the Higgs

- H → neutrinos, "invisible"
- H → $b\bar{b}b\bar{b}$, thru pseudoscalar decays

Strategy

- Train BDT to identify VBF jet pair,
  i.e., train BDT on Multijet vs. VBF H → neutrinos
- Apply that BDT to Multijet vs. VBF H → $b\bar{b}b\bar{b}$

Why

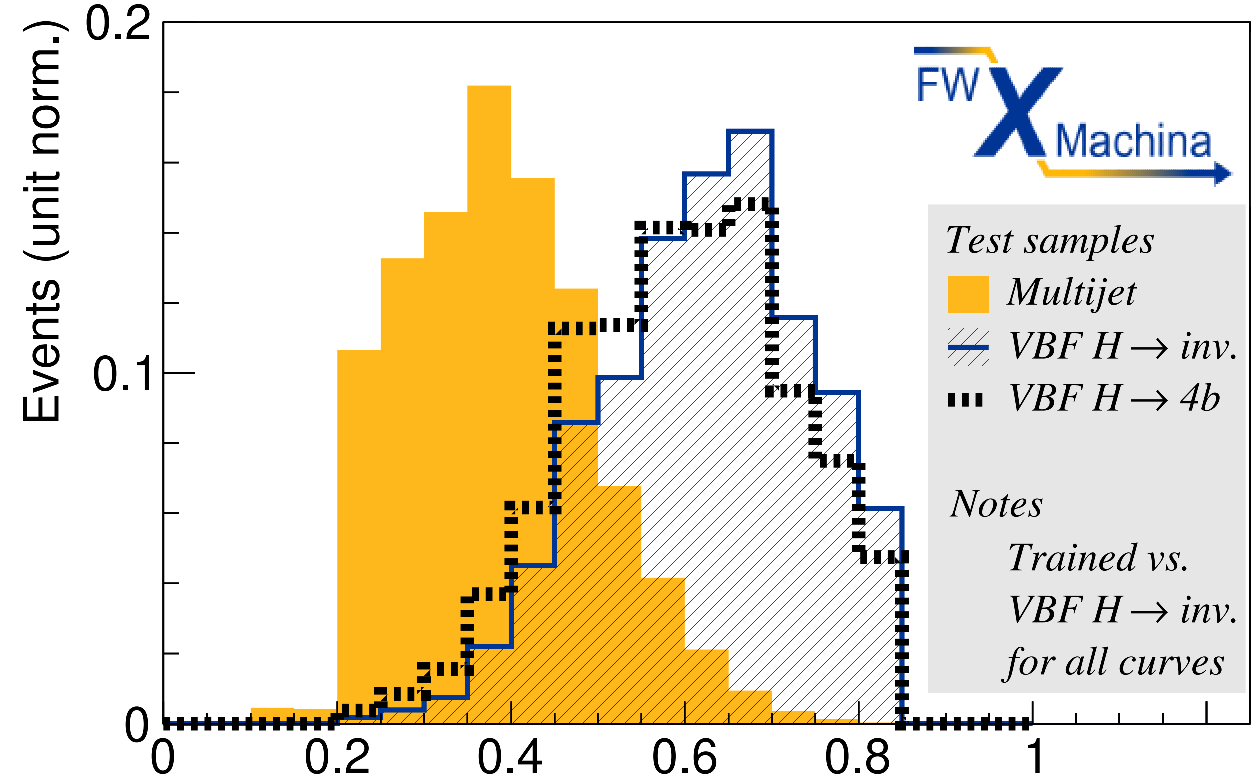
- If it works for VBF H → $b\bar{b}b\bar{b}$, then it can be a trigger for VBF independent of the Higgs decay
- Does it work? Next slide

Physics: VBF Higgs vs. multijet (2)

TM Hong

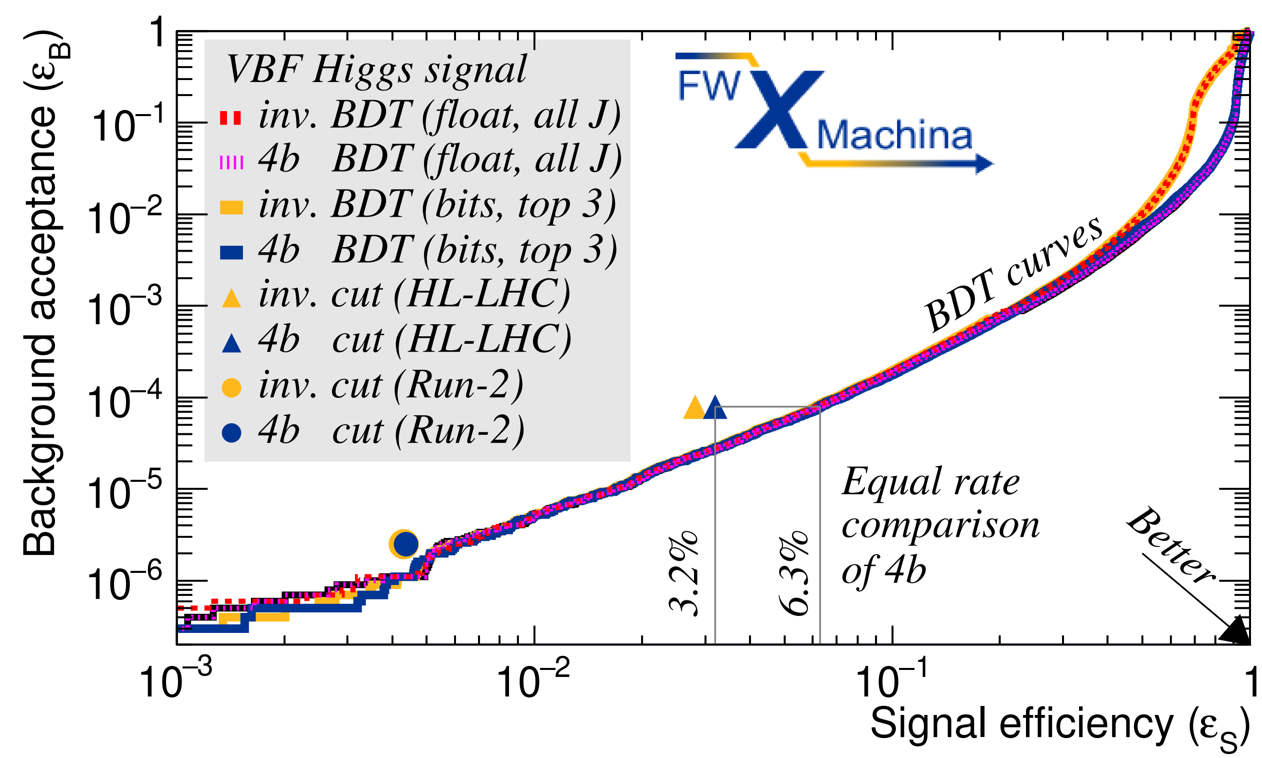
It works!

- Reminder. Did *not* train on VBF H → $b\bar{b}b\bar{b}$
- Subtlety re: jet selection (see paper)
- Distributions given on the right



Performance comparison

- Try to mimic ATLAS HL-LHC cuts as best we can using Madgraph + Delphes
- Two-fold signal efficiency improvement from ATLAS-inspired → fwX results



Details

- We validated our setup to reproduce the signal efficiency in the ATLAS Run-2 paper
- Comparison using bit integers, not floats

Firmware: VBF Higgs vs. multijet

Ran two configurations

- Optimized version
- Non-optimized version (for comparison)
- Both using 100 trees, max depth of 4
- Results given on the right

Performance

- 5 clock ticks = 16 ns
- Negligible resource usage

	VBF H Optimized	VBF H Non-opt
N_{var}	5	7
$N_{bit\text{-}var}$	8	12
$N_{bit\text{-}score}$	16	16
N_{bin}	40k	1M
Latency	5 ticks	6 ticks
LUT	1%	1.5%
Flip Flops	~0	~0
BRAM	2%	30%
DSP	0	~0

Benchmark using e^+ vs. γ

- In the paper, we also define one set of parameters to scale up one param. at a time
- Uses 4 variables, 8 bits & same as above
- 3 clock ticks = 10 ns
- Negligible resource usage

Where can we put fwX?

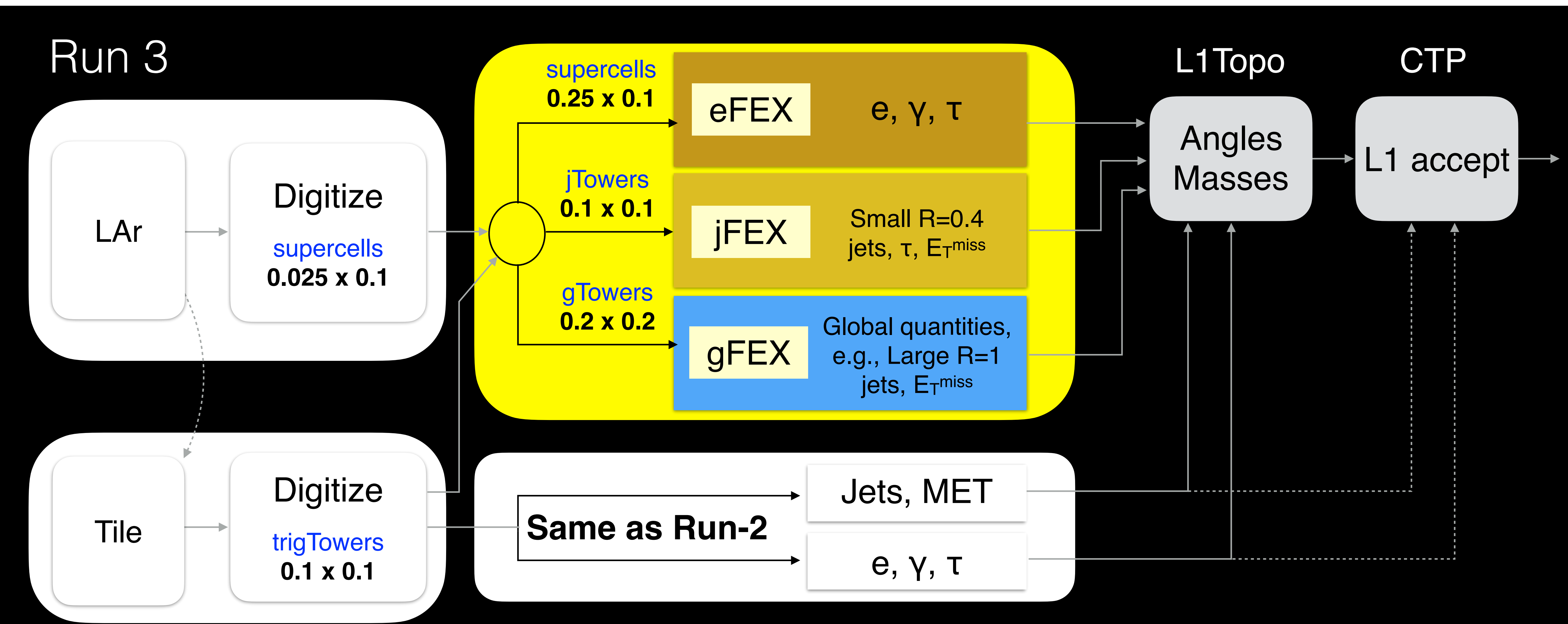
L1Calo TDR
https://cds.cern.ch/record/1602235

TM Hong

My guess

More speculative

- L1Topo to do combo algorithms
- gFEX to develop new algorithms
- jFEX to develop new algorithms
- eFEX to develop new algorithms



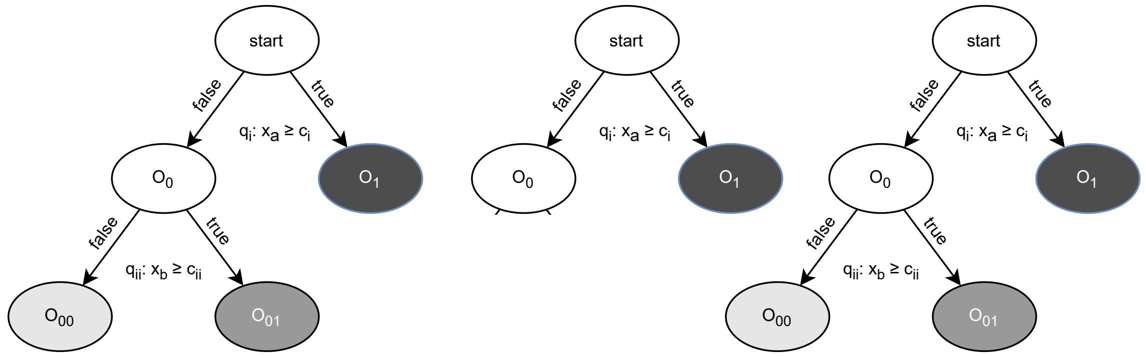
Source: L1Calo TDR, https://cds.cern.ch/record/1602235. This diagram courtesy B. Carlson.

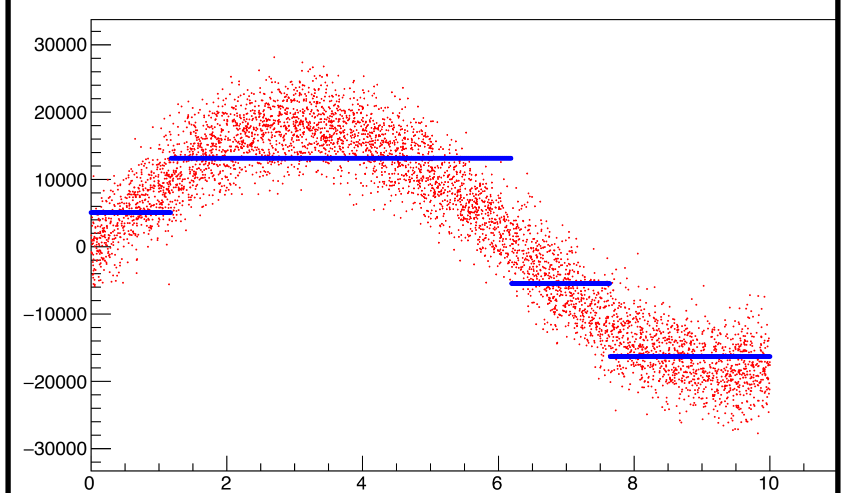
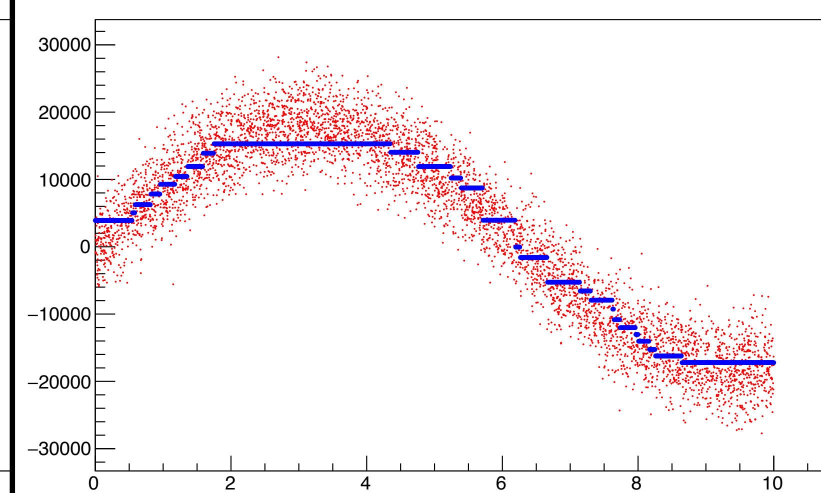
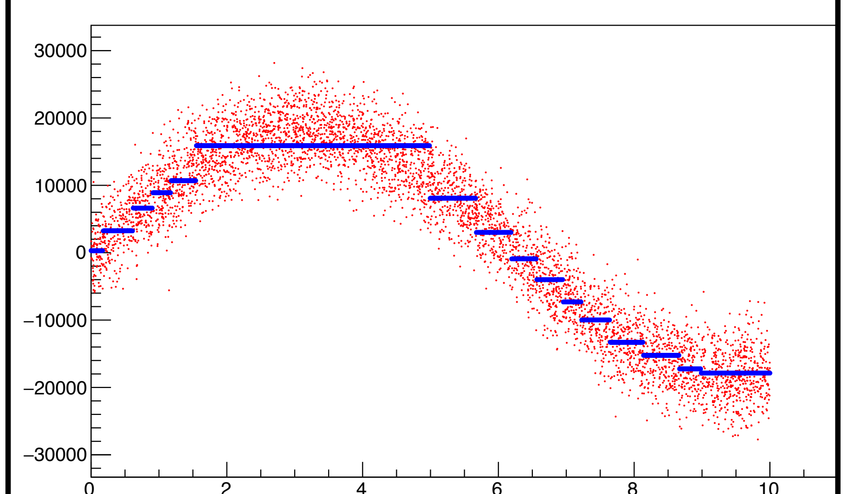
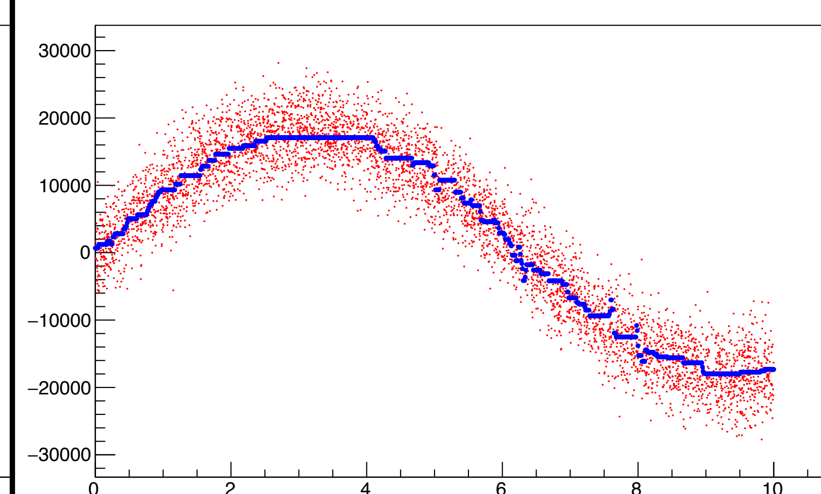
What about other ML? (we're working on it)

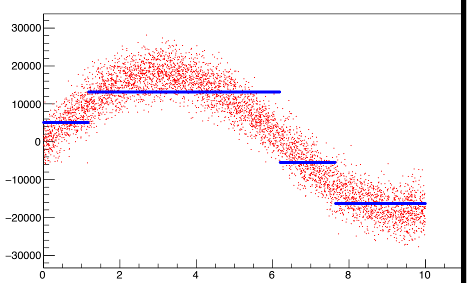
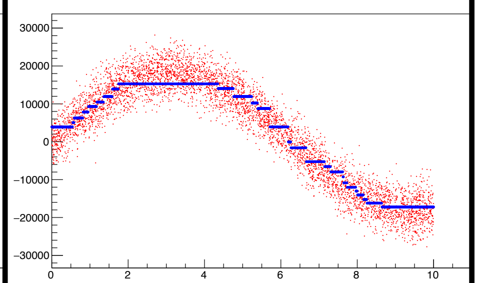
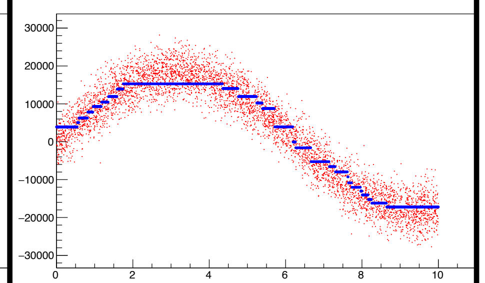
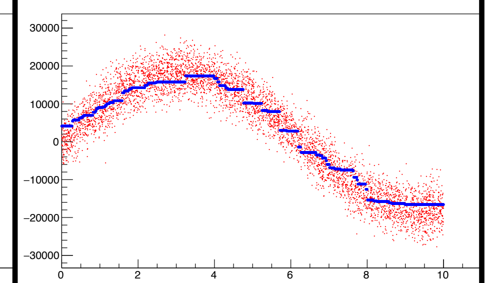
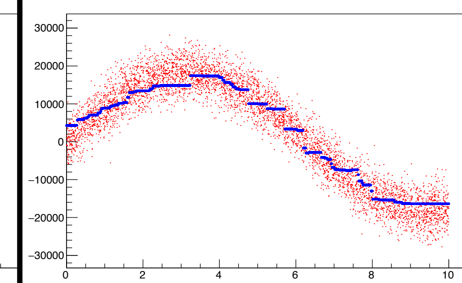
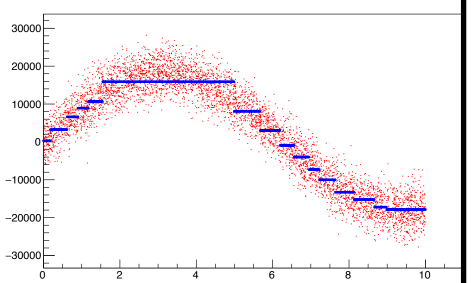
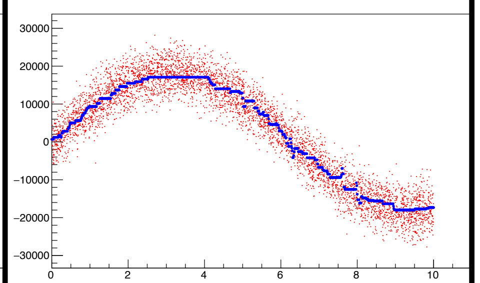
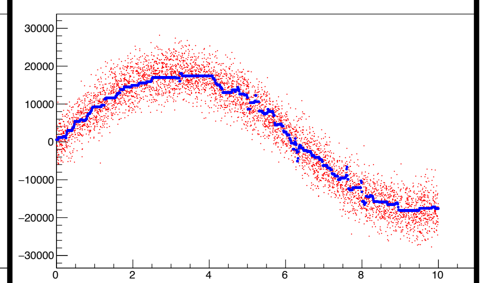
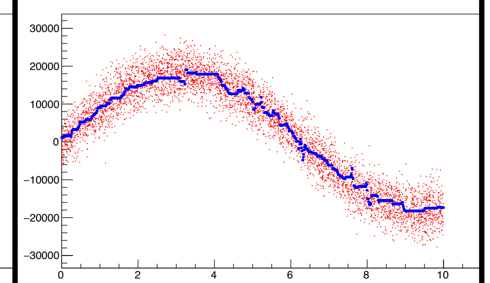
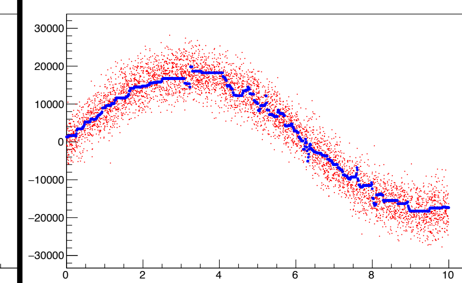
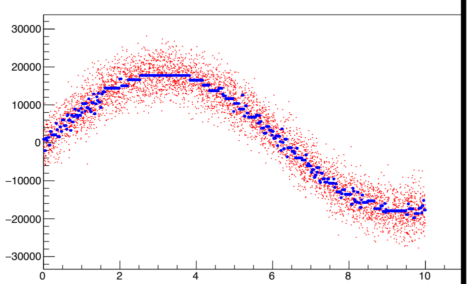
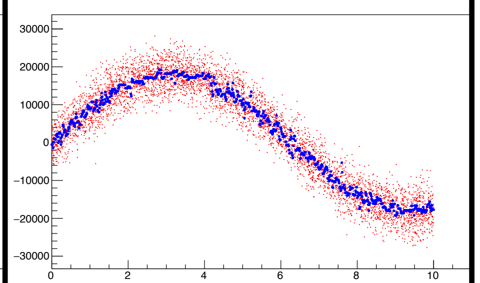
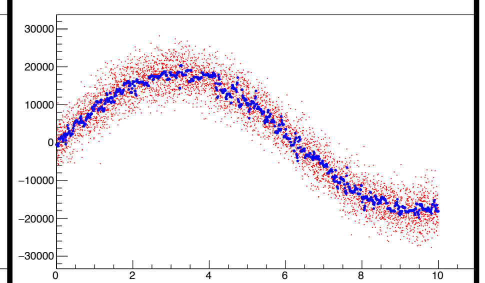
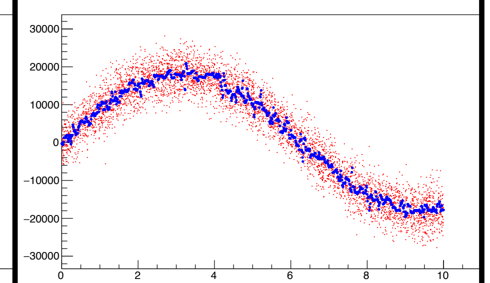
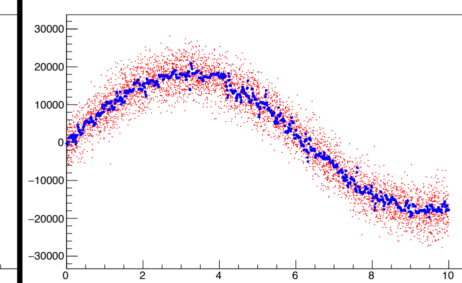
TM Hong

Regression (using BDT)

- Toy problem in 1-d
- Train / test on $f(x) = \sin(x) + \text{Gaussian}(x)$
- For sample of x: $y = f(x)$ in 16 bits



	$N_{tree} = 1$	$N_{tree} = 10$
$N_{depth} = 2$		
$N_{depth} = 4$		

$(x,y) = (N_{tree}, N_{depth})$	Ntree = 1	10	25	100	400
2					
4					
8					

Summary

Why (10 min)

- Higgs boson
- Dark sector

Method (20 min)

- $H_{125} \rightarrow$ Dark matter pair $B_{invisible} < 0.13$
- $H_{125} \rightarrow$ Dark photon $+ \gamma$ $B_{dark\ photon} < 0.014$

Trigger (30 min)

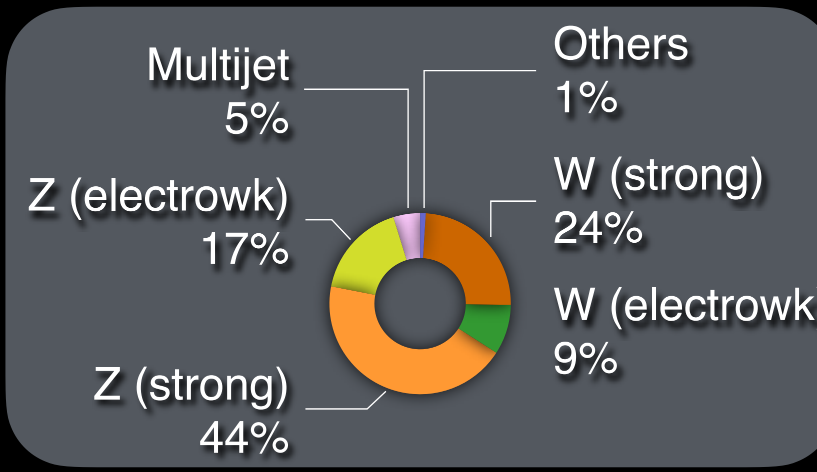
- E_T^{miss}, VBF
- ML on FPGA http://fwx.pitt.edu

Abstract

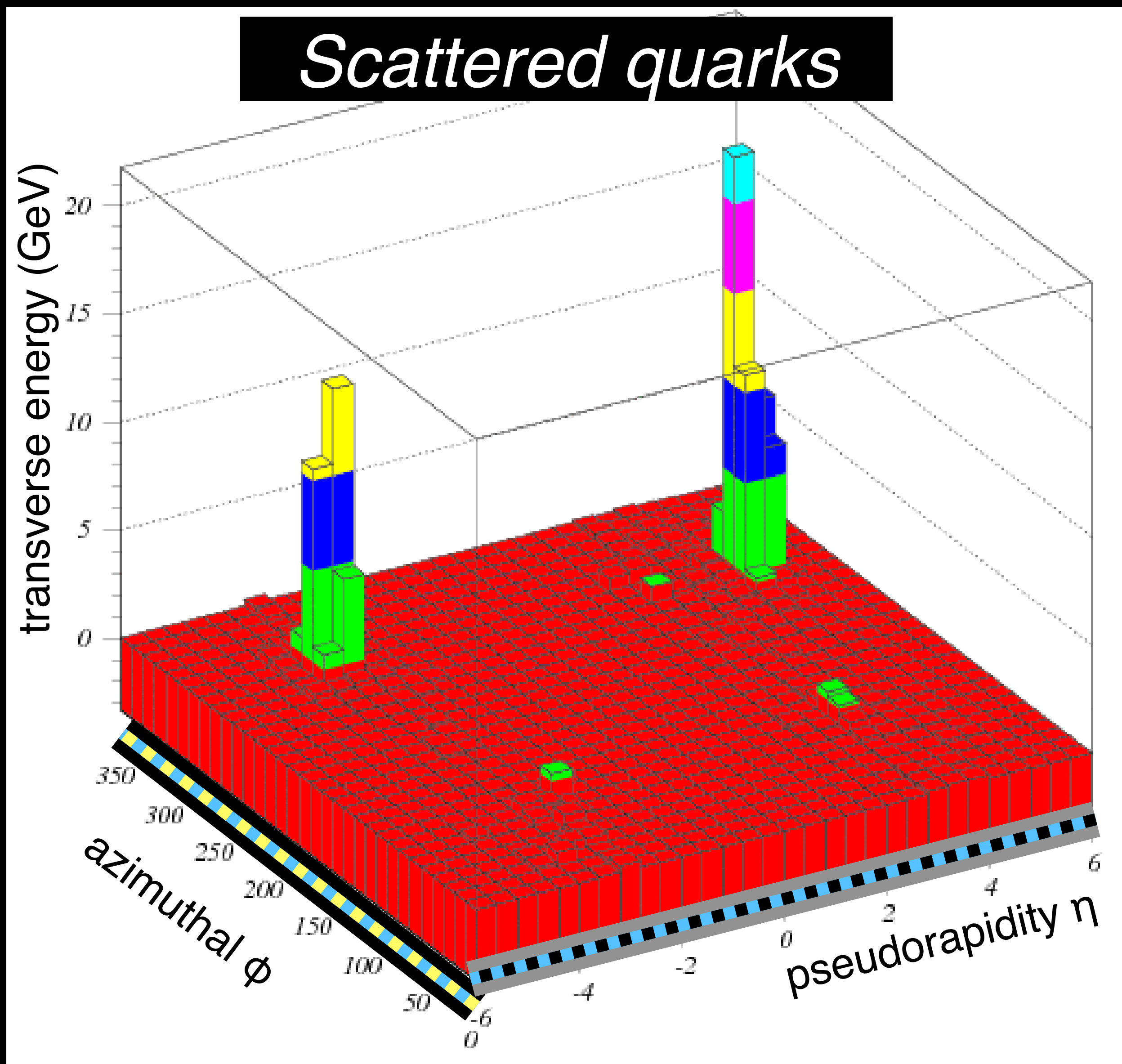
With more data coming from LHC collisions, detailed measurements of Higgs boson properties allow us to probe whether it communicates with the unknown and/or undiscovered sector beyond the Standard Model. One motivation is weakly interacting dark matter, which are invisible to the detecting apparatus, through a Higgs portal. I will discuss the latest ATLAS results of the search for Higgs bosons decaying to invisible particles. I will also describe the technical challenges of triggering on such events using missing energy from the Higgs boson decay and/or hadronic jets from the Higgs boson production, including the potential use of machine learning methods on FPGA boards in real-time level-1 trigger systems. will discuss how such interactions produce the recently discovered Higgs boson, and how it may serve as a portal to unknown sectors of elementary particles, such as dark matter. I will also describe the technical challenges of saving such minuscule fractions of weak force collisions, including the use of artificial intelligence in real-time trigger systems.

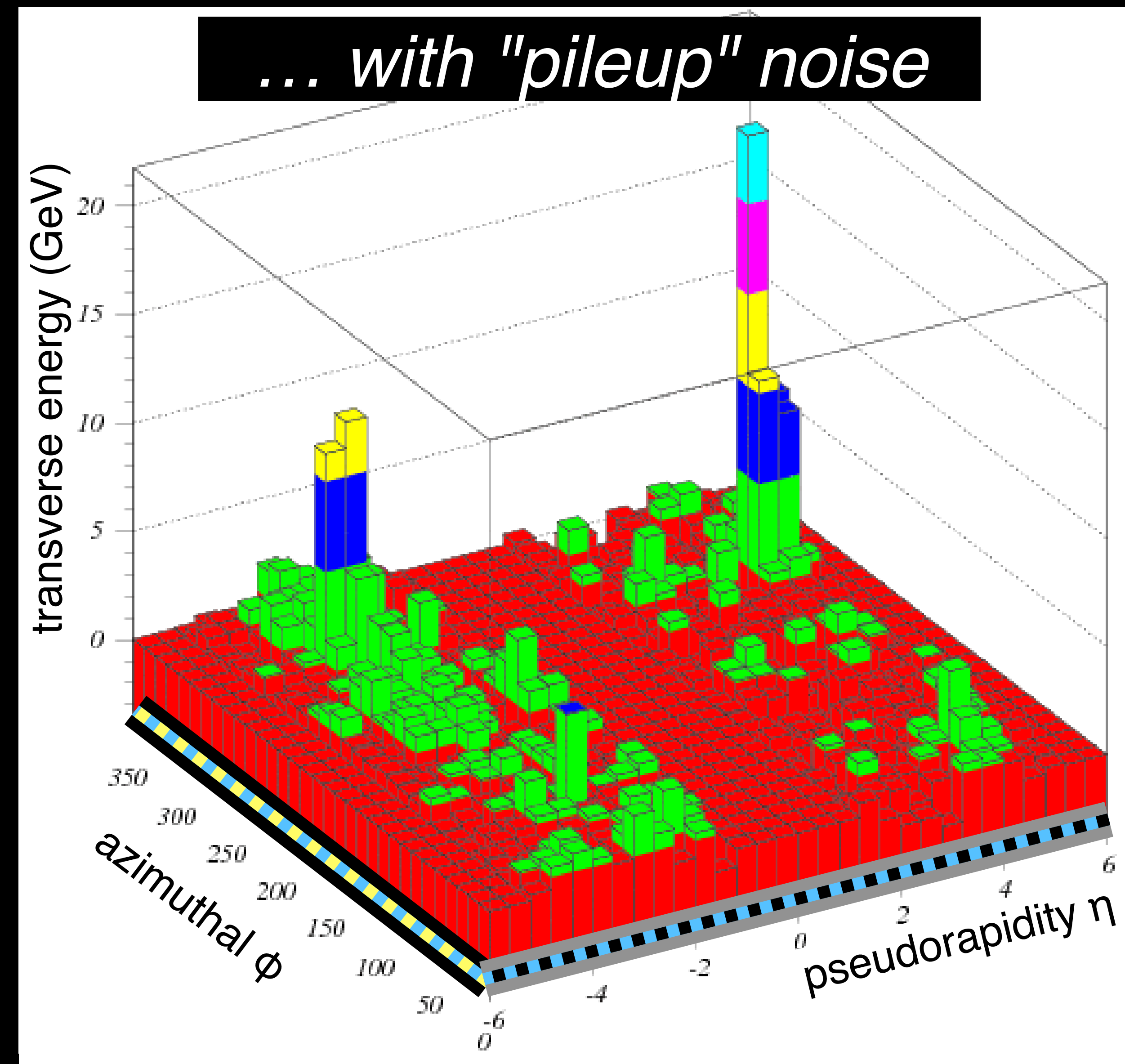
Distribution of energy

Simulation of the polar angle for one collision



Ellis, Huston, Hatakeyama, Loch, Tönnesmann, Prog. in Part. & Nucl. Phys. 60 (2008) 484

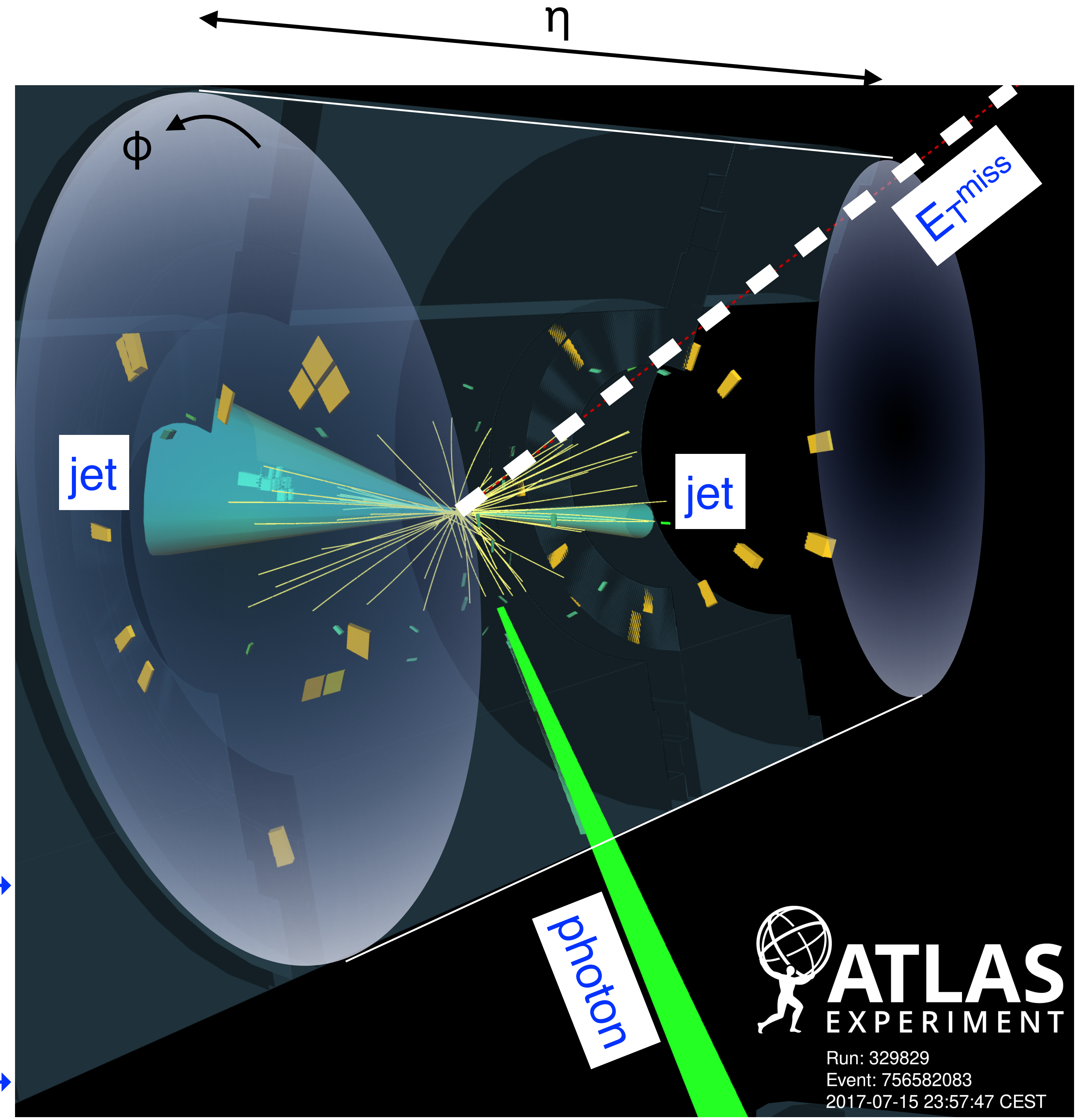




Large Δη between the scattered quark jets

Detector signature

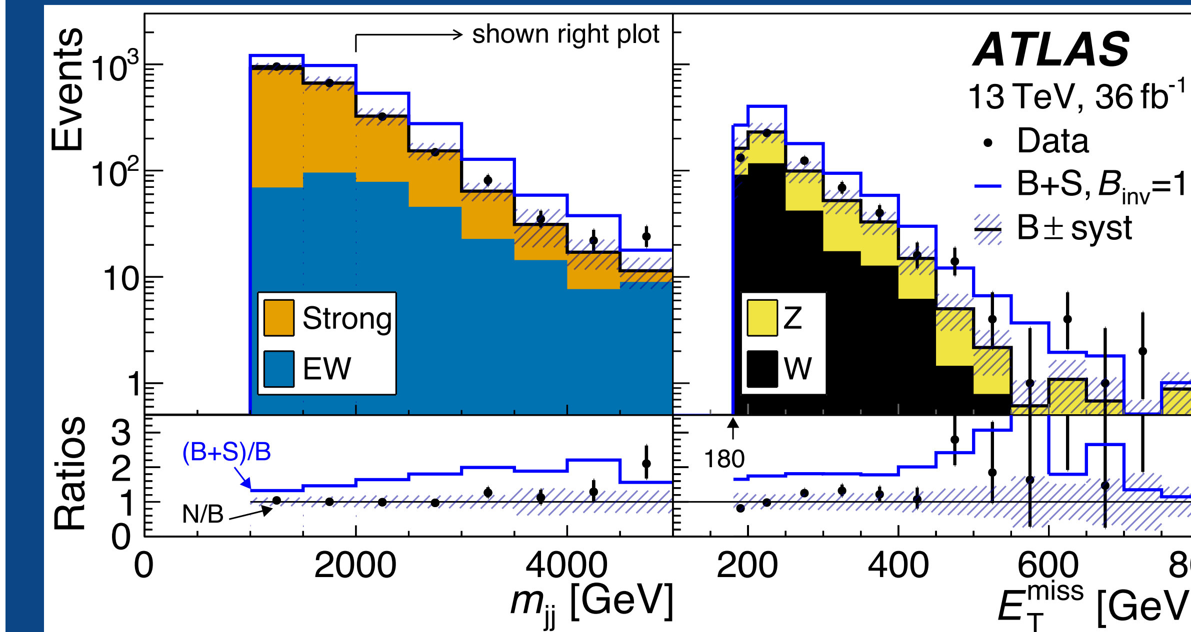
- ATLAS geometry
  - η along the beam direction
  - ϕ azimuthal angle
- VBF jet pair
  - High p_T
  - Wide gap in η
  - Not back-to-back in ϕ
  - Large $m_{jet\text{-}jet}$ 2 TeV →
  - Low hadronic activity in between
- E_T^{miss}
  - p_T imbalance 840 GeV →
- For +γ
  - High-p_T photon 540 GeV →
  - $m_T(E_T^{miss}, \gamma)$ 1.1 TeV →
  - Its η in between jets



http://cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2021-004/fig_19.pdf

Signal models

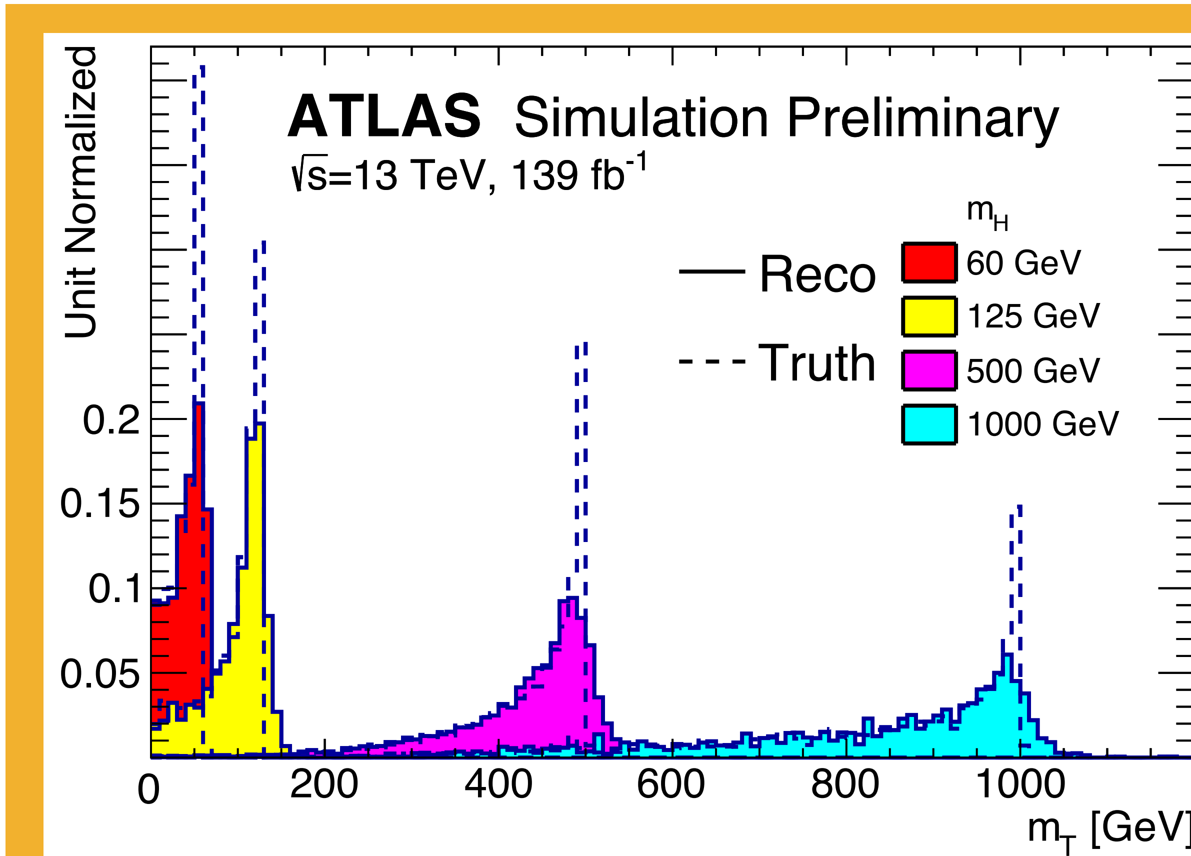
- H portal to χ
  - VBF H_{125} w/ POWHEG NLO
  - VBF H_{125} + γ_{ISR} w/ MG5_aMC@NLO
  - S-to-B is higher with m_{jj}, E_T^{miss}, see →



Dijet invariant mass ***p_T imbalance***

http://cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/EXOT-2016-37/fig_05.pdf

- H portal to γ_d
  - VBF H_{125} → $\gamma\gamma_{dark}$ w/ POWHEG v2
  - $m_T(E_T^{miss}, \gamma)$ as proxy for m_H, see →



Transverse mass of E_T^{miss} and γ

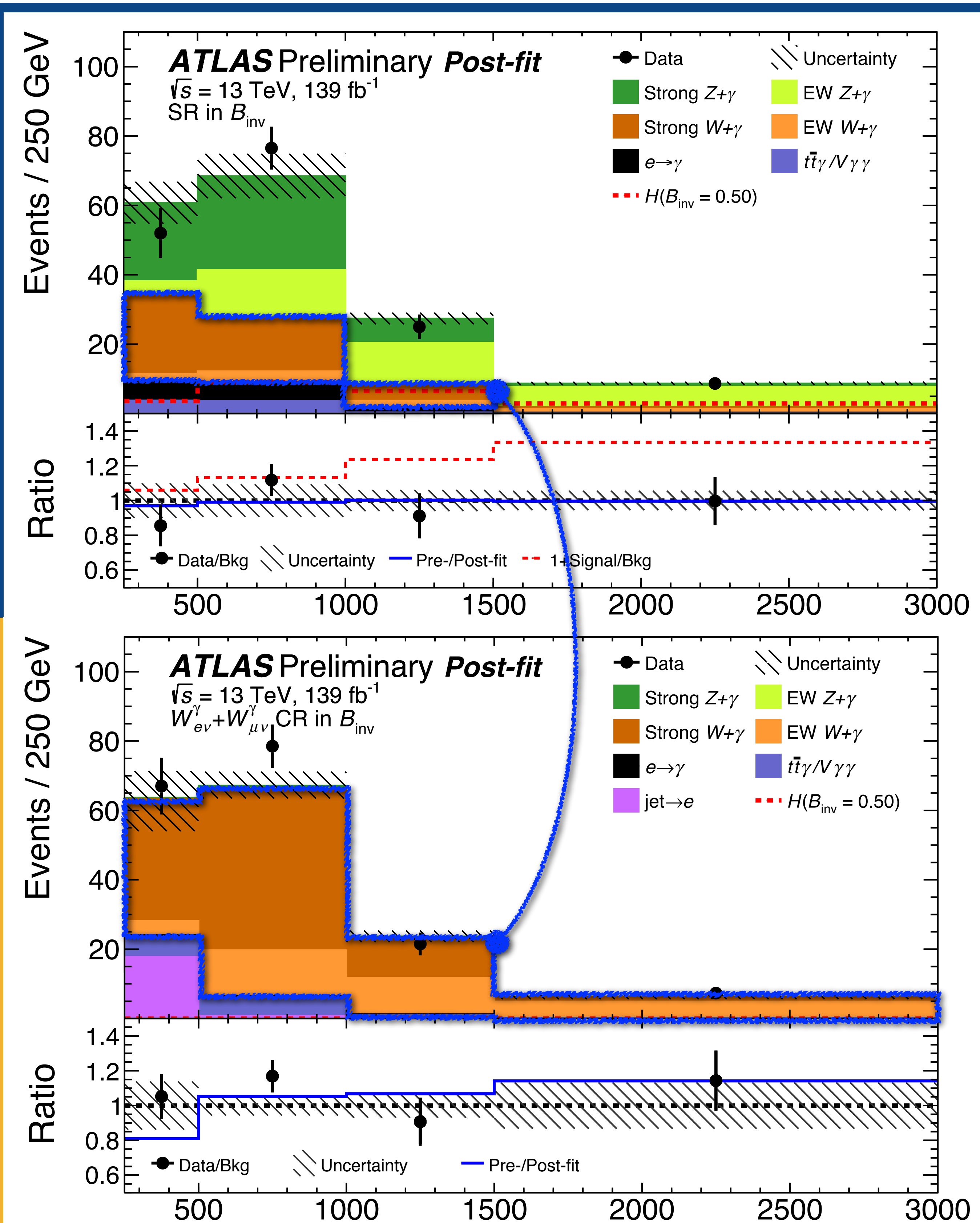
http://cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2021-004/fig_13.pdf

Backgrounds, for +γ

- Weak boson bkg'd
  - $Z \to \nu\nu$ No leptons
  - $W \to \ell\nu$ Loses a lepton
- Signal Region
  - E_T^{miss} trigger, > 150 GeV
  - "Centrality" of γ, 3rd jet
  - For $+\gamma_{ISR}$, $15 < p_T^{\gamma} < 110$ GeV
  - For $+\gamma_{dark}$, max(110,0.7 m_T)
- Control Region
  - For $W \to \ell\nu$, Require a lepton
  - Lepton trigger, > 30 GeV
  - Reverse γ centrality cut

SR

CR



Dijet invariant mass

http://cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2021-004/fig_04a.pdf, fig_04b.pdf

Uncertainties, for +γ

- Statistical
  - √N
  - MC
- Theoretical
  - Wγ, Zγ theory
- Experimental
  - JES, JER

	1σ Uncertainty on $\mathcal{B}_{\text{inv}}$	on $\mathcal{B}(H \to \gamma\gamma_{\text{d}})$
Data stats.	0.106	0.0051
$V\gamma$+jets theory	0.056	0.0028
MC stats.	0.045	0.0026
Jet Scale and Resolution	0.045	0.0011
Photon	0.032	0.0011
$e \to \gamma$, jet$\to e, \gamma$ Bkg.	0.026	0.0024
Pileup	0.025	0.0004
$W\gamma$+jets/$Z\gamma$+jets Norm.	0.021	0.0005
$E_{\text{T}}^{\text{miss}}$	0.012	0.0003
Signal theory	0.004	0.0010
Lepton	0.002	0.0008
Total	0.148	0.0071

http://cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2021-004/tab_05.pdf

Evaluated by fixing parameters to their best-fit values and quadratically subtracting from the total nominal systematic uncertainty

to next slide

Table 1: Summary of generators used for simulation. The details and the corresponding references are provided in the body of the text. The V in V+jets represents either a W or a Z boson.

Process	Generator	ME Order	PDF	Parton Shower	Tune
		Signal Samples			
ggF Higgs	Powheg v2 NNLOPS	NNLO	PDF4LHC15	Pythia8.230	AZNLO
VBF Higgs+γ	MadGraph5_aMC@NLO 2.6.2	NLO	PDF4LHC15	Herwig 7.1.3p1	A14
ggF Higgs$\rightarrow \gamma\gamma_{\rm d}$	Powheg v2 NNLOPS	NNLO	PDF4LHC15	Pythia8.244p3	AZNLO
VBF Higgs$\rightarrow \gamma\gamma_{\rm d}$	Powheg v2	NLO	CTEQ6L1	Pythia8.244p3	AZNLO
		Background Samples			
Strong $V\gamma$+ jets	Sherpa v2.2.8	NLO (up to 1-jets), LO (up to 3-jets)	NNPDF3.0nnlo	Sherpa MEPS@NLO	Sherpa
EW $V\gamma$+ jets	MadGraph5_aMC@NLO 2.6.5	LO	NNPDF3.1lo	Pythia8.240	A14
EW VV+jets	Sherpa v2.2.1 or Sherpa v2.2.2	LO	NNPDF3.0nnlo	Sherpa MEPS@LO	Sherpa
VV+jets	Sherpa v2.2.1 or Sherpa v2.2.2	NLO (up to 1-jet), LO (up to 3-jets)	NNPDF3.0nnlo	Sherpa MEPS@NLO	Sherpa
EW V+ jets	Herwig 7.1.3 or Herwig 7.2.0	NLO	MMHT2014nlo68cl	Herwig 7.1.3	Herwig 7
Strong $W(\rightarrow \mu\nu)$ + jets/ $W(\rightarrow \tau\nu)$ + jets	Sherpa v2.2.8	NLO (up to 2-jets), LO (up to 4-jets)	NNPDF3.0nnlo	Sherpa MEPS@NLO	Sherpa
$t\bar{t}\gamma$	MadGraph5_aMC@NLO 2.2.3	NLO	NNPDF2.3lo	Pythia8.186	A14
$t\bar{t}$	PowhegBox v2	NLO	NNPDF3.0nlo	Pythia8.230	A14
γ + jet	Sherpa v2.2.2	NLO (up to 2-jets), LO (up to 4-jets)	NNPDF3.0nnlo	Sherpa MEPS@NLO	Sherpa
		Systematic Samples			
$V\gamma$+ jets α^4 interference	MadGraph5_aMC@NLO 2.6.2	LO	NNPDF3.1lo	Pythia8.240	AZNLO

http://cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2021-004/tab_01.pdf

Signal region for +γ

TM Hong

Table 3: Summary of the requirements defining the different regions considered in this analysis. Where present, the values in squared brackets are referring to the regions defined in the search for $H \to \gamma\gamma_d$ signal. The leading and subleading jets must satisfying the fJVT requirements mentioned in Sec. 5. In the SR and $Z^{\gamma}_{\text{Rev.Cen.}}$ CR definitions $E_{\text{T}}^{\text{miss,lep-rm}} \equiv E_{\text{T}}^{\text{miss}}$ since no lepton is present.

Variable	SR	$W^{\gamma}_{\mu\nu}$ CR	$W^{\gamma}_{e\nu}$ CR	$Z^{\gamma}_{\text{Rev.Cen.}}$ CR	Fake−e CR
$p_{\text{T}}\ (j_1)$ [GeV]	> 60				
$p_{\text{T}}\ (j_2)$ [GeV]	> 50				
N_{jet}	2,3				
$N_{\text{b-jet}}$	< 2				
$\Delta\phi_{\text{jj}}$	< 2.5 [2.0]				
$\lvert\Delta\eta_{\text{jj}}\rvert$	> 3.0				
$\eta(j_1) \times \eta(j_2)$	< 0				
C_3	< 0.7				
m_{jj} [TeV]	> 0.25				
$E_{\text{T}}^{\text{miss}}$ [GeV]	> 150	–	> 80	> 150	< 80
$E_{\text{T}}^{\text{miss,lep-rm}}$ [GeV]	–	> 150	> 150	–	> 150
$E_{\text{T}}^{\text{jets,no-jvt}}$ [GeV]	> 130				
$\Delta\phi(j_i, E_{\text{T}}^{\text{miss,lep-rm}})$	> 1.0				
N_{γ}	1				
$p_{\text{T}}\ (\gamma)$ [GeV]	$> 15, < 110$ [$> 15, <\max(110, 0.733 \times m_{\text{T}})$]				
C_{γ}	> 0.4	> 0.4	> 0.4	< 0.4	> 0.4
$\Delta\phi(\gamma, E_{\text{T}}^{\text{miss,lep-rm}})$	> 1.8 [–]				
N_{ℓ}	0	1 μ	1 e	0	1 e
$p_{\text{T}}\ (\ell)$ [GeV]	> 30				

http://cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2021-004/tab_03.pdf

8 variables fed to DNN

η_{γ}, η_{j2}

centrality C_{γ} [102] is defined as

$$C_{\gamma} = \exp\left(-\frac{4}{(\eta_1 - \eta_2)^2}\left(\eta_{\gamma} - \frac{\eta_1 + \eta_2}{2}\right)^2\right), \quad (1)$$