Search for High Mass States Decaying to Tau Pairs with the CMS Experiment

Andrew A. Johnson

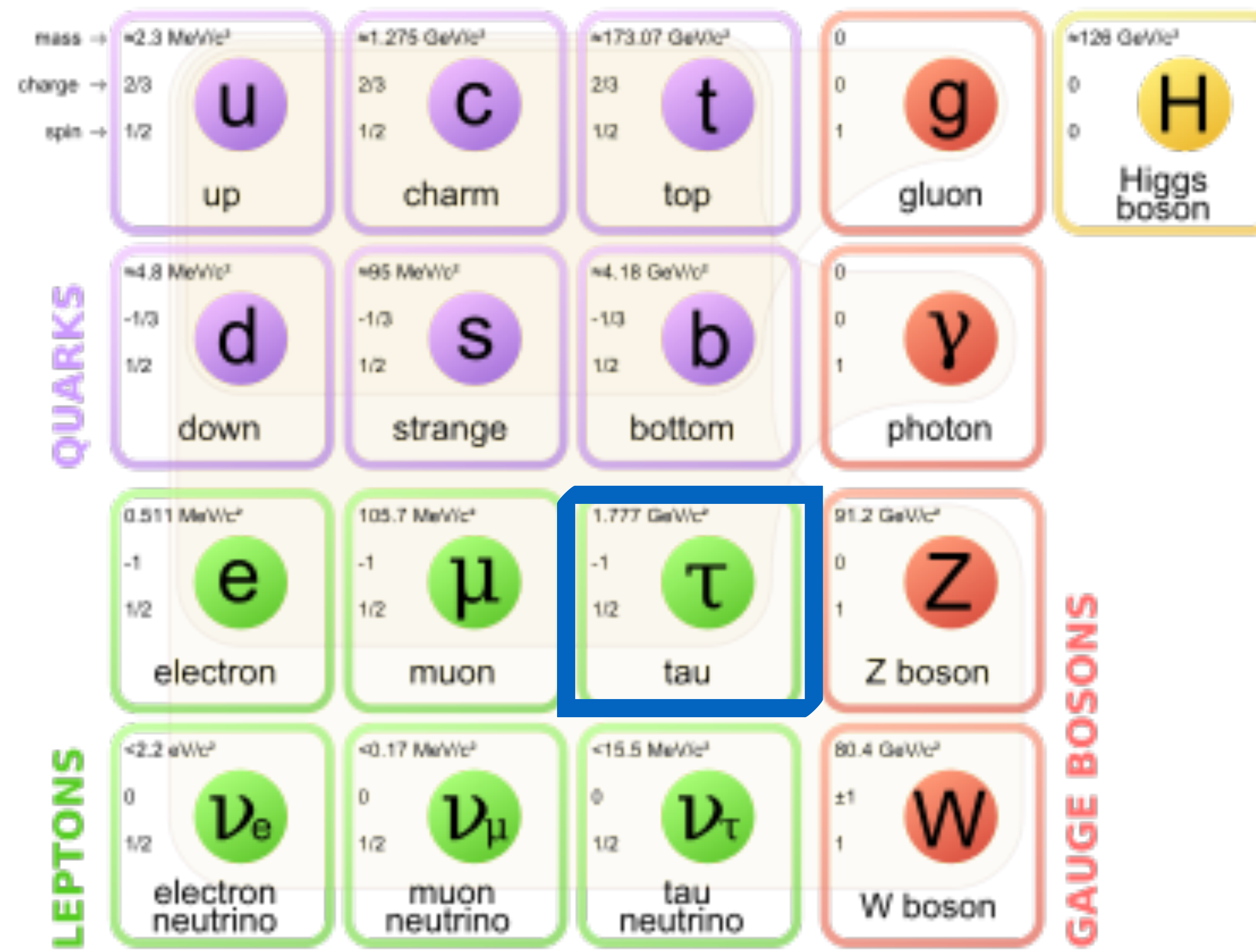
November 4th, 2016

Overview

- Physics motivation
- The LHC & CMS
- Analysis strategy & object selection
- Background estimation
- Results and limits
- Future impact studies
- Summary and outlook

The Standard Model

- Theory governing the electromagnetic, strong, and weak interactions
- Includes all (known) fundamental matter particles (**fermions**) and force carriers (**bosons**)
- Quarks combine to form composite **hadrons**
- Clusters of hadrons form **jets**



Key players in this thesis: taus, electrons, muons, neutrinos, jets

Physics Motivation

- Standard Model is incomplete!
    - No gravity, no dark matter candidate, hierarchy problem…
- Many Beyond-Standard Model (BSM) theories add an extra symmetry group to the Standard Model
- Expect to see an extra neutral boson, Z', as a result
- Searches have been done for Z'->ee, μμ (CMS limit: 3.2 TeV)
- New physics could arise in ττ channel

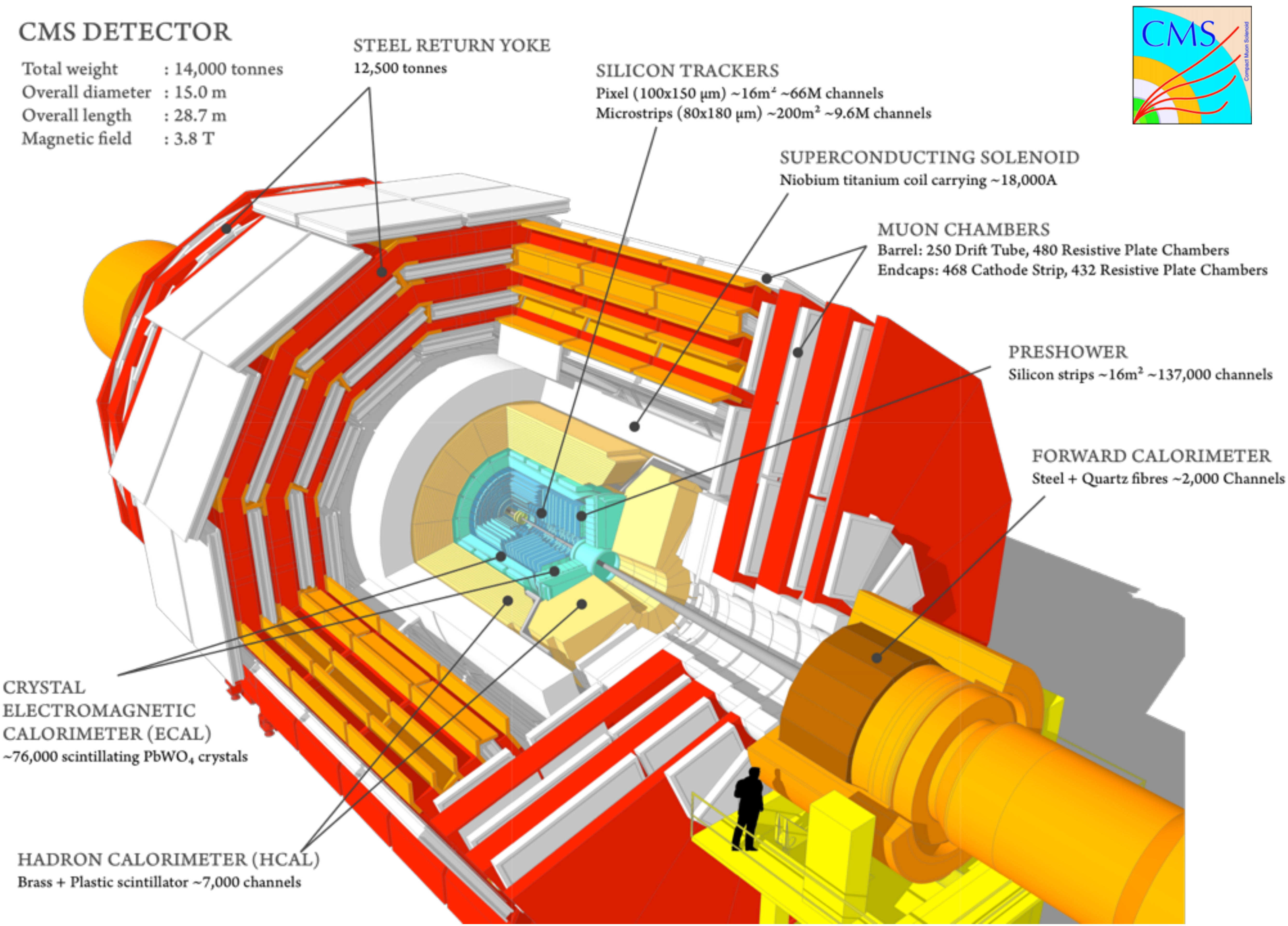
Search is really for high mass, back-to-back tau pairs

CERN

- Largest particle physics laboratory in the world
- Employs over 10^4 staff, engineers, scientists, and research personnel + thousands of graduate students (from over 80 countries)
- Located in Geneva, Switzerland
- Main function: constructing and operating particle accelerators, particle physics research

Large Hadron Collider

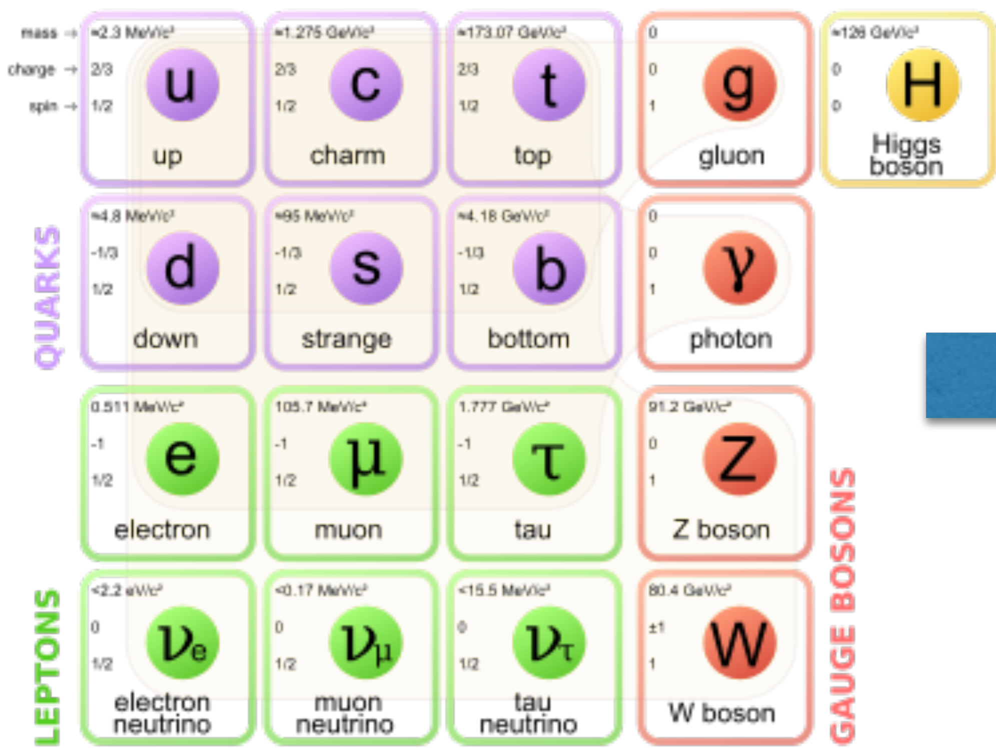
- World's largest and highest-energy particle accelerator
- 27 km in circumference, 100-200m underground
- Collides protons (and sometimes ions) together at 13 TeV
- Four detectors study particles created in pp collisions
- I work on the Compact Muon Solenoid (CMS) Experiment



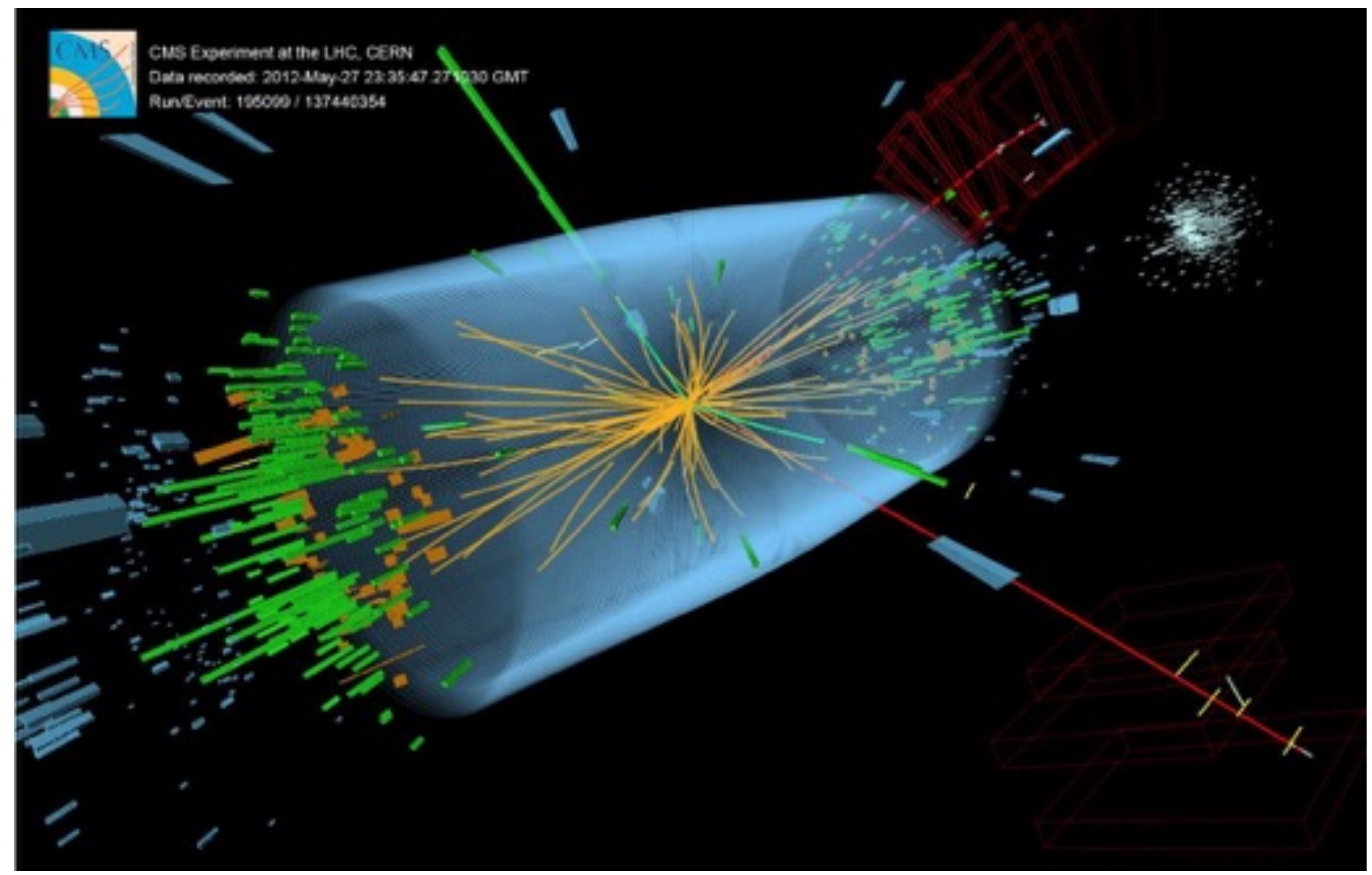
CMS DETECTOR
Total weight : 14,000 tonnes
Overall diameter : 15.0 m
Overall length : 28.7 m
Magnetic field : 3.8 T
STEEL RETURN YOKE
12,500 tonnes
SILICON TRACKERS
Pixel (100x150 μm) ~16m² ~66M channels
Microstrips (80x180 μm) ~200m² ~9.6M channels
SUPERCONDUCTING SOLENOID
Niobium titanium coil carrying ~18,000A
MUON CHAMBERS
Barrel: 250 Drift Tube, 480 Resistive Plate Chambers
Endcaps: 468 Cathode Strip, 432 Resistive Plate Chambers
PRESHOWER
Silicon strips ~16m² ~137,000 channels
FORWARD CALORIMETER
Steel + Quartz fibres ~2,000 Channels
CRYSTAL ELECTROMAGNETIC CALORIMETER (ECAL)
~76,000 scintillating $PbWO_4$ crystals
HADRON CALORIMETER (HCAL)
Brass + Plastic scintillator ~7,000 channels
CMS

Physics at CMS

Standard Model



CMS Event

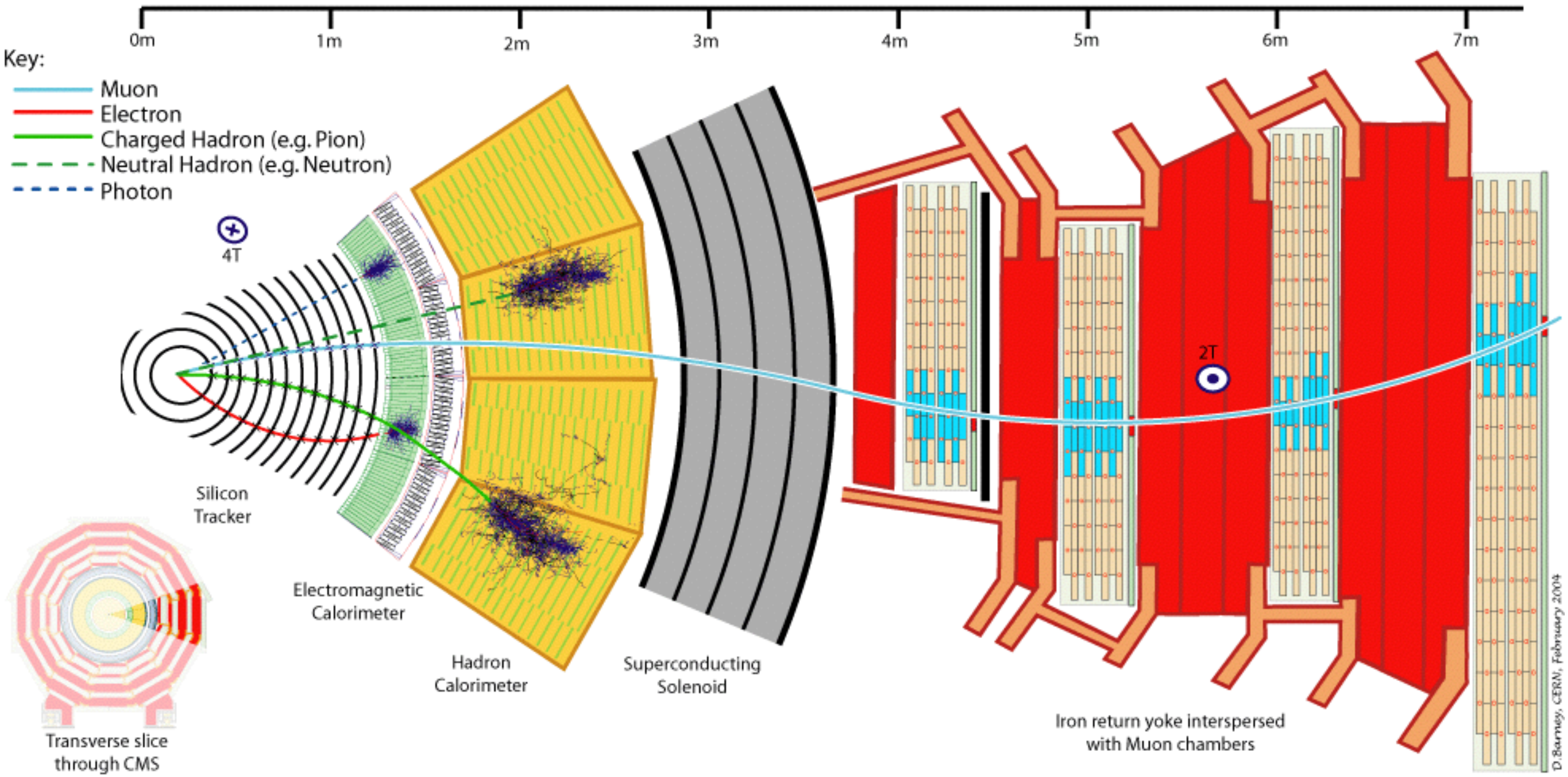


LHC delivers proton-proton collisions (**events**)
at a rate of 40 million per second

CMS takes "snapshots" of each one

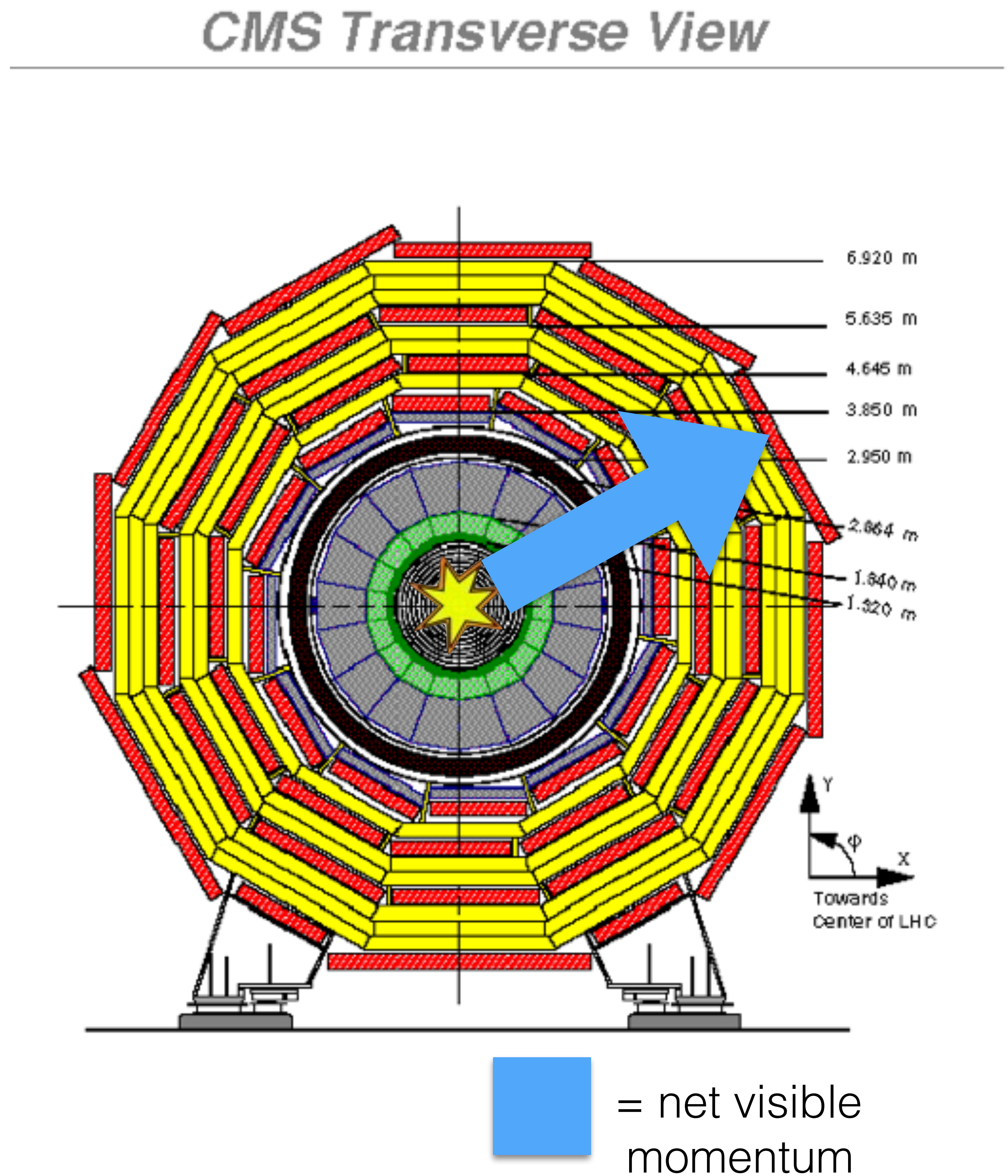
Triggering system throws out > 99.99%

How CMS Detects Particles



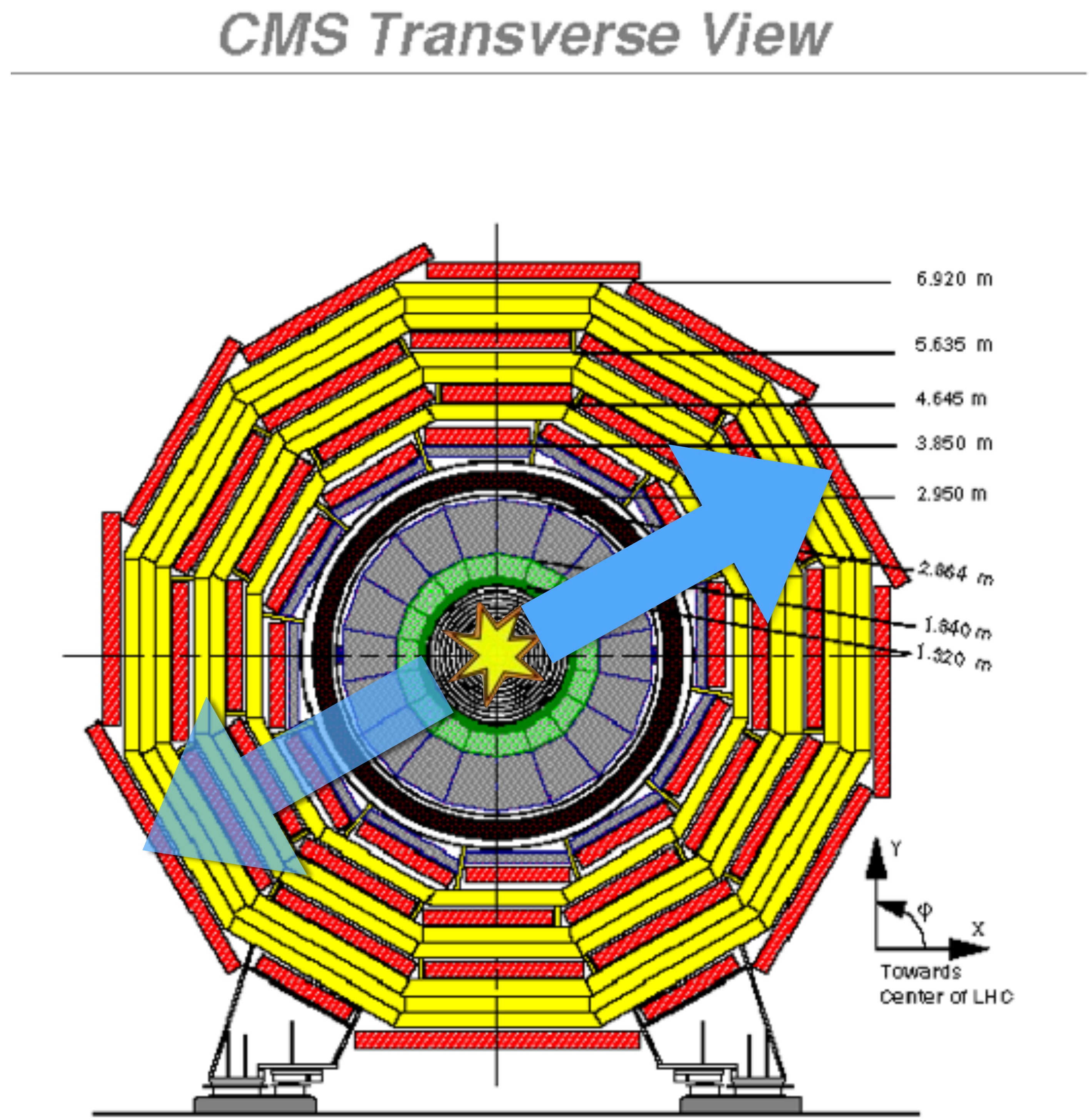
Measuring neutrinos

- Neutrinos go through everything
- Protons start with zero "transverse" momentum (**pT**)
- What if we see net visible momentum in the transverse direction?
- There must be something missing with net "invisible"momentum to cancel it out
- We call this "missing transverse energy" (**MET**)



Measuring neutrinos

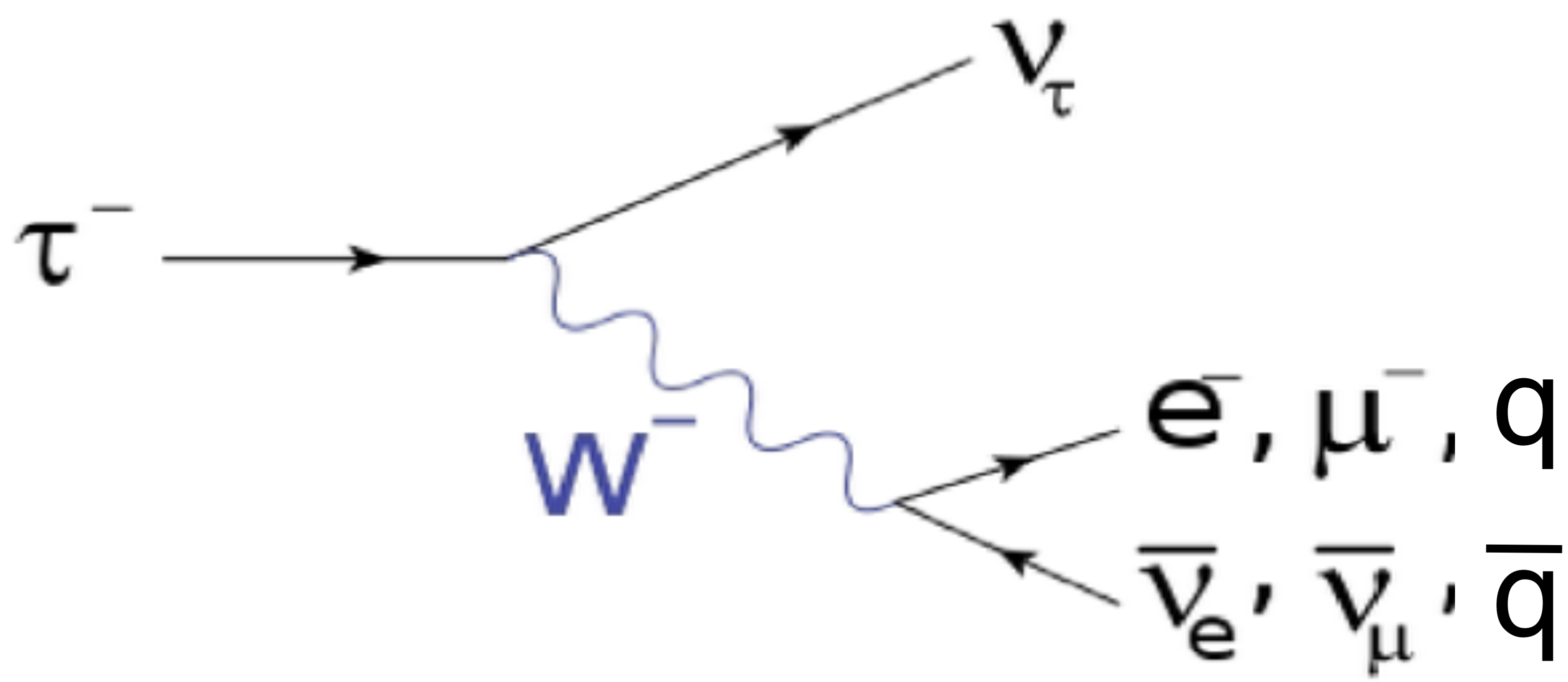
- Neutrinos go through everything
- Protons start with zero "transverse" momentum (**pT**)
- What if we see net visible momentum in the transverse direction?
- There must be something missing with net "invisible"momentum to cancel it out
- We call this "missing transverse energy" (**MET**)



= net invisible momentum

= net visible momentum

τ Physics

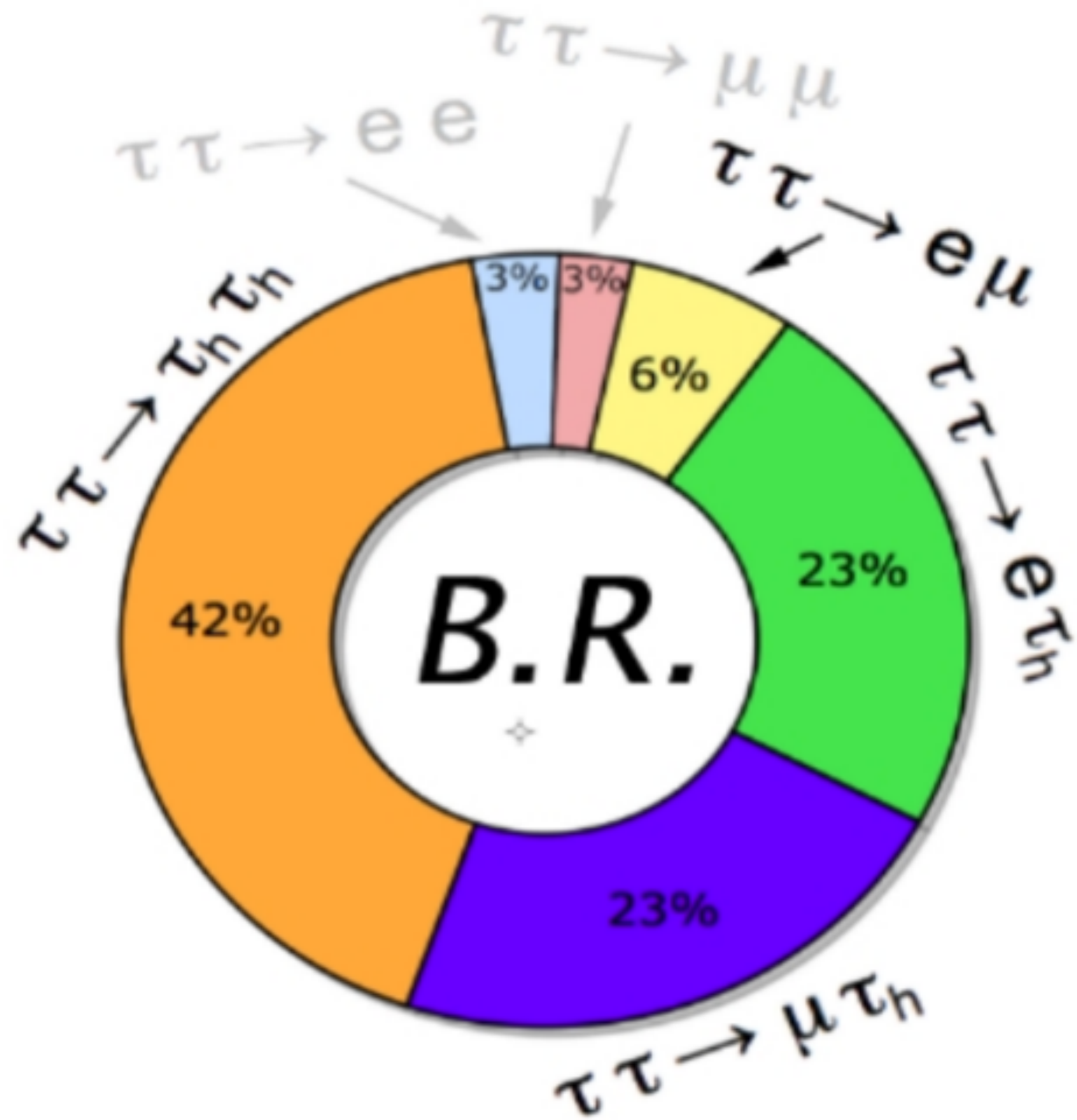


- The **Tau** is a third-generation lepton
- Spin 1/2, charge -1e, M = 1.78 GeV
- Lifetime = 2.9×10^{-13} s
- Three decay channels: τ_e, τ_μ, τ_h (hadrons)
- Each tau decay produces 1 or 2 neutrinos (MET)

13 TeV Z' → ττ Search

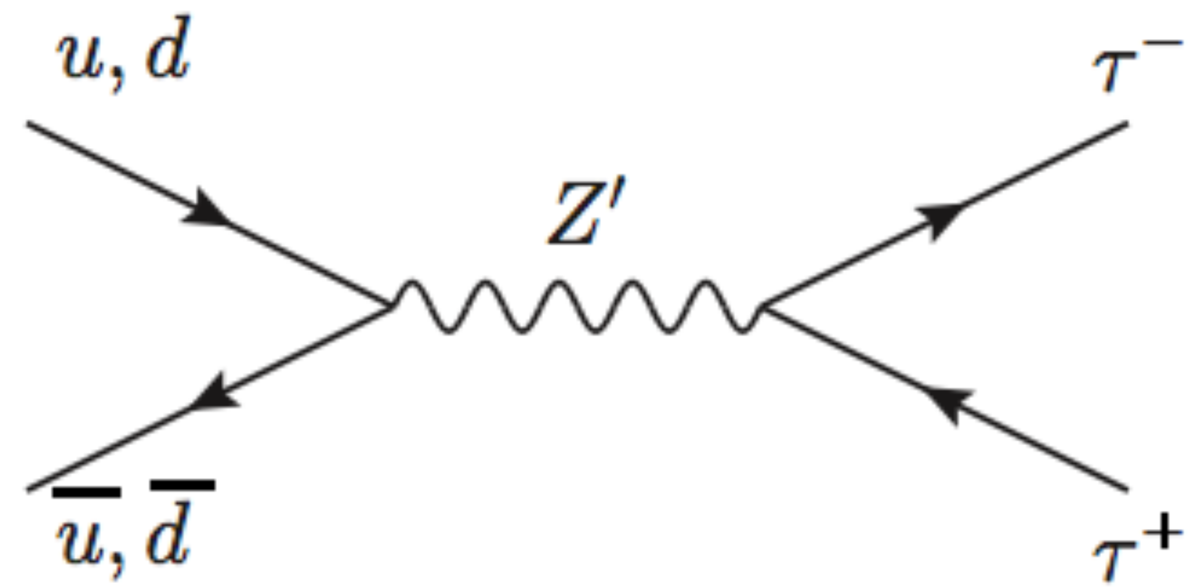
τ Physics

- Four channels considered in Z'->ττ analysis
  - $\tau_e\tau_\mu$, $\tau_e\tau_h$, $\tau_\mu\tau_h$, $\tau_h\tau_h$
  - 2-4 missing neutrinos in each channel
- "τ_h" denotes hadronic decay
- ee, μμ channels not explored due to large background from Z->ee/μμ decay
- $\tau_e\tau_\mu$ is "cleanest"



Analysis Strategy

- Look for events with two τ candidates
  - High pT (momentum)
  - Back-to-back (in transverse plane)
  - Opposite sign
  - Lots of missing energy (MET)
- Use Sequential Standard Model $Z' \to \tau\tau$ as benchmark
  - Identical to SM Z, except heavier



Selection cuts optimized for best signal-to-background ratio: $s/\sqrt{s+b}$

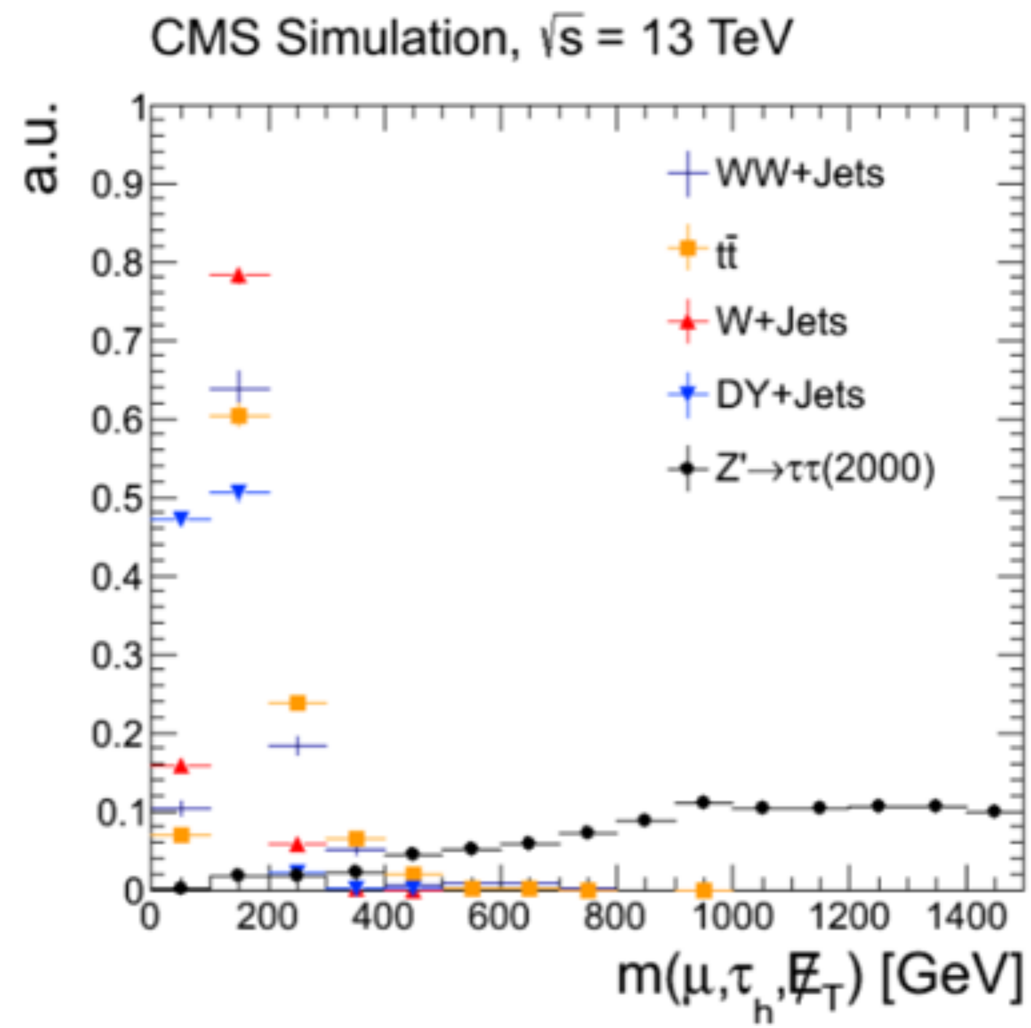
Analysis Strategy

Reconstructed mass definition:

$$M(\tau_1, \tau_2, \not{E}_T) = \sqrt{(E_{\tau_1} + E_{\tau_2} + \not{E}_T)^2 - (\overrightarrow{p_{\tau_1}} + \overrightarrow{p_{\tau_2}} + \overrightarrow{\not{E}_T})^2}$$

Invariant mass of all visible components + MET

Search most sensitive in high-reconstructed-mass region



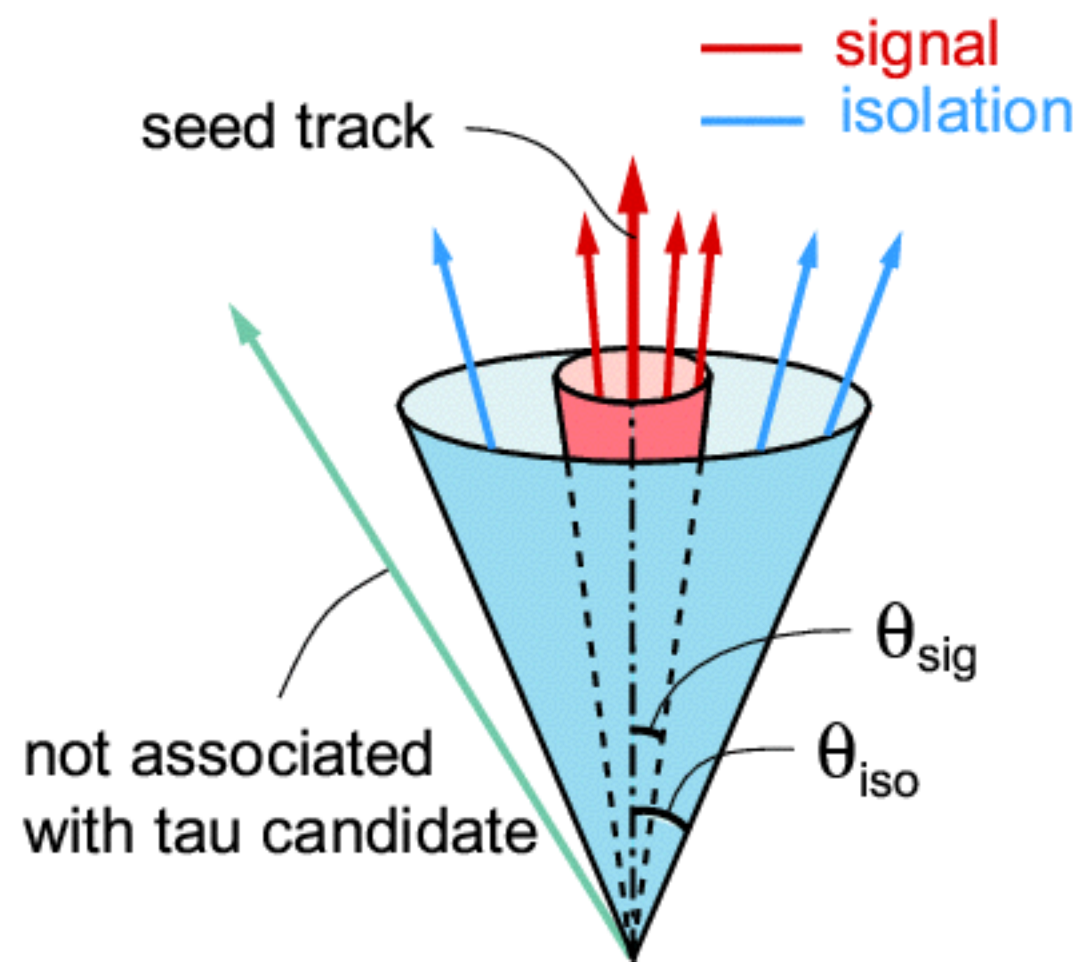
Signal shape broad, peaked at high mass

Analysis Strategy

- Blind signal selection region (SR) to minimize bias
- Estimate background contributions via data-driven analysis
- Unblind and study reconstructed mass distributions
- Look for excesses/set limits

Object Selection

- Want clean, well-reconstructed electrons, muons, and hadronic taus
- CMS Physics Object Groups (POGs) recommend particle ID criteria
- Apply pT, η cuts
- Isolation is a good tool for killing noisy background



Object Selection

Electrons:

pT > 35 GeV ($\tau_e\tau_h$ and $\tau_e\tau_\mu$)

Within detector acceptance ($|\eta| < 2.1$)

Relative Isolation < 0.15

"loose" ID criteria

Muons:

pT > 30 GeV ($\tau_\mu\tau_h$), pT > 10 GeV ($\tau_e\tau_\mu$)

Within detector acceptance ($|\eta| < 2.1$)

Relative Isolation < 0.15

"medium" ID criteria

Object Selection

Hadronic Taus (τ_h):

pT > 60 GeV ($\tau_h\tau_h$), pT > 20 GeV ($\tau_e\tau_h$, $\tau_\mu\tau_h$)

Within detector acceptance ($|\eta| < 2.1$)

1 or 3 charged hadrons ("prongs")

MVA-based muon, electron vetoes

"Tight" isolation recommended for τ_h ID

Event Selection

Topological cuts:

Opposite sign $\tau^+\tau^-$ pair

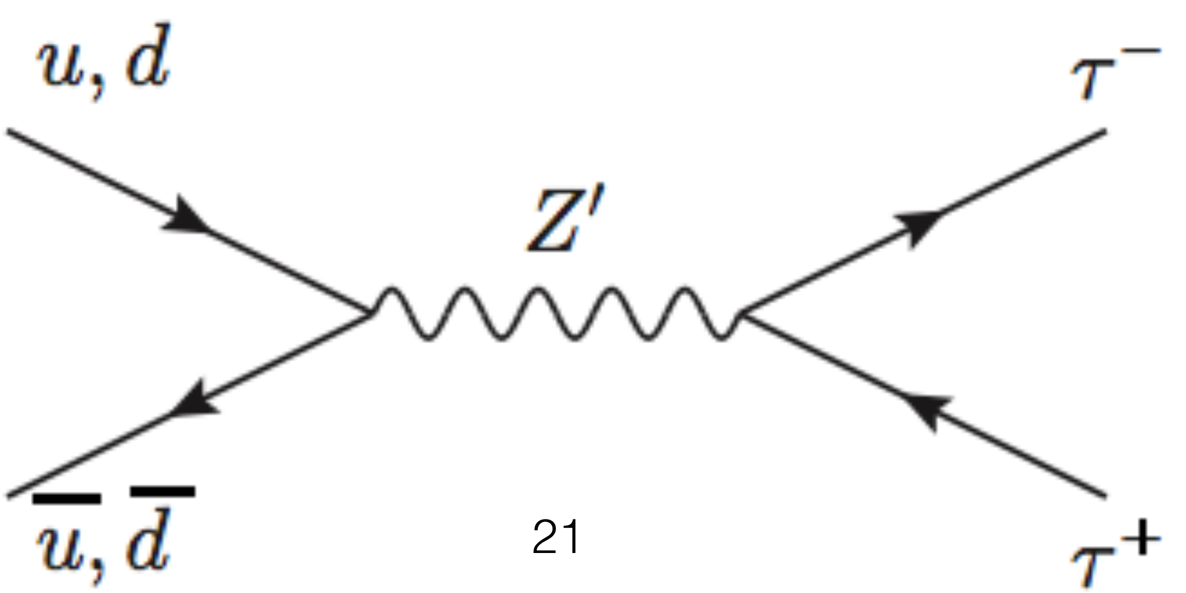
$\cos \Delta\phi(\tau_1, \tau_2) < -0.95$

MET > 30 GeV

$p_\zeta - 3.1 * p_{\zeta,vis} > -50$

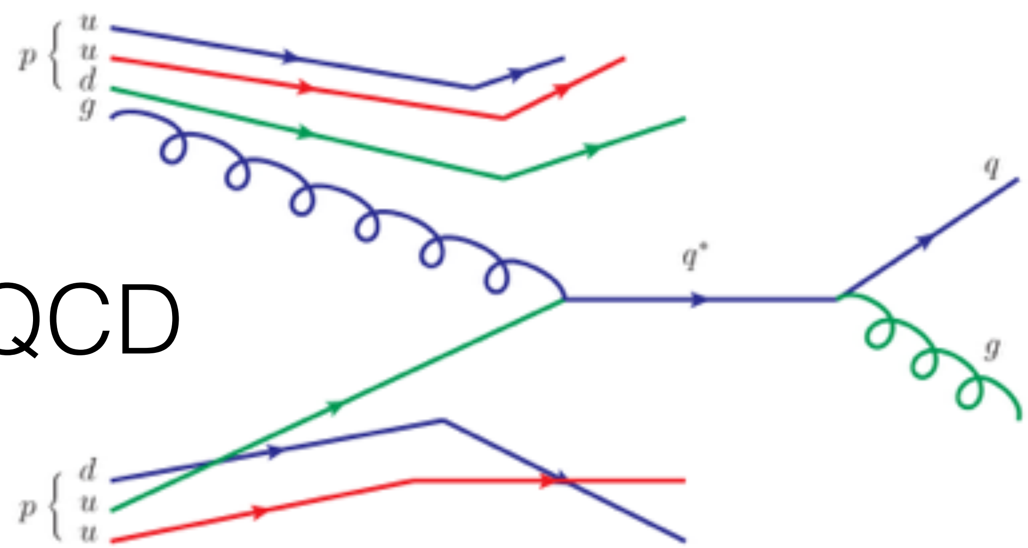
0 jets from b quark decays

Designed to select back-to-back tau pairs

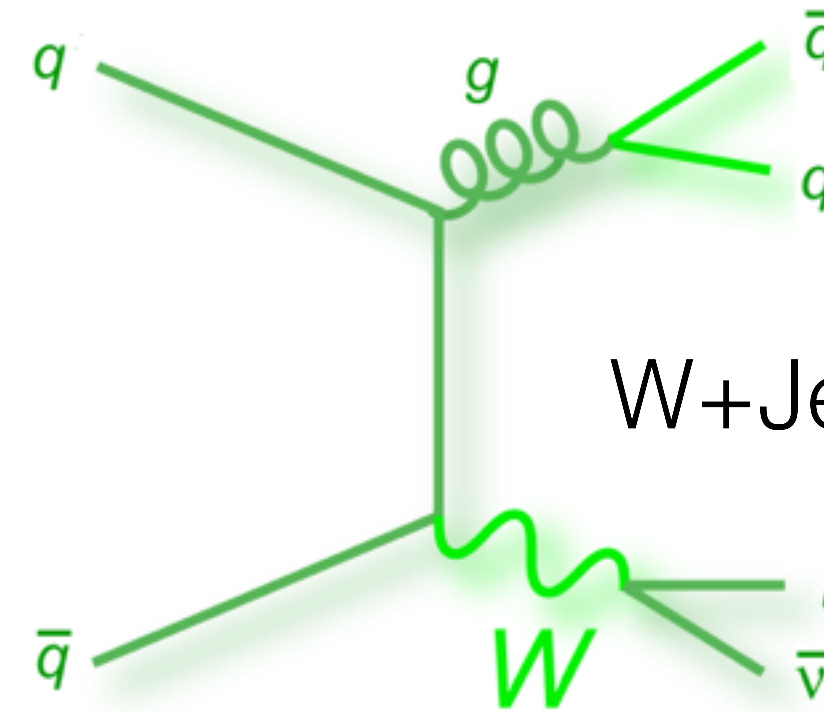


Background Estimation

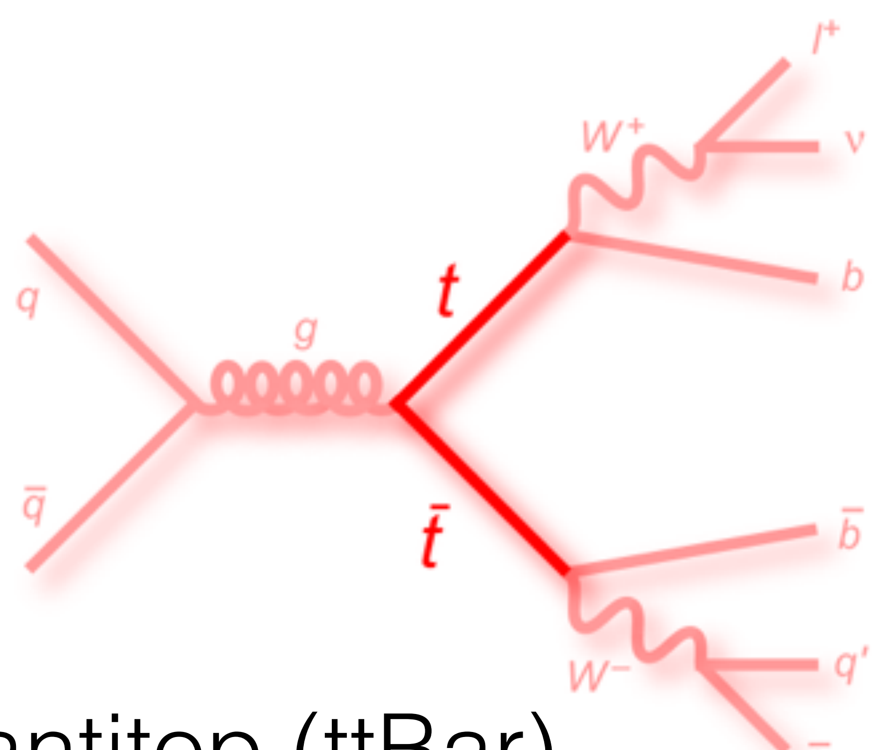
Main Backgrounds



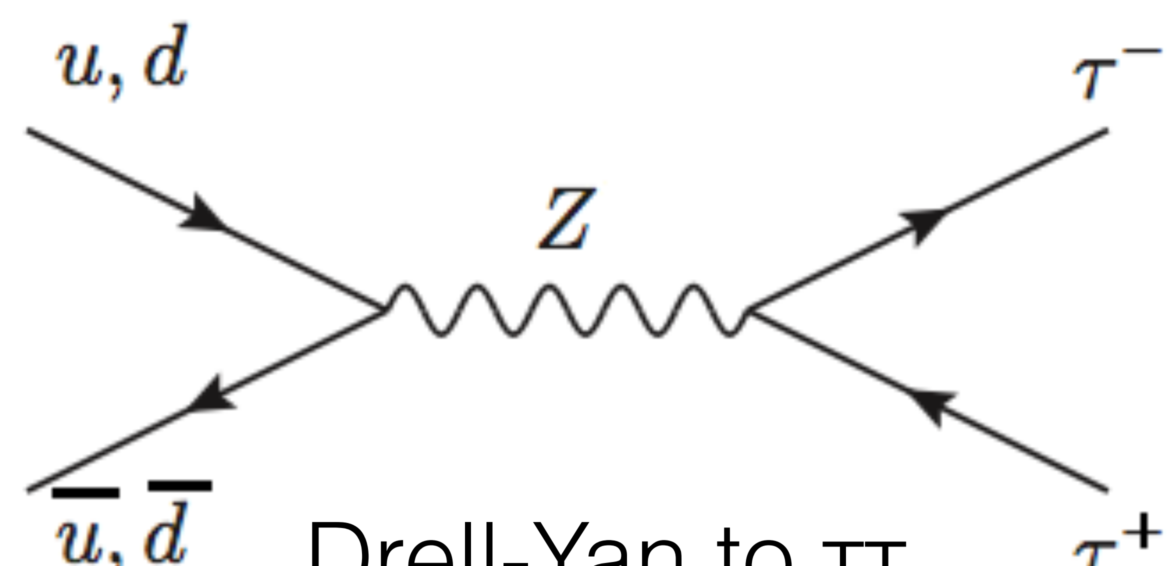
QCD



W+Jets



top-antitop (ttBar)

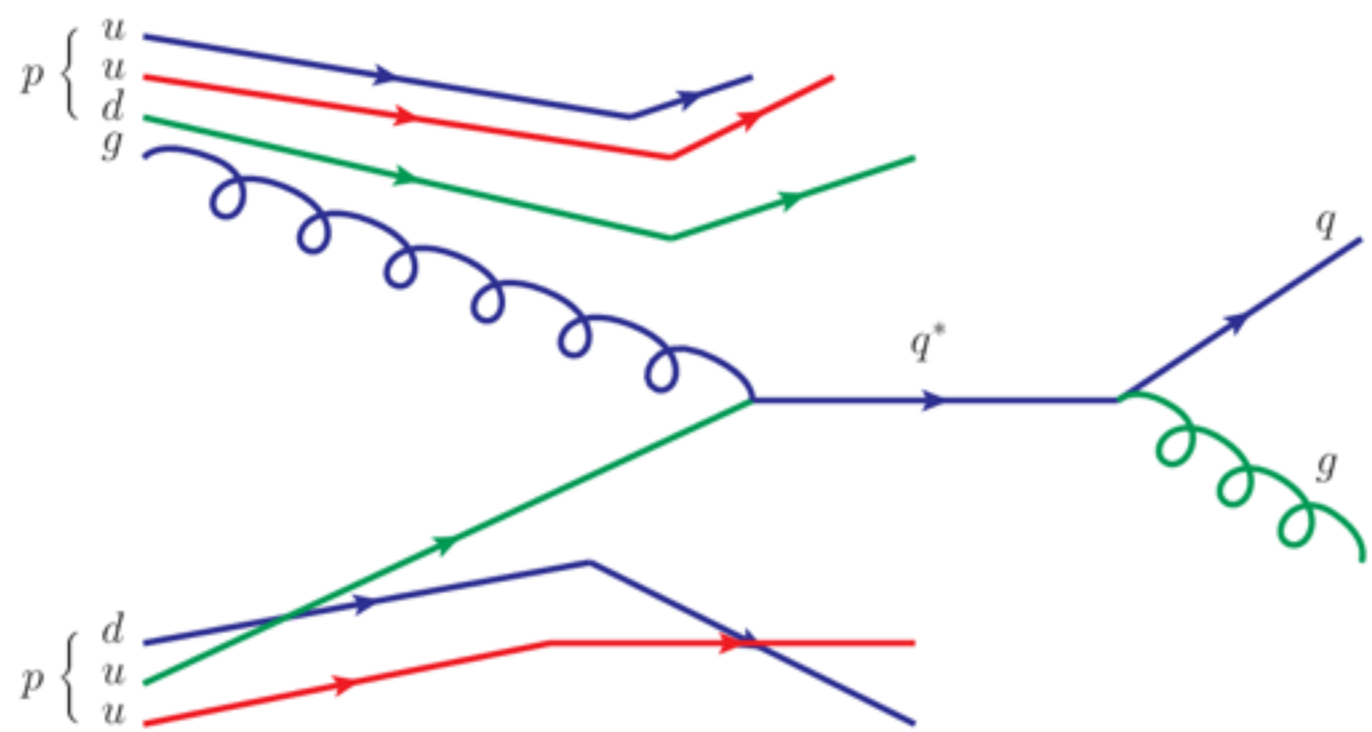


Drell-Yan to ττ

QCD

Quarks and gluons hadronizing into jets

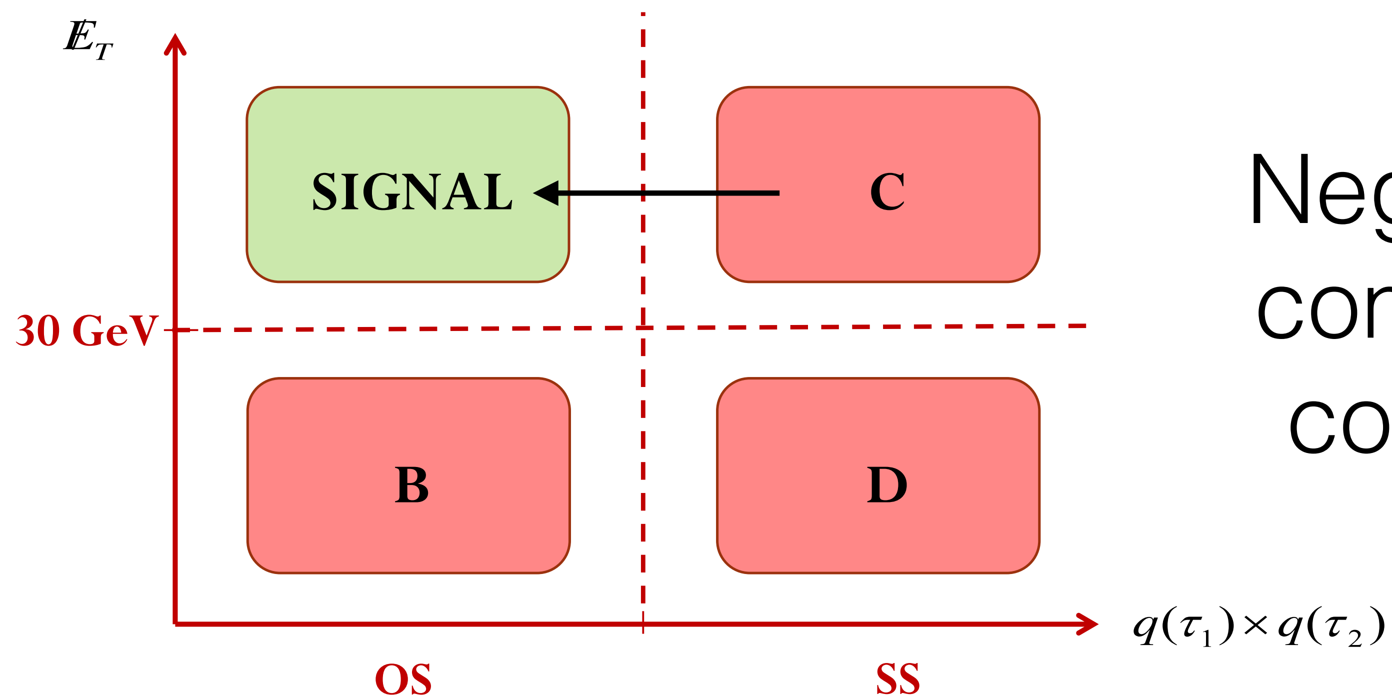
- Dominant (~85%) BG in $\tau_h\tau_h$
- ~20% of BG in $\tau_e\tau_h$
- ~5% of BG in $\tau_e\tau_\mu$
- ~6% of BG in $\tau_\mu\tau_h$
- all estimations data driven



QCD Estimation in $\tau_h\tau_h$

Data driven estimation using "ABCD" method

Take shape from SS region (**C**)



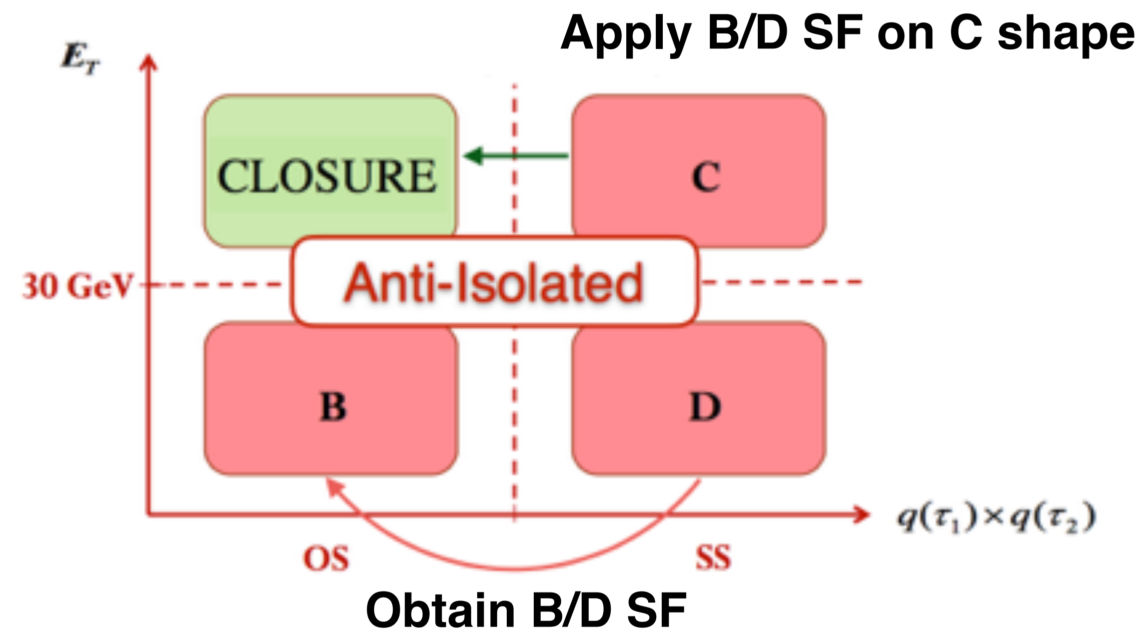
Negligible signal contamination in control regions (**CRs**)

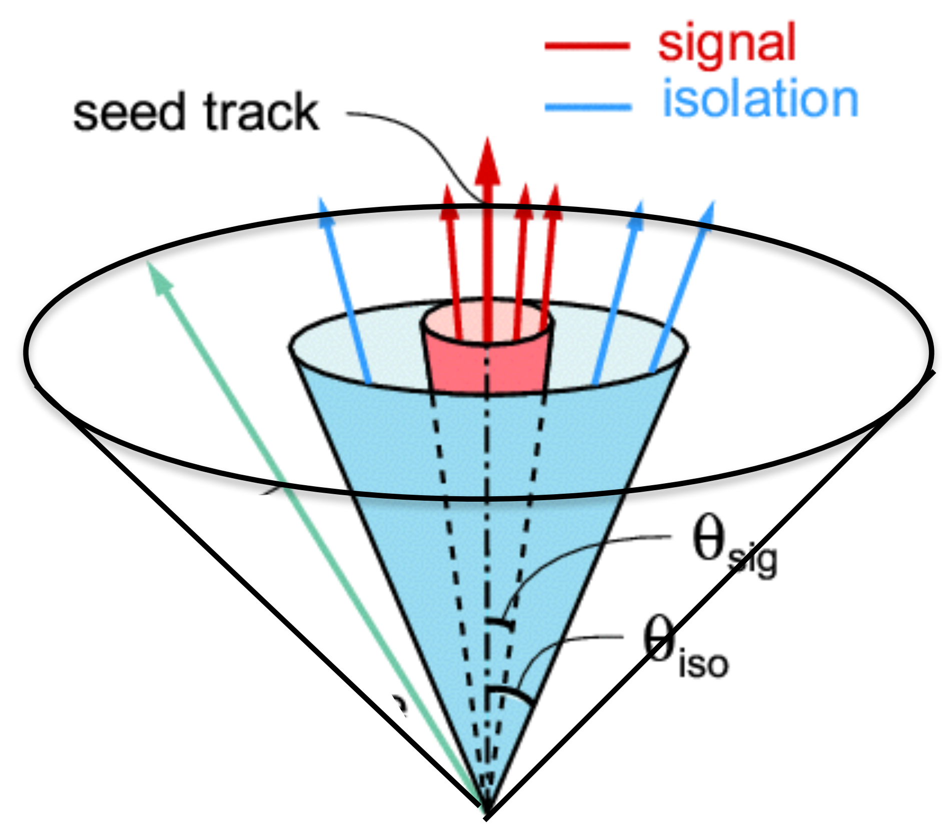
Derive OS/SS scale factor using low-MET CRs (**B**/**D**)

$$N_{QCD}^{Signal}(m_{\tau\tau\not{E}_T}) = N_{QCD}^{C}(m_{\tau\tau\not{E}_T})\times\frac{N_{QCD}^{B}}{N_{QCD}^{D}} = N_{QCD}^{C}(m_{\tau\tau\not{E}_T})\times R_{OS/SS}$$

QCD Closure in $\tau_h\tau_h$

- QCD MC for $\tau_h\tau_h$ has insufficient statistics
- Repeat ABCD estimation in **anti-isolated** regime where both τ_h's pass "loose" isolation but fail "tight"
- Check data/estimation in **CLOSURE** region



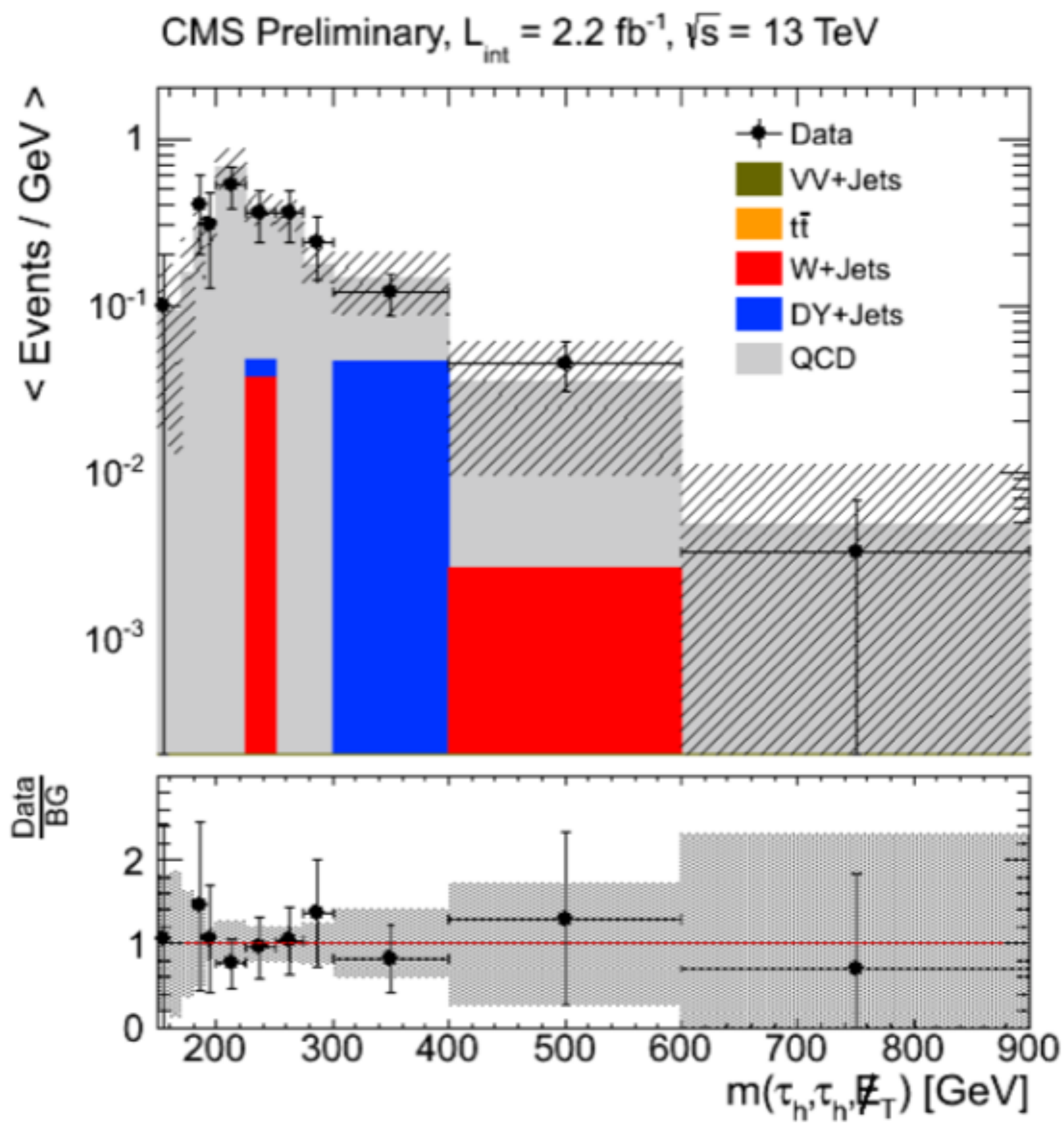


QCD Closure in $\tau_h\tau_h$

ABCD Estimation in CLOSURE: 427+-26

Data - non-QCD MC in CLOSURE: 429

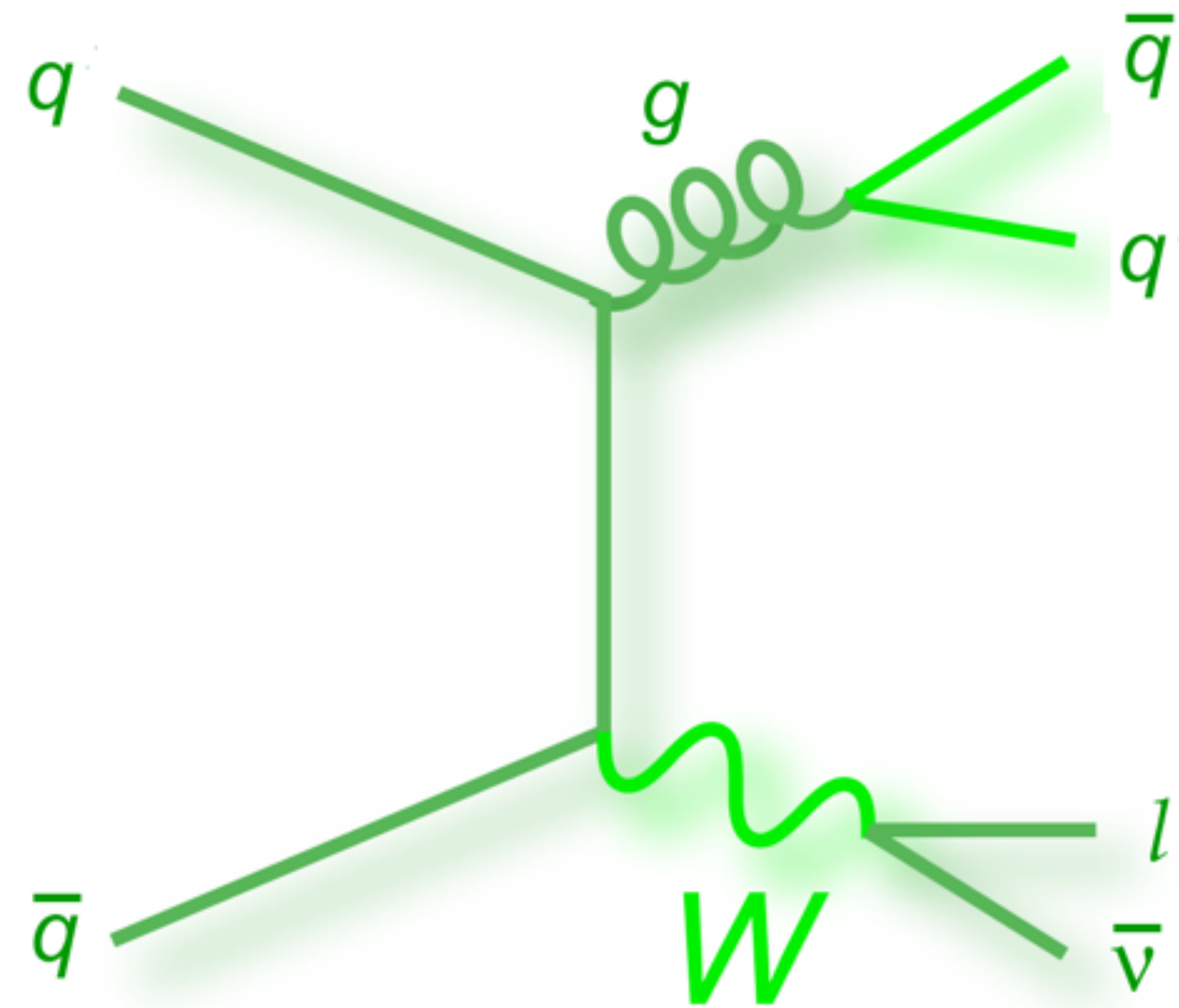
Good agreement!



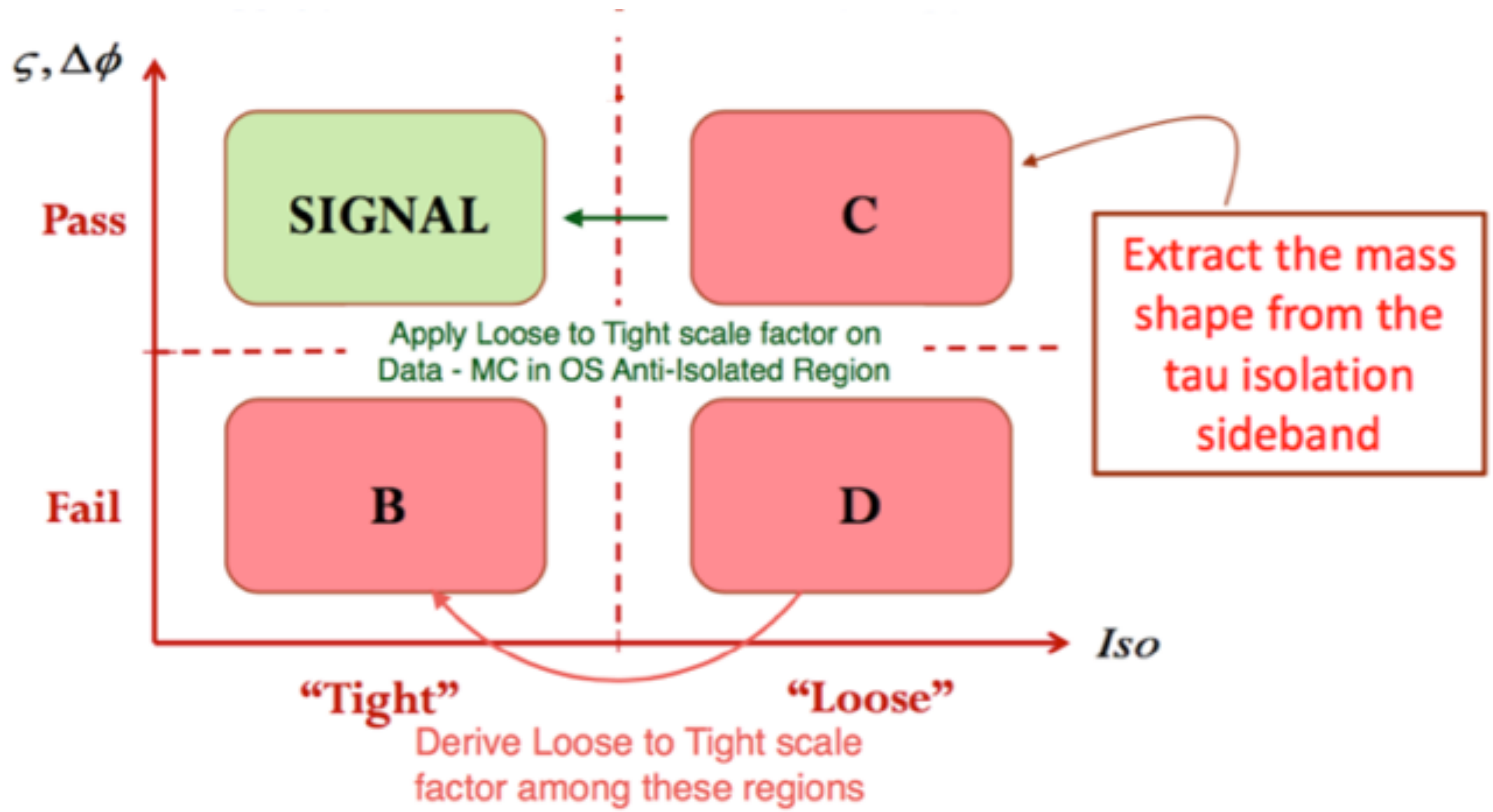
W+Jets Estimation

W faking leptonic tau, jet faking hadronic tau

- <1% of the background in $\tau_h\tau_h$ (MC based)
- ~40% of the background in $\tau_e\tau_h$ (data driven)
- ~3% of the background in $\tau_e\tau_\mu$ (MC based)
- ~45% of the background in $\tau_\mu\tau_h$ (data driven)



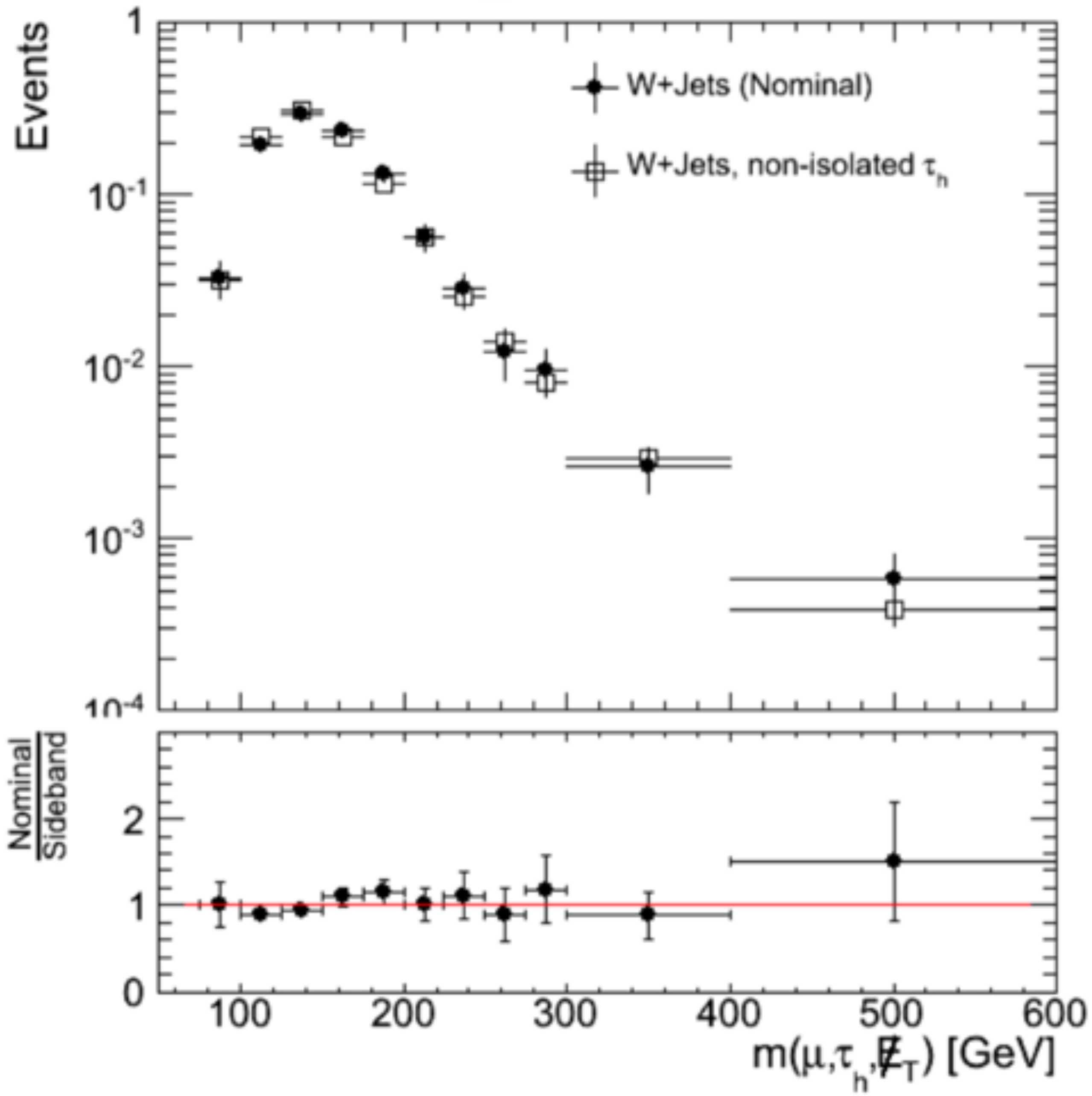
W+Jets Estimation in $\tau_\mu\tau_h$ and $\tau_e\tau_h$



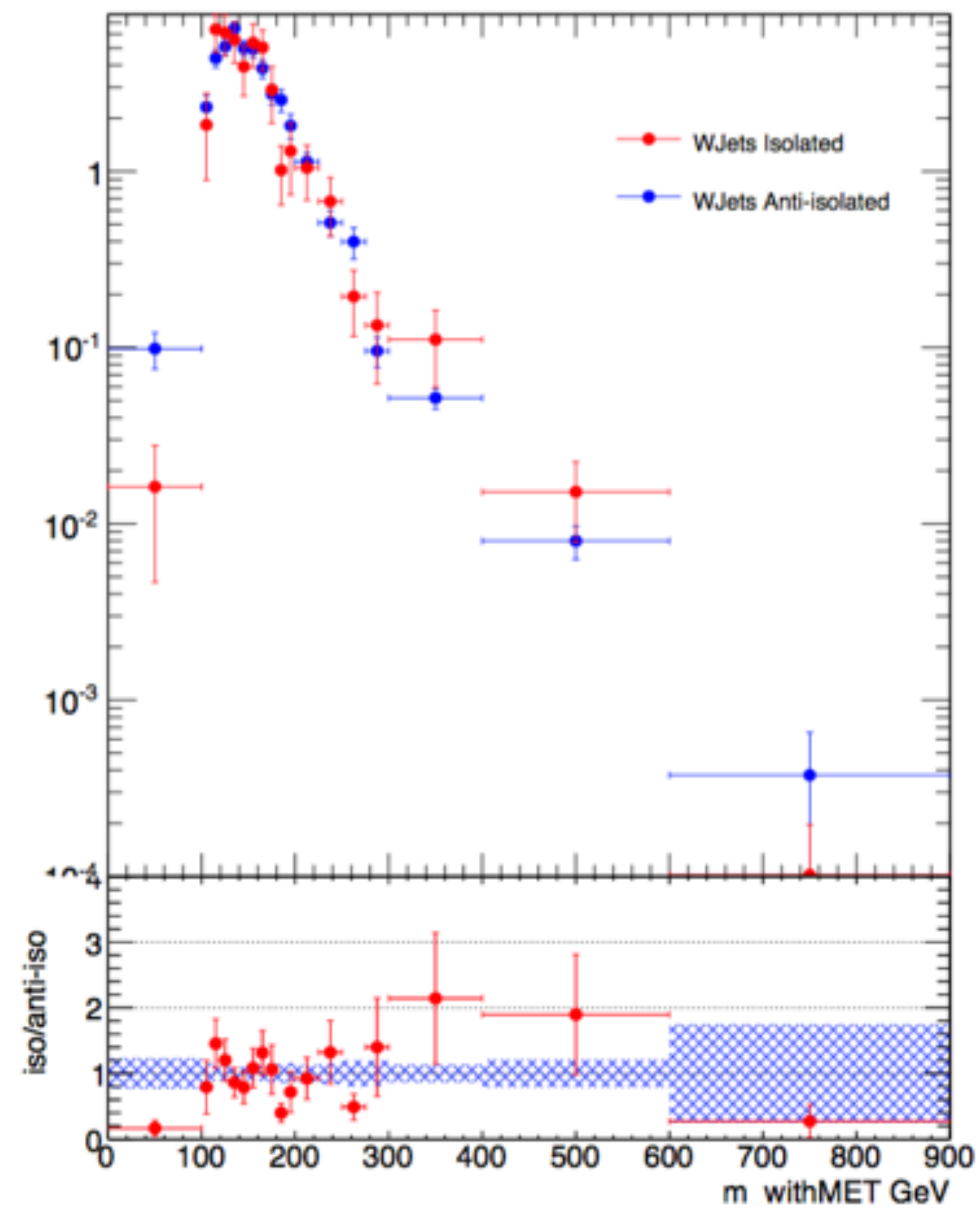
$$N^{\text{Signal}}_{\text{W+Jets}} = \frac{N^{\text{B}}_{\text{W+Jets}}}{N^{\text{D}}_{\text{W+Jets}}} \times N^{\text{C}}_{\text{W+Jets}}$$

W+Jets Closure in $\tau_\mu\tau_h$ and $\tau_e\tau_h$

$\tau_\mu\tau_h$



$\tau_e\tau_h$

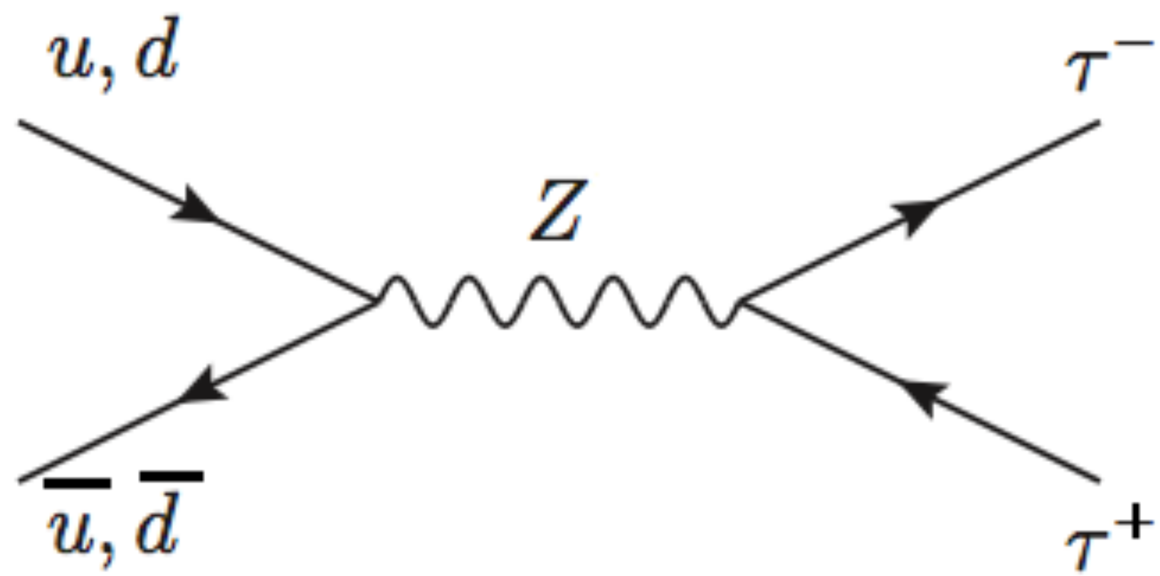


Mass shapes correctly modeled in loose iso sidebands

Drell-Yan Validation

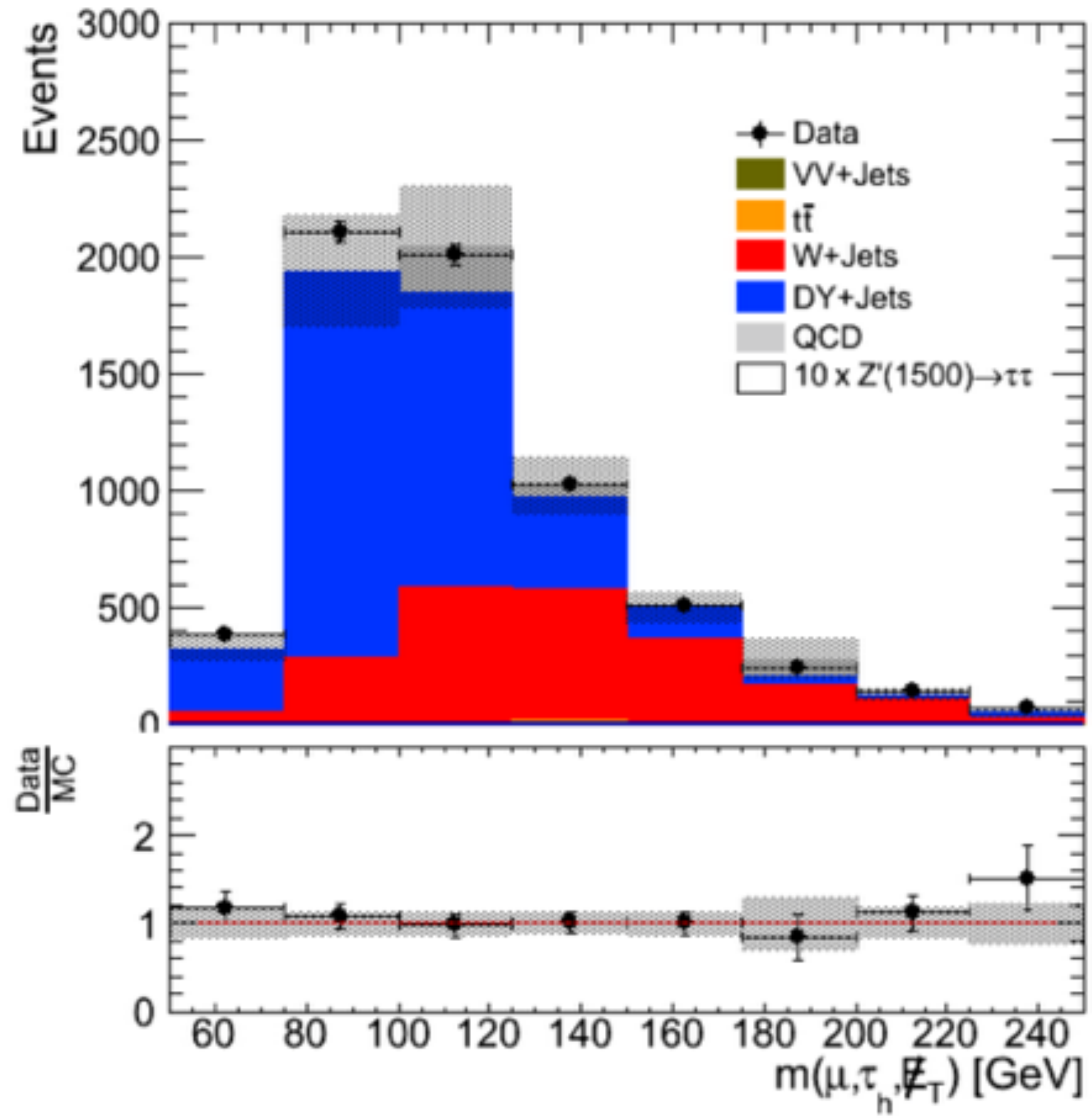
SM Z boson decaying to tau pairs

- ~15% BG in $\tau_h\tau_h$
- ~30% of BG in $\tau_e\tau_h$
- ~45% of BG in $\tau_e\tau_\mu$
- ~45% of BG in $\tau_\mu\tau_h$
- Mostly low-mass



DY Validation in $\tau_{\mu}\tau_{h}$

- Expect DY to be well-modeled by simulation
- Validate in low mass (< 300 GeV) region
- Data/MC consistent with unity
- Estimate DY from MC
- Take systematic uncertainty from validation (control) region



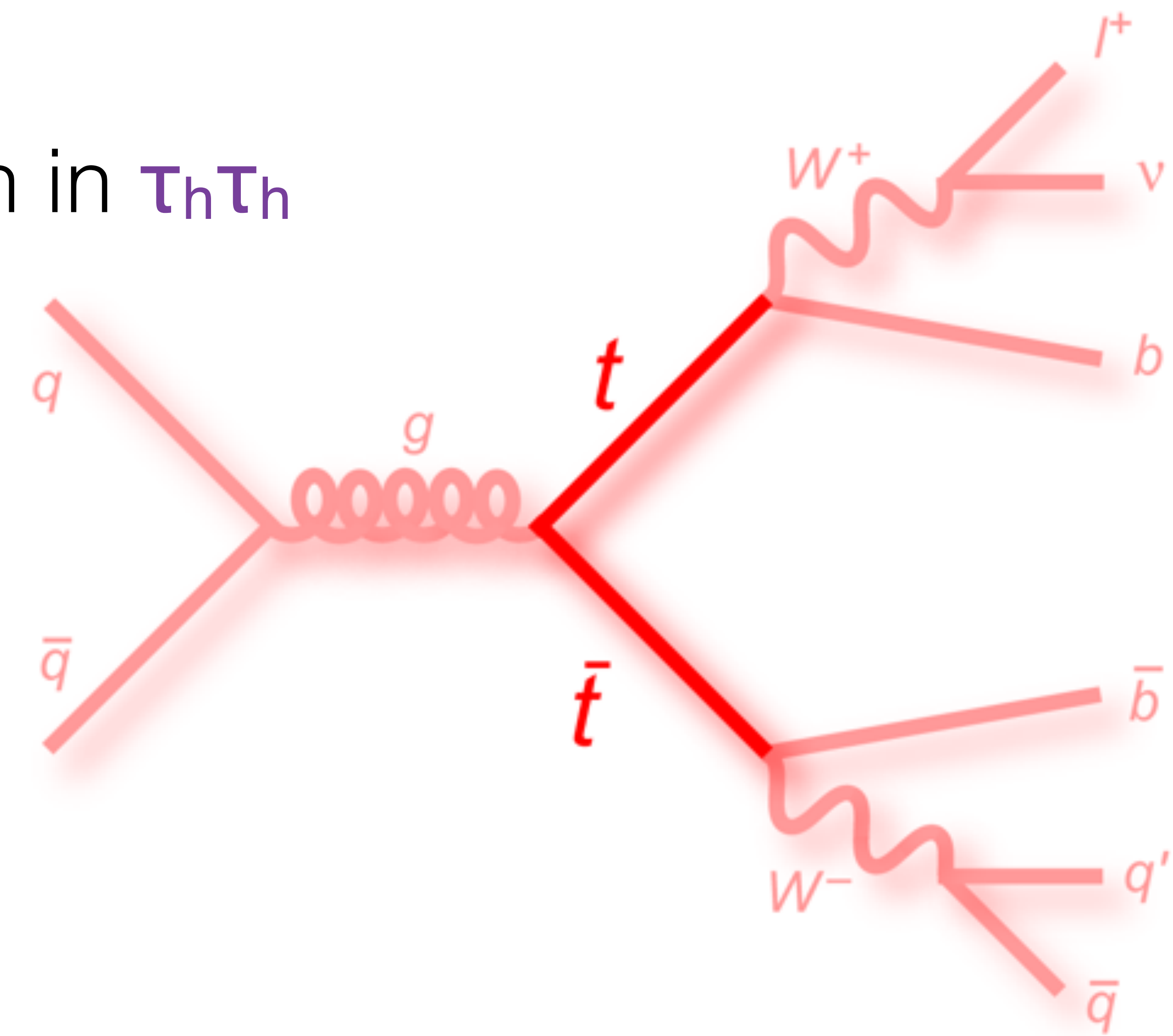
(Data - bkg)/(DY MC) = 0.96 +- 0.06

Systematic on DY estimation

ttBar Validation

W bosons from top quarks faking taus

- Negligible contribution in $\tau_h\tau_h$
- ~3% of BG in $\tau_e\tau_h$
- ~30% of BG in $\tau_e\tau_\mu$
- ~2% of BG in $\tau_\mu\tau_h$



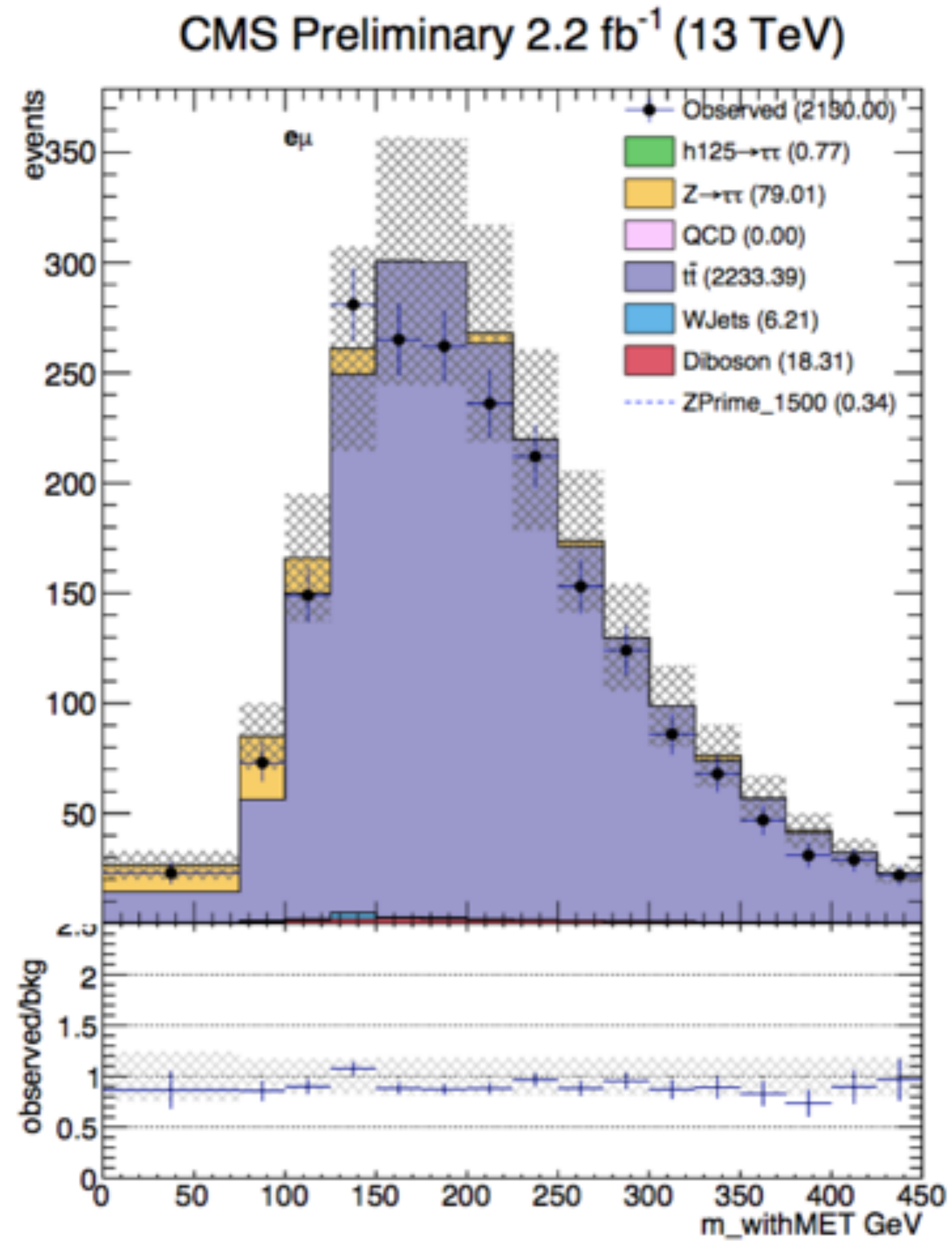
ttBar Validation in $\tau_e\tau_\mu$

- Shape and normalization taken from MC
- MC validated in b-rich CR
- Systematic extracted from CR

Control Region:

- Inverted p_ζ cut
- # b-jets > 0

Data/MC consistent with unity

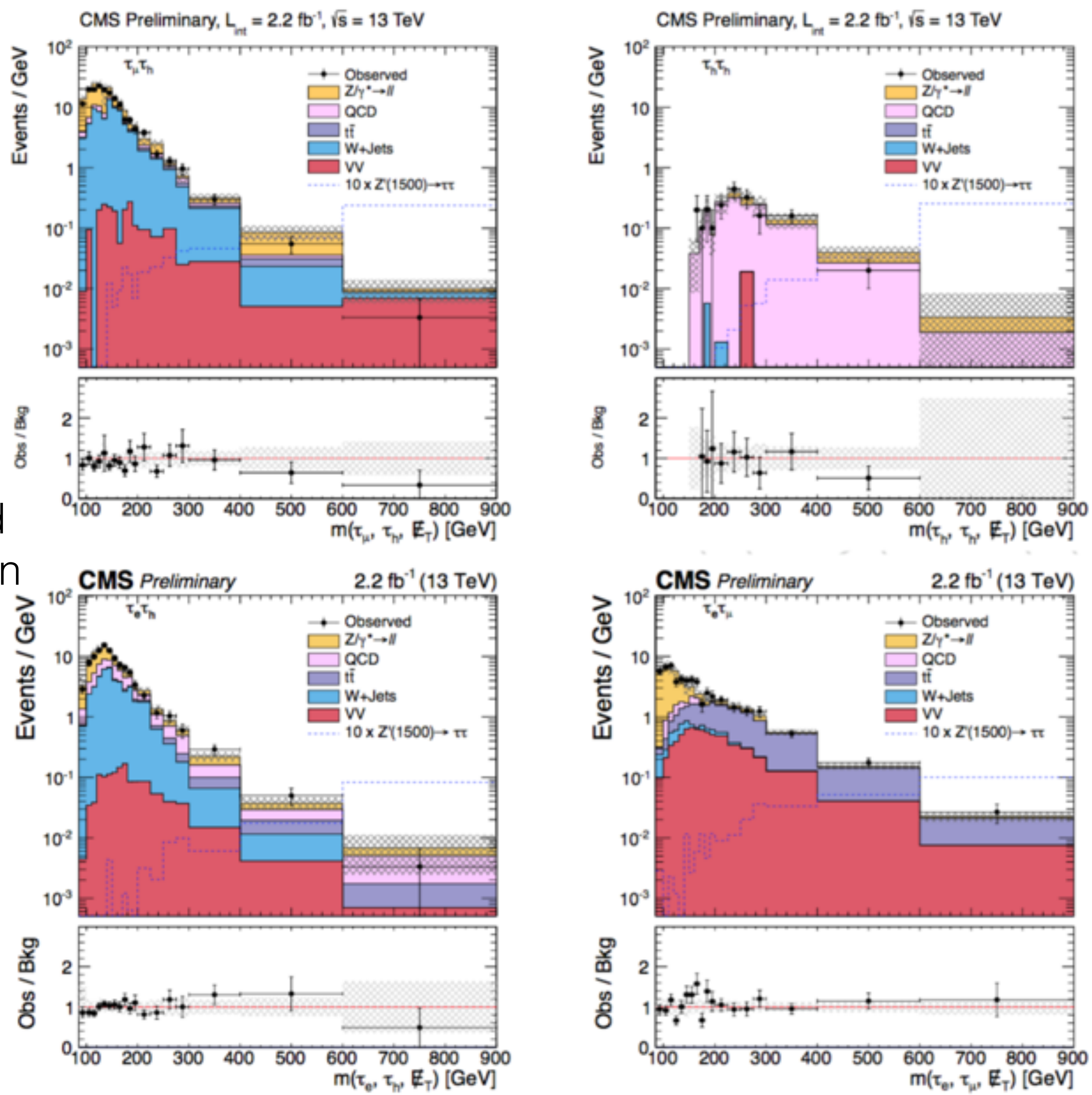


$$\text{ttbar Systematics} = |1 - \frac{\text{ttbar MC}}{\text{Data} - \text{other backgrounds}}| = 8\%$$

Results

Unblinded Results

Fair agreement between observation and SM expectation in all channels

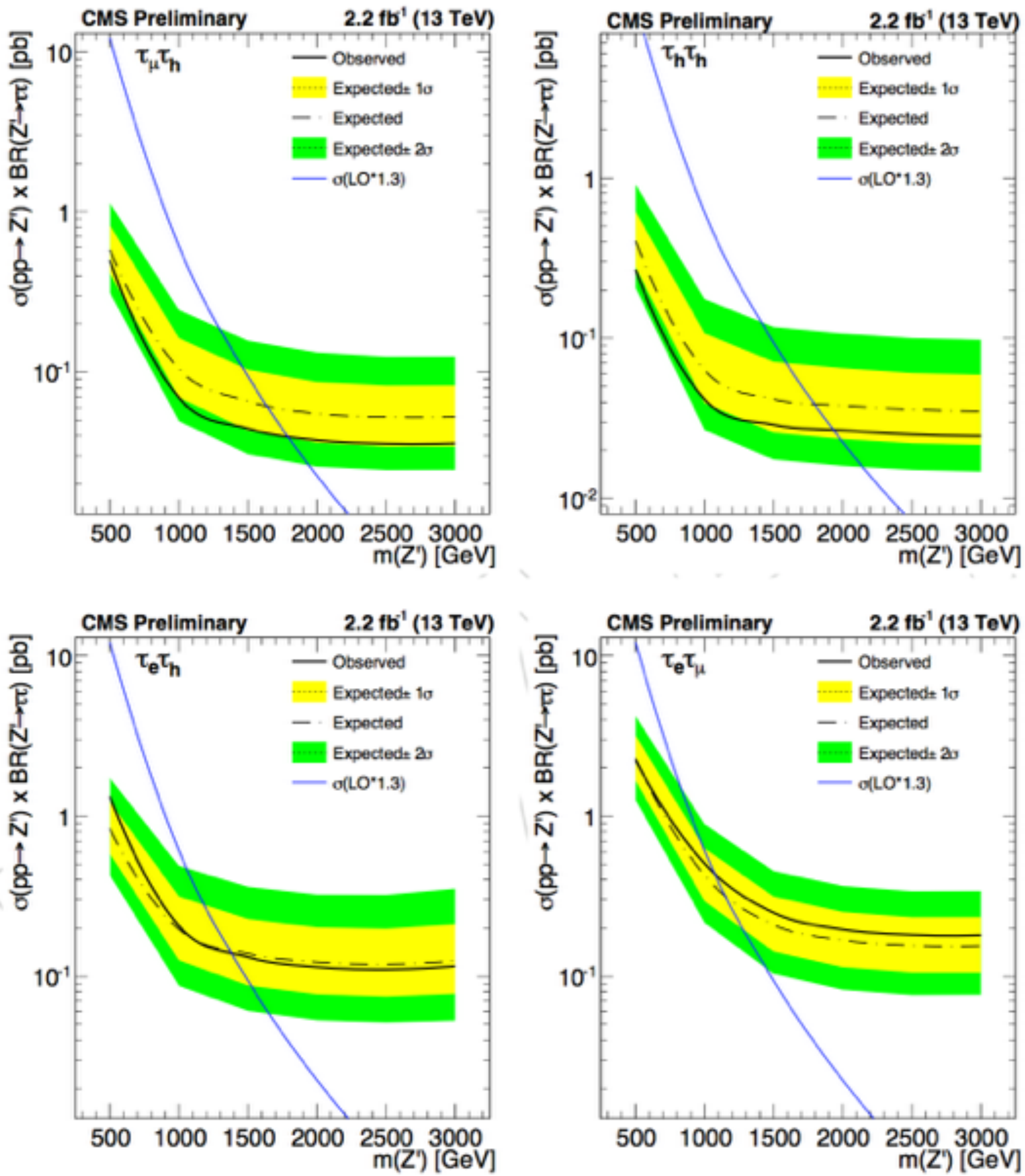


No excesses observed, so we set limits

Setting Limits

- Look at difference between data and background
  - See how much potential signal is “allowed”
- Do this for each bin of reconstructed mass distribution
  - Different Z’ masses have different distributions
- Comparison quantified by “Higgs combine” tool
  - Takes BG, signal, data rates as input
  - Performs *profile likelihood fit*, returns limit for each bin

Individual Channel Limits

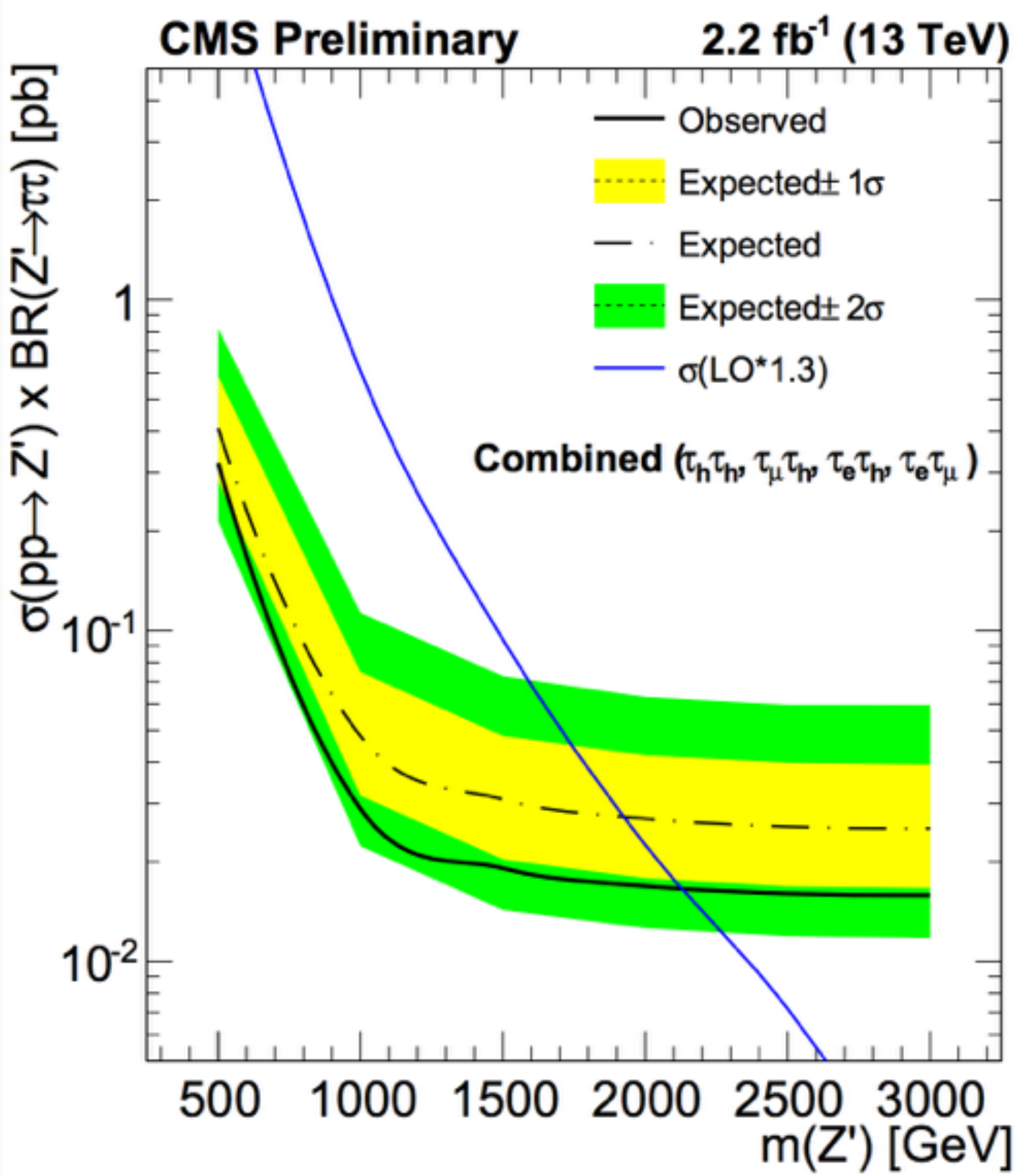


Blue line = theory

Exclude Z' at 95% confidence where observed limit < theory

$\tau_h\tau_h$ is main driver of overall limit

Combined Limit



8 TeV limit was 1.3 TeV
($\tau_e \tau_\mu$ only)

Exclude **m(Z') < 2.12 TeV**
(1.9 TeV expected)

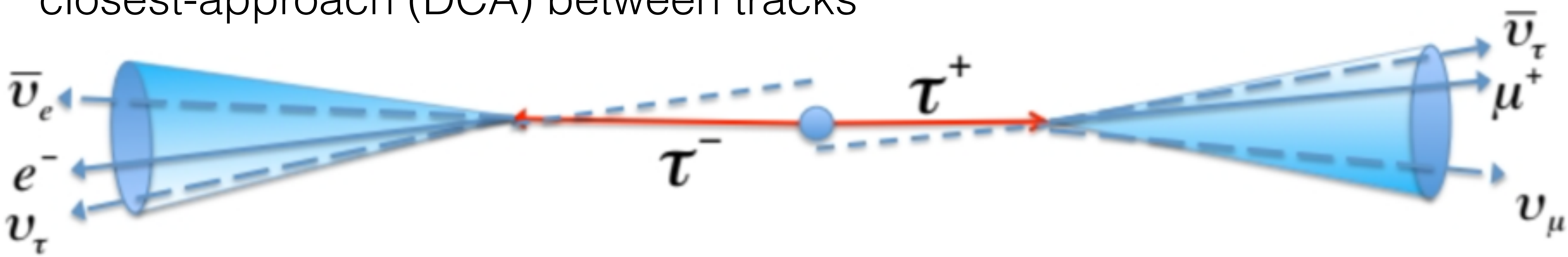
New record limit!!

Paper cleared for publication in JHEP this Wednesday

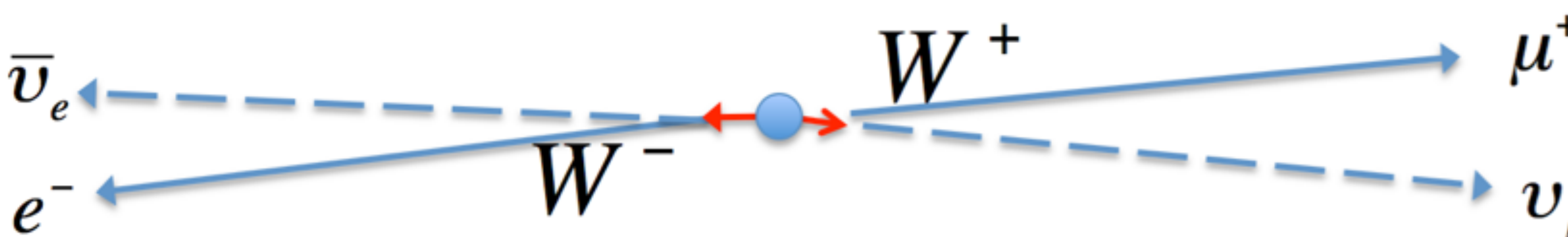
Tau Lifetime Studies

Tau Lifetime Studies

- Tau decay distance resolvable in the detector
- Electron, muon, pion tracks from diTau decays should have nonzero impact parameter (IP) with respect to beam spot, large distance-of-closest-approach (DCA) between tracks



- Most backgrounds in diTau analyses are prompt:



- Cutting on IP/DCA should be an effective discriminant

Impact Parameter

- Impact parameter is the transverse distance between the reconstructed track and the PV/beam spot

- We sum the e, μ, π± IPs, divide by error in order to suppress poorly-reconstructed tracks

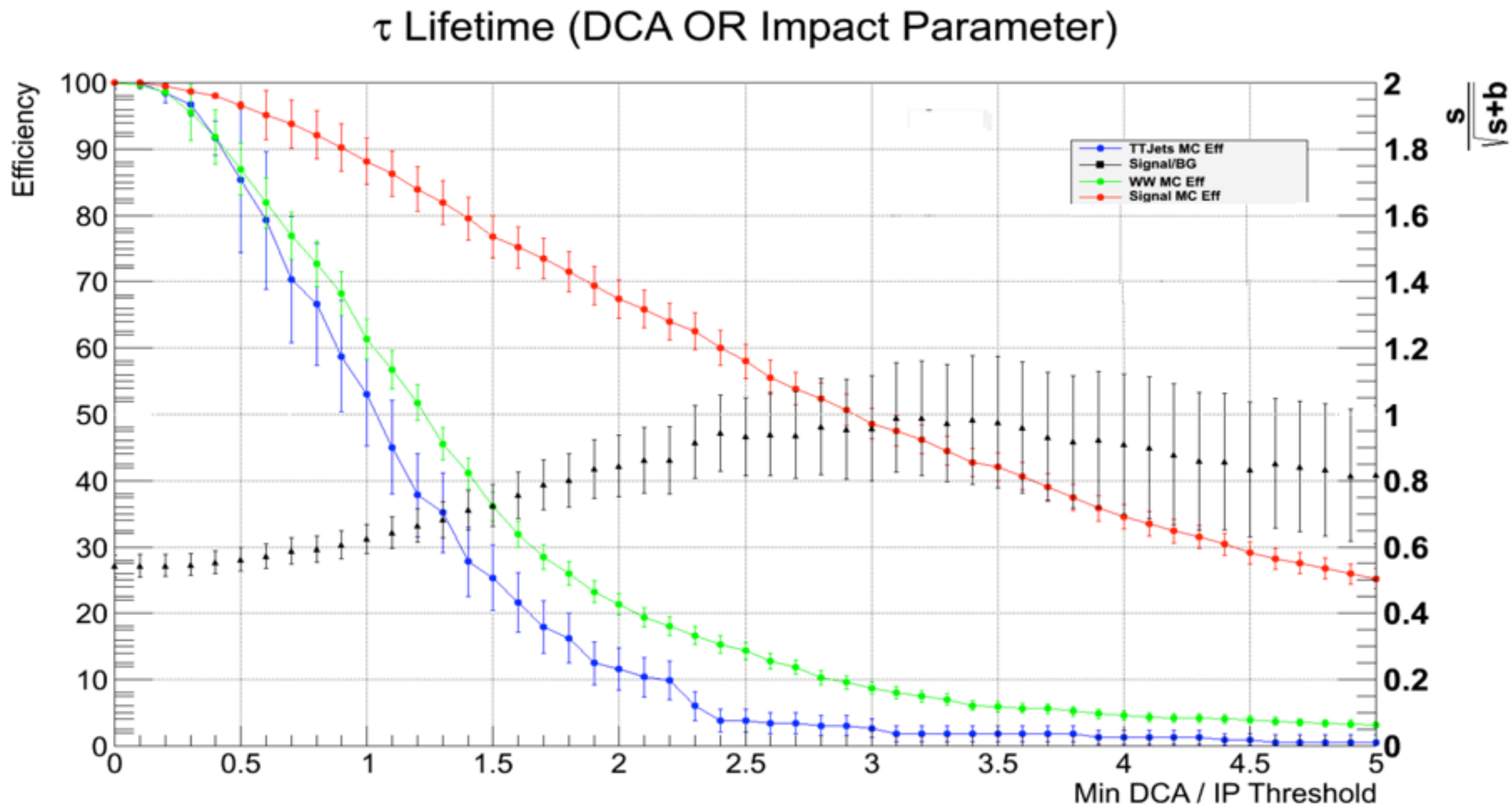
$$\sqrt{\frac{\left(|IP_e|+|IP_\mu|\right)^2}{\sigma^2_{IP_e}+\sigma^2_{IP_\mu}}}$$

- IP ~50 μm for signal, ~20 μm for prompt BG, σ_{IP} ~25 μm

Distance of Closest Approach

- DCA is the transverse distance between two tracks at their point of closest approach
- We also divide DCA by its error to suppress poorly-reconstructed tracks
- Not dependent on beam spot
- One variable instead of two!
- DCA ~70 μm for signal, ~30 μm for BGs, σ_{DCA} ~30 μm
- Best performance when both IP and DCA applied simultaneously

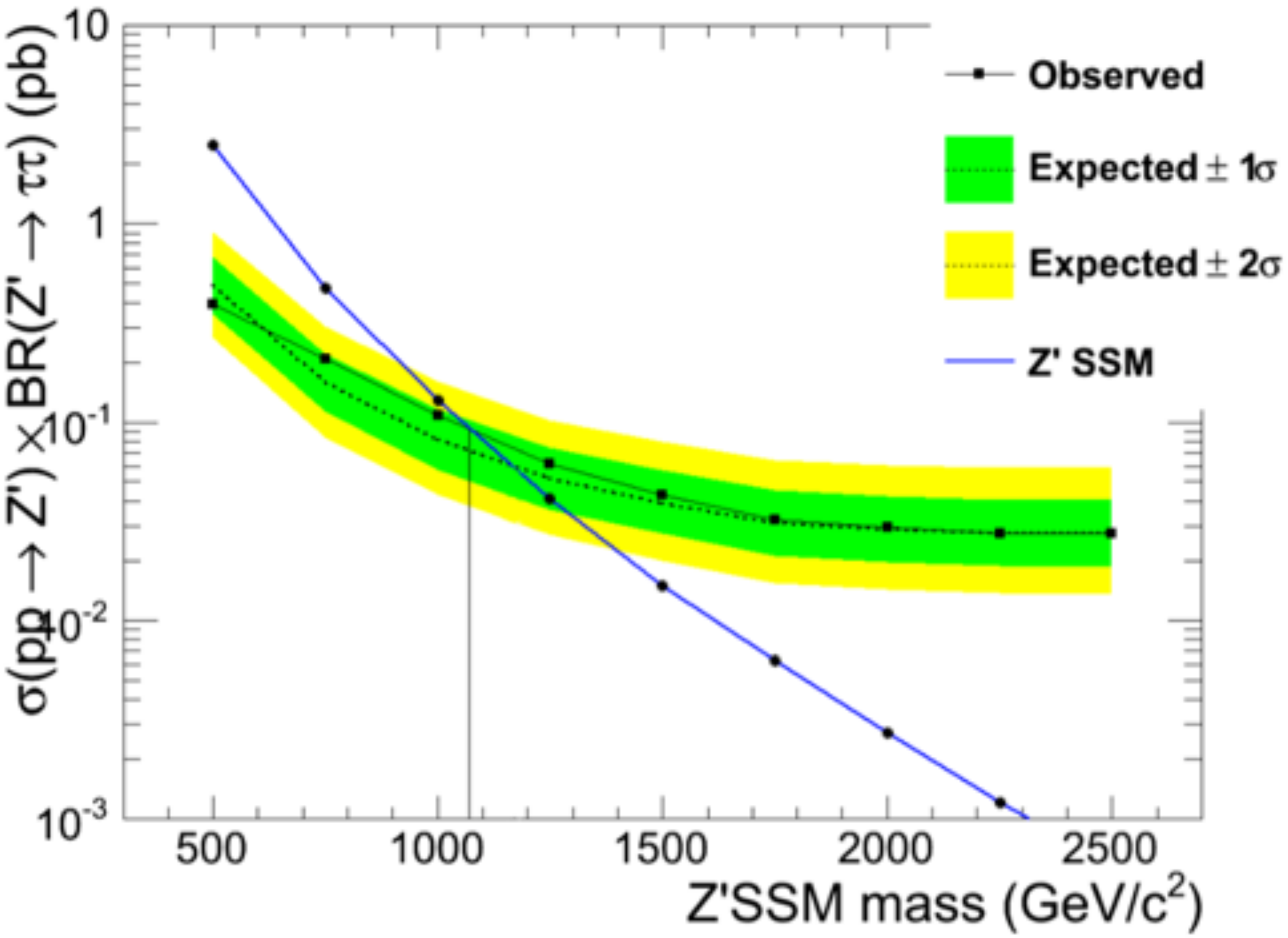
Performance of Lifetime Cuts



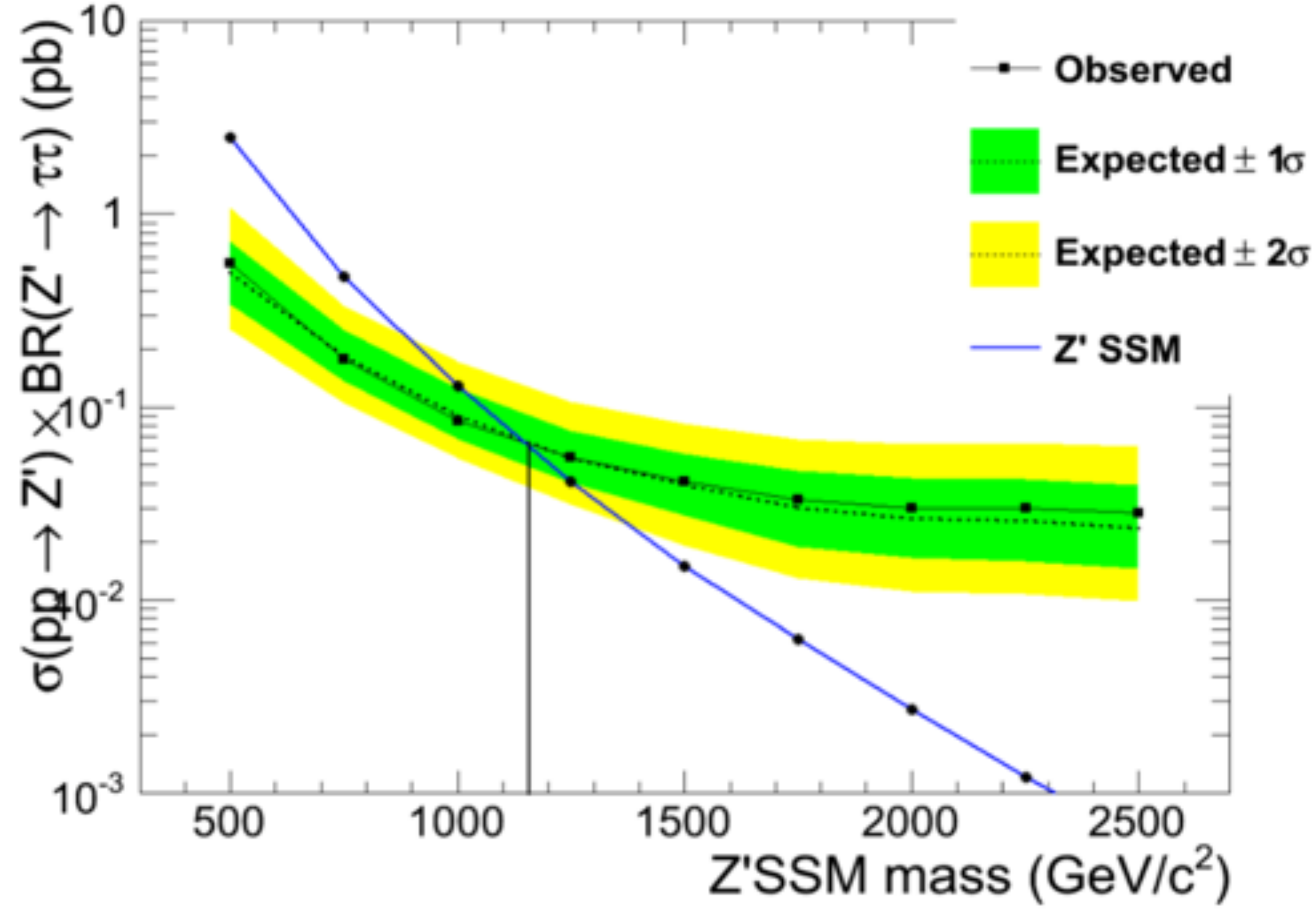
- Backgrounds show Gaussian behavior
- Background tails due to tau decays in MC

Impact on Limit: 8 TeV

Standard Limit



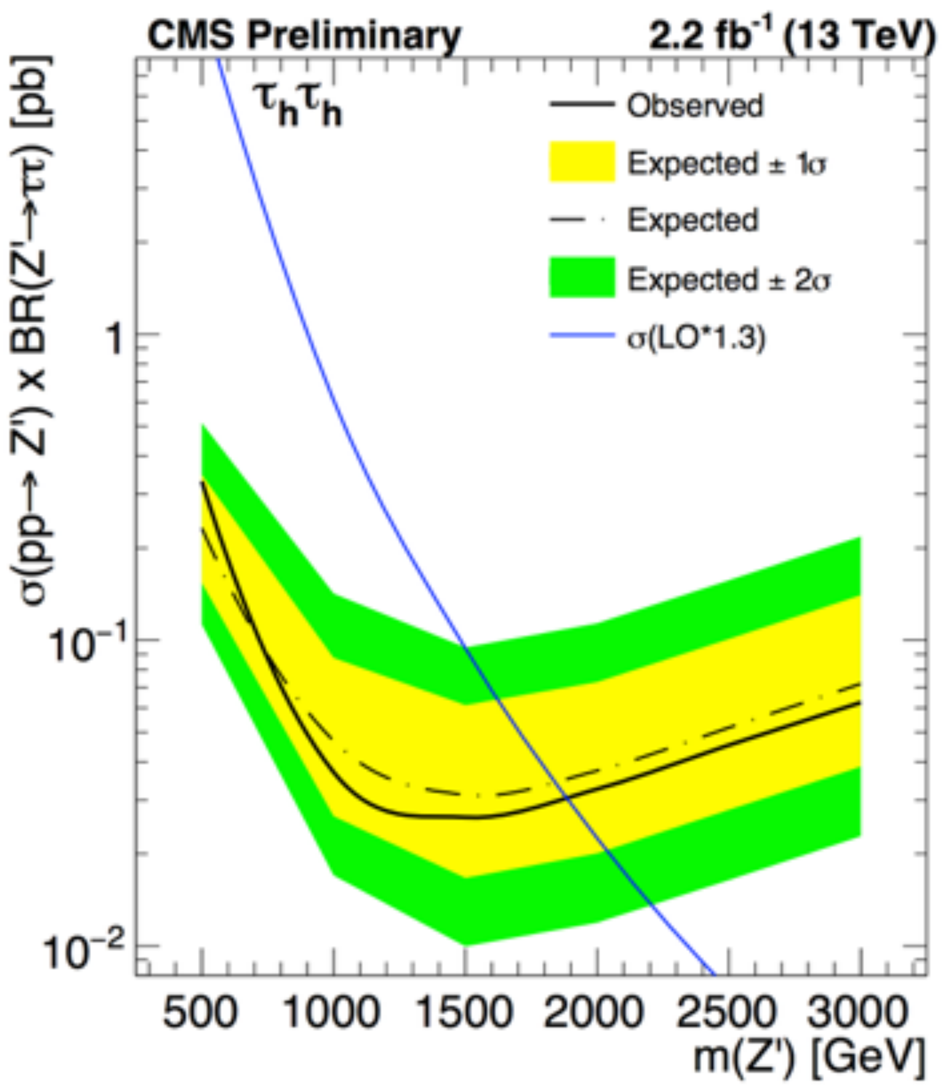
Limit with lifetime > 2



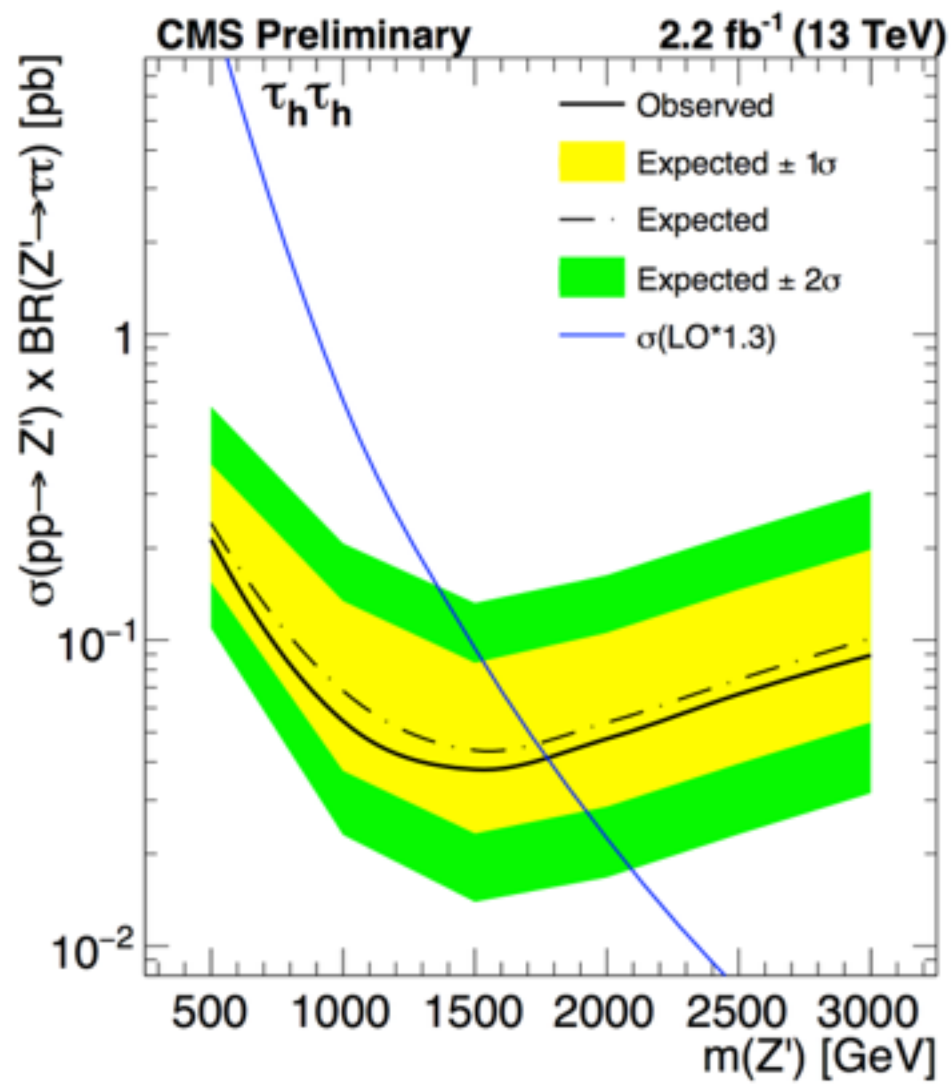
- ~70 GeV improvement
- Low statistics in search region

Impact on Limit: 13 TeV

Standard Limit



Limit with lifetime > 3



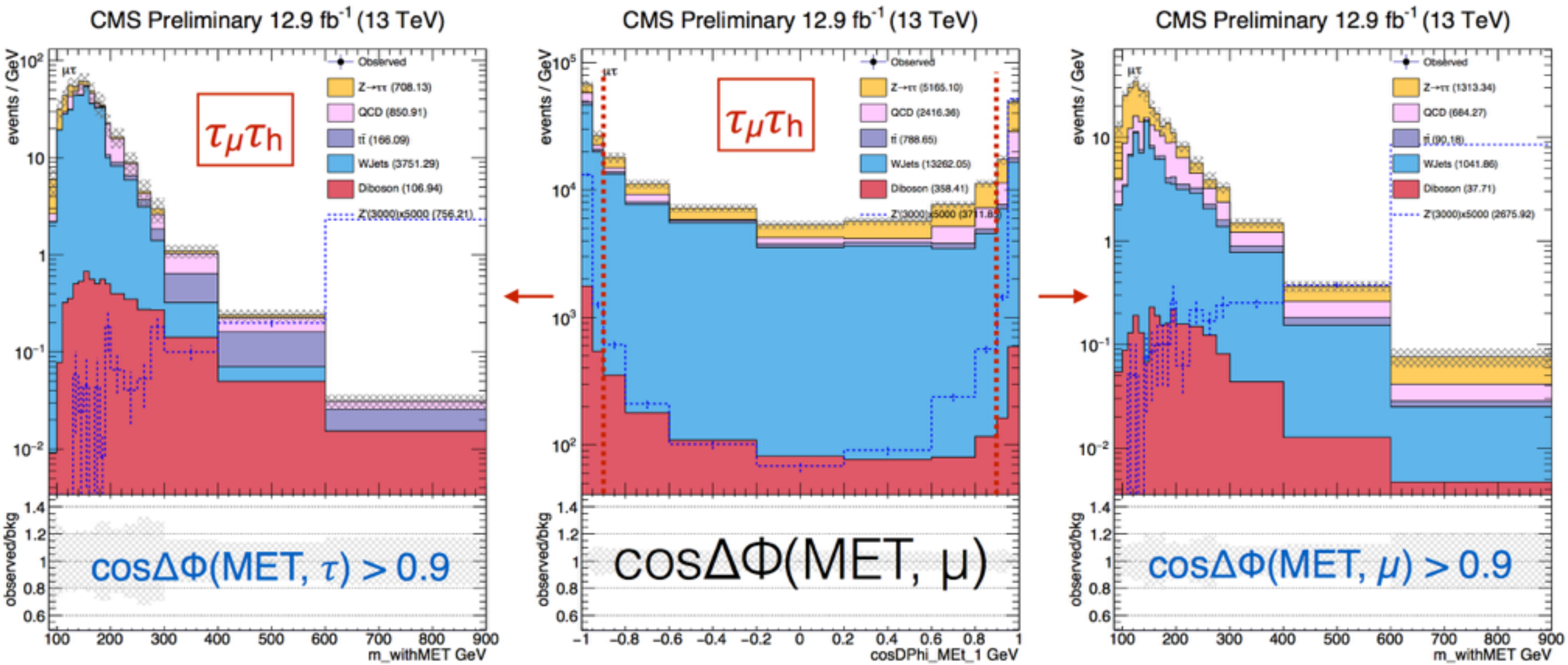
- Limit essentially unaffected due to low statistics
- 1/10th the data compared to 8 TeV
- Expect improvement with more data

Summary

- Search for high ττ+MET mass performed in pp collisions at $\sqrt{s}$ = 13 TeV
- All channels show agreement with SM
- We exclude a Z' decaying to ττ below 2.12 TeV - world's best limit
- Lifetime algorithms expected to have lasting impact with larger data samples
- Thanks for everything!

BACKUP

Additional Topology Studies

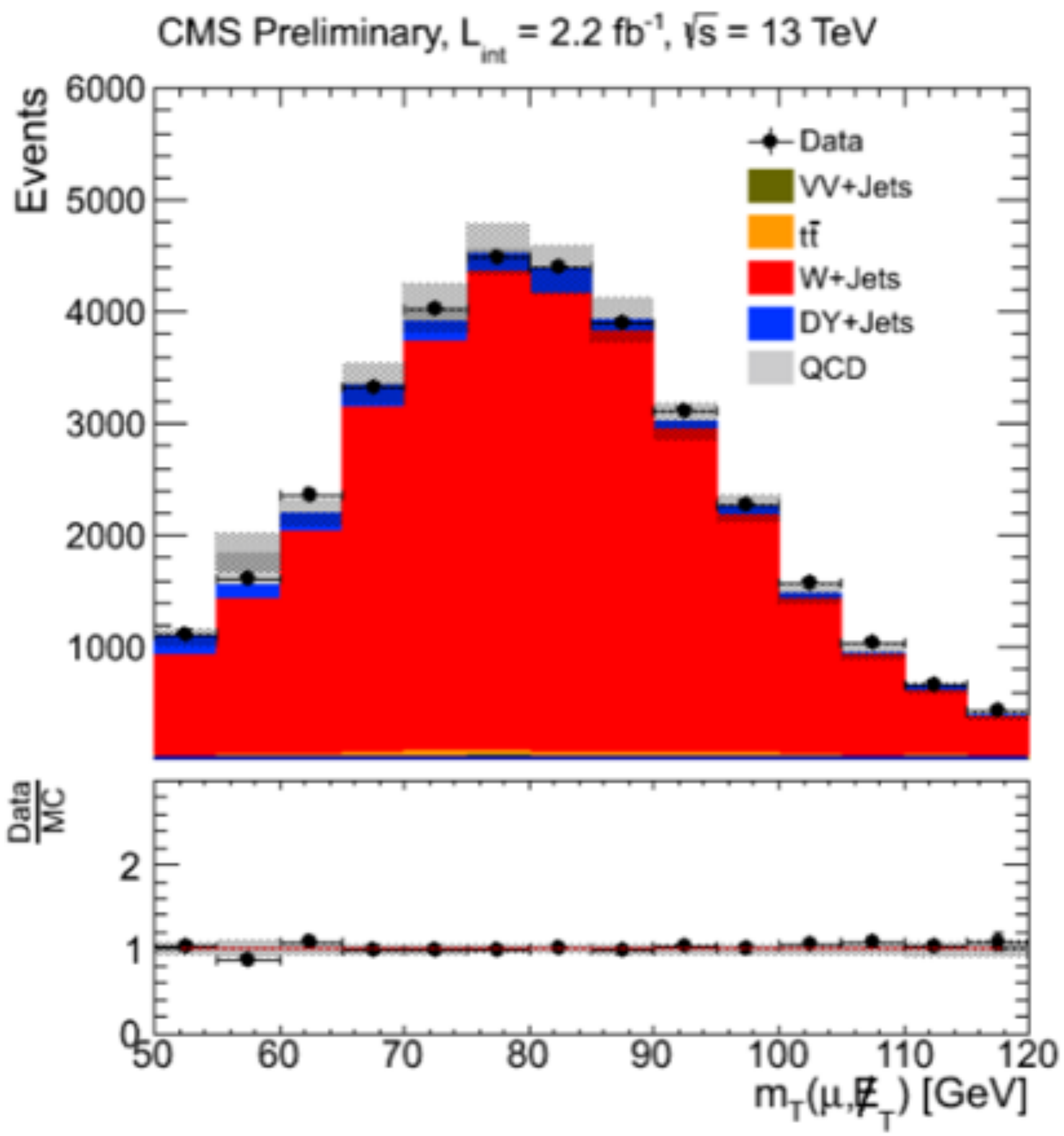


MET aligned with τ

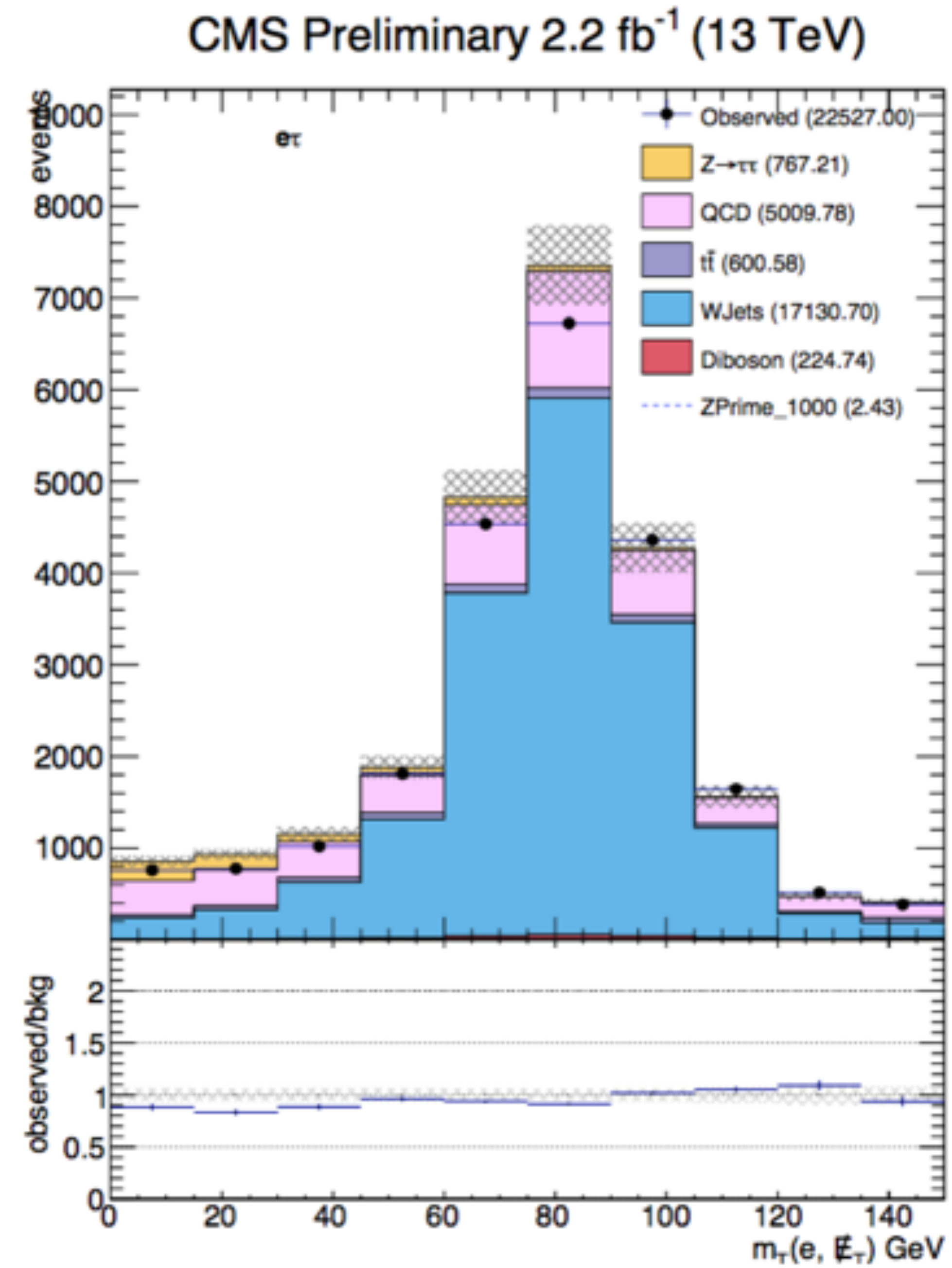
MET aligned with μ

W+Jets Closure in $\tau_\mu\tau_h$ and $\tau_e\tau_h$

$\tau_\mu\tau_h$



$\tau_e\tau_h$



CR **D** (anti-isolated, failing cosΔϕ, p_ζ cuts) in $\tau_\mu\tau_h$ and $\tau_e\tau_h$

Background Estimation

Expected events in full mass range

smaller yields in $\tau_h\tau_h$ due to higher pT threshold

Mix of MC & data-driven estimates

Process	$\tau_h\tau_h$	$\mu\tau_h$	$e\tau_h$	$e\mu$
Drell-Yan	8 ± 3	882 ± 127	375 ± 40	321 ± 37
W+jets	0.1 ± 0.1	916 ± 96	456 ± 35	19 ± 6
Diboson	0.5 ± 0.5	29 ± 7	18 ± 4	108 ± 11
$t\bar{t}$	–	26 ± 7	26 ± 6	223 ± 20
(QCD) Multijet	49 ± 13	122 ± 84	250 ± 50	36 ± 16
Total	58 ± 13	1976 ± 180	1125 ± 73	707 ± 47
Observed	55	1807	1113	728

Expected events at high mass ($m(\tau_1,\tau_2,\mathrm{MET}) > 300$ GeV)

Process	$\tau_h\tau_h$	$\mu\tau_h$	$e\tau_h$	$e\mu$
Drell-Yan	5 ± 2	16 ± 4	9 ± 4	4 ± 3
W+jets	0.004 ± 0.004	23 ± 9	7 ± 5	0.2 ± 0.5
Diboson	0.02 ± 0.02	6 ± 3	3 ± 2	23 ± 5
$t\bar{t}$	–	4 ± 2	5 ± 3	65 ± 12
(QCD) Multijet	18 ± 6	4 ± 3	9 ± 3	0.8 ± 1.0
Total	23 ± 6	51 ± 11	33 ± 8	93 ± 13
Observed	20	42	40	96

Systematics

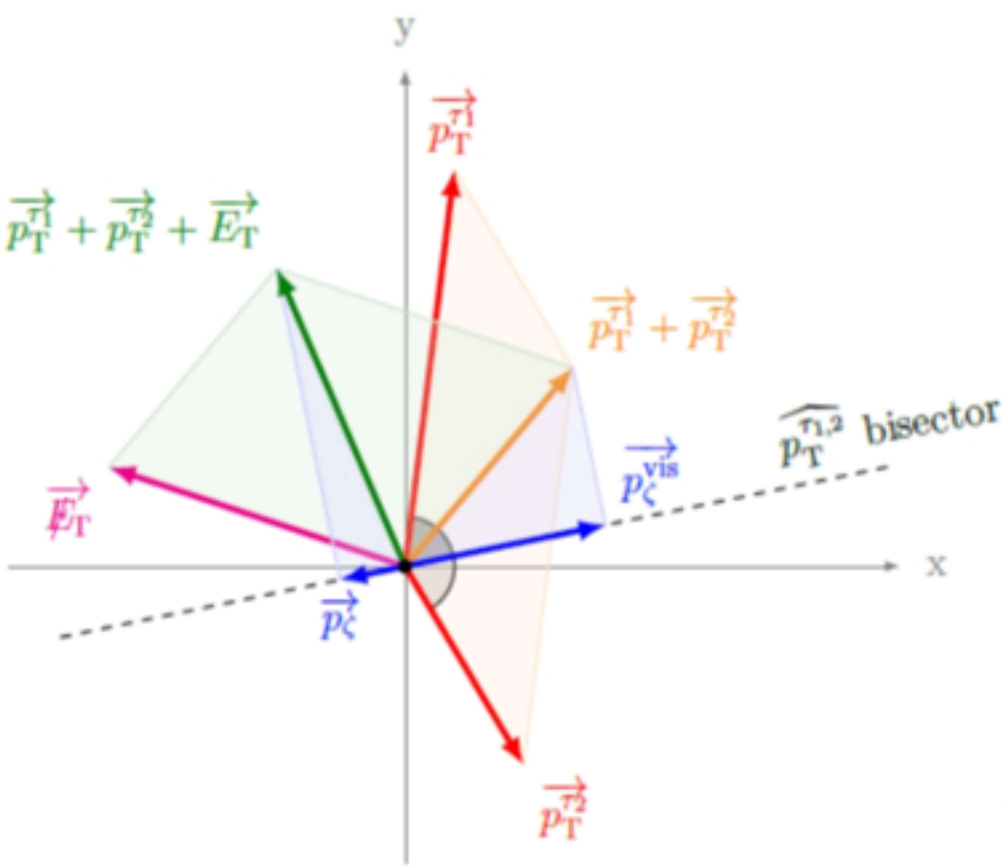
"L" indicates lumi systematic (2.5%)

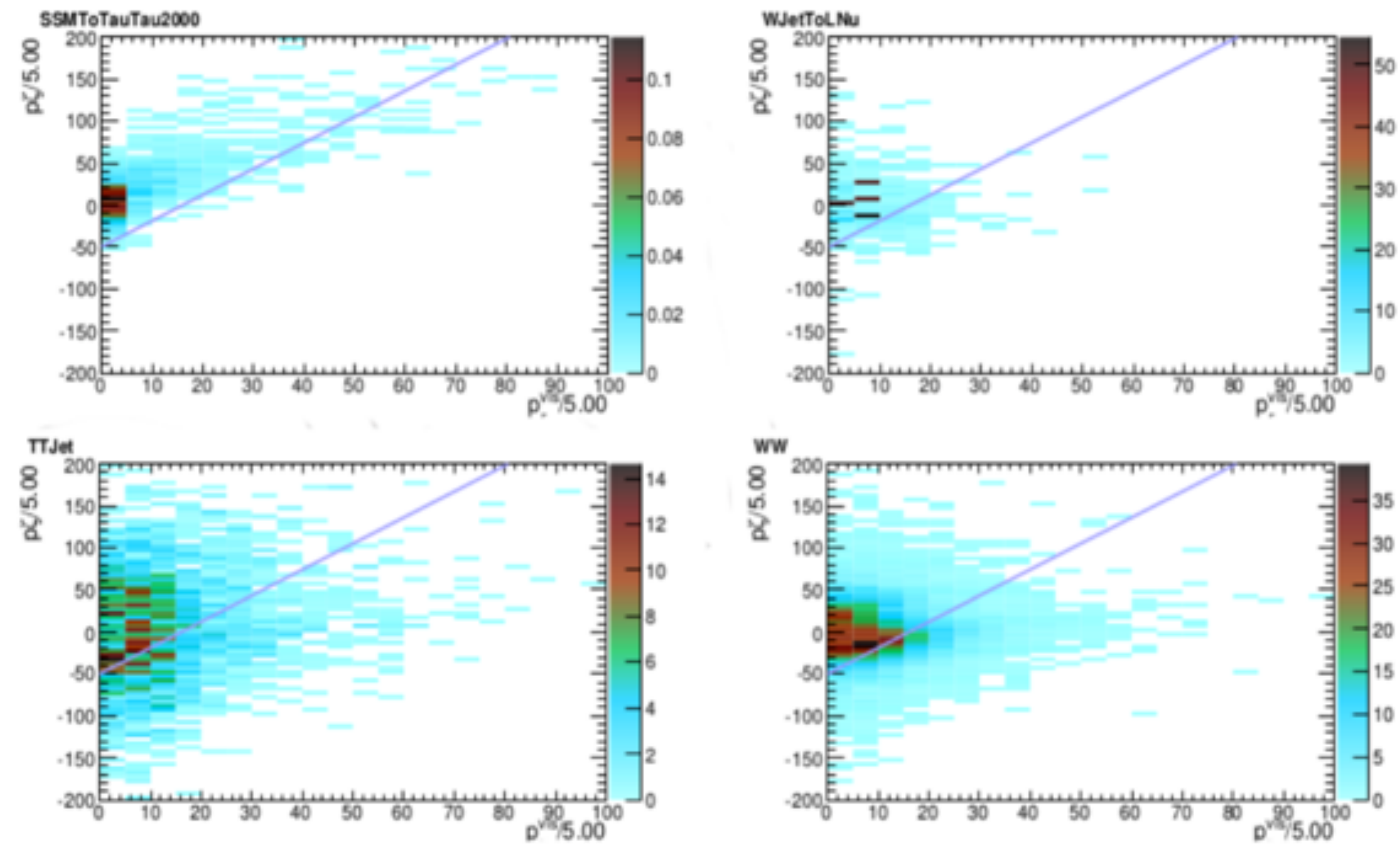
Source ,,,:	QCD hh, μh, eh, eμ	W hh, μh, eh, eμ	DY hh, μh, eh, eμ	$t\bar{t}$ hh, μh, eh, eμ	VV hh, μh, eh, eμ	Signal hh, μh, eh, eμ
Lumi	–,–,–,–	L,–,–,L	L,L,L,L	L,L,L,L	L,L,L,L	L,L,L,L
μ ID + Trig	–,–,–,–	–,–,–,7	–,7,–,7	–,7,–,7	–,7,–,7	–,7,–,7
e ID + Trig	–,–,–,–	–,–,–,6	–,–,6,6	–,–,6,6	–,–,6,6	–,–,6,6
τ_h Trig	–,–,–,–	10,–,–,–	10,–,–,–	10,–,–,–	10,–,–,–	10,–,–,–
τ_h ID	–,–,–,–	30,–,–,–	12,6,6,–	12,6,6,–	12,6,6,–	12,6,6,–
b ID	–,–,–,s	10,–,–,s	3,3,s,s	10,12,s,s	3,3,s,s	3,3,s,s
JES	–,–,–,s	12,–,–,s	8,s,s,s	12,s,s,s	8,s,s,s	2,2,s,s
TES	–,–,–,s	11,–,–,s	11,s,s,s	11,s,s,s	8,s,s,s	3,3,s,s
MMS	–,–,–,–	–,–,–,1	–,1,–,1	–,1,–,1	–,1,–,1	–,1,–,1
EES	–,–,–,–	–,–,–,1	–,–,1,1	–,–,1,1	–,–,1,1	–,–,1,1
top p_T	–,–,–,–	–,–,–,–	–,–,–,–	–,–,–,s	–,–,–,–	–,–,–,–
pdf	–,–,–,–	–,–,–,–	–,–,–,–	–,–,–,–	–,–,–,–	(1–12)
bin-by-bin stat.	s,s,s,s	s,s,s,s	s,s,s,s	s,s,s,s	s,s,s,s	s,s,s,s
Closure+Norm.	20,67,18,37	5,8,9,41	19,6,10,10	8,8,8,8	15,15,15,15	
W+Jets MC Norm.	–,–,–,–	–,–,7,–	–,–,–,–	–,–,–,–	–,–,–,–	–,–,–,–

"s" indicates shape systematics - treated as uncorrelated across bins

all other systematics treated as 100% correlated

p_ζ Definition



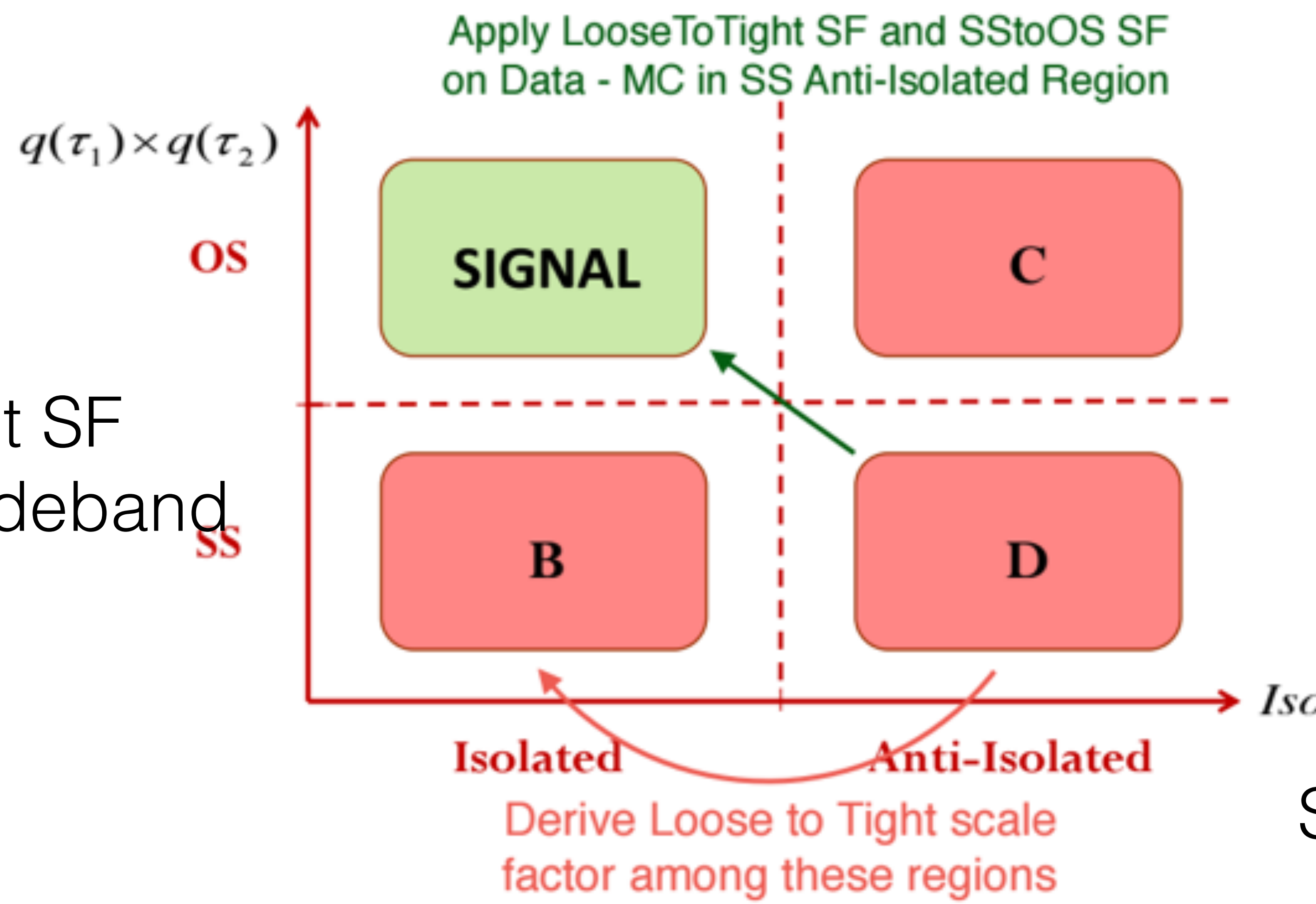


- "Back-to-back" taus pushes visible component lower
- Neutrinos from tau decays pushes p_ζ higher
- Selection of $p_\zeta - 3.1p_\zeta^{vis} > -50$ optimized from MC

QCD Estimation in $\tau_\mu\tau_h$ and $\tau_e\tau_h$

$N_{sig} = N_D \times SF_{SSToOS} \times SF_{LooseToTight}$

SStoOS SF derived in low-MET sideband (**C'**/**D'**)



LooseToTight SF derived in SS sideband (**B**/**D**)

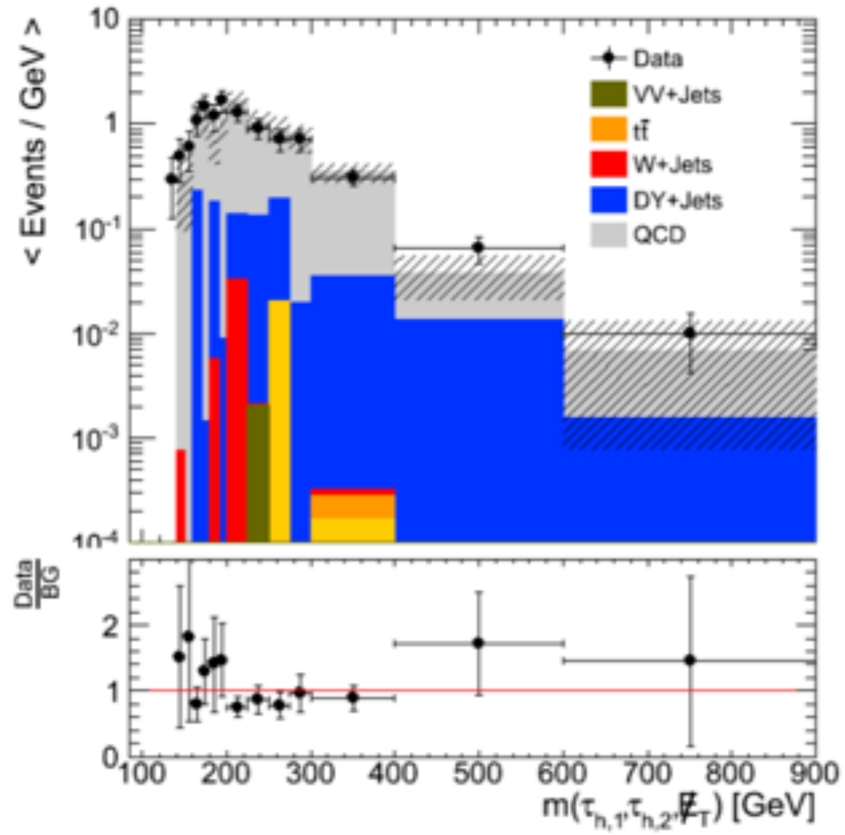
Shape taken from **D**

Anti-isolated sideband: τ_h failing "tight" WP, passing "loose" WP

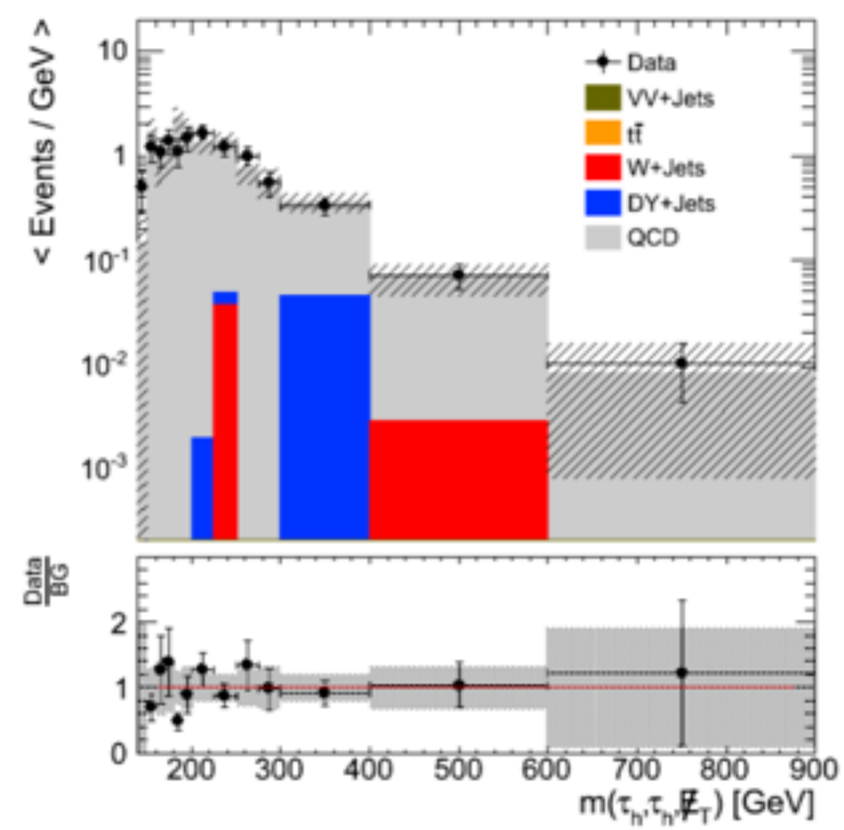
Muon isolation: 0.15 < relIso < 0.95

QCD Shape Validation

- In both isolated and anti-isolated regimes, shape taken from SS sideband, normalized to OS yield
- Check data/estimation in low-MET sidebands for both regimes
- Good agreement between data and estimation
- Shapes provide robust estimation
- Similar shape checks done for other preselection variables (backup)



Isolated

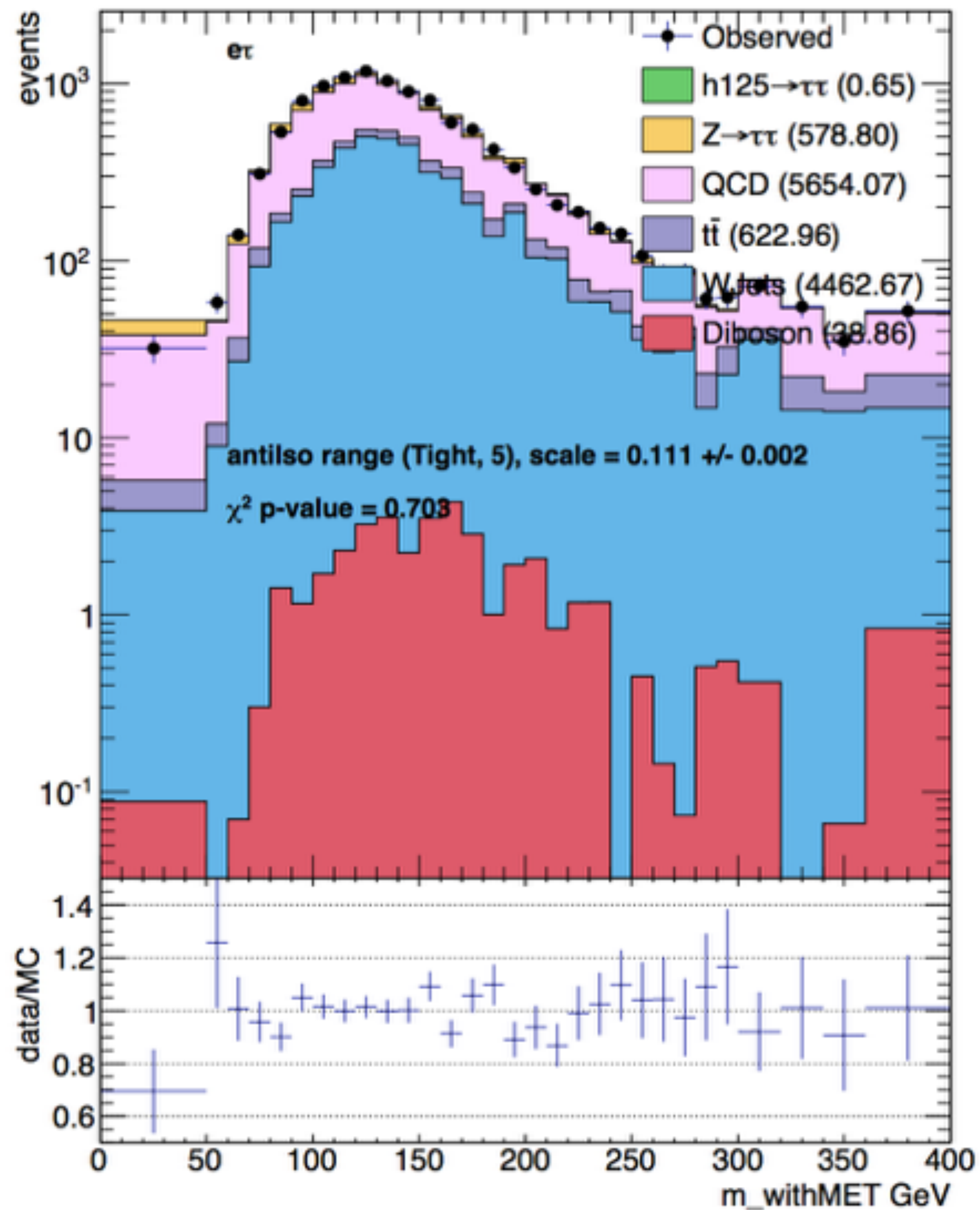


Anti-isolated

No additional systematics due to closure test

QCD Estimation in $\tau_\mu\tau_h$ and $\tau_e\tau_h$

Shape in anti-isolated sideband provides good estimator of isolated regions

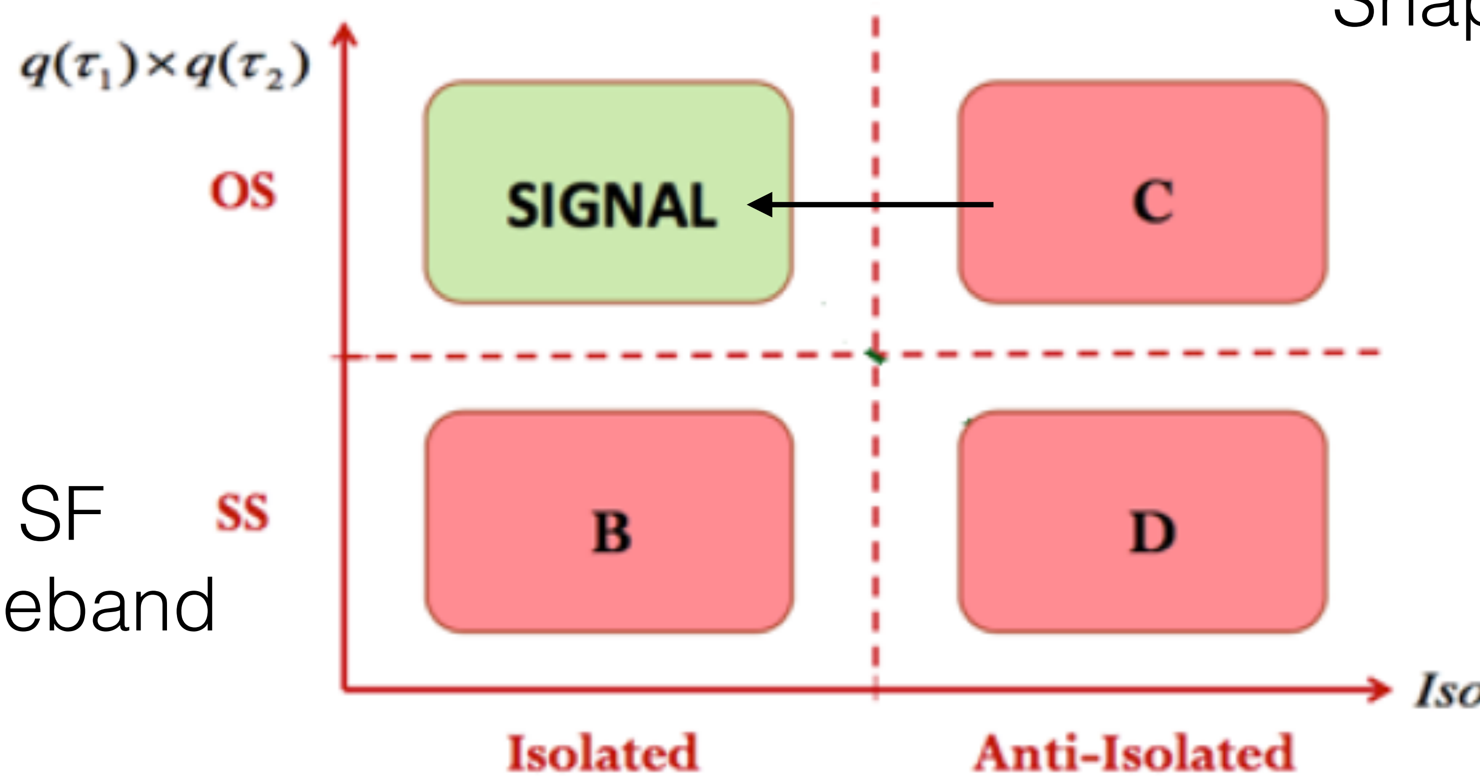


CR **B** (SS, isolated) in $\tau_e\tau_h$

QCD Estimation in $\tau_e\tau_\mu$

$N_{sig} = N_C \times SF_{LooseToTight}$

Shape taken from **C**



LooseToTight SF derived in SS sideband (**B**/**D**)

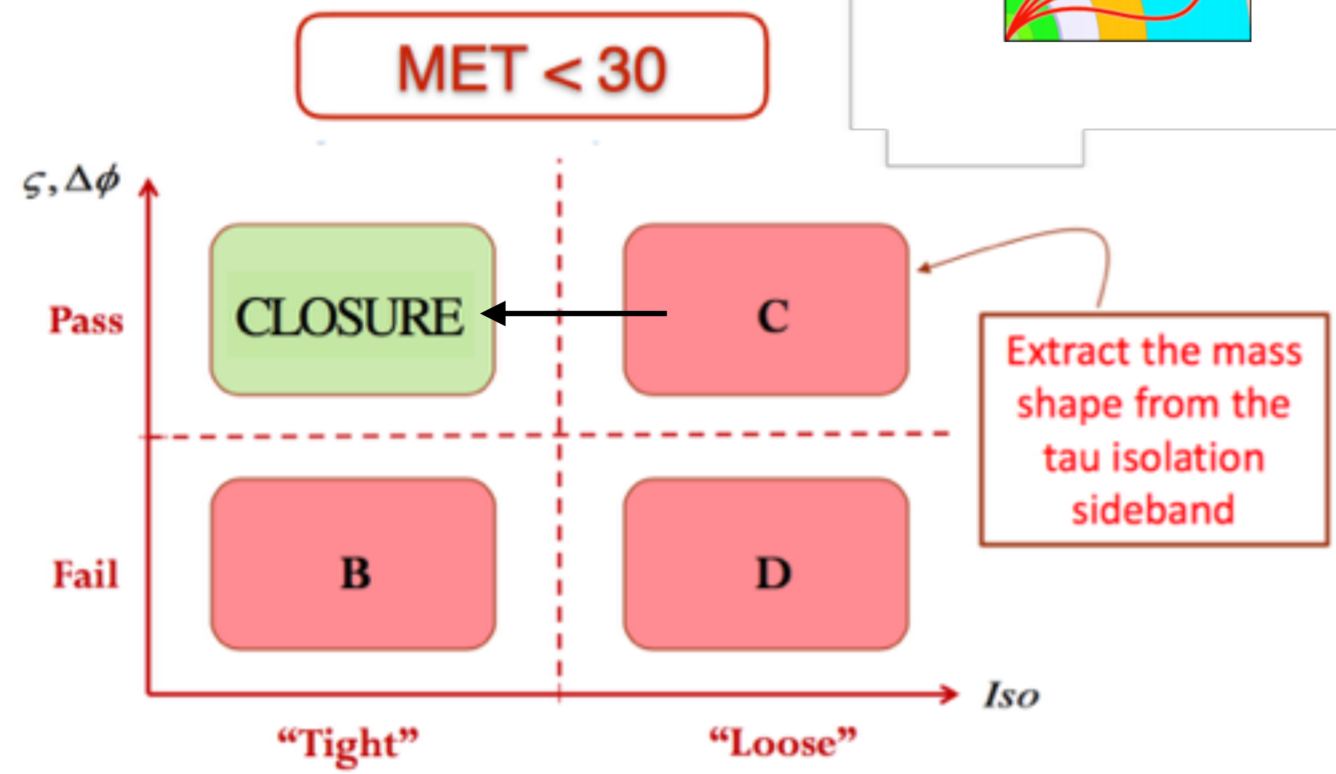
Anti-isolated sideband: Muon isolation: 0.15 < relIso < 0.95

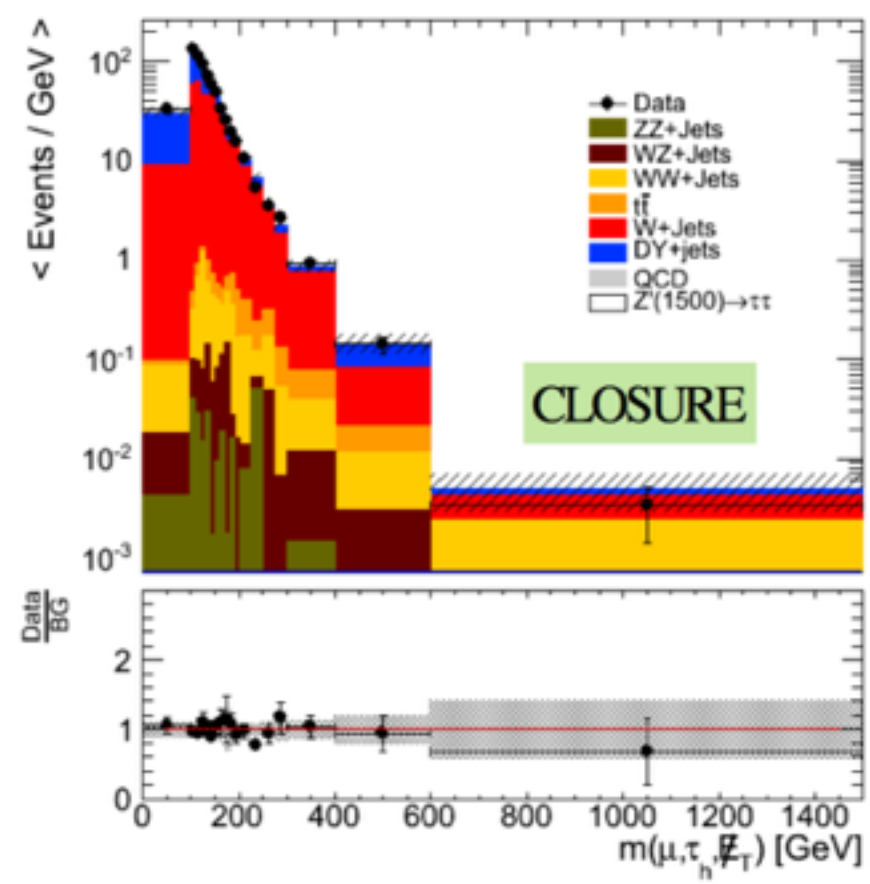
Simpler than $\tau_\mu\tau_h$ and $\tau_e\tau_h$: easier to disentangle other BGs

Further W+Jets Closure in $\tau_\mu\tau_h$

Repeat "ABCD" estimation in low-MET regime

- Shape taken from "C"
- Yield taken as "C" multiplied by B/D SF

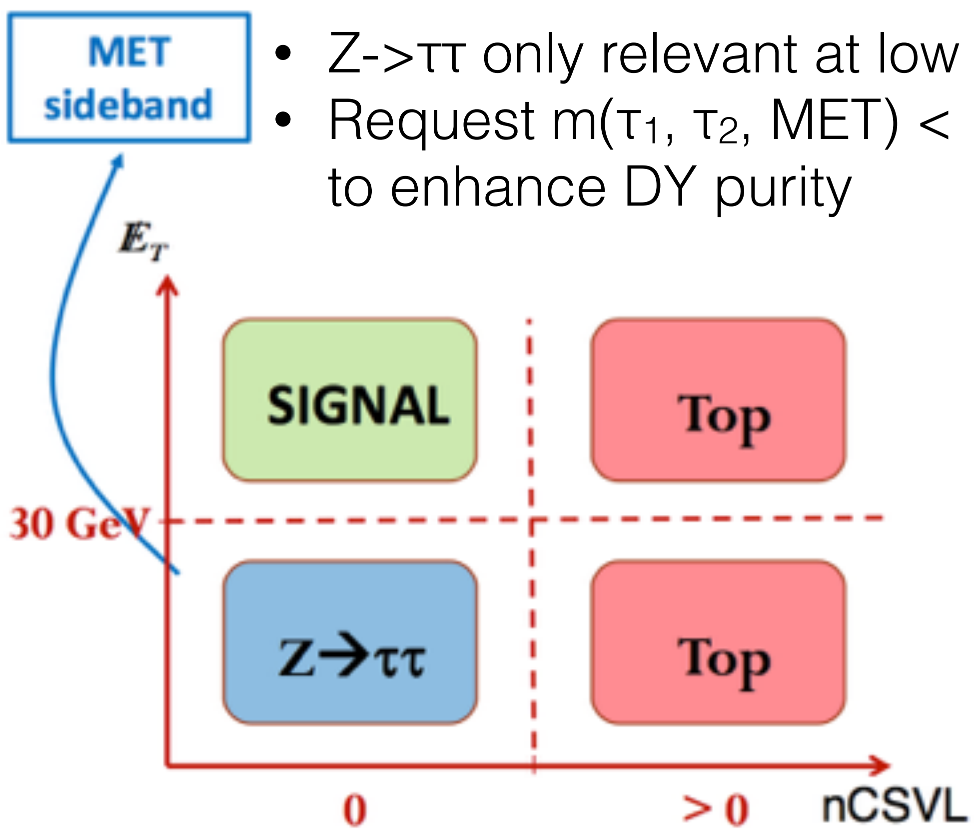




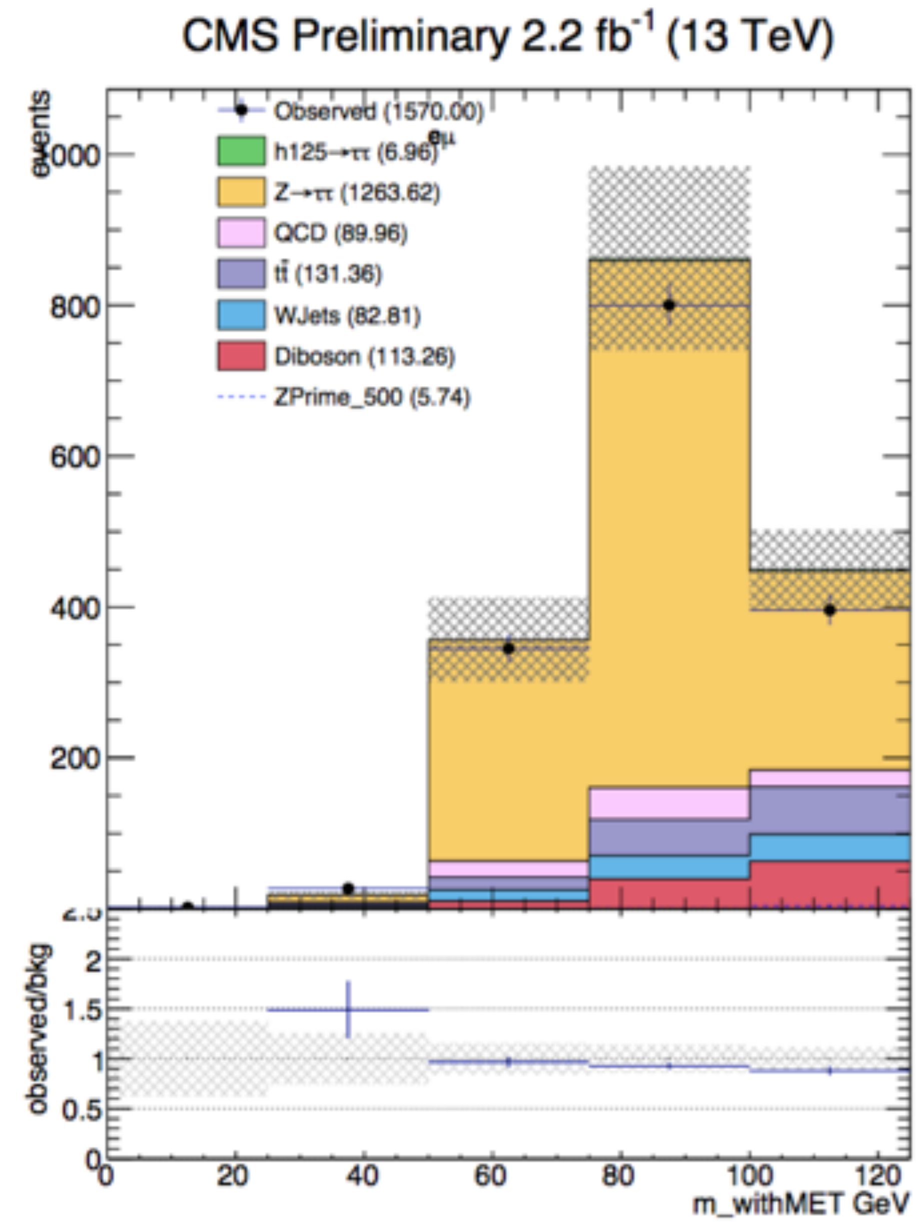
Good agreement between data/MC in closure region

DY+Jets Validation in $\tau_e\tau_h$ and $\tau_e\tau_\mu$

- Take shape + normalization from MC
- Extract systematic from low-MET sideband



- Z->ττ only relevant at low mass
- Request m(τ1, τ2, MET) < 125 GeV to enhance DY purity



(Data - non-DY MC)/(DY MC) = 0.90 +- 0.10

Systematic on DY estimation

Setting Limits

- Limits on signal strength set by “combine” tool used by Higgs analysis
- Input: predicted BG rates, Z’SSM rates, data rates for each reconstructed mass bin, systematic uncertainties
- Performs CL_s profile likelihood fitting
  - “Signal strength” ($0 < \mu < 1$) hypotheses compared with predicted and observed rates
- Output: 95% confidence upper limit on μ for each mass bin
- Plotted vs. Z’ mass to get confidence band

Setting Limits

$$L(\mu, \boldsymbol{\theta}) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} \quad \prod_{k=1}^{M} \frac{u_k^{m_k}}{m_k!} e^{-u_k}$$

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})}$$

$$q_\mu = \begin{cases} -2\ln\lambda(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases}$$

$$p_\mu = \int_{q_{\mu,\mathrm{obs}}}^{\infty} f(q_\mu|\mu)\, dq_\mu$$

Samples Used

MiniAODv2 generated in 74x

- Backgrounds:
  - **DY+Jets: DYJetsToLL M-XtoY_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/RunIISpring15MiniAODv2-74X_mcRun2 asymptotic_v2-v1**
  - **ttbar: /TTJets_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/RunIISpring15MiniAODv2-74X_mcRun2_asymptotic_v2-v1(_ext3-v1)/MINIAODSIM**
  - **W+Jets: /WJetsToLNu_HT-XtoY_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/RunIISpring15MiniAODv2-74X_mcRun2_asymptotic_v2-v1/MINIAODSIM**
  - **Diboson: /VV_TuneCUETP8M1_13TeV-pythia8/RunIISpring15MiniAODv2-74X_mcRun2_asymptotic_v2-v1/MINIAODSIM**
  - **Single top:/ST tW (anti)top 5f inclusiveDecays 13TeV-powheg-pythia8 TuneCUETP8M1/RunIISpring15MiniAODv2-74X mcRun2 asymptotic v2-v2**
- Data:
  - **/SingleMuon/Run2015(C,D)-(05Oct2015,PromptReco-v4)/MINIAOD**
  - **/SingleElectron/Run2015(C,D)-(05Oct2015,PromptReco-v4)/MINIAOD**
  - **/Tau/Run2015(C,D)-(05Oct2015,PromptReco-v4)/MINIAOD**
  - **Golden JSON: Cert_246908-260627_13TeV_PromptReco_Collisions15_25ns_JSON.txt (2.2 fb-1)**
- Signal Samples
  - **/ZprimeToTauTau_M_X_TuneCUETP8M1_τola_13TeV_pythia8/RunIISpring15MiniAODv2-74X_mcRun2_asymptotic_v2-v1/MINIAODSIM**

Triggers

Hadronic Taus (τ_h):

HLT_DoubleMediumIsoPFTau(35,40)_Trk1_eta2p1_Reg (D,C/MC)

SF = 0.9

apply trigger eff as weight (data)

Electrons:

HLT_Ele27_eta2p1_WP(75,Loose)_Gsf_v (MC, data)

SF = 1.0 (barrel), 0.94 (endcap)

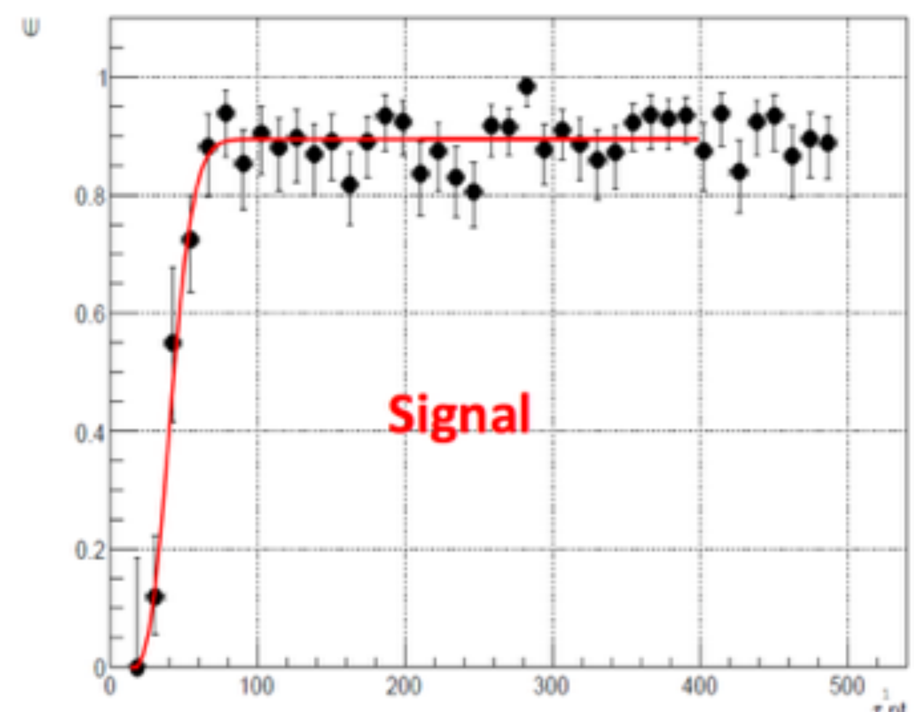
Muons:

HLT_IsoMu(17,18,24)_eta2p1 (MC,D,C)

SF = 0.991 ($|\eta| <= 1.2$), 0.986 ($|\eta| > 1.2$)

Use trigger eff SFs to match data to MC

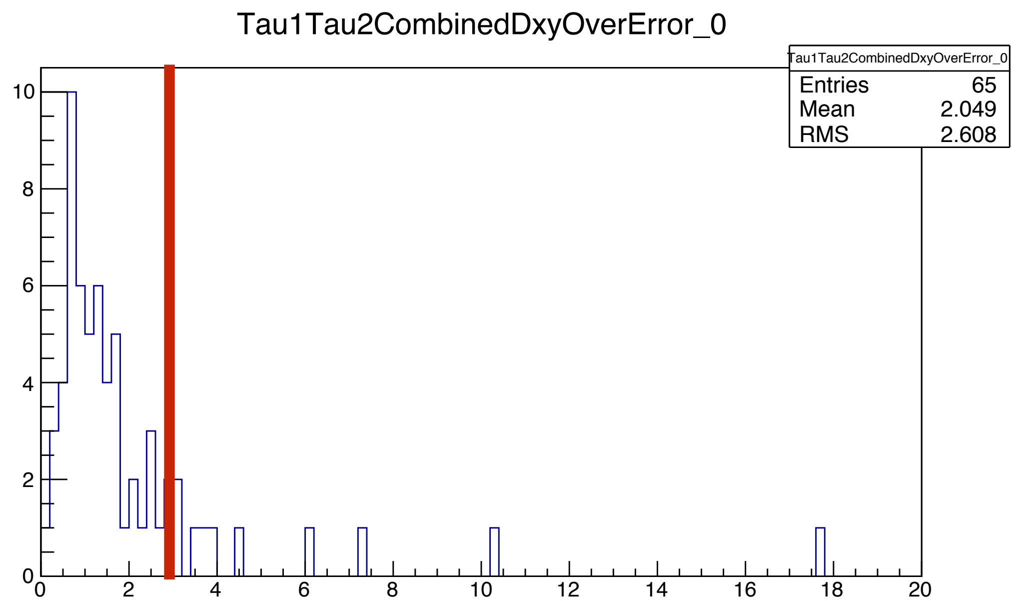
Measure trigger efficiencies using tag-and probe method



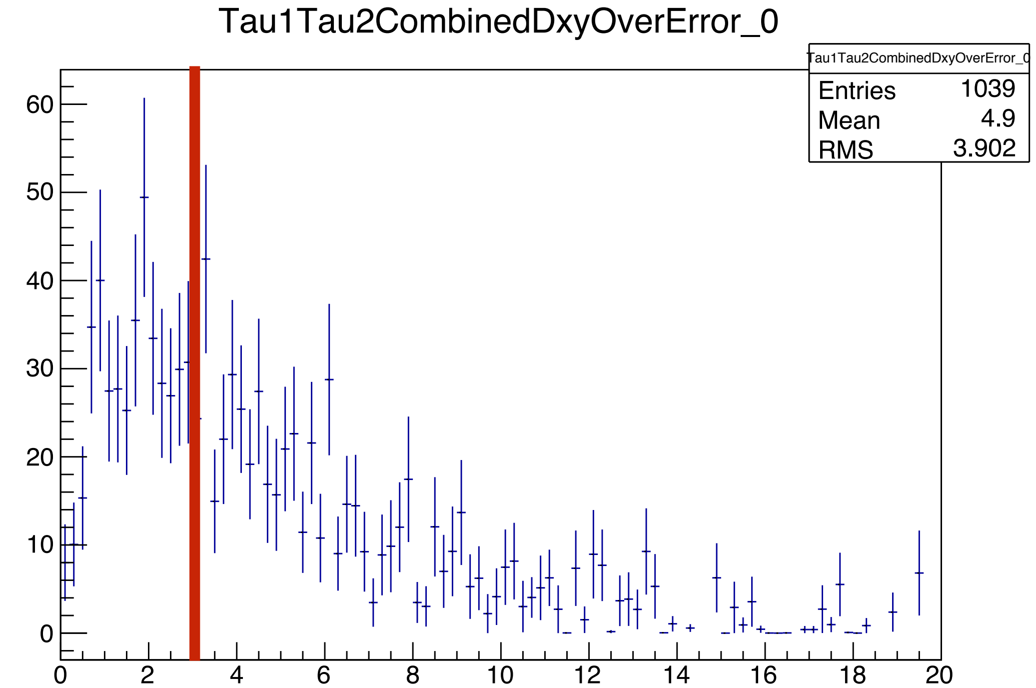
τ_h trigger turn-on curve

Lifetime Performance at 13 TeV

Data (QCD)

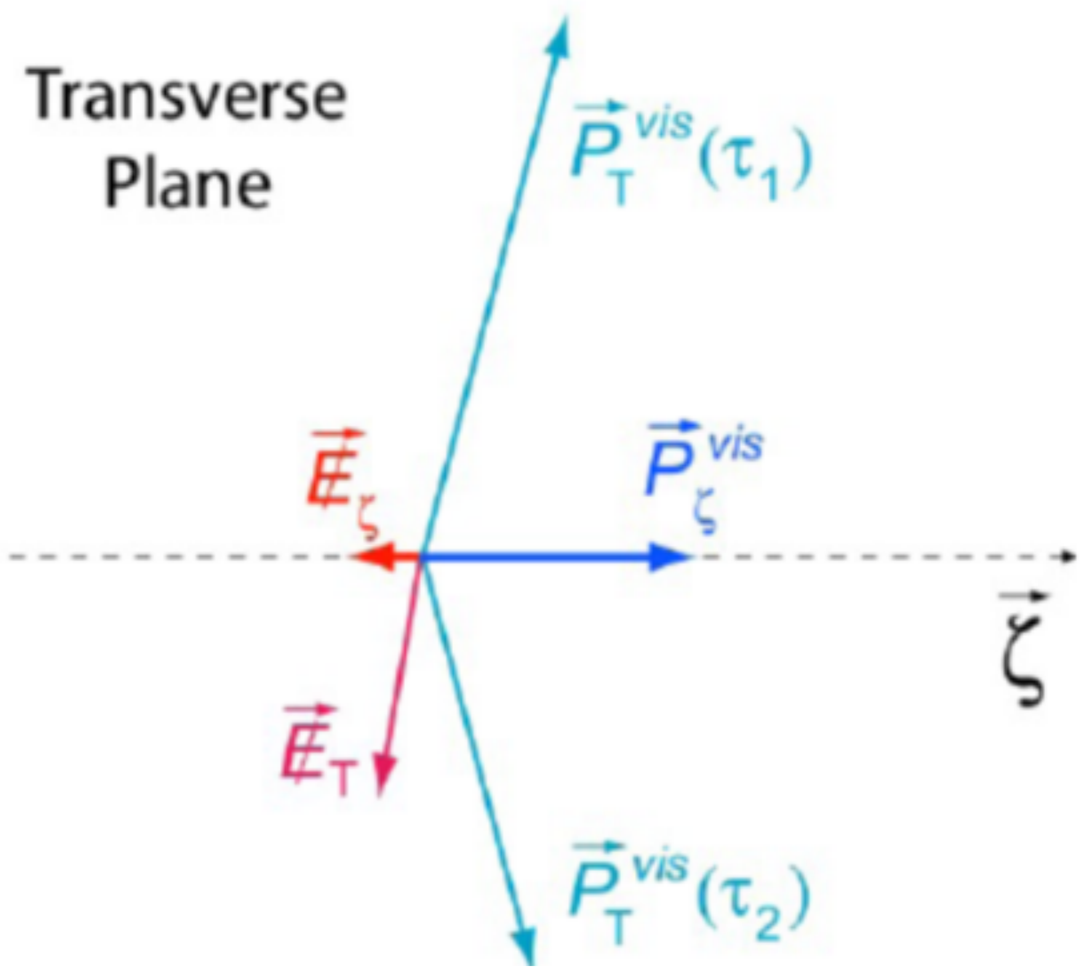


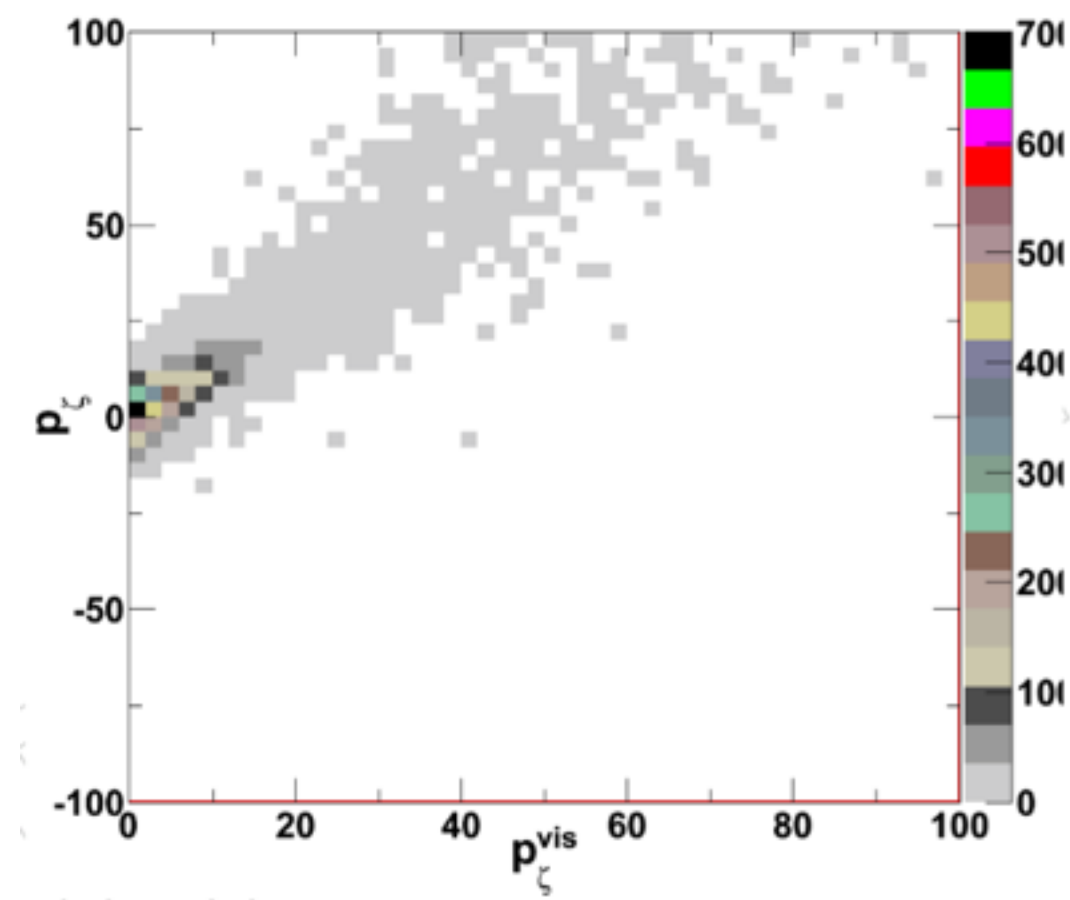
2.5 TeV Z' MC



- Leading charged pion tracks used in fully-hadronic channel
- Data used as proxy for QCD since MC not available at 13 TeV
- Lifetime cut of > 3 would eliminate most BG, kill ~50% of signal

pZeta definition



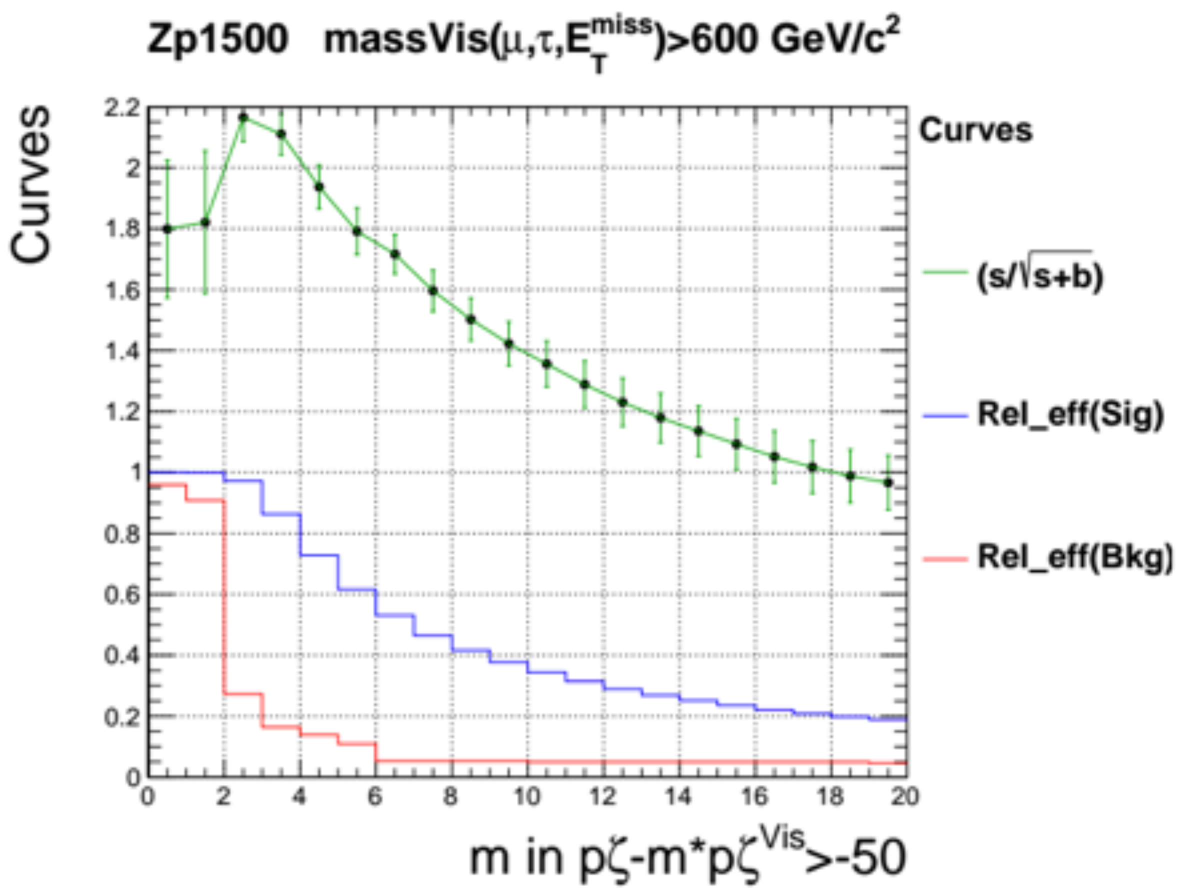


$$p_\zeta^{vis} = \overrightarrow{p}_{\tau_1}^{vis}\hat{\zeta} + \overrightarrow{p}_{\tau_2}^{vis}\hat{\zeta}$$

$$p_\zeta = p_\zeta^{vis} + \overrightarrow{\not{E}}_T\hat{\zeta}$$

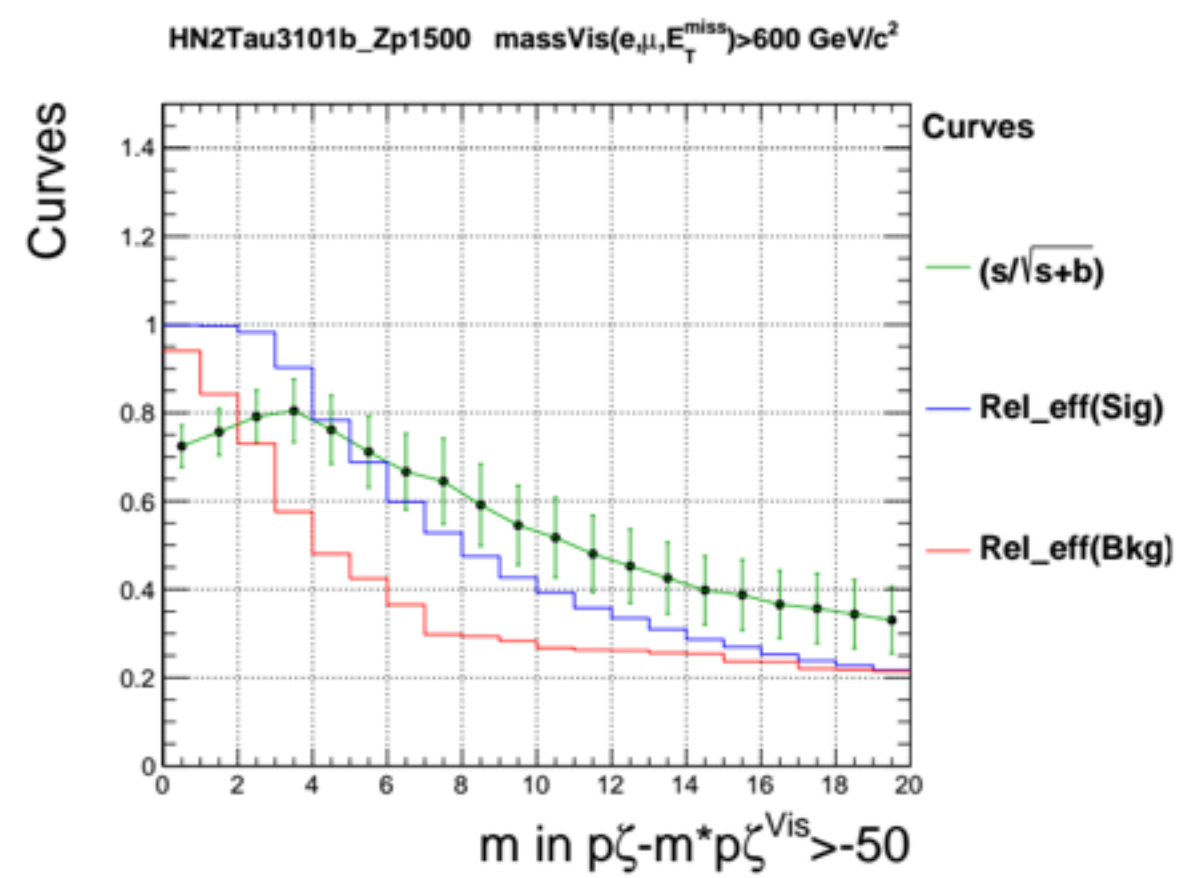
pZeta optimization

Z' 1500 GeV.



$\tau_\mu \tau_h$

Z' 1500 GeV.



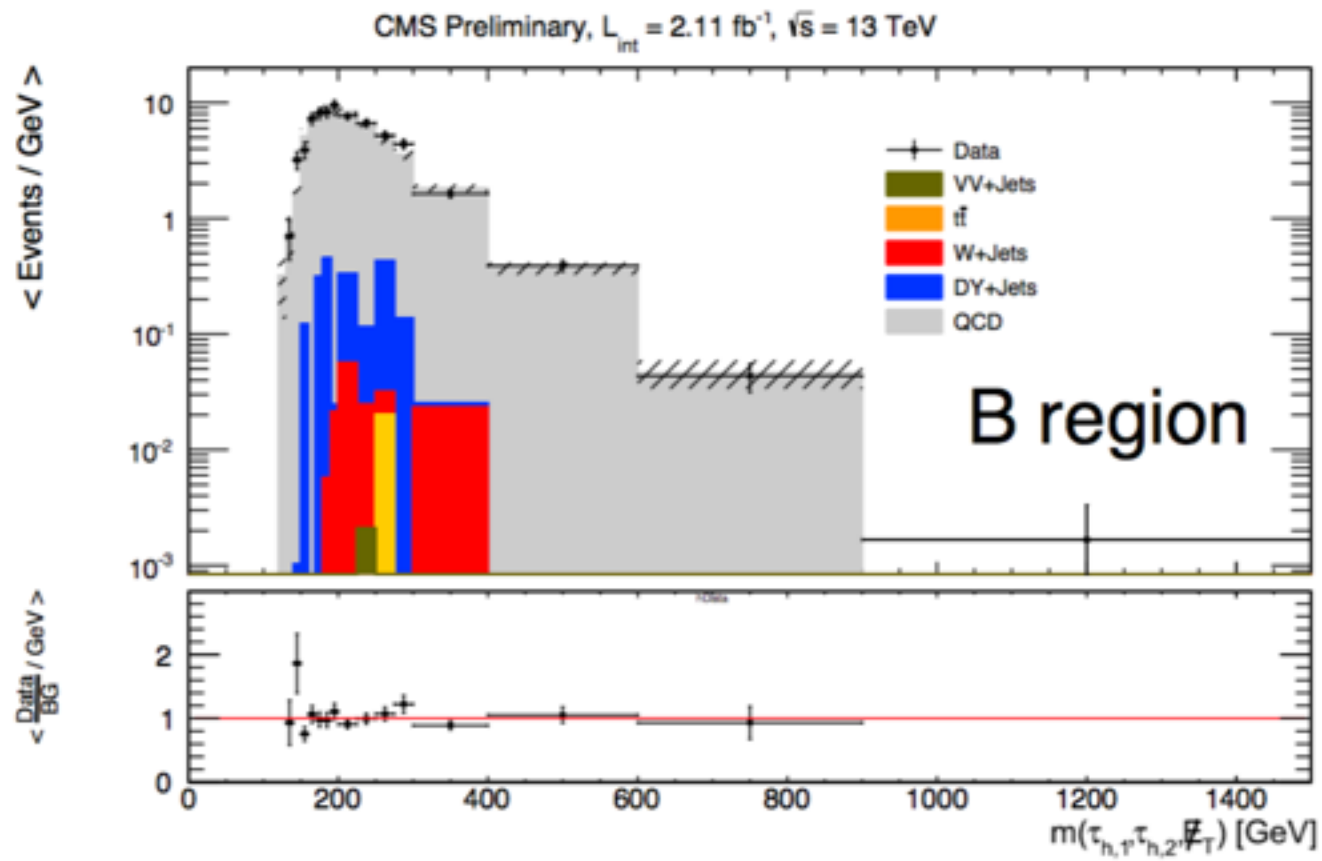
$\tau_e \tau_\mu$

Max around m = 3.1

QCD Closure (1) in $\tau_h \tau_h$

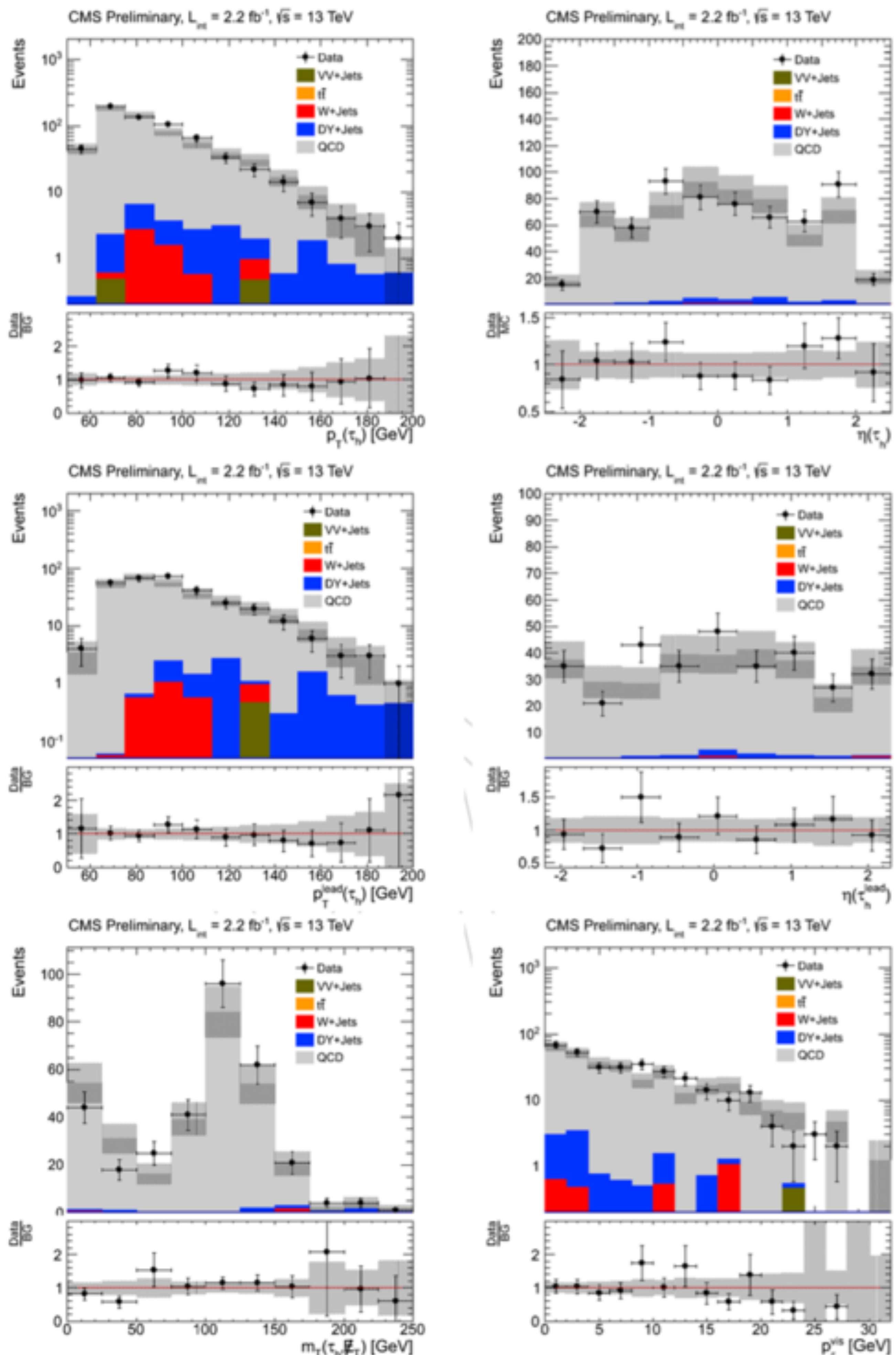
Using SS to model OS shape

- Take SS shape (D region), apply to OS (B region) yield
  - Shape is data - nonQCD MC in D
  - Yield is data - nonQCD MC in B
- SS shape nicely models OS shape



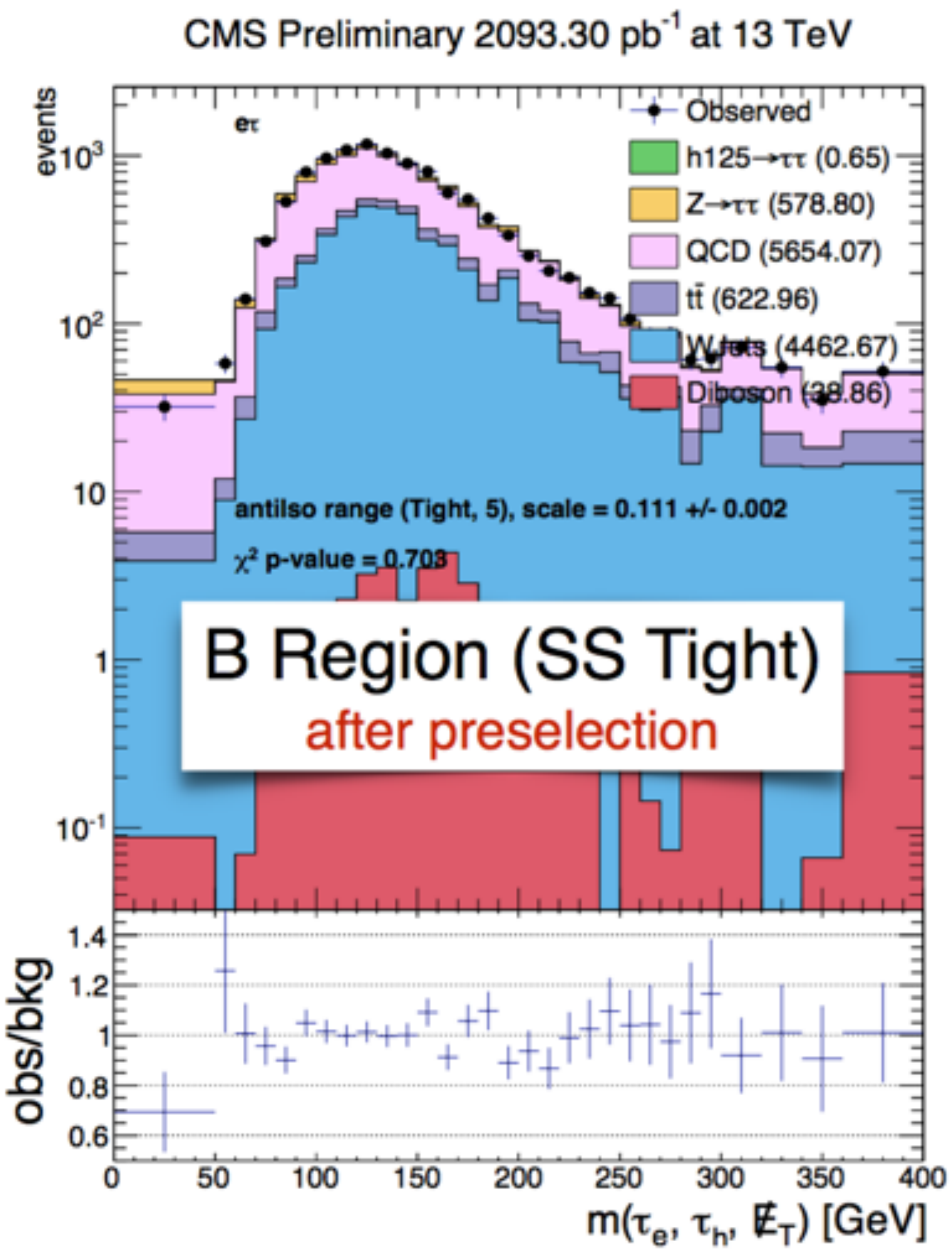
Validity of Estimation

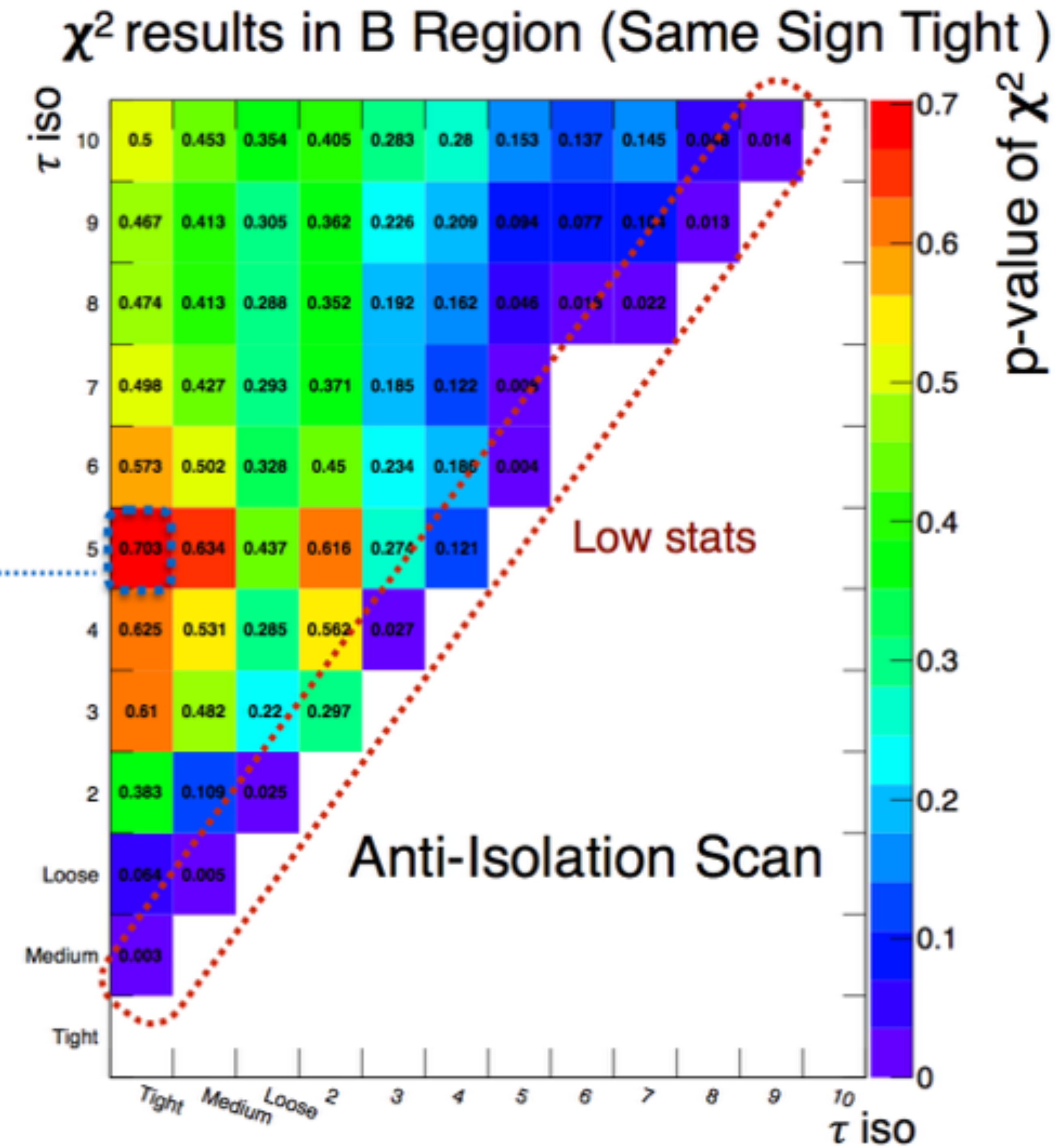
- Repeat ABCD estimation in “anti-isolated” control region
- Compare data yield/estimation yield in control region


CMS Preliminary, L_int = 2.2 fb^-1, √s = 13 TeV
Events
Data
VV+Jets
tt̄
W+Jets
DY+Jets
QCD
Data/BG
Data/MC
p_T(τ_h) [GeV]
η(τ_h)
p_T^lead(τ_h) [GeV]
η(τ_h^lead)
m_T(τ_h, E̸_T) [GeV]
p_ζ^vis [GeV]

QCD Closure in $\tau_e\tau_h$

$\tau_e\tau_h$



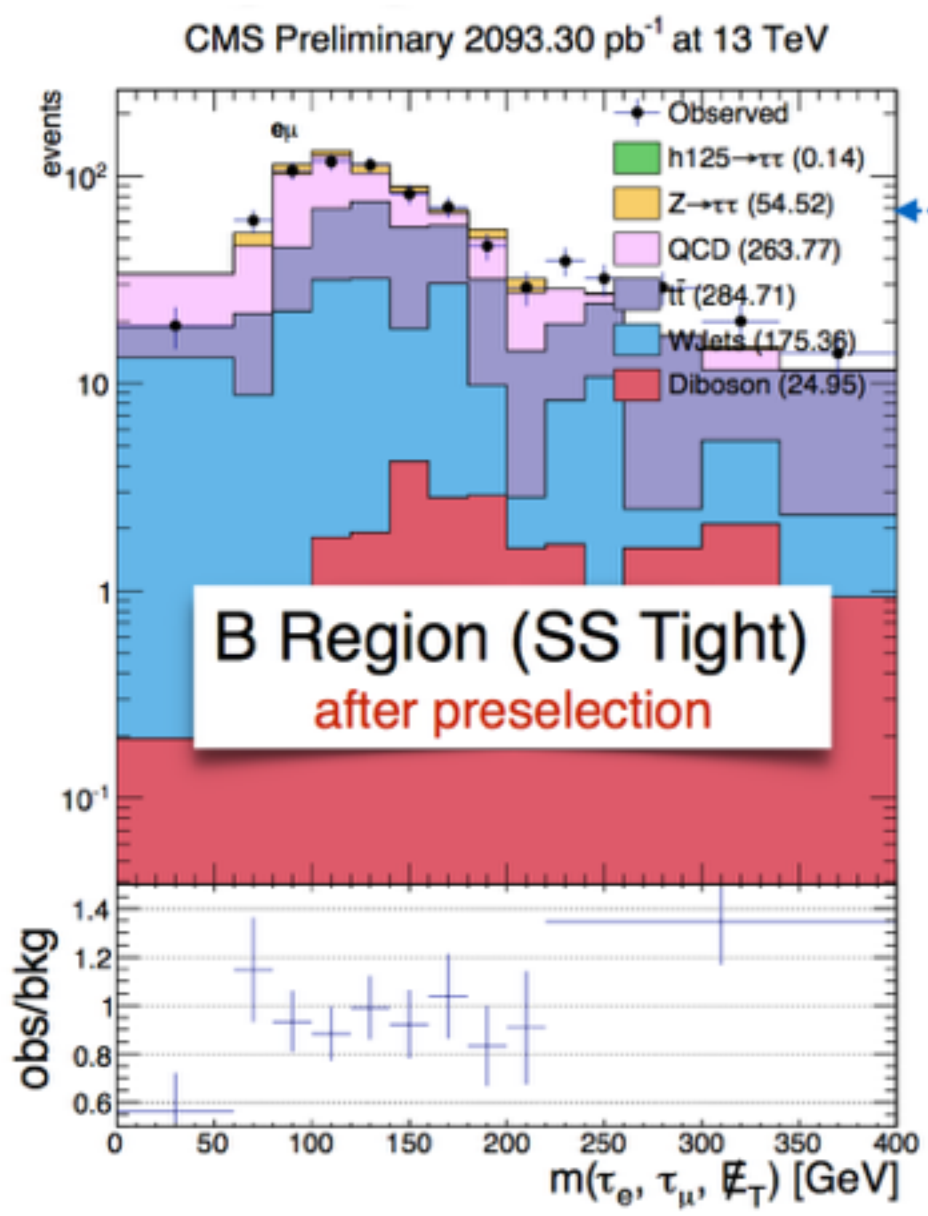


Choose anti-isolation region (tight, 5)

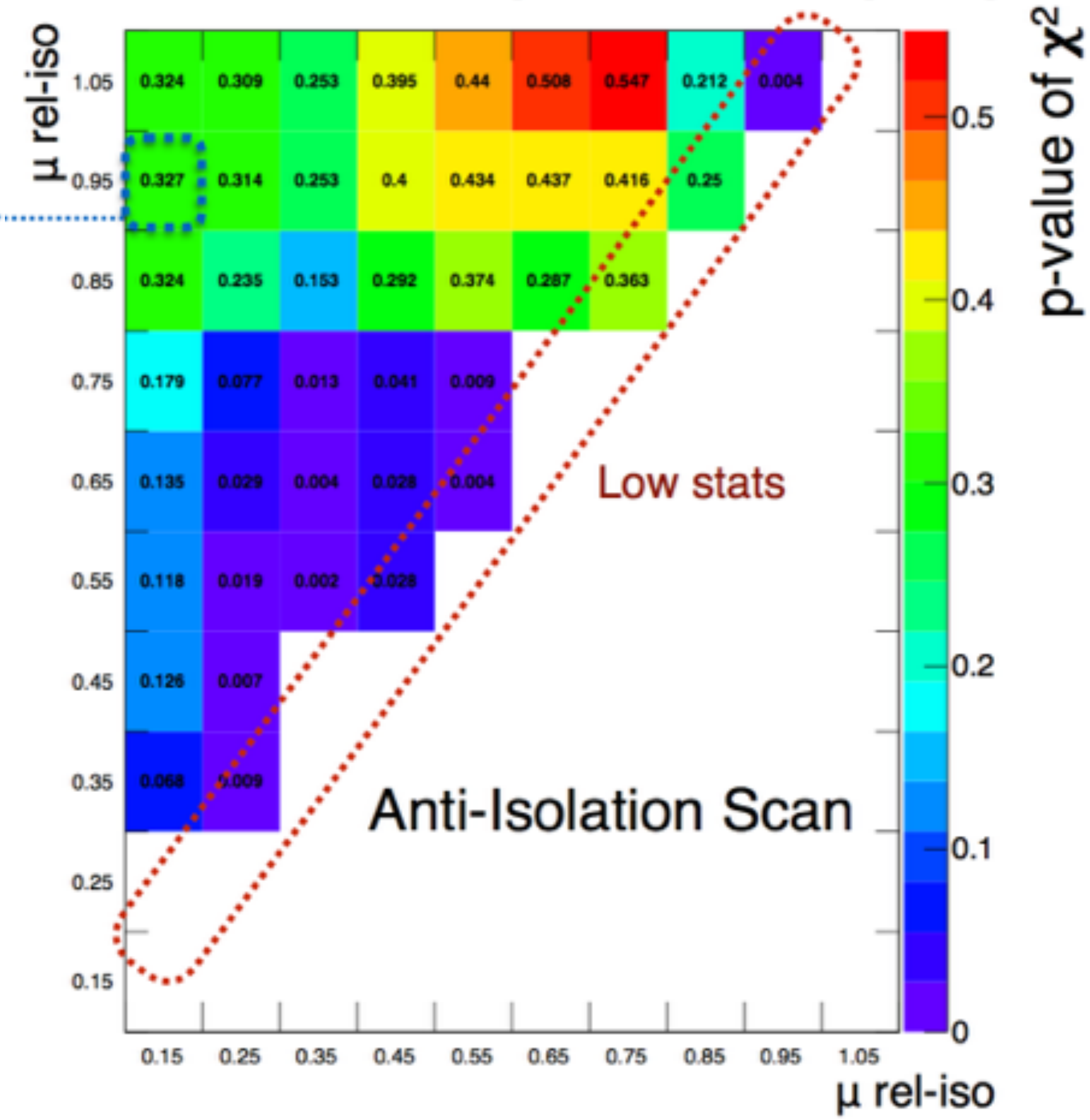
- good modeling (p-value of χ^2) of QCD in isolated region
- high statistics

QCD Closure in $\tau_e\tau_\mu$

$\tau_e\tau_\mu$



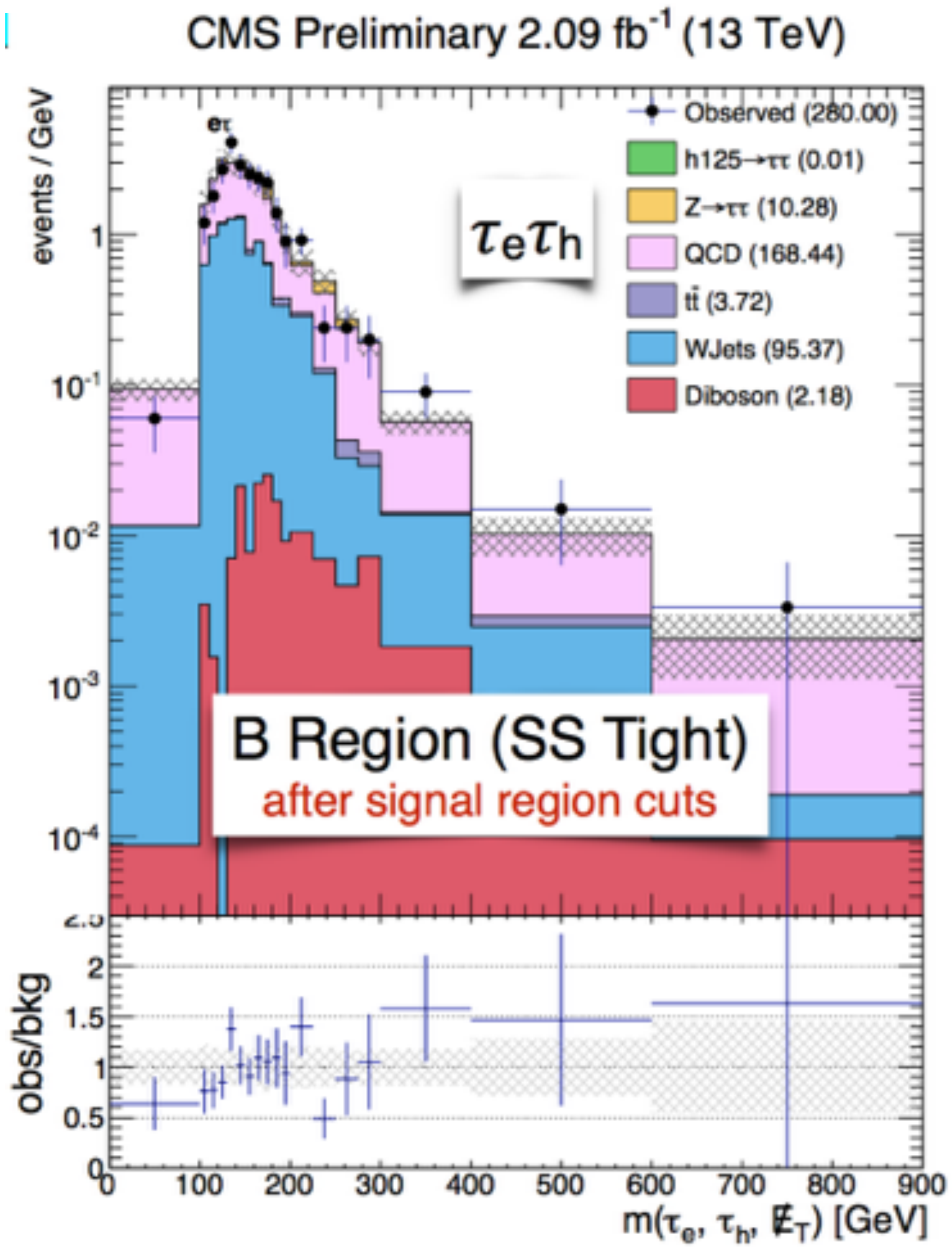
χ^2 results in B Region (Same Sign Tight)



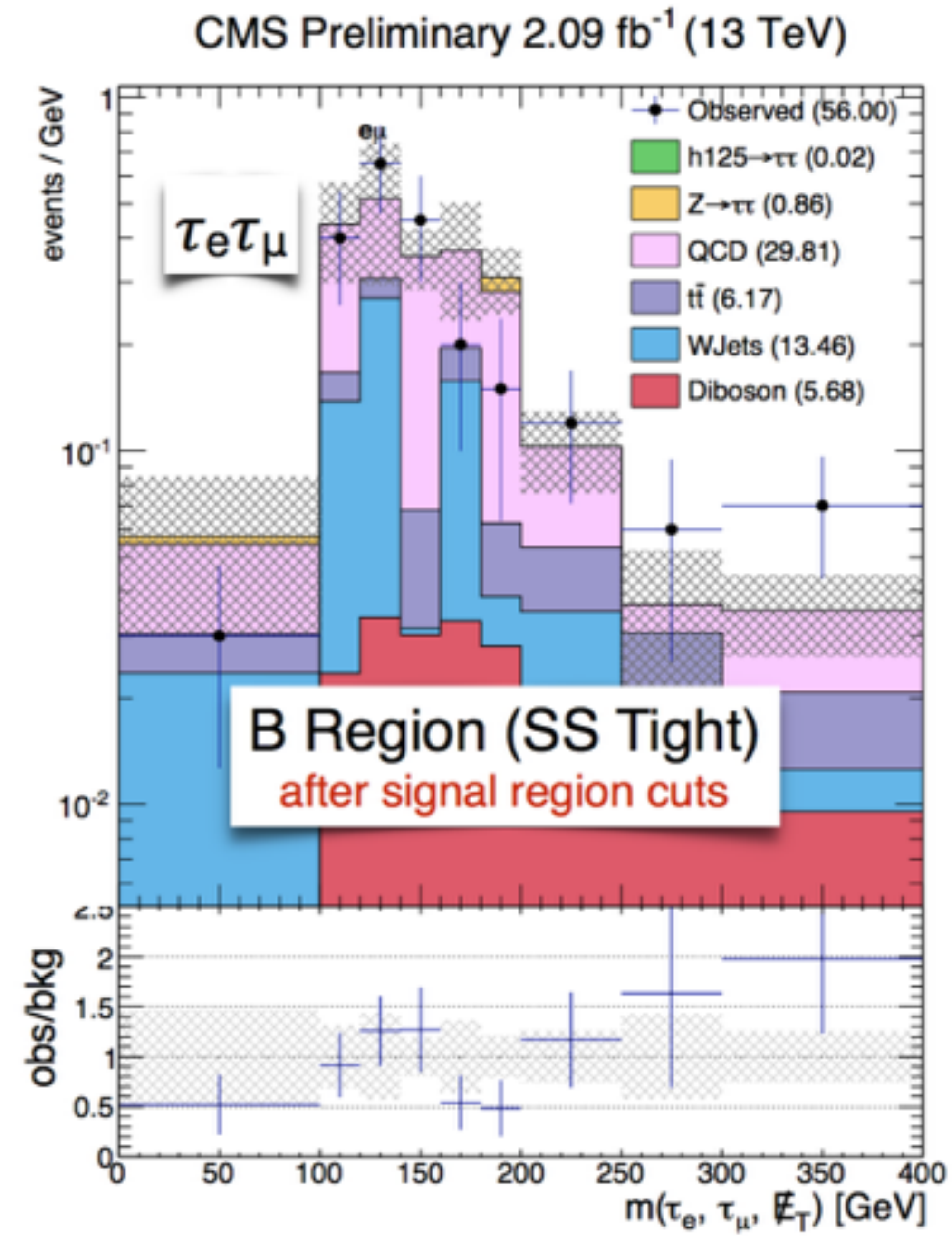
Choose anti-isolation region (0.15, 0.95)

- good modeling (p-value of χ^2) of QCD in isolated region
- high statistics

QCD Closure in $\tau_e\tau_h$, $\tau_e\tau_\mu$



Tight/Loose = 0.13 ± 0.02



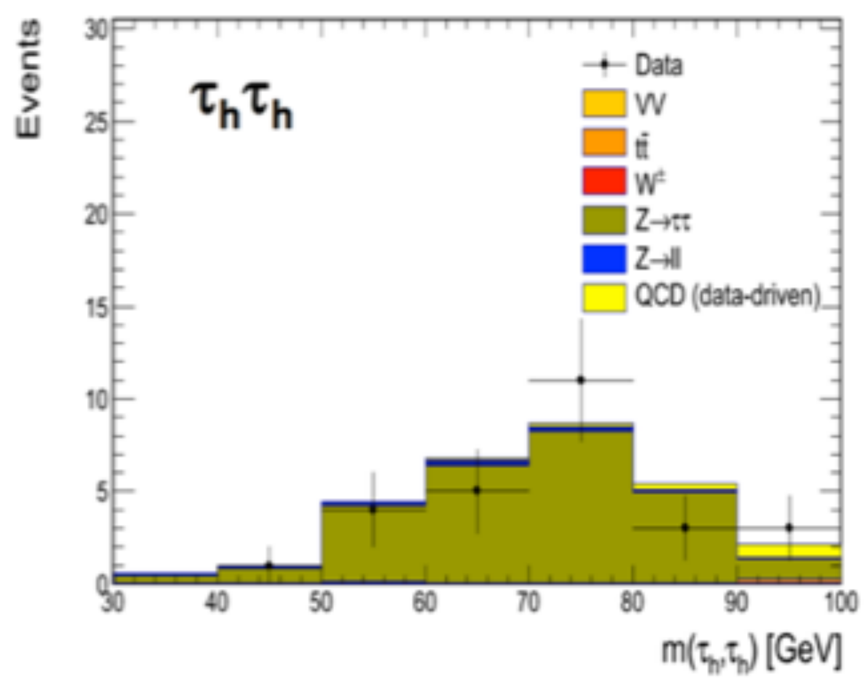
Tight/Loose = 0.20 ± 0.08

Good agreement between data/estimation in SS tight move to backup - methods have changed

DY Validation in $\tau_h\tau_h$ and $\tau_\mu\tau_h$

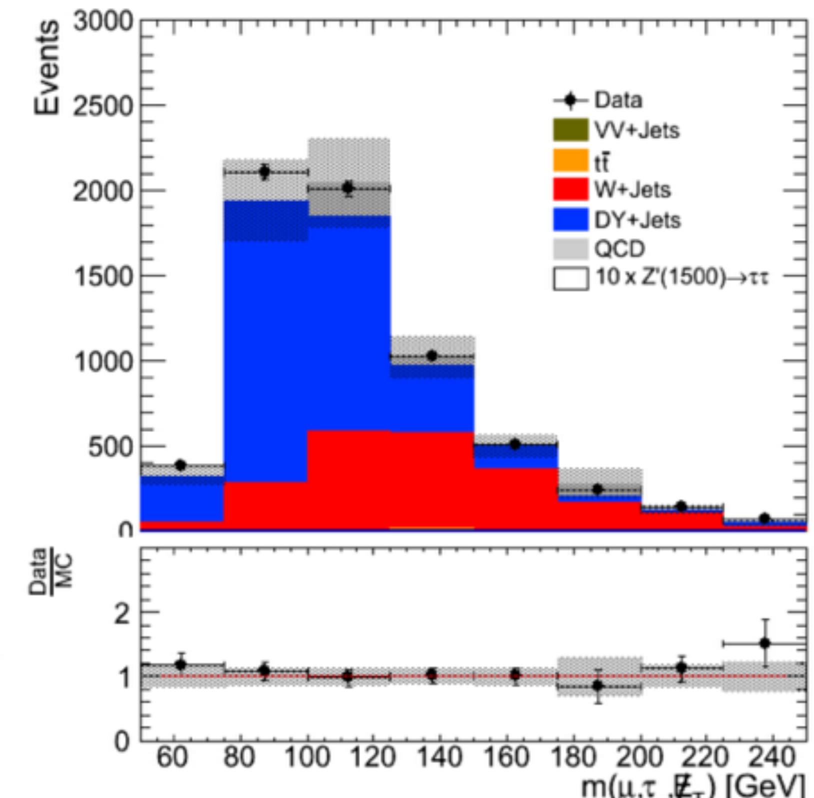
- Expect DY to be well-modeled by simulation
- Validate in low mass region
- Data/MC consistent with unity
- Estimate DY from MC

$\tau_h\tau_h$



(Data - bkg)/(DY MC) = 0.97 +- 0.19

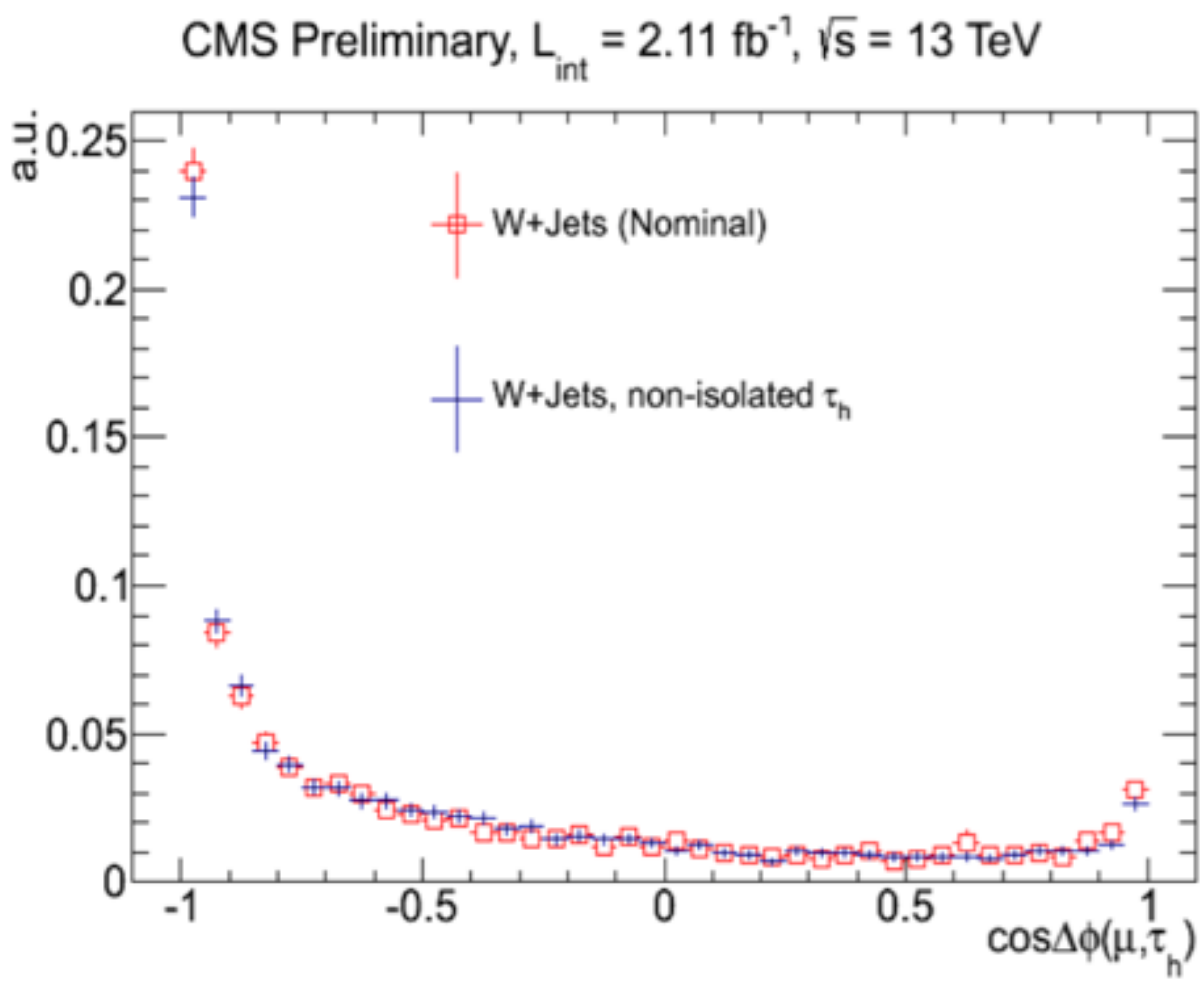
$\tau_\mu\tau_h$

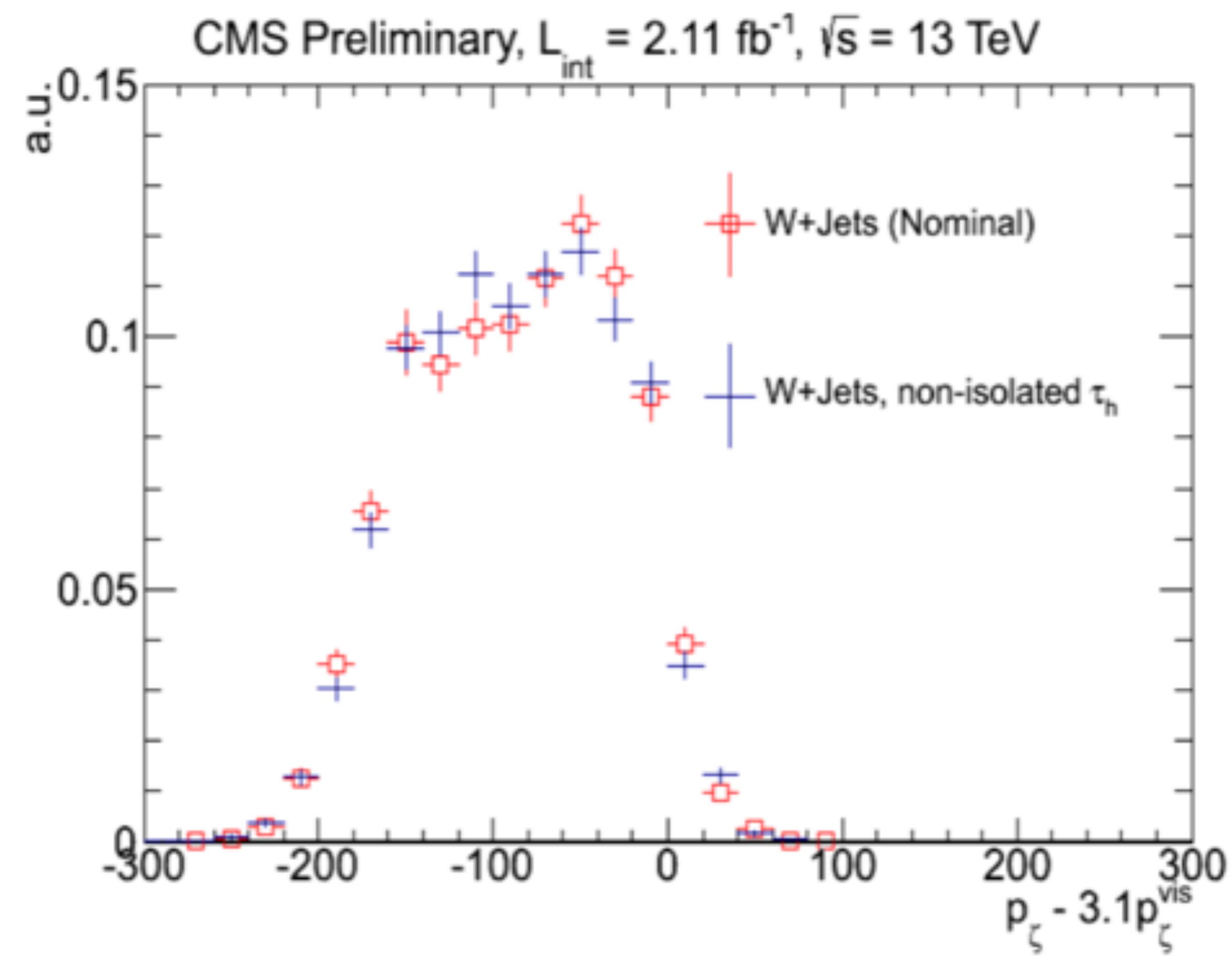


(Data - bkg)/(DY MC) = 0.96 +- 0.06

Systematic on DY estimation

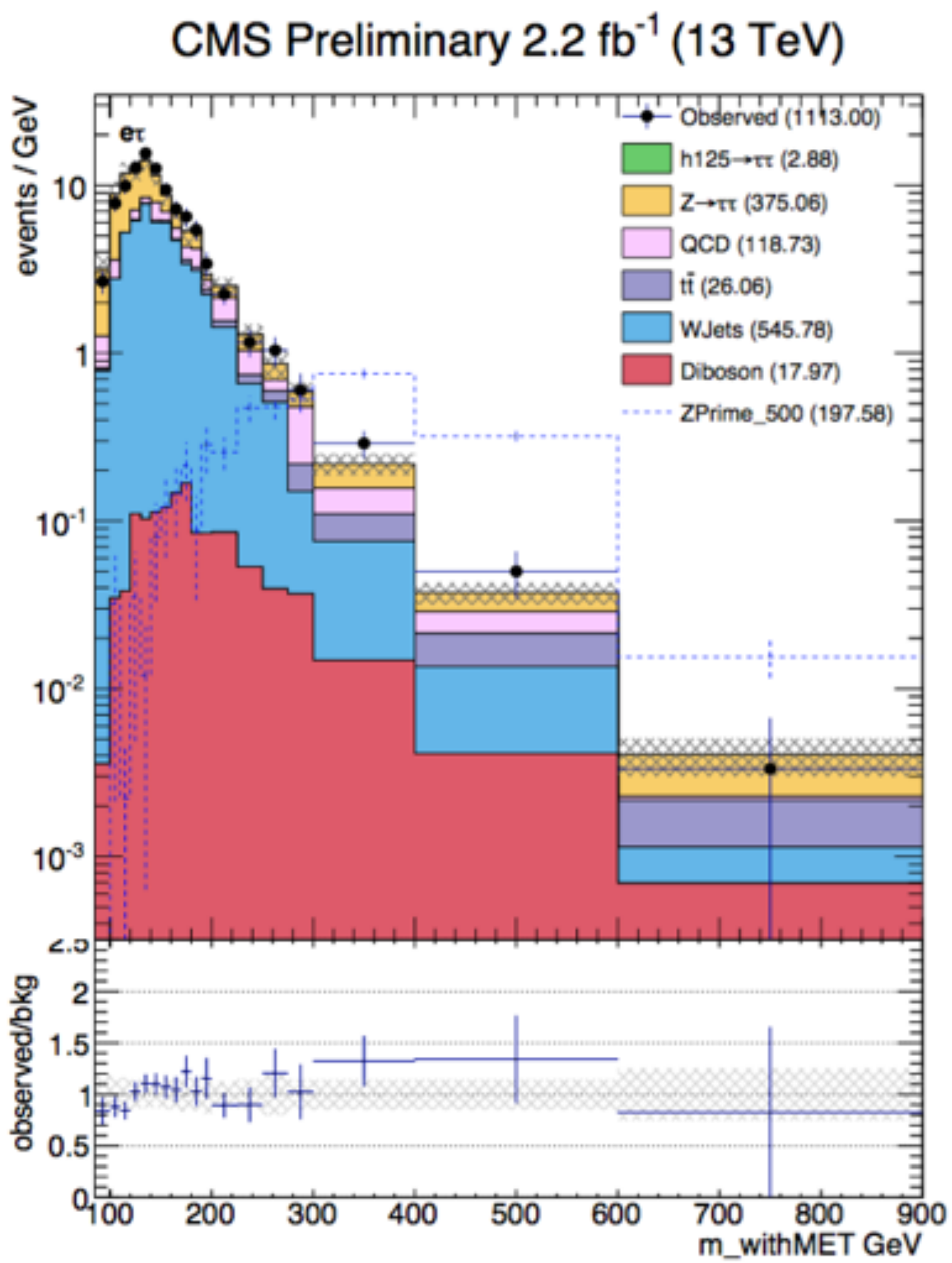
W+Jets Closure in $\tau_\mu\tau_h$



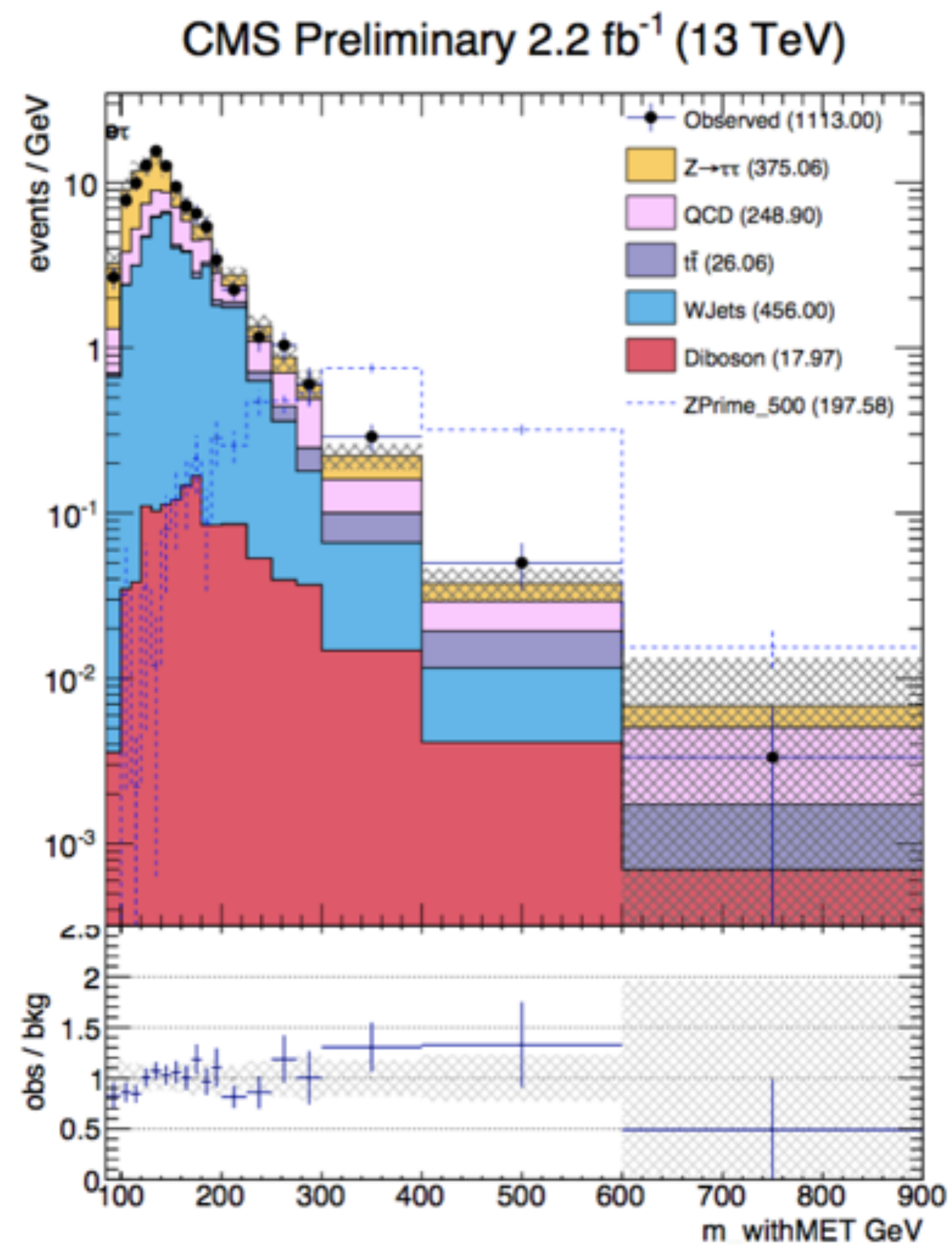


pZeta, Cos Δϕ correctly modeled in loose iso sidebands

MC vs data driven W+Jets in $\tau_e\tau_h$



MC based



Data driven

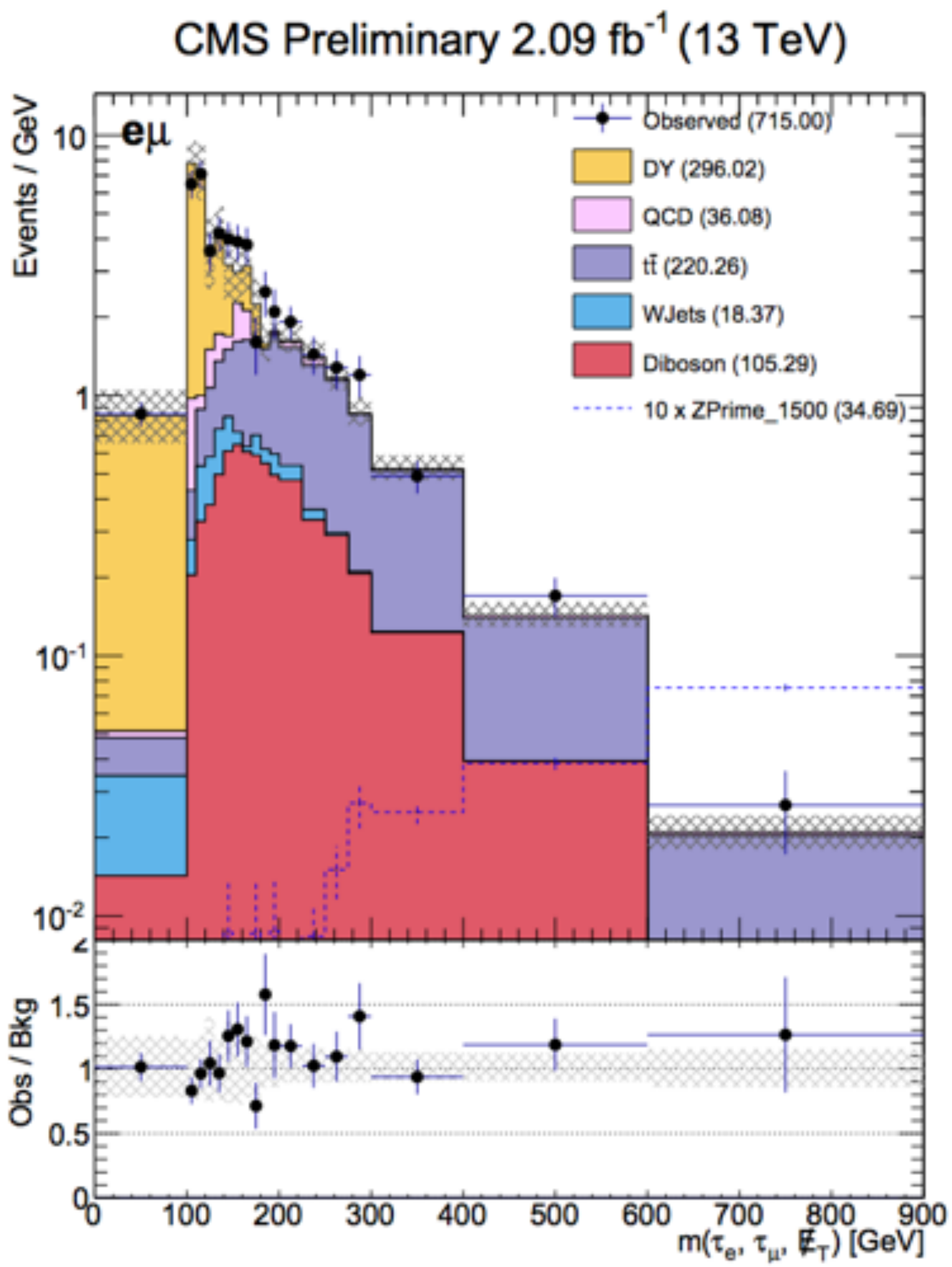
HT-binned DY vs mass-binned DY

- Mass-binned samples improve statistics in high-mass tail

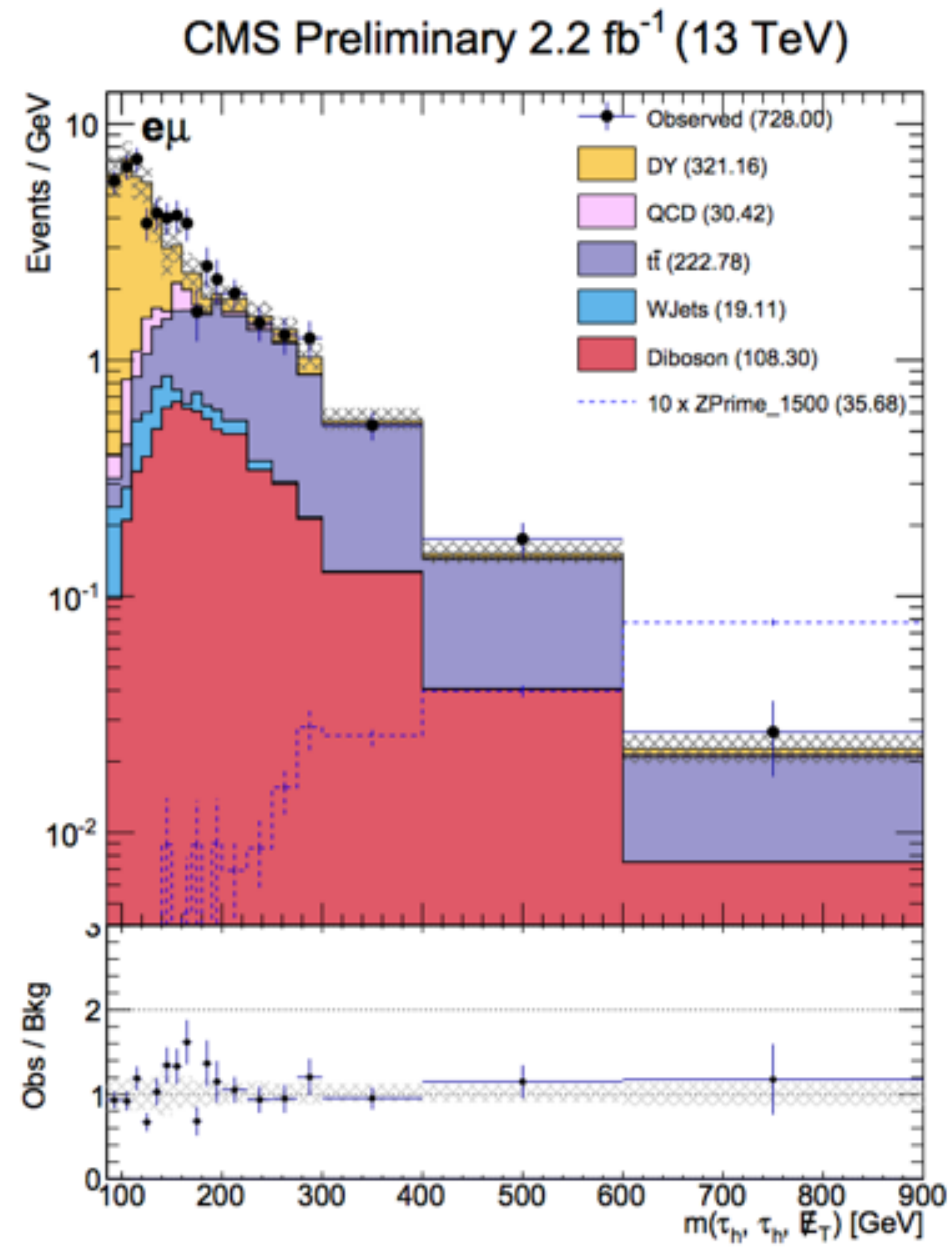
Process	cross-section (pb)	Official CMS Datasets (MINIAODSIM)
$Z \to LL$ (old)	6025.2	/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
HT binned	147.4*1.23	/DYJetsToLL_M-50_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
LO samples	40.99*1.23	/DYJetsToLL_M-50_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	5.678*1.23	/DYJetsToLL_M-50_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	2.198*1.23	/DYJetsToLL_M-50_HT-600toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
$Z \to LL$ (new)	6025.2	/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8
mass binned	7.67*0.987	/DYJetsToLL_M-200to400_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8
NLO samples	0.423*0.987	/DYJetsToLL_M-400to500_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8
	0.24*0.987	/DYJetsToLL_M-500to700_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8
	0.035*0.987	/DYJetsToLL_M-700to800_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8
	0.03*0.987	/DYJetsToLL_M-800to1000_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8
	0.016*0.987	/DYJetsToLL_M-1000to1500_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8

HT-binned DY vs mass-binned DY

$\tau_e \tau_\mu$



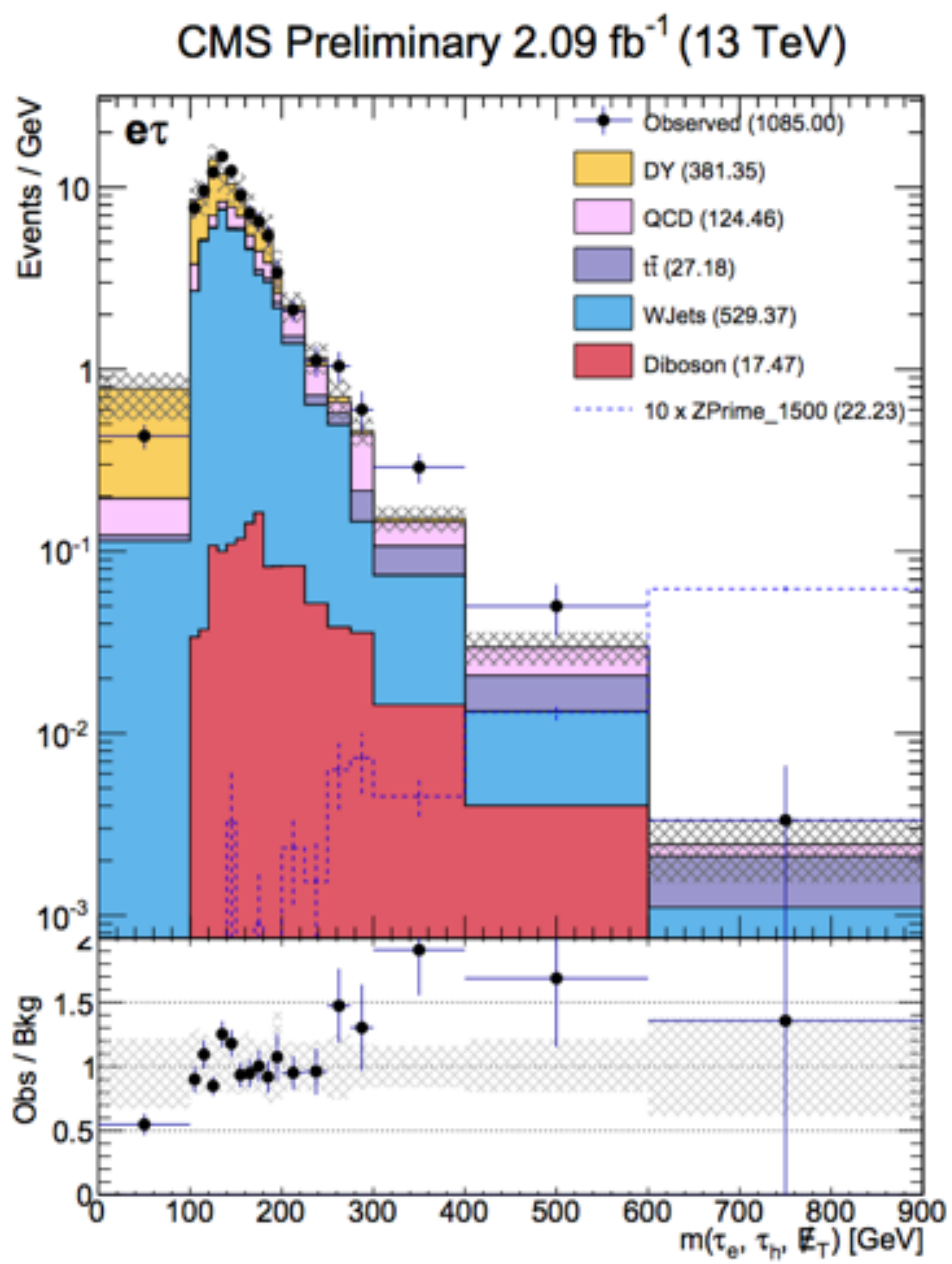
HT-binned



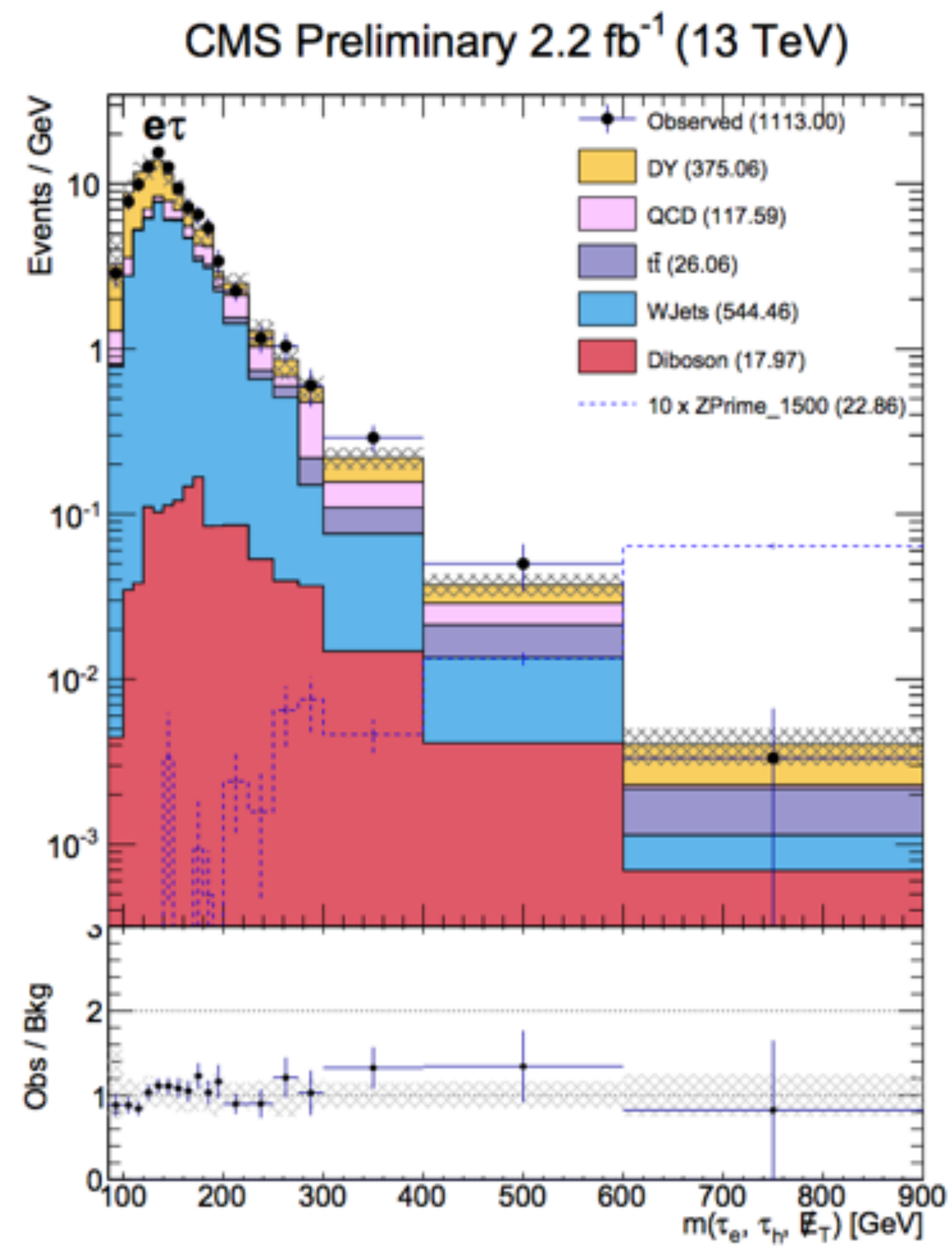
mass-binned

HT-binned DY vs mass-binned DY

$\tau_e\tau_h$



HT-binned



mass-binned

1+2+3 prongs vs 1+3 prongs

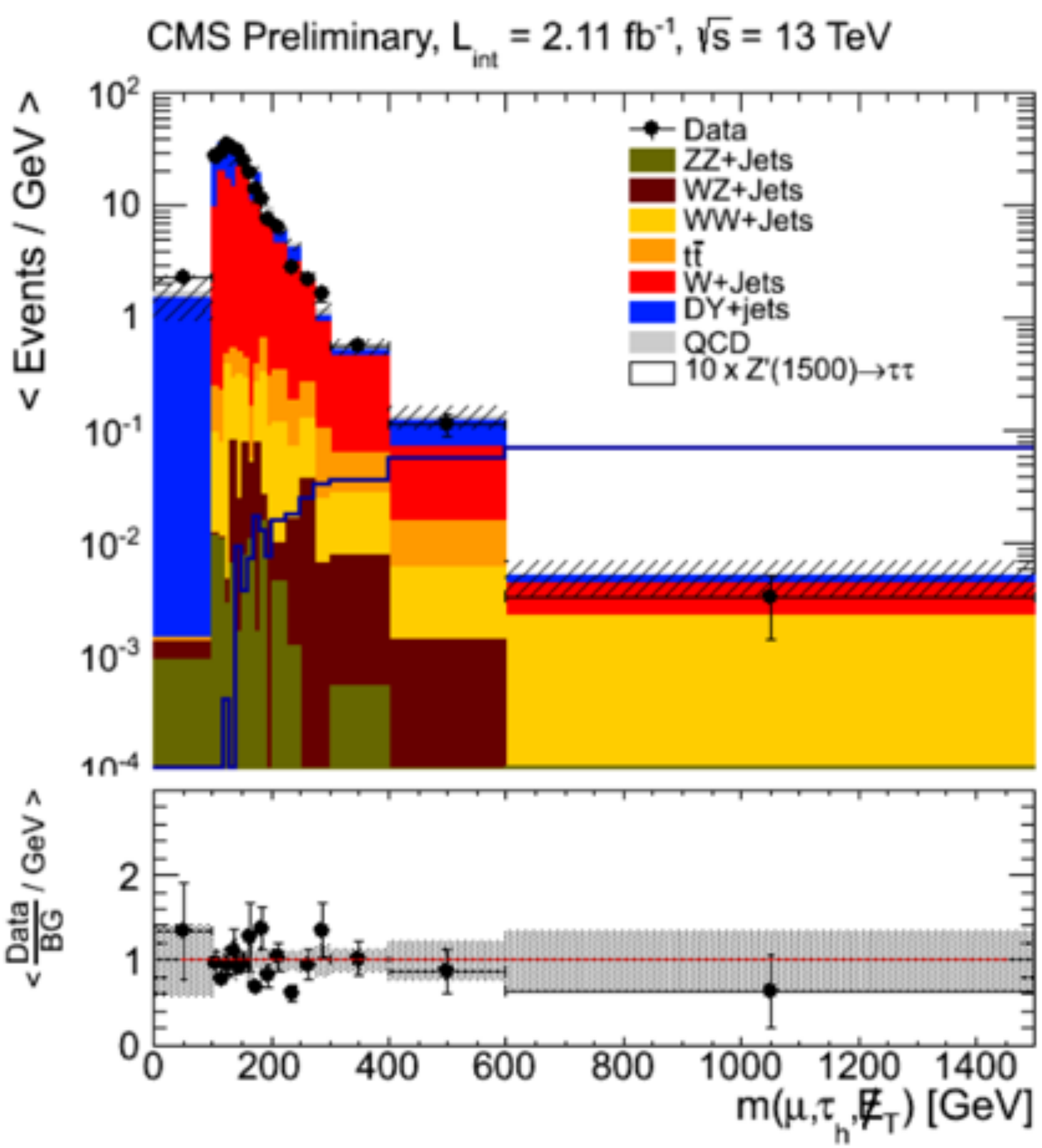
- $\tau_h\tau_h$ and $\tau_\mu\tau_h$ include 2-prong taus and $\tau_e\tau_h$ does not
- removing 2-prong taus from search region improves sensitivity
- 2-prong taus dropped from $\tau_h\tau_h$ and $\tau_\mu\tau_h$

$M_{Z'}$ (GeV)	500	1000	1500	2000	2500	3000
$\tau_e\tau_h$ (1,3)	0.713	0.173	0.122	0.107	0.104	0.105
$\tau_e\tau_h$ (1,2,3)	0.871	0.229	0.163	0.146	0.140	0.143
$\tau_\mu\tau_h$ (1,3)	0.588	0.105	0.065	0.055	0.052	0.052
$\tau_\mu\tau_h$ (1,2,3)	0.705	0.121	0.075	0.062	0.060	0.061
$\tau_h\tau_h$ (1,3)	0.428	0.069	0.045	0.042	0.039	0.038
$\tau_h\tau_h$ (1,2,3)	0.725	0.093	0.061	0.055	0.052	0.051

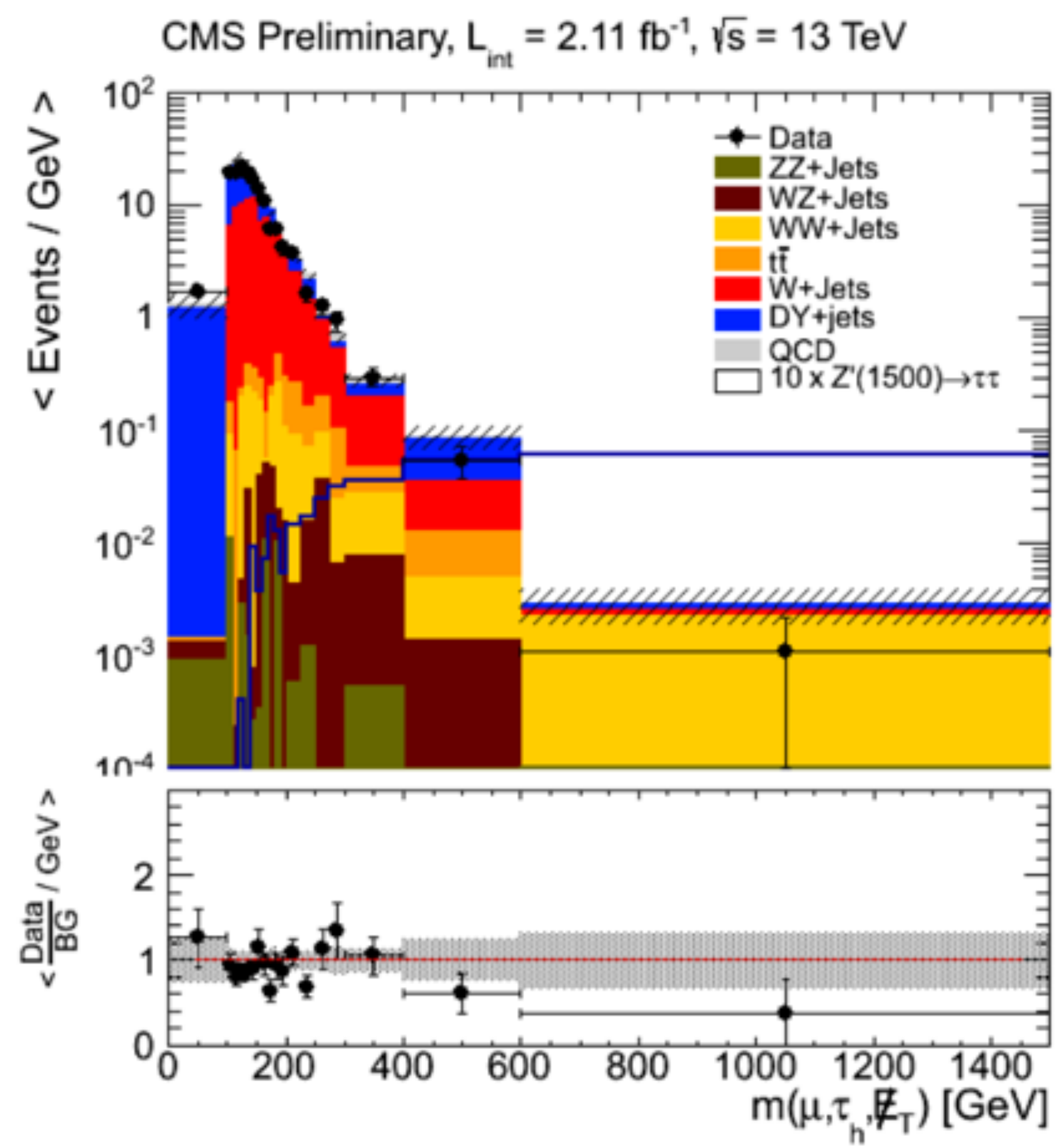
Expected limits (pb) of 1+3 prong taus vs 1+2+3 prong taus

1+2+3 prong vs 1+3 prong taus

$\tau_\mu \tau_h$



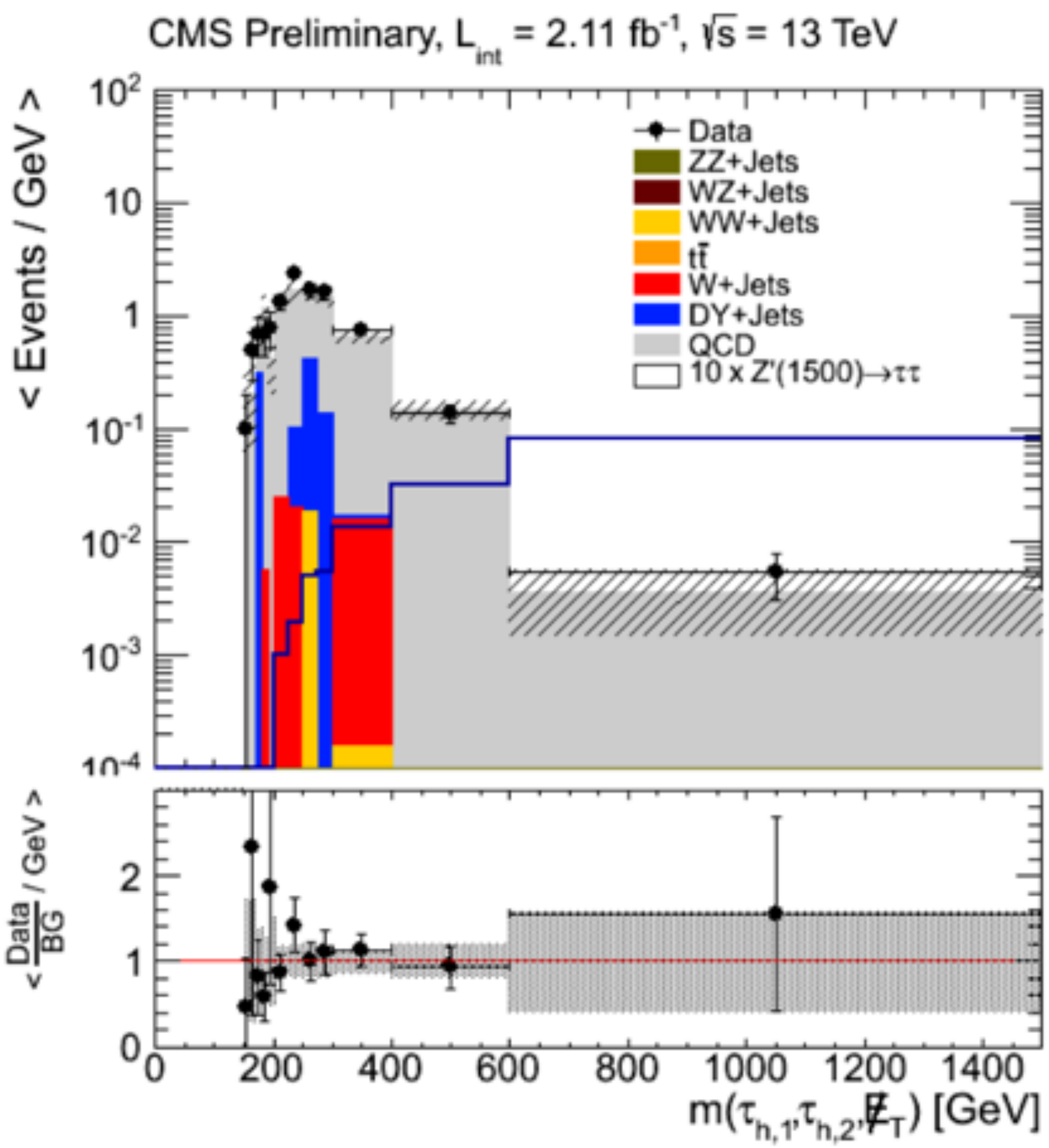
HT-binned, 1+2+3 prong



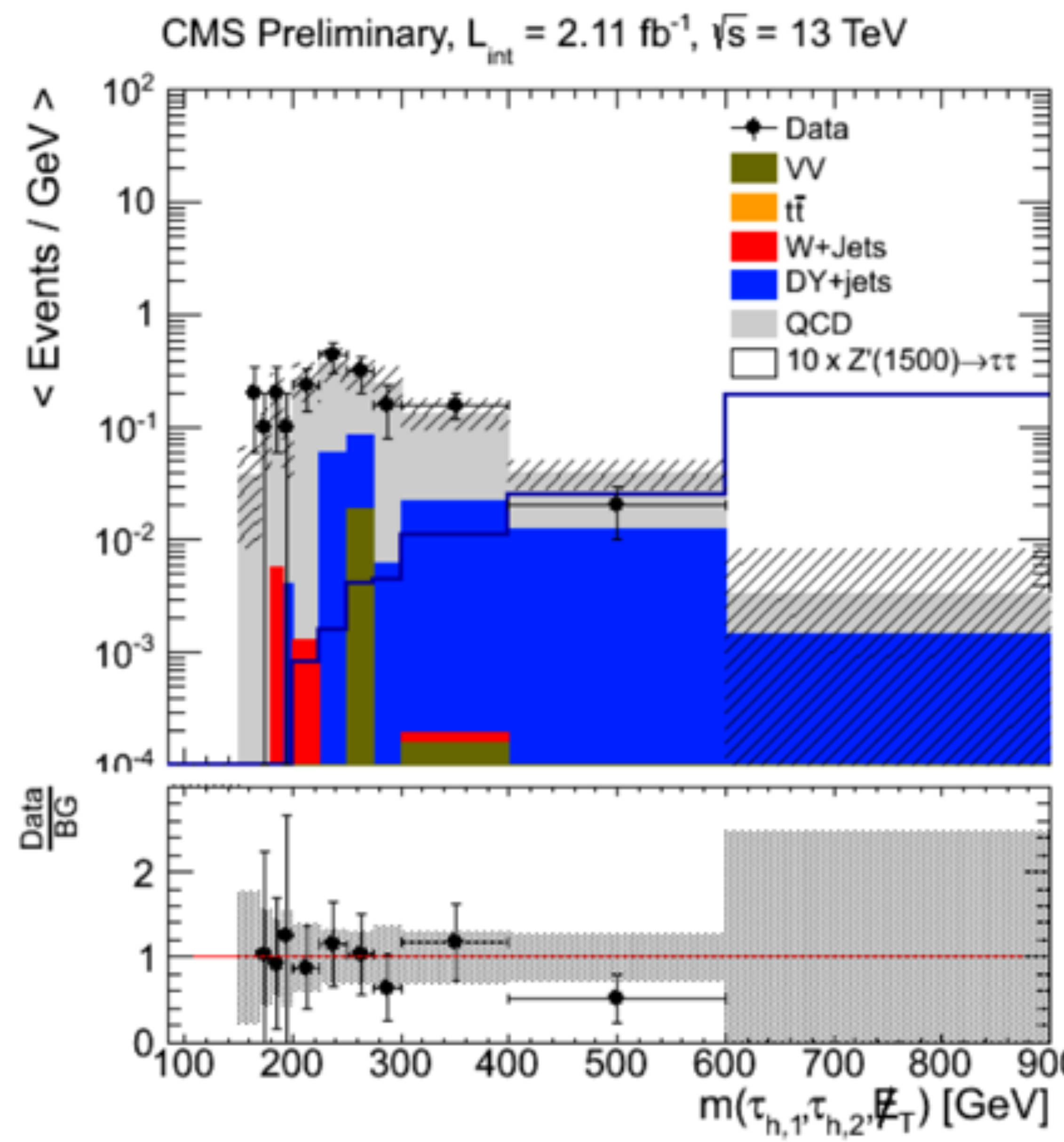
mass-binned, 1+3 prong

1+2+3 prong vs 1+3 prong taus

$\tau_h\tau_h$



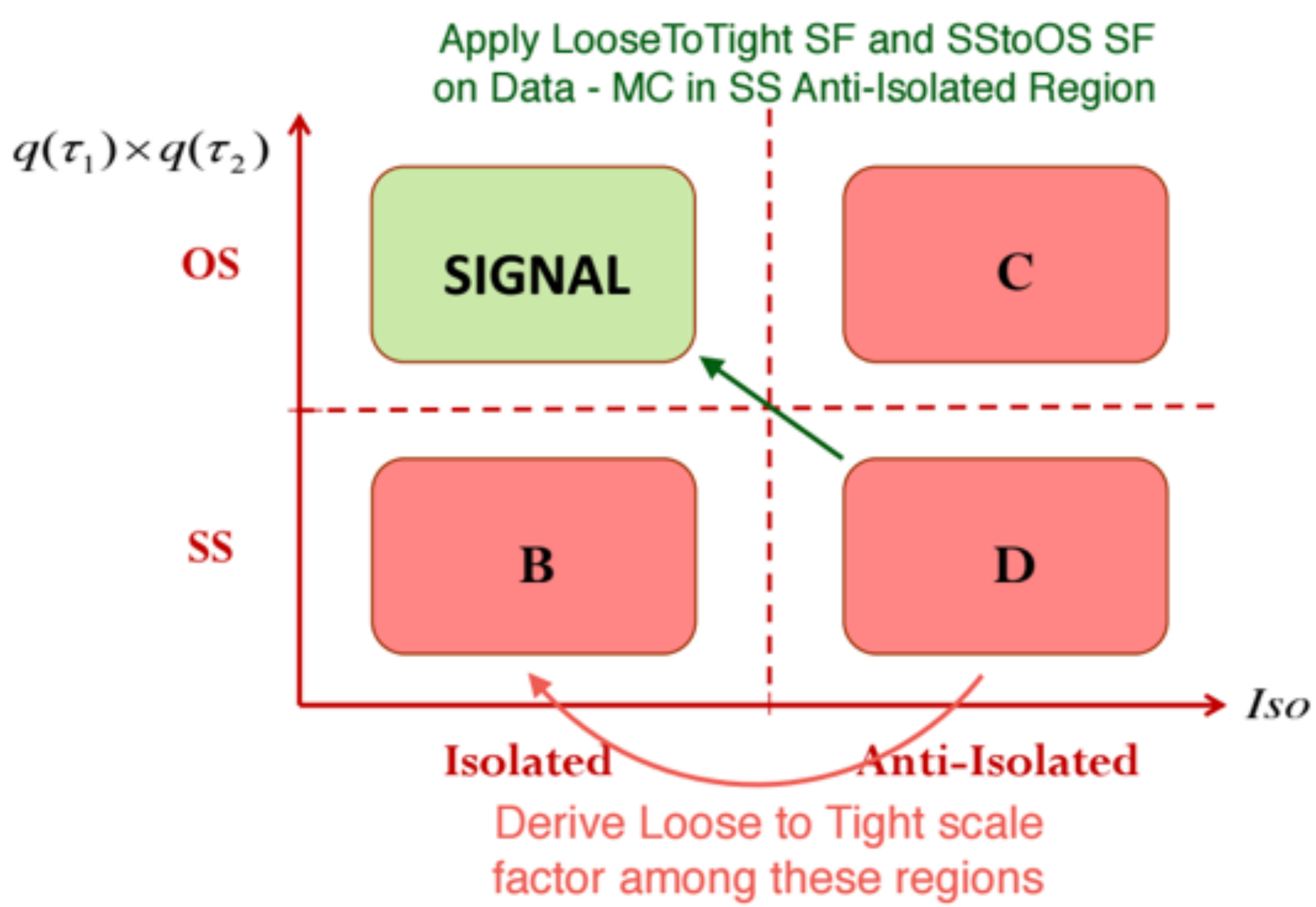
HT-binned, 1+2+3 prong



mass-binned, 1+3 prong

QCD in $\tau_e\tau_h$

- QCD shape now taken from SS anti-isolated region
- Signal estimation obtained by multiplying SS anti-isolated region by anti-isolated/isolated SF and SStoOS SF



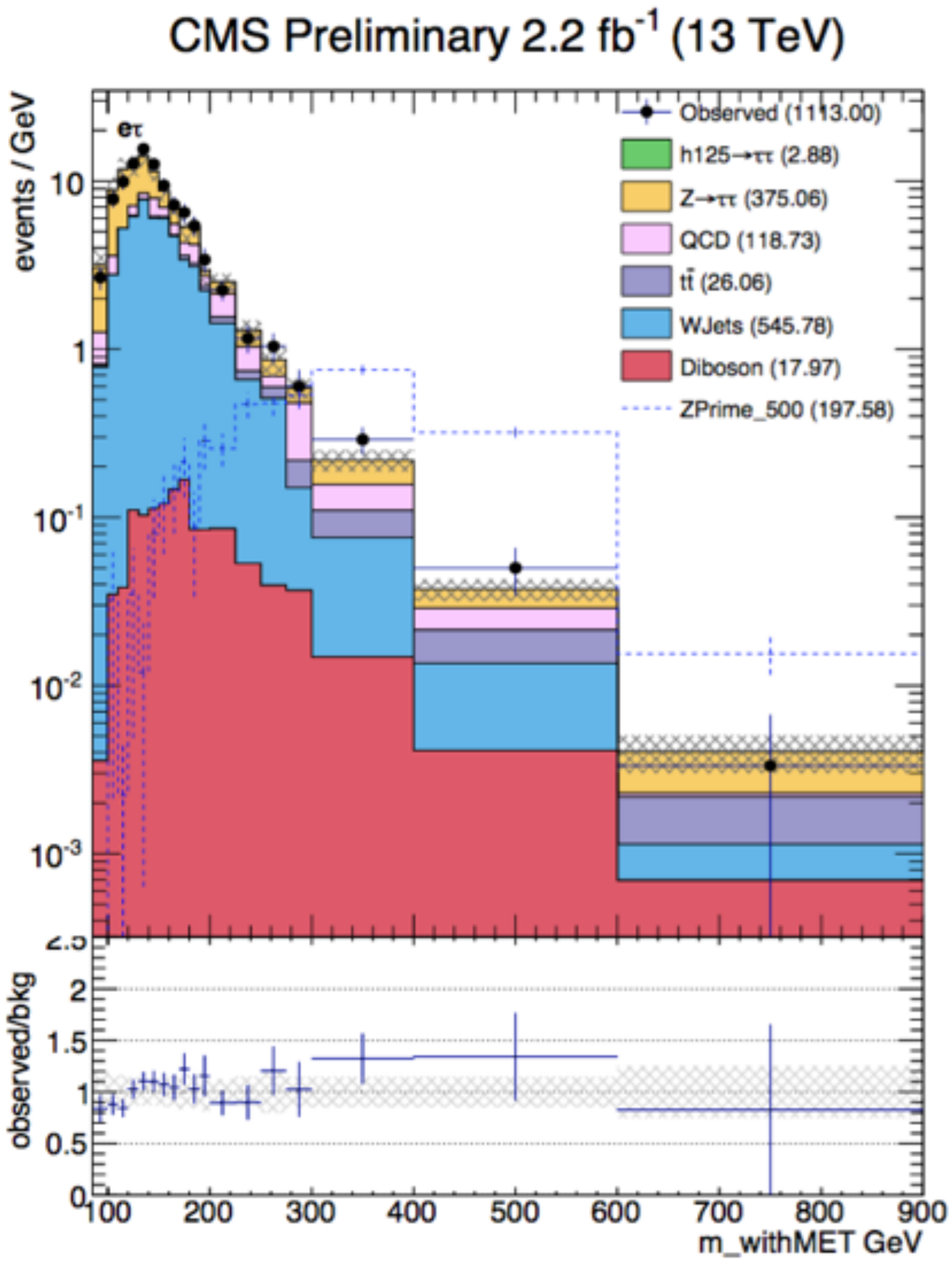
W+Jets in $\tau_e\tau_h$

- $\tau_e\tau_h$ switched from MC-based to data driven approach for W+Jets estimation
- A (signal) region: pass pZeta and cosDphi requirement, isolated tau
- B region: fail pZeta or cosDphi, isolated tau
- C region: pass pZeta and cosDphi, anti-isolated tau
- D region: fail pZeta or cosDphi, anti-isolated tau

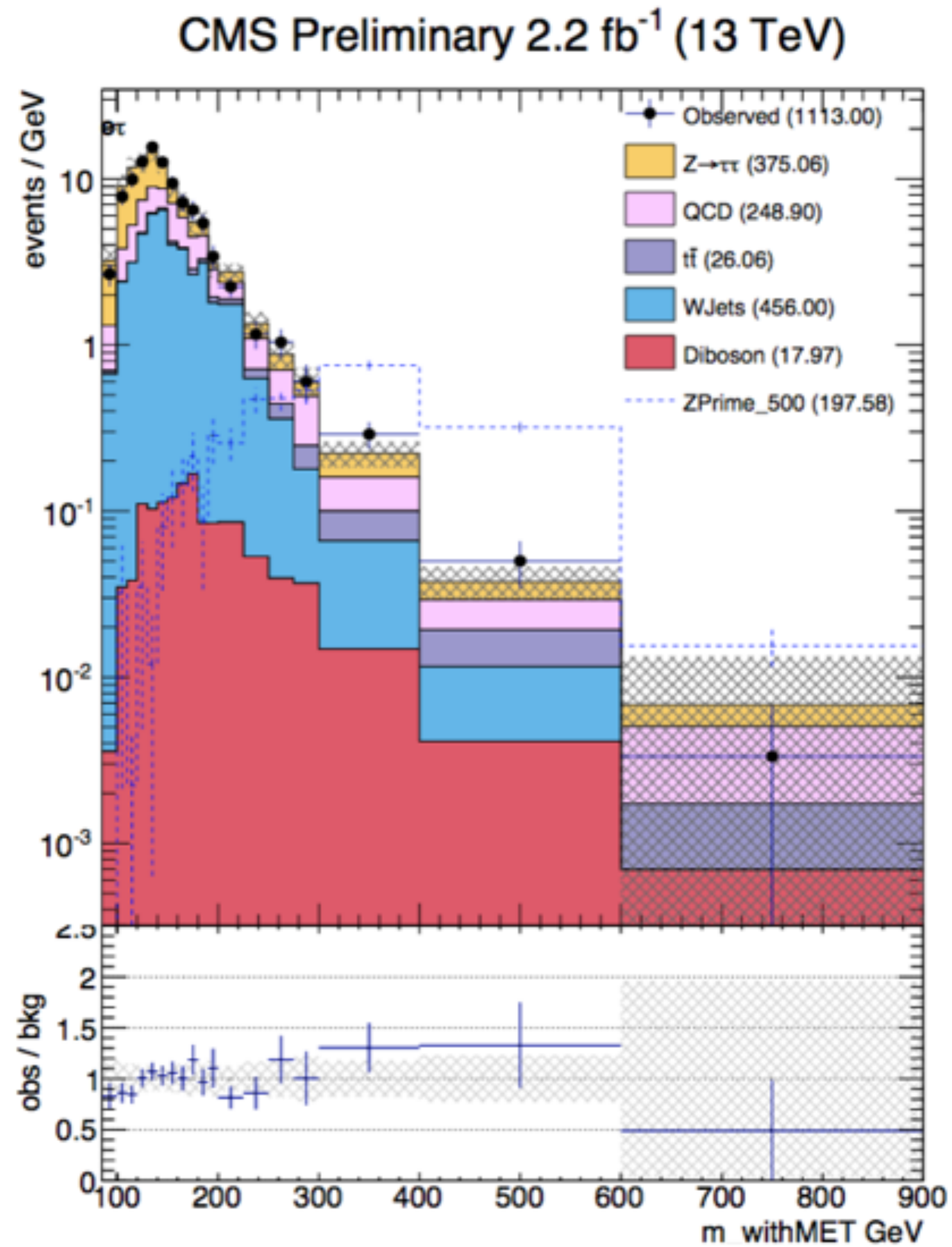
W+Jets estimation given as:

$$N^{\text{signal}}_{\text{W+jets}} = N^{\text{C}}_{\text{data - other bkg}} \times \frac{N^{\text{B}}_{\text{data - other bkgs}}}{N^{\text{D}}_{\text{data - other bkg}}}$$

W+Jets in $\tau_e\tau_h$



MC-based



Data driven

Systematics

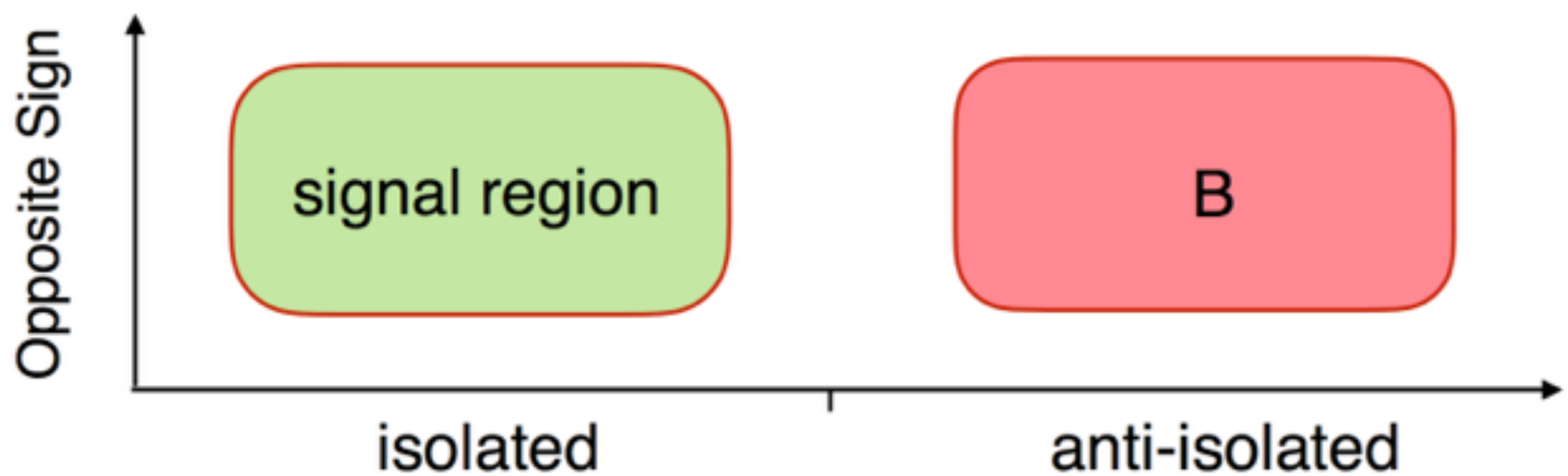
- Updated systematics to reflect new QCD, W+Jets estimation in $\tau_e\tau_h$
- Updated muon/electron ID systematics per ARC's suggestions
- "s" indicates shape-based uncertainty

Source , , , :	QCD $hh, \mu h, eh, e\mu$	W $hh, \mu h, eh, e\mu$	DY $hh, \mu h, eh, e\mu$	$t\bar{t}$ $hh, \mu h, eh, e\mu$	VV $hh, \mu h, eh, e\mu$	Signal $hh, \mu h, eh, e\mu$
Lumi	–,–,–,–	L,–,–,L	L,L,L,L	L,L,L,L	L,L,L,L	L,L,L,L
μ ID + Trig	–,–,–,–	–,–,–,7	–,7,–,7	–,7,–,7	–,7,–,7	–,7,–,7
e ID + Trig	–,–,–,–	–,–,–,6	–,–,6,6	–,–,6,6	–,–,6,6	–,–,6,6
τ_h Trig	–,–,–,–	10,–,–,–	10,–,–,–	10,–,–,–	10,–,–,–	10,–,–,–
τ_h ID	–,–,–,–	30,–,–,–	12,6,6,–	12,6,6,–	12,6,6,–	12,6,6,–
b ID	–,–,–,s	10,–,–,s	3,3,s,s	10,12,s,s	3,3,s,s	3,3,s,s
JES	–,–,–,s	12,–,–,s	8,s,s,s	12,s,s,s	8,s,s,s	2,2,s,s
TES	–,–,–,s	11,–,–,s	11,s,s,s	11,s,s,s	8,s,s,s	3,3,s,s
MMS	–,–,–,–	–,–,–,1	–,1,–,1	–,1,–,1	–,1,–,1	–,1,–,1
EES	–,–,–,–	–,–,–,1	–,–,1,1	–,–,1,1	–,–,1,1	–,–,1,1
top p_T	–,–,–,–	–,–,–,–	–,–,–,–	–,–,–,s	–,–,–,–	–,–,–,–
pdf	–,–,–,–	–,–,–,–	–,–,–,–	–,–,–,–	–,–,–,–	(1–12)
bin-by-bin stat.	s,s,s,s	s,s,s,s	s,s,s,s	s,s,s,s	s,s,s,s	s,s,s,s
Closure+Norm.	20,67,18,37	5,8,9,41	19,6,10,10	8,8,8,8	15,15,15,15	
W+Jets MC Norm.	–,–,–,–	–,–,7,–	–,–,–,–	–,–,–,–	–,–,–,–	–,–,–,–

MC-based W+Jets estimation in $\tau_e\tau_h$

W+Jets Estimation in $\tau_e\tau_h$ and $\tau_e\tau_\mu$

Method: take shape from MC in anti-isolated “B” region (same as defined in QCD study), normalize to yield in SR

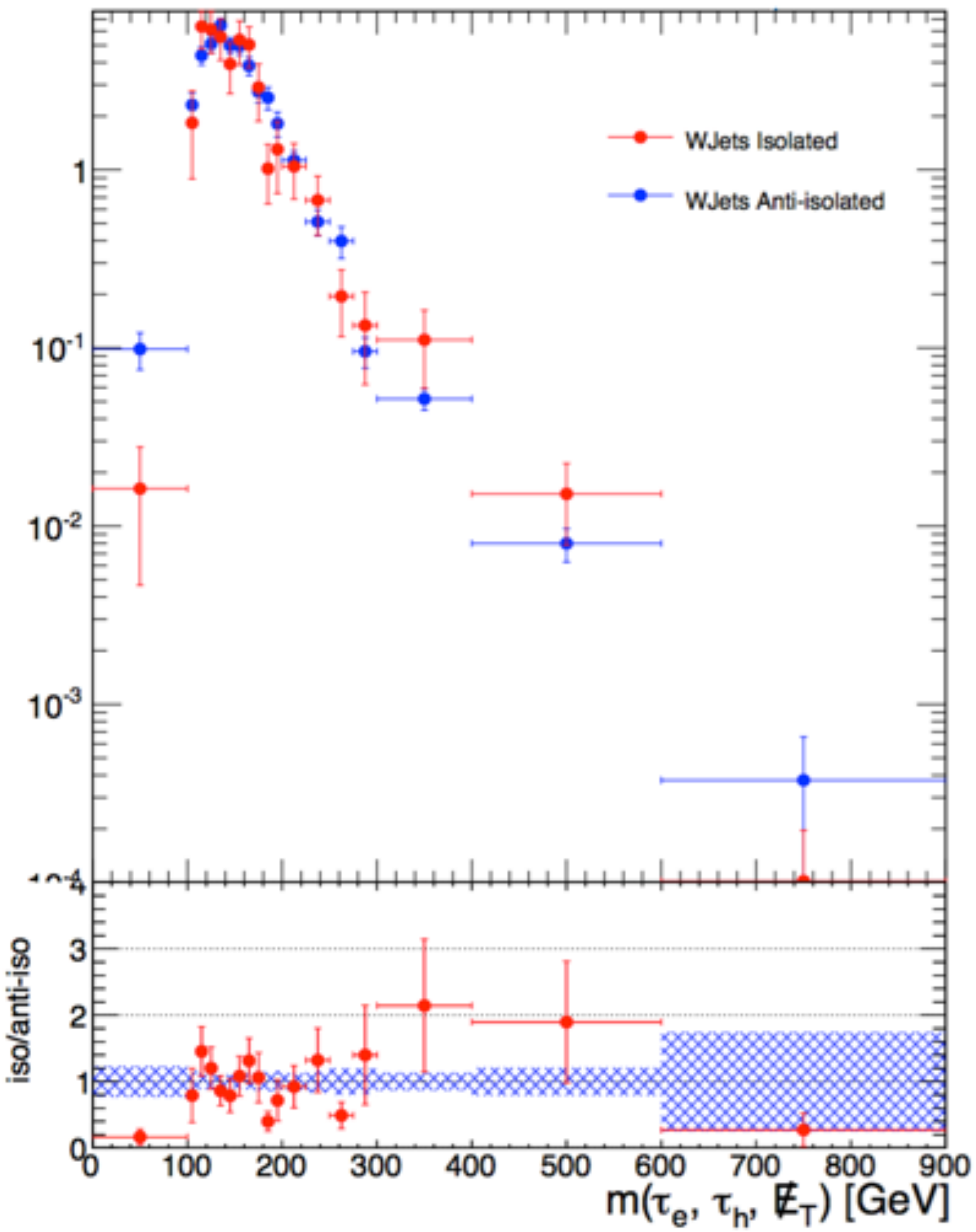


- Due to finite MC statistics, shape in anti-isolated region is much smoother
- QCD contamination in anti-isolated region limits effectiveness of fully data-driven estimation

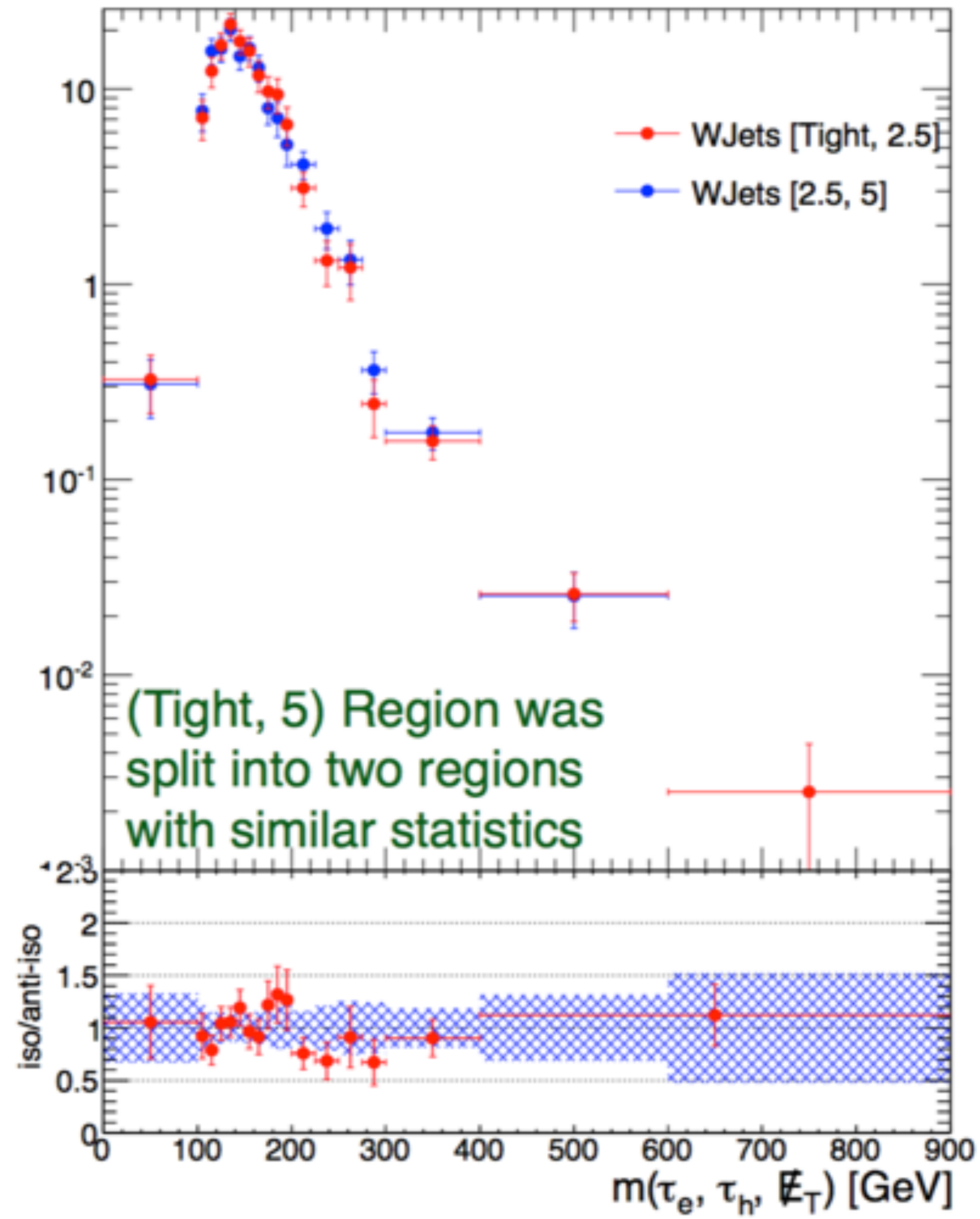
Checks:

- Does anti-isolated shape correctly model isolated shape?
- Extraction of SF for W+Jets yield in CR

W+Jets Closure in $\tau_e\tau_h$



Good agreement between TauIsoTight and Tight < TauIso < 5



Good agreement between Tight < TauIso < 2.5 and 2.5 < TauIso < 5

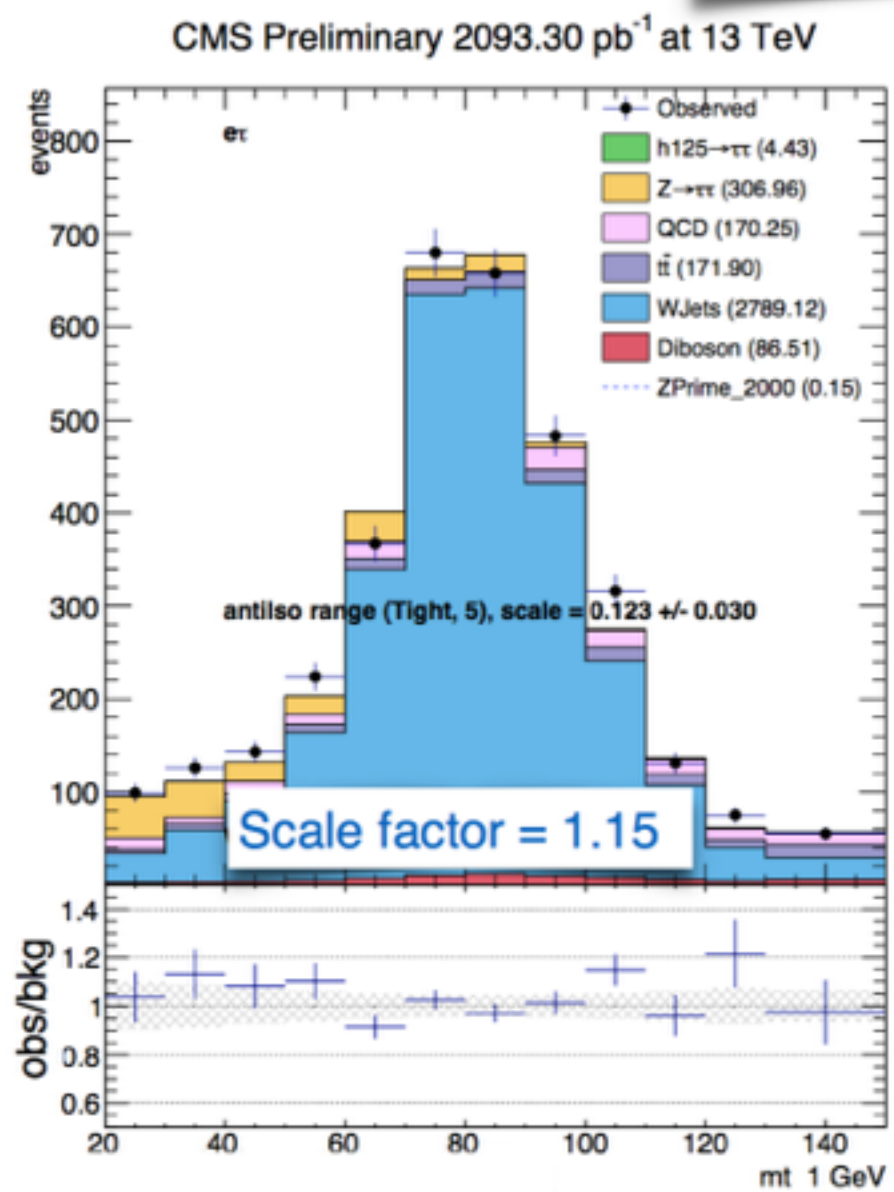
W+Jets shape is compatible between iso regions

W+Jets Estimation in $\tau_e\tau_h$

Control region: inverted pZeta cut

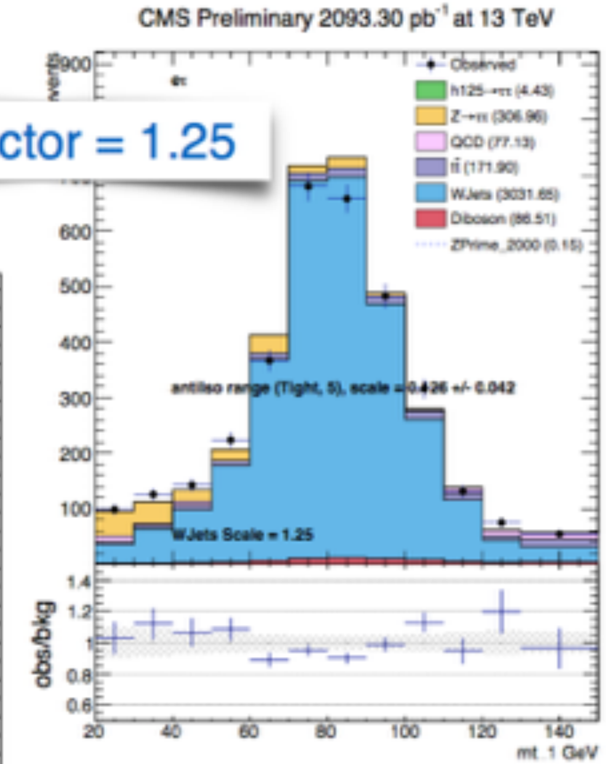
$p_\zeta - 3.1^* p_{\zeta,vis} < -50$

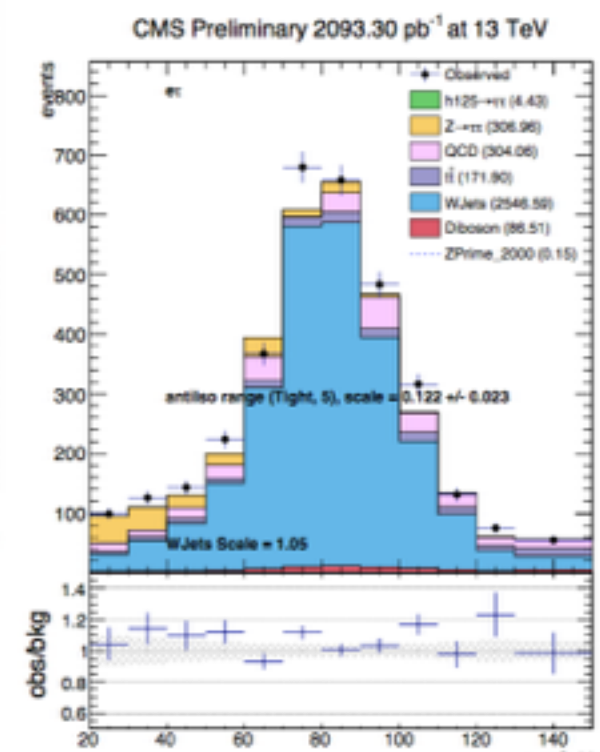
Run scan of W+Jets event rate from 0.8 to 1.25 in steps of 0.05. For each rate, perform QCD estimation



Scale factor = 1.15

Scale factor = 1.25





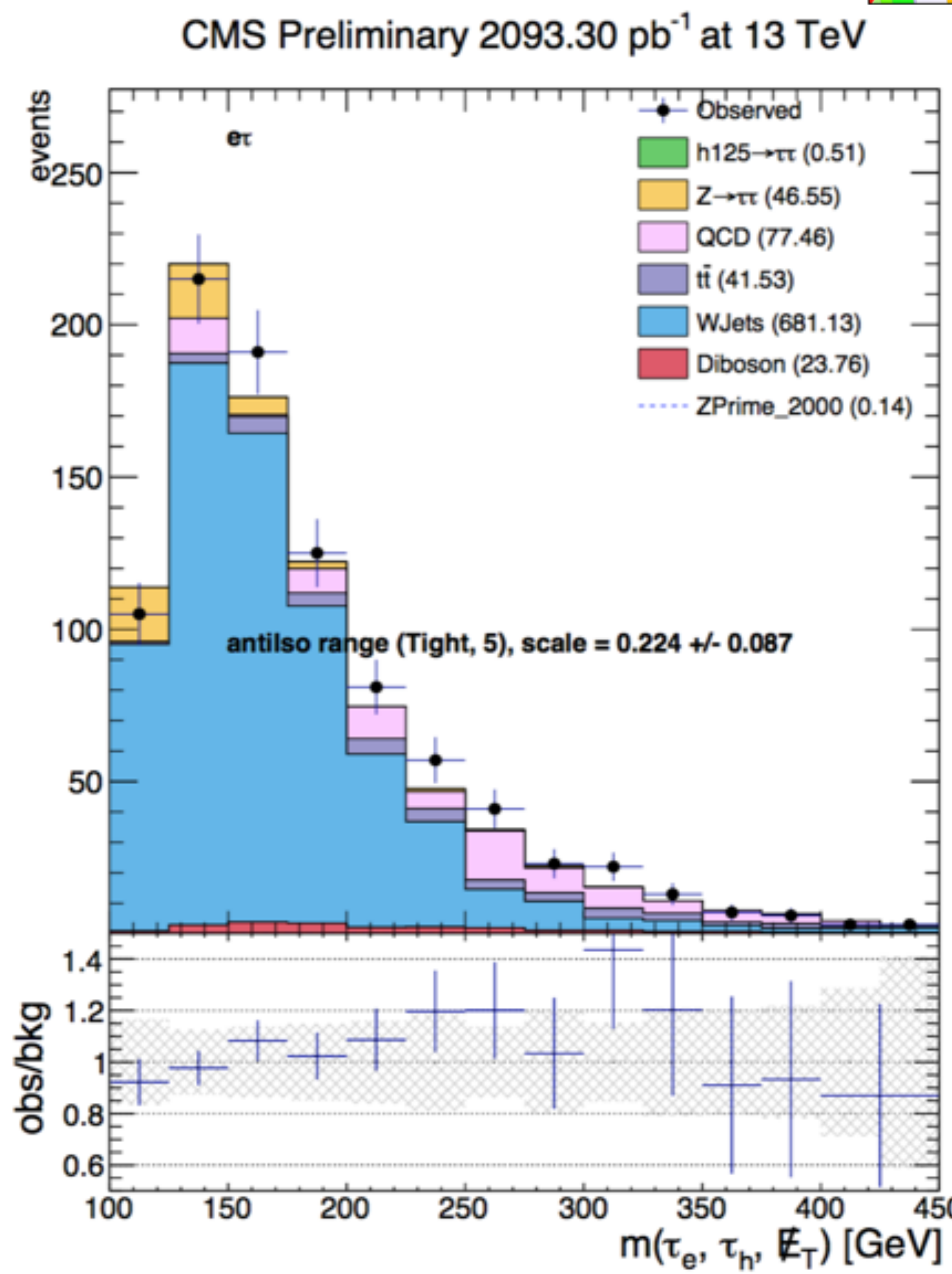
Scale factor = 1.05

SF = 1.15+-0.10 gives best agreement

W+Jets Closure in $\tau_e\tau_h$

W-enriched CR:

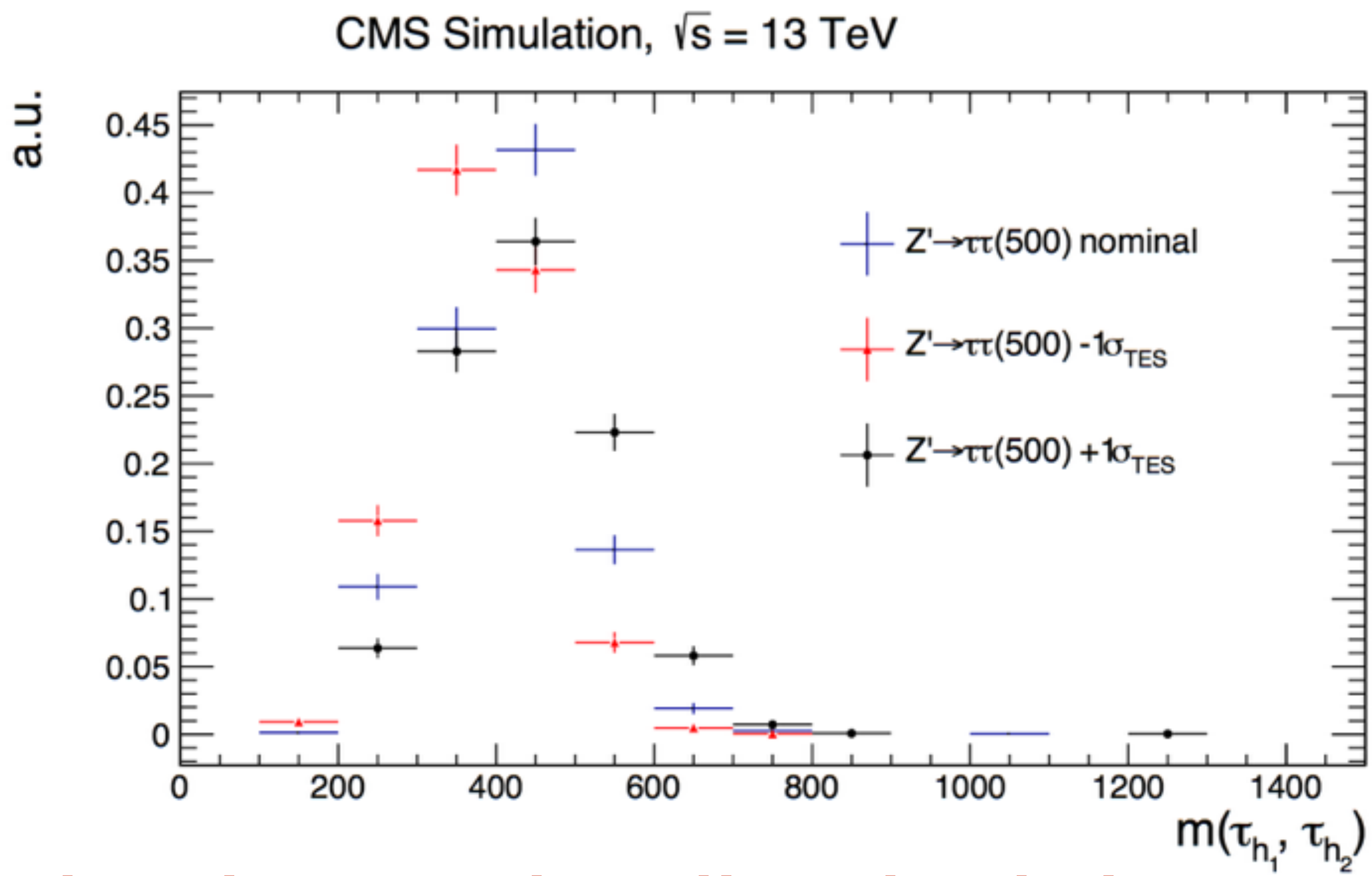
- p_ζ cut inverted
- cos $\Delta\phi$ cut loosened (< -0.8)



Good agreement between data and BG estimation

Shape Systematics

TES in 500 GeV Signal Sample



"s" indicates shape systematics - treated as uncorrelated across bins

all other systematics treated as 100% correlated

W+Jets Estimation in $\tau_e\tau_h$: post-unblinding

- As discussed, moved to data-driven W+Jets estimation
- Updated SFs shown in at the bottom of the table

Process	A Region	B Region	C Region	D Region
Drell-Yan	375.1 ± 31.4	550.5 ± 55.2	444.0 ± 41.3	767.2 ± 66.6
W+jets	474.6 ± 42.1	2688.6 ± 99.6	3242.8 ± 140.6	17130.7 ± 580.3
Diboson	18.0 ± 1.1	100.6 ± 4.3	32.8 ± 2.4	224.7 ± 13.4
$t\bar{t}$	26.1 ± 1.5	192.8 ± 10.5	74.5 ± 3.3	600.6 ± 22.4
Multijet	248.9 ± 13.7	553.9 ± 104.4	2025.9 ± 238.8	5009.8 ± 571.9
Observation	1113	4159	5203	22527
Purity	42%	66%	56%	72%
Observation - $\sum_{i \neq W} BG_i$	-	2761.2 ± 118.6	2625.8 ± 242.4	15924.7 ± 576.4
Data-driven W+Jets estimation	456.0 ± 27.6	-	-	-
SF	0.96 ± 0.10	1.03 ± 0.06	0.81 ± 0.08	0.93 ± 0.05