Tau Reconstruction at CMS

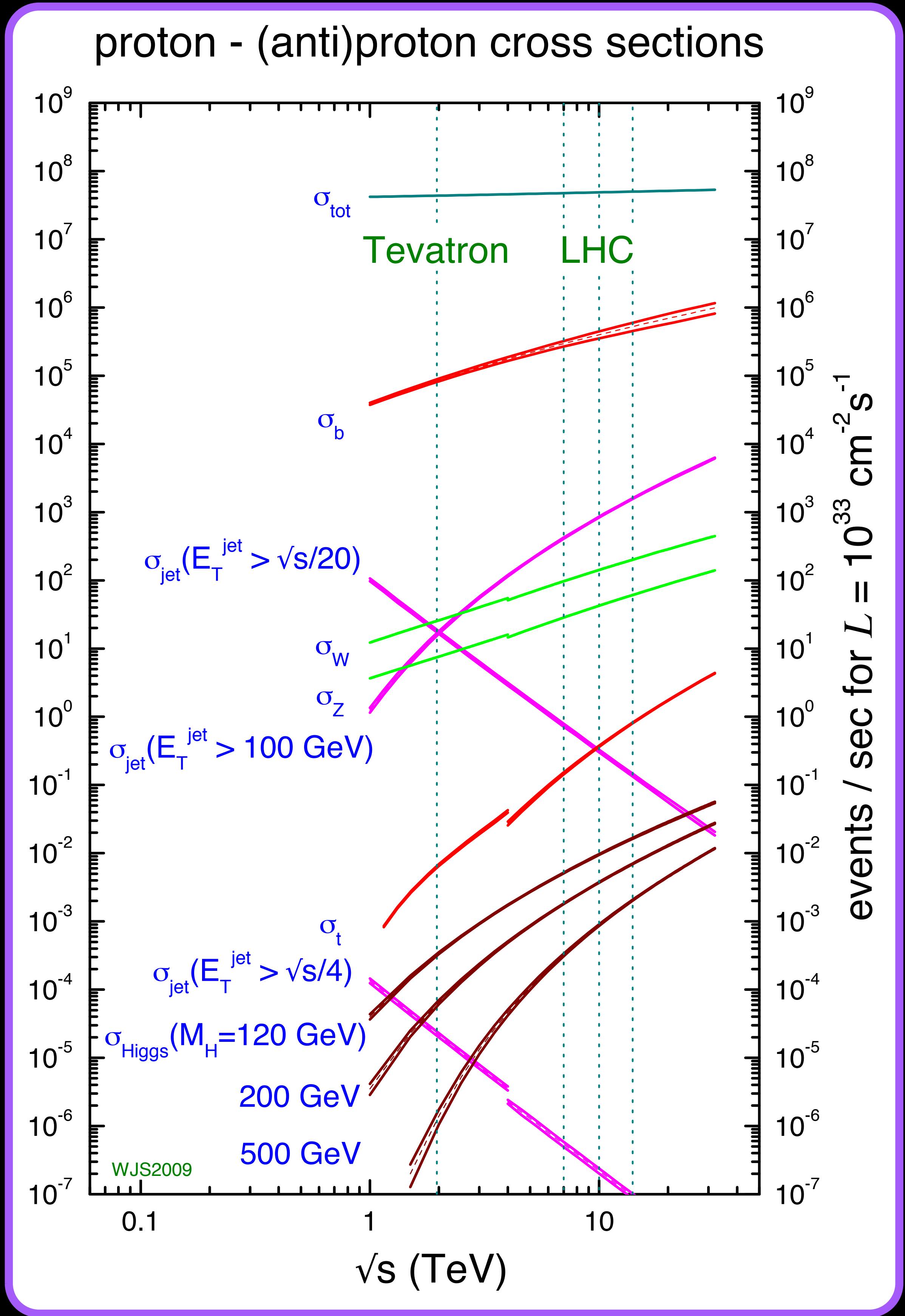
Evan Friis

Outline

- Taus at hadron colliders
- Tau ID algorithms at CMS
- Tau ID measurements in 2010
- $Z \rightarrow \tau\tau$ standard candle
- SVfit: τ pair mass reconstruction
- MSSM $H \rightarrow \tau\tau$

Tau ID at hadron colliders

- Exciting new physics signals could appear in tau channels
- QCD Jets can fake taus
- QCD background is very large!
- Dominant background in many analyses is fake taus



http://projects.hepforge.org/mstwpdf/plots/plots.html

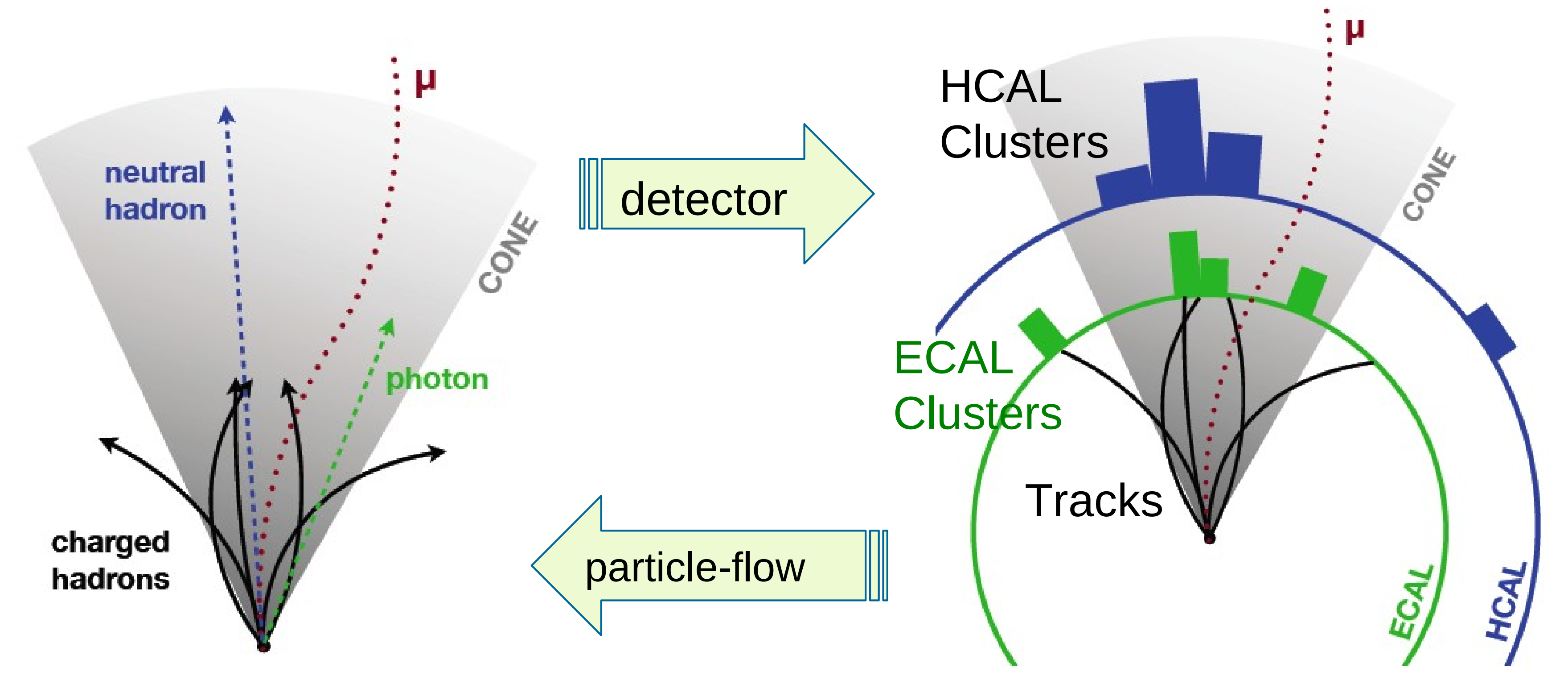
Particle Flow Algorithm

taus in CMS are built using Particle Flow objects

- Clusters and links signals from all subdetectors
- Produces a list of particle candidates

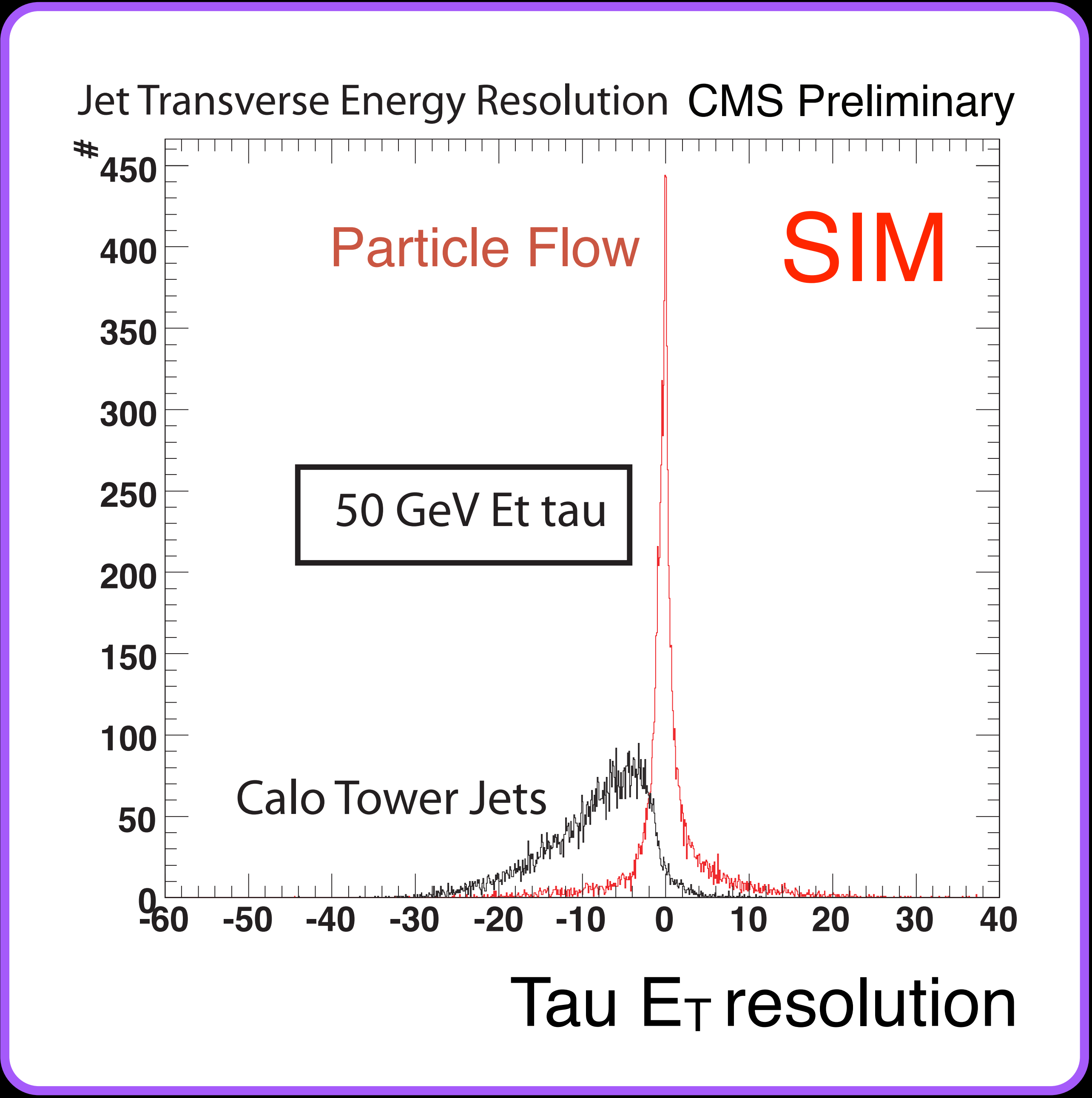
$$h\ h^0\ e\ \mu\ \gamma$$

- To the user looks just like Monte Carlo



F. Beaudette

Particle Flow Performance



Particle Flow resolves overlaps in subdectors

Improves resolution

Traditional CMS Tau ID

geometrically defined isolation

define geometric region around tau
candidate and require low detector activity

relies on the fact that taus
are more collimated than QCD

CMS Physics Technical Design Report (PTDR)
results use these algorithms

see CMS PAS PFT-08-001

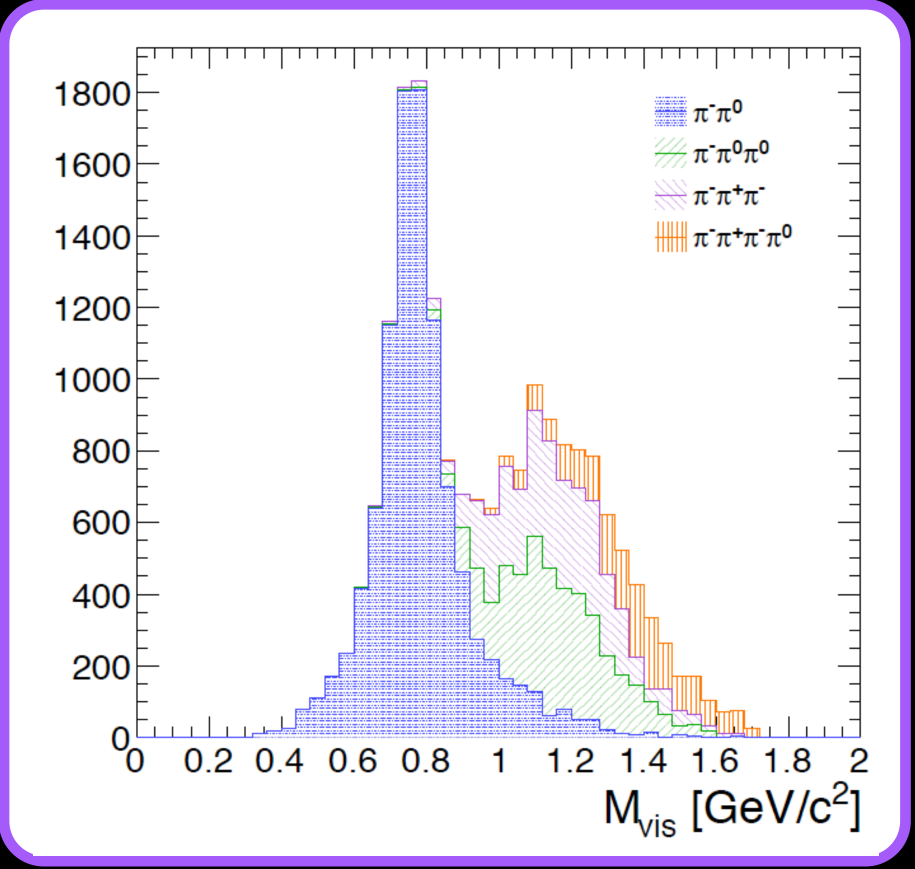
Decay Mode CMS Tau ID

Particle Flow algorithm allows examination of meson content

two new algorithms:

Hadrons Plus Strips (HPS) algorithm

Tau Neural Classifier (TaNC) algorithm

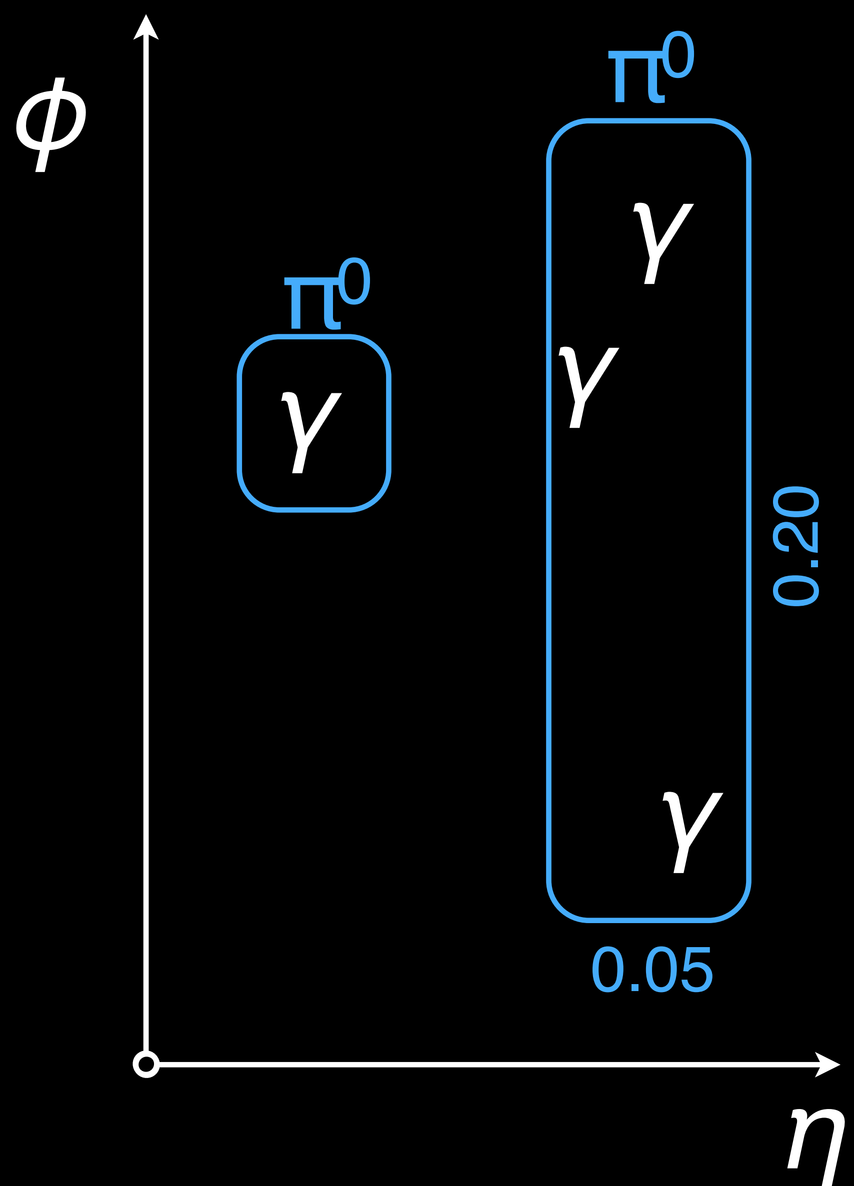


GOAL:
optimize tau identification for individual tau decay modes

Hadrons Plus Strips Algorithm

build signal components combinatorially

cluster gammas into π^0 candidates using η-φ strips



build all possible taus that have a 'tau-like' multiplicity from the seed jet

π^+

$\pi^+ \pi^0$

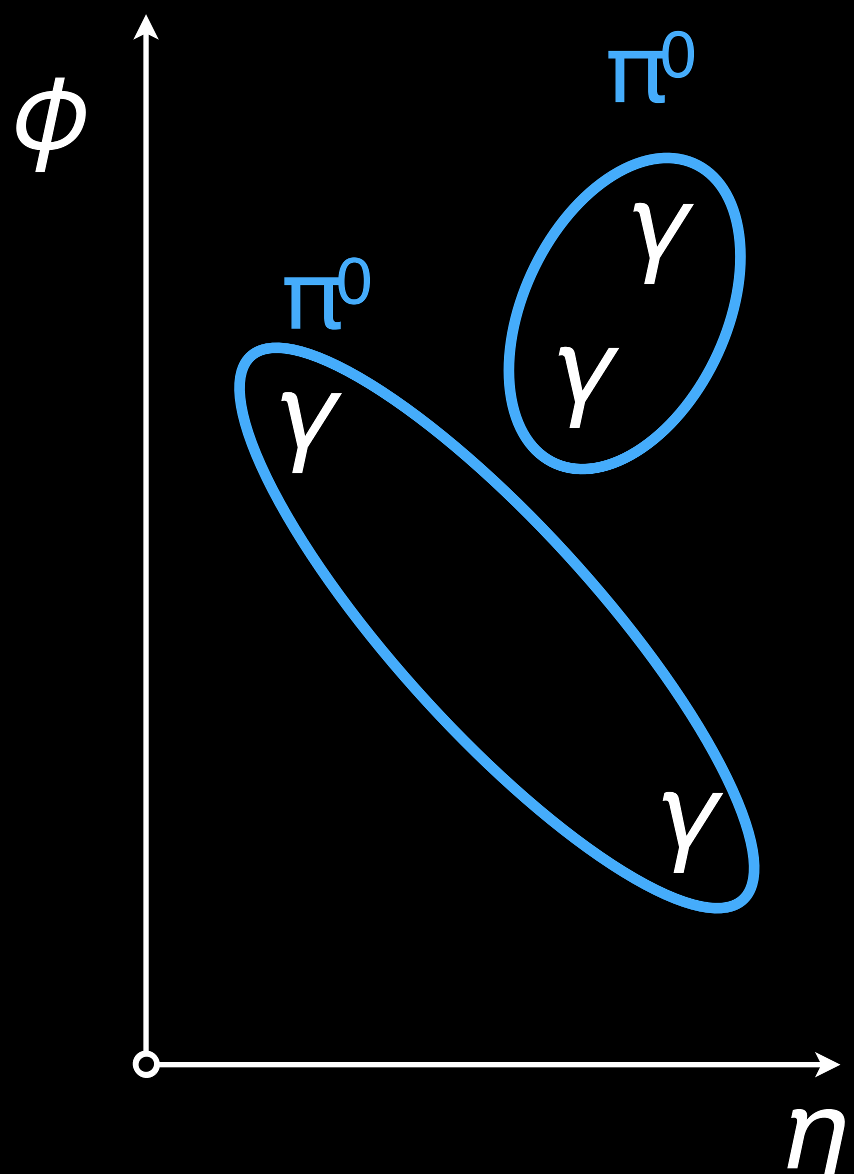
$\pi^+ \pi^+ \pi^-$

tau that is 'most isolated' with compatible m_{vis} is the final tau candidate associated to the seed jet

Tau Neural Classifier

a neural network for each decay mode

cluster gammas into π^0 candidates by combinatoric pairs compatible with $m_{\pi 0}$



signal objects are defined using shrinking cone

depending on decay mode

π^+

$\pi^+ \pi^0$

$\pi^+ \pi^0 \pi^0$

$\pi^+ \pi^+ \pi^-$

$\pi^+ \pi^+ \pi^- \pi^0$

a different neural network is applied!

Tau Performance

measuring the fake rate is easy
(no lack of background at the LHC)

fake rate is **different**
for W+jets, multi-jet, heavy flavor

independently measuring signal efficiency is *hard*

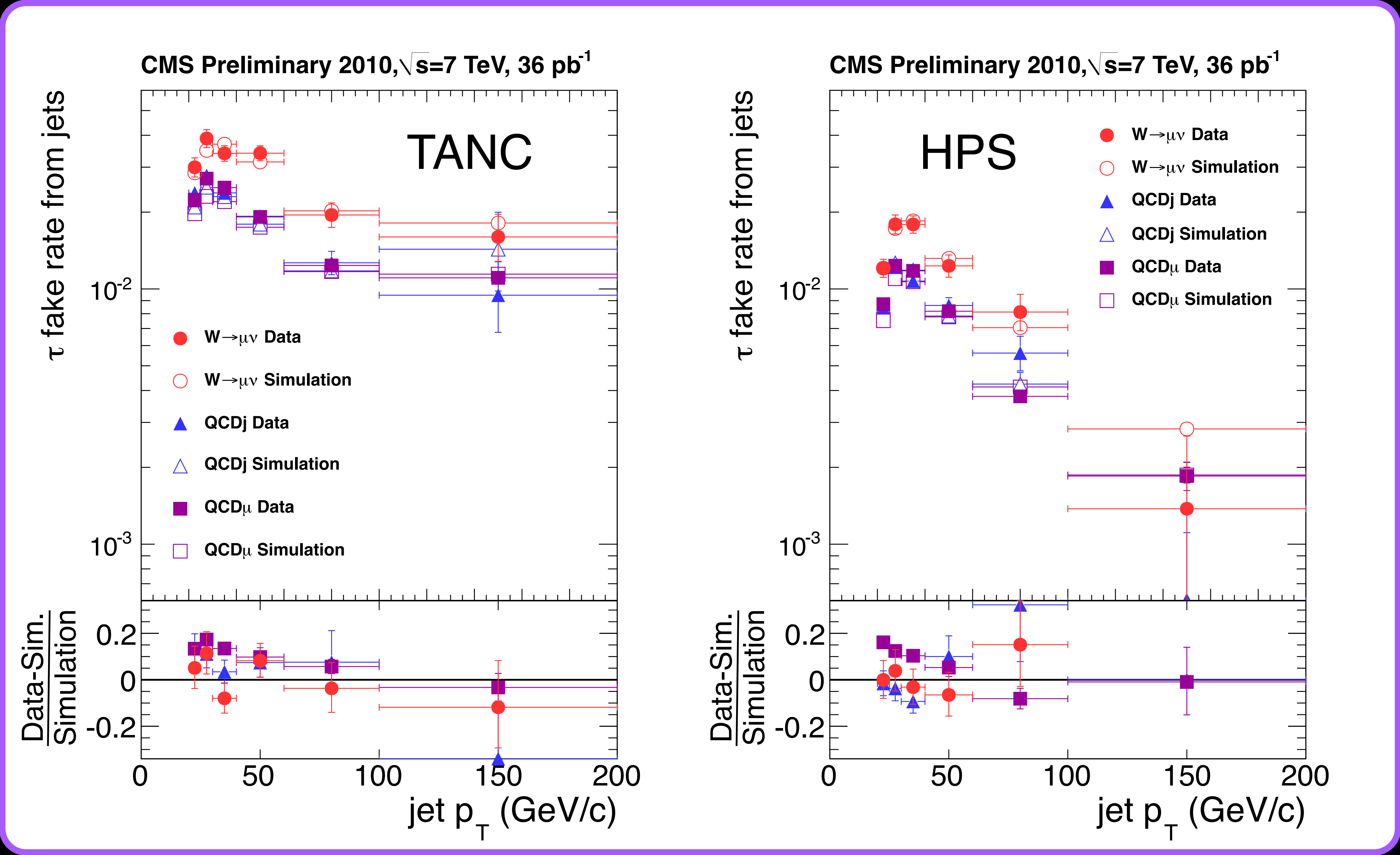
in practice efficiency extracted *in situ*
in Z and Higgs analyses

Fake Rate measurement

TAU-11-001

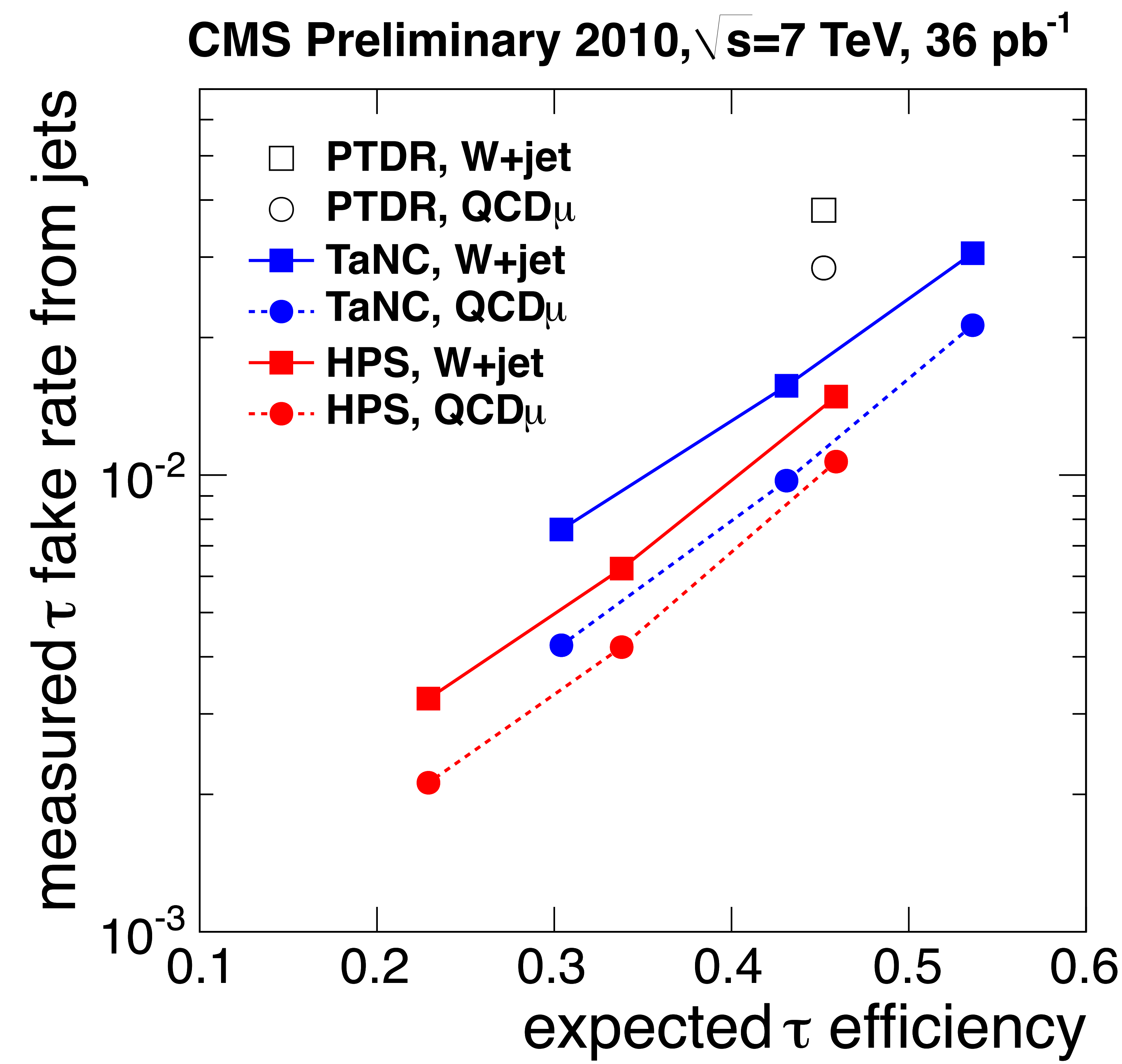
open points = simulation closed points = data

red is W+jets purple is heavy flavor



Tau ID Performance

TAU-11-001

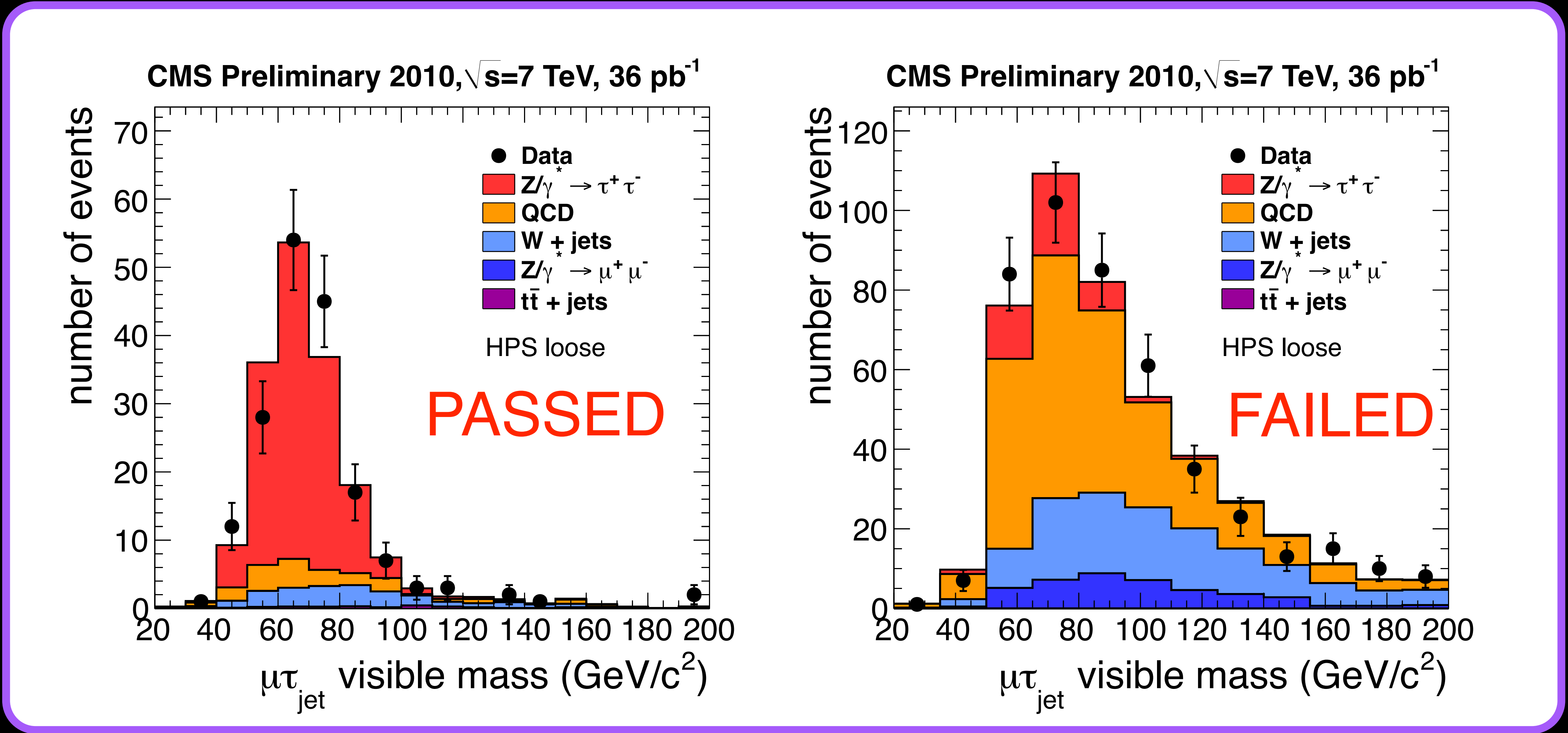


3X improvement since PTDR

TAU-11-001

Tag & Probe efficiency

- Preselect $Z \rightarrow \tau_\mu \tau_{had}$ using loose isolation + Z analysis cuts
- Fit ττ M_{vis} passing/failing spectra composition
- Method robust against Higgs contamination in Z peak



result: DATA/MC ≅ 1.0 ± 25%

Zττ cross section

$$\sigma(pp \to ZX) \times \mathcal{B}(Z \to \tau^+\tau^-) = \frac{N}{\mathcal{A} \cdot \epsilon \cdot \mathcal{B}' \cdot \mathcal{L}}$$

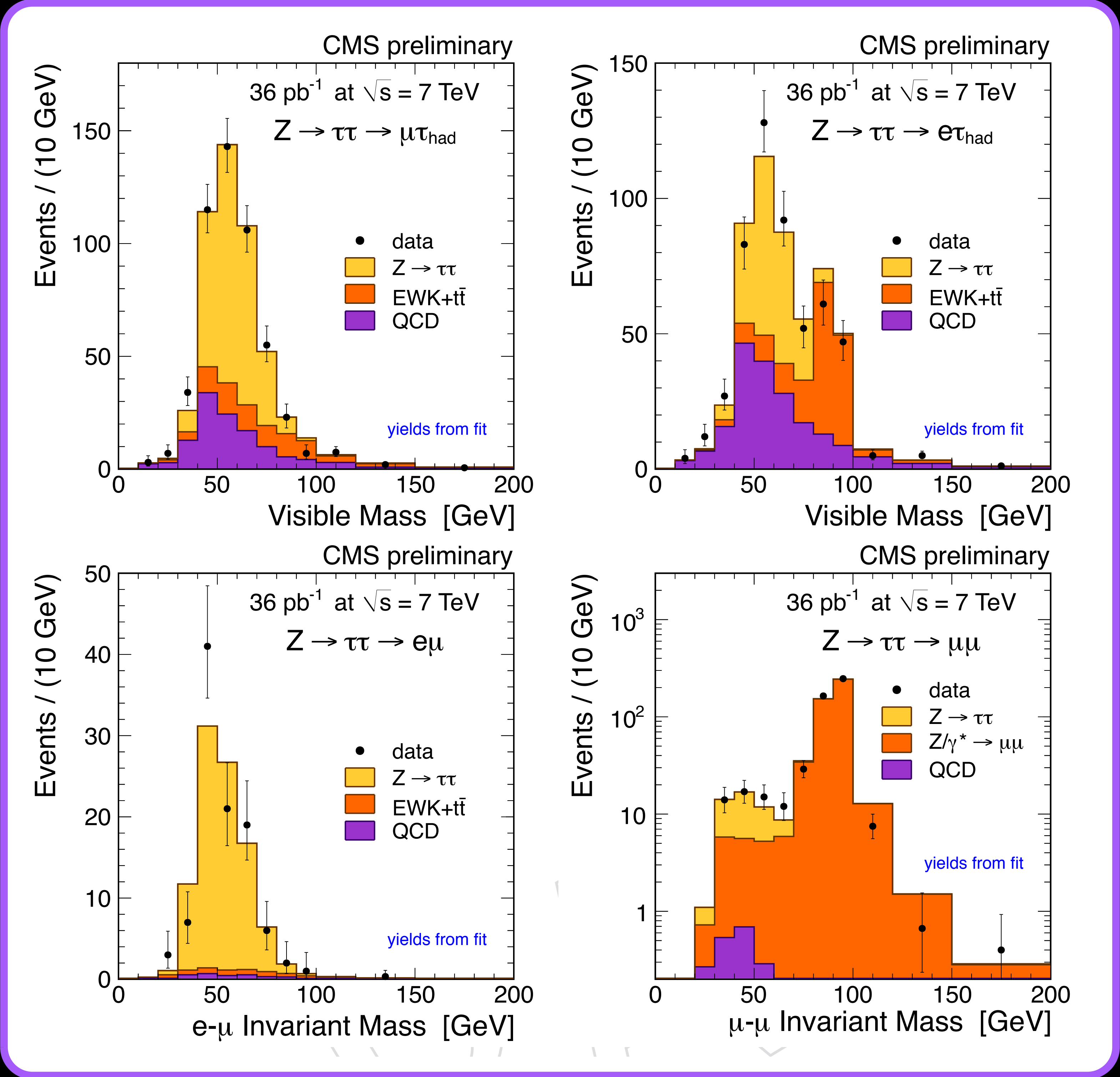
!

- Analysis goal: understand taus at CMS
- Four channels: μ-had, e-had, e-μ, μ-μ
- ~25% τ-ID uncertainty limits utility as test of Standard model
- Instead: believe in lepton universality and use Z peak to measure τ-ID
    - Combined fit of all channels

EWK-10-013

Z→ττ cross section

final Zττ events selected

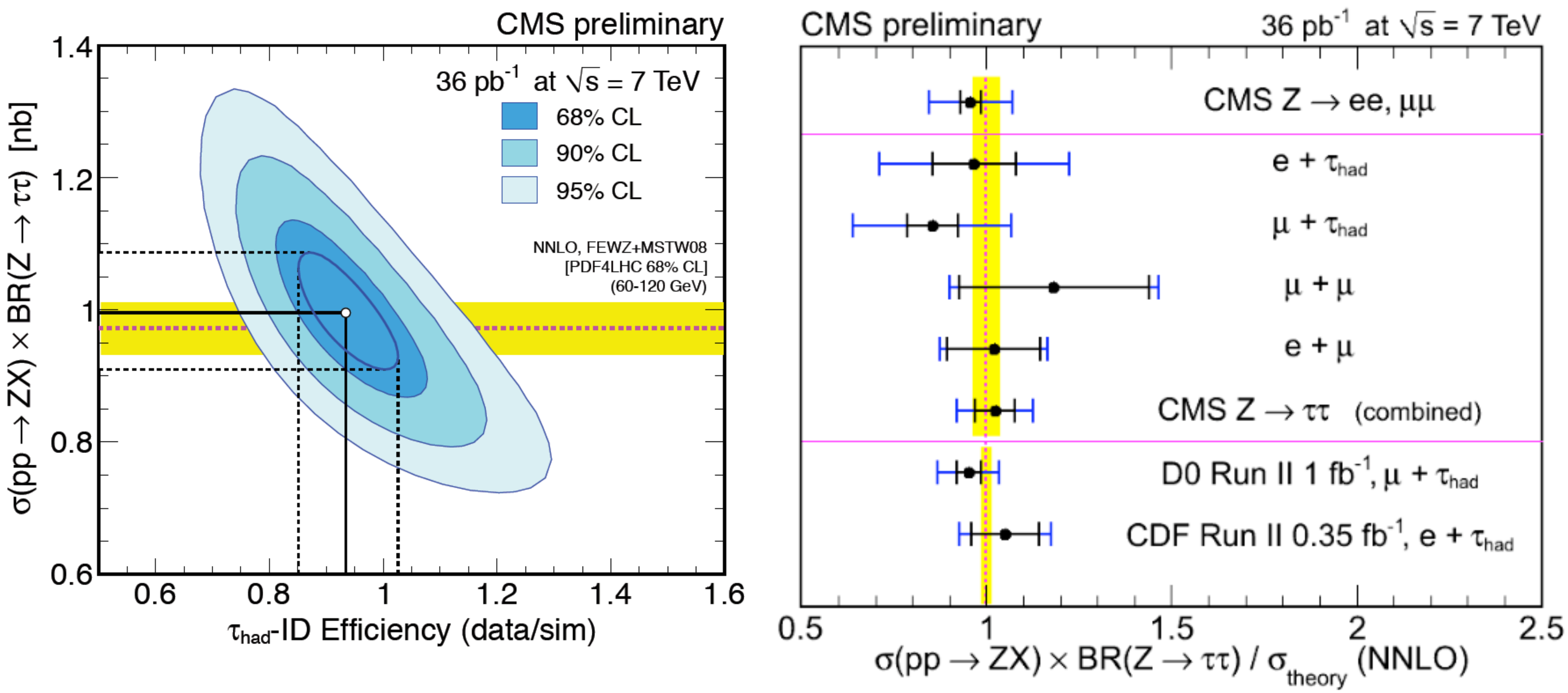


non-tau backgrounds measured from data

Combined Zττ Fit

simultaneously fit all channels for σ_Z AND $\varepsilon_{\tau ID}$

e-μ and μ-μ channels drive $\sigma_{Z\tau\tau}$ central value



$$\sigma\,(\mathrm{pp} \to \mathrm{ZX}) \times \mathcal{B}\,(\mathrm{Z} \to \tau^+\tau^-) = 1.00 \pm 0.05\,(\mathrm{stat.}) \pm 0.08\,(\mathrm{syst.}) \pm 0.04\,(\mathrm{lumi.})\,\mathrm{nb},$$

fitted $\varepsilon_{\tau ID}$ MC/DATA = 0.93±0.09

assume $\sigma_{Z\tau\tau} = \sigma_{Zee/\mu\mu} \rightarrow \varepsilon_{\tau ID}$ MC/DATA = 0.96±0.07

Taus are well understood objects at CMS!

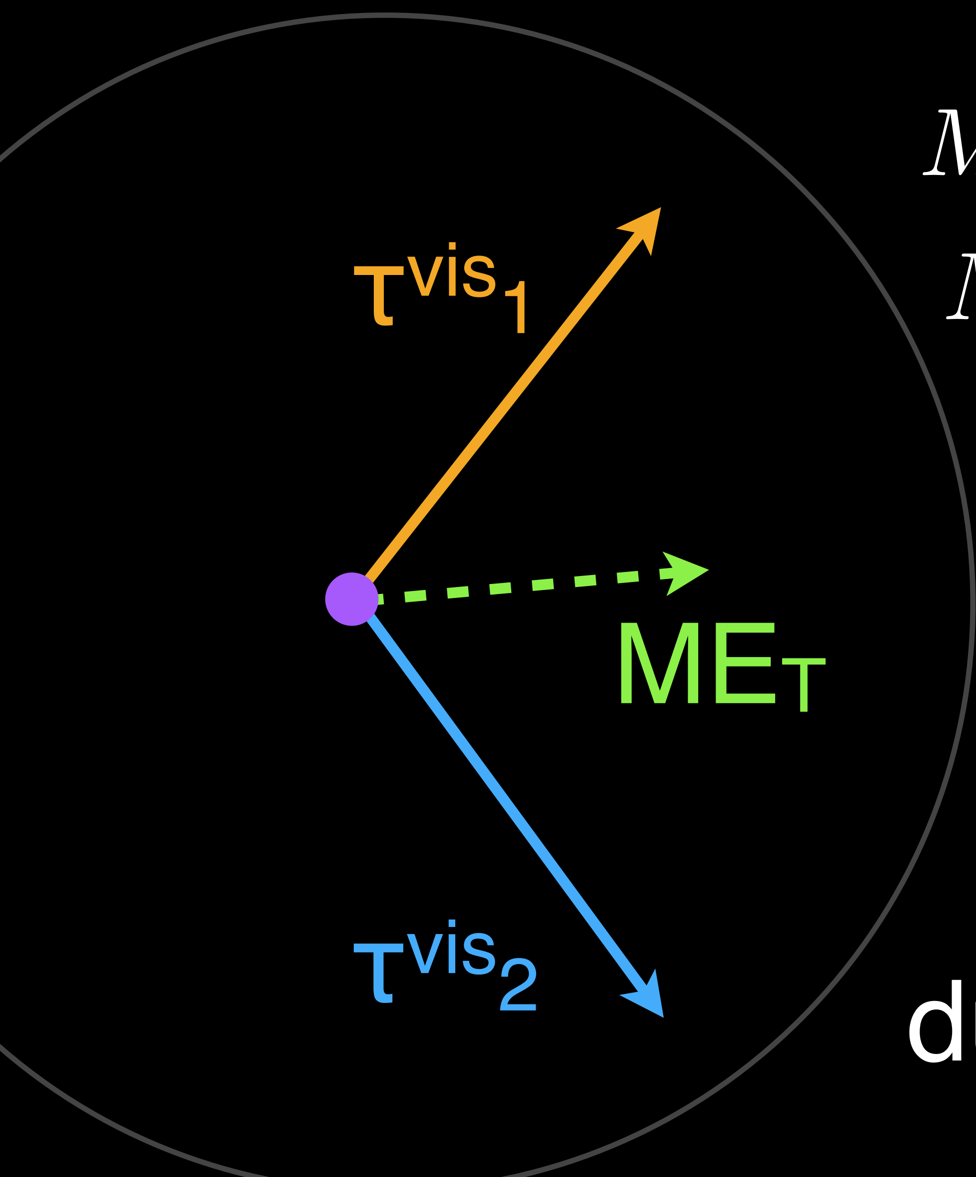
searches with taus

Tau pair mass reconstruction

- Searches for new resonances to taus
  - Tau pair invariant mass is the most natural observable
- Tau decays have invisible component
- Recovery methods:
  - Collinear approximation
  - SVfit algorithm
  - Missing Mass algorithm (see Alexei's talk)

The Collinear Approximation

- Assume neutrinos collinear with visible products
- Approximates true ττ mass



$$MET_x = P_1^{\nu} \sin\theta_1 \cos\phi_1 + P_2^{\nu} \sin\theta_2 \cos\phi_2$$
$$MET_y = P_1^{\nu} \sin\theta_1 \sin\phi_1 + P_2^{\nu} \sin\theta_2 \sin\phi_2$$

Very sensitive to MET angular resolution!

~45% of events thrown away due to unphysical negative energies

SVfit mass reconstruction

parameterize the physics of tau decays

maximize a likelihood function
w.r.t. the tau decay parameters

likelihood terms:

- decay phase-space
- reconstructed neutrinos and measured ME_T
- “p_T-balance” regularization
- compatibility with tracker information (optional)

Decay Parameterization

a τ decay can be completely described by:

θ	Rest frame angle between boost and visible decay productions
ϕ	Azimuthal angle of visible decay products *about* boost direction
m_{invis}	Mass of invisible decay product system ($m_{invis} = 0$ for τ_{had})
R	Flight distance in lab frame

Decay Parameterization

reconstructing the tau energy

rest frame visible energy:

$$E_{vis} = \frac{m_\tau^2 + m_{vis}^2 - m_{\nu\nu}^2}{2m_\tau}$$

Lorentz invariant component of visible momentum

$$\sin\theta^{\mathrm{LAB}} = \frac{p_{vis}\sin\theta}{p_{vis}^{\mathrm{LAB}}}$$

Solve Lorentz transformation along boost for γ

$$p_{vis}^{\mathrm{LAB}}\cos\theta^{\mathrm{LAB}} = \gamma\beta E_{vis} + \gamma p_{vis}\cos\theta$$

$$E_\tau^{\mathrm{LAB}} = \gamma m_\tau$$

Decay Parameterization

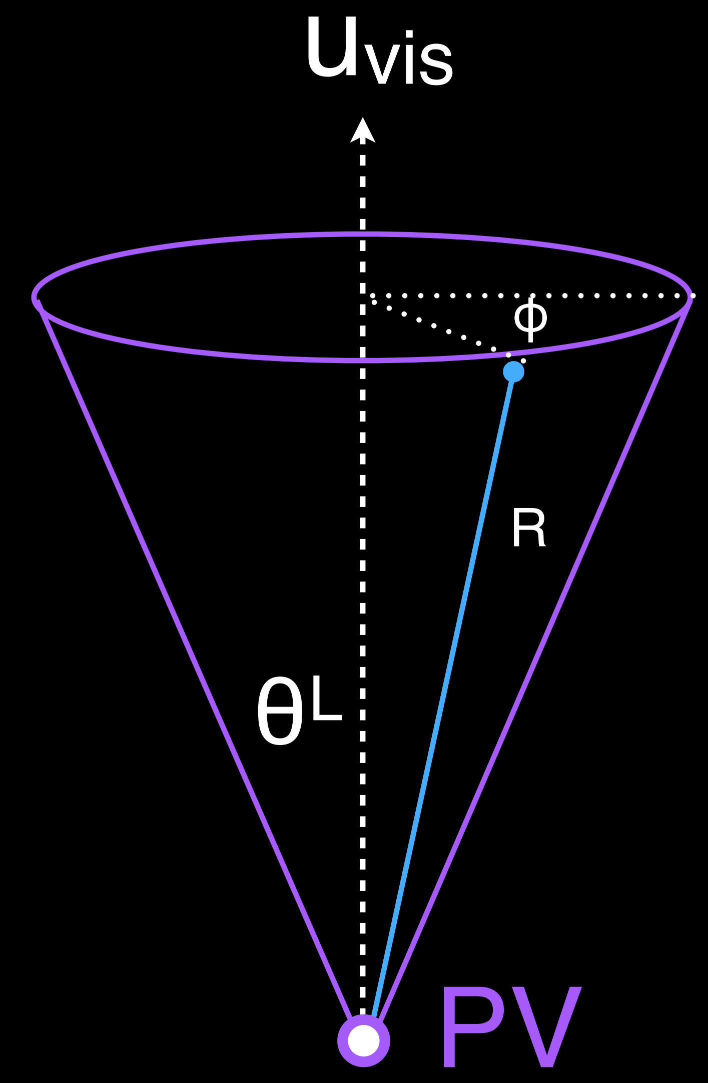
reconstructing the tau direction

$$\sin\theta^{\mathrm{LAB}} = \frac{p_{vis}\sin\theta}{p_{vis}^{\mathrm{LAB}}}$$

constrains τ direction to lie on cone
with angle θ^{LAB} about visible momentum

azimuthal fit parameter ϕ
determines direction

flight path R determines
location of decay vertex
(optional)



Phase Space Likelihood

- Two or three-body decay
- Hadronic decays trivial
- Leptonic decays depend on $m_{\nu\nu}$

$$\mathrm{d}\Gamma \propto |\mathcal{M}|^2 \frac{((m_\tau^2 - (m_{\nu\nu} + m_{vis})^2)(m_\tau^2 - (m_{\nu\nu} - m_{vis})^2))^{1/2}}{2m_\tau} m_{\nu\nu} \mathrm{d}m_{\nu\nu} \sin\theta d\theta$$

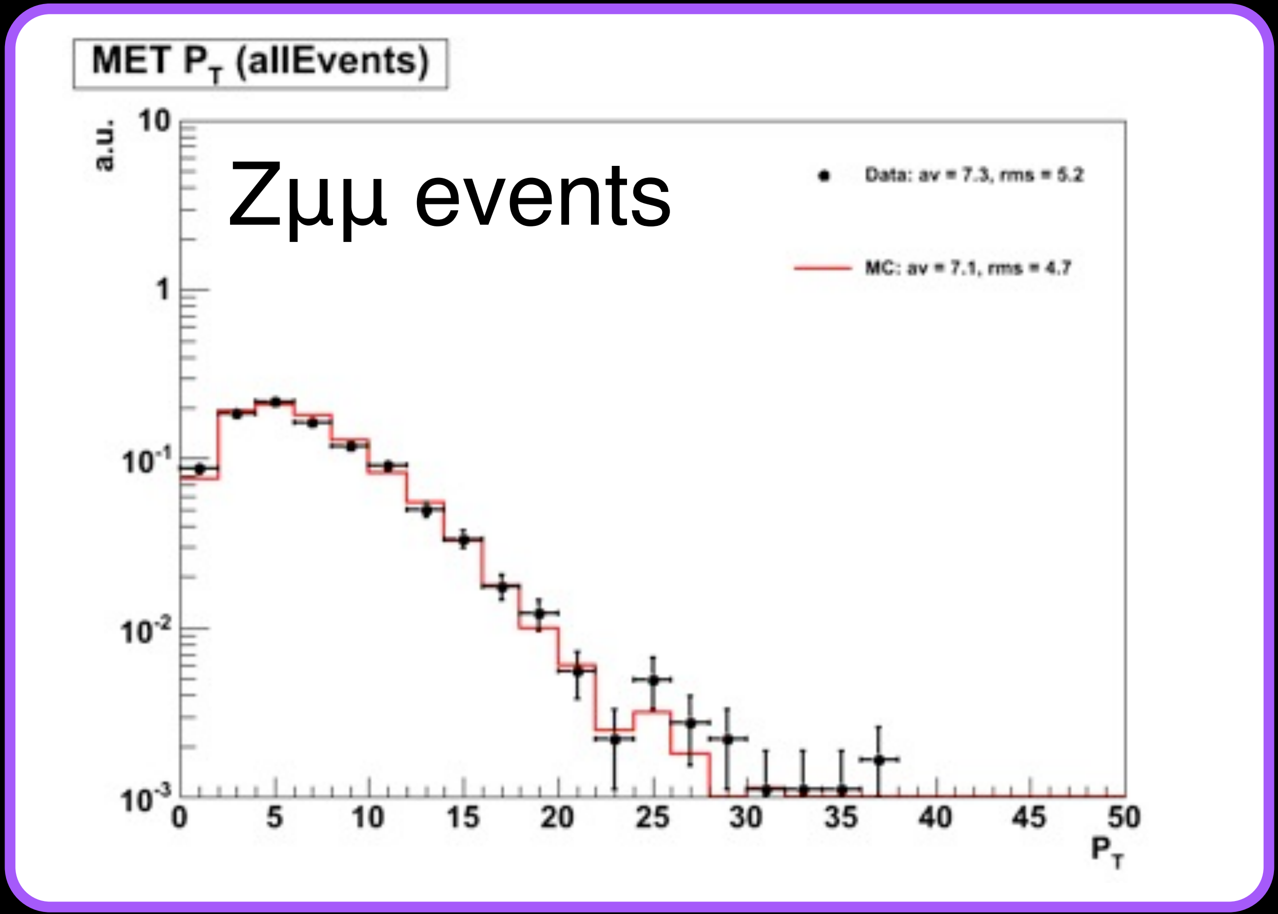
Decays are assumed to be isotropic: $|M|^2 = 1$

Investigating extending term to account for polarization correlations between taus

Possible handle to separate Higgses from Zs!

ME_T likelihood

- Compare SV-fitted neutrinos to measured missing transverse energy
- Resolution of ME_T parameterized by ΣE_T of the event
- Resolution measured in $Z \rightarrow \mu\mu$ events



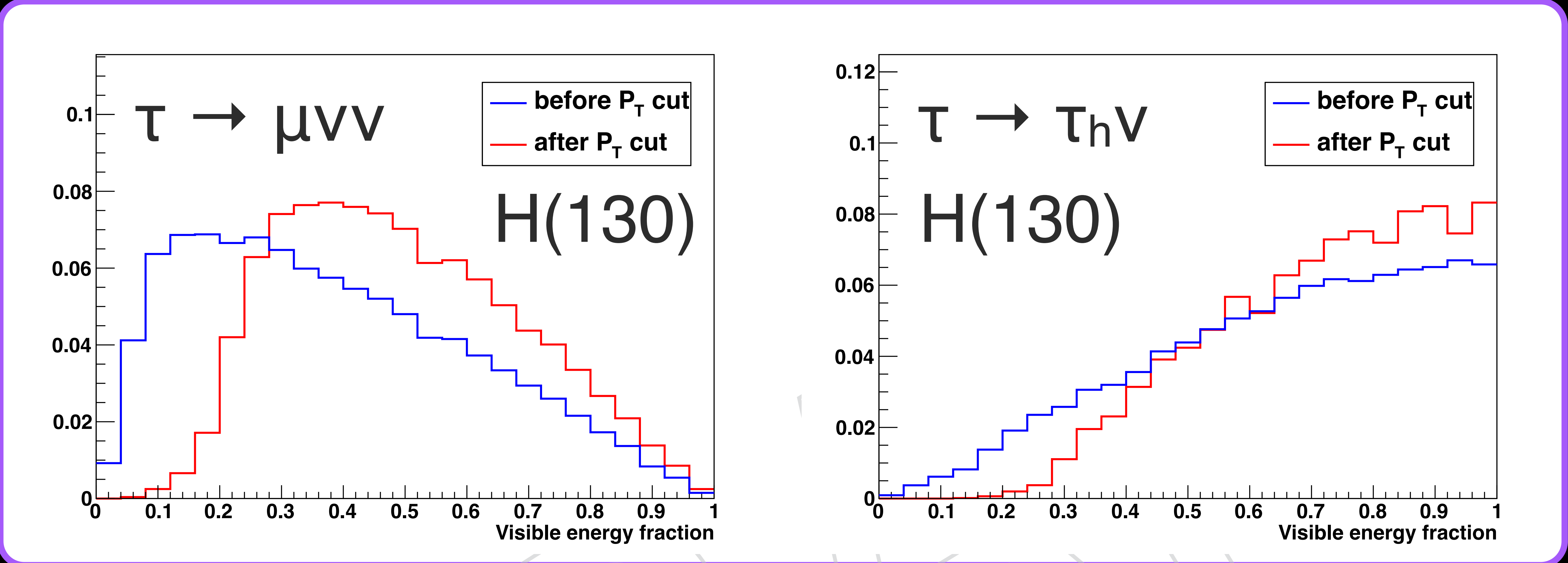
P_T - balance likelihood

to reduce background one must apply
kinematic thresholds on visible decay products

for Z/H μ-had channel: $p_T^{\mu} > 15$ GeV, $p_T^{had} > 20$ GeV

significantly biases the kinematic distributions

p_T balance term: probability for decay p_T from mass M



P_T - balance likelihood

term which represents the probability $P(p_T \mid M_{\tau\tau})$

first assume isotropic decay of resonance M at rest

each decay product has $P(p_T)$

$$P_{iso}(p_T) \propto \frac{p_T}{\sqrt{1-(2p_T/M)^2}}$$

smear with Gaussian and add Gamma
to account for resolution and non-zero boost

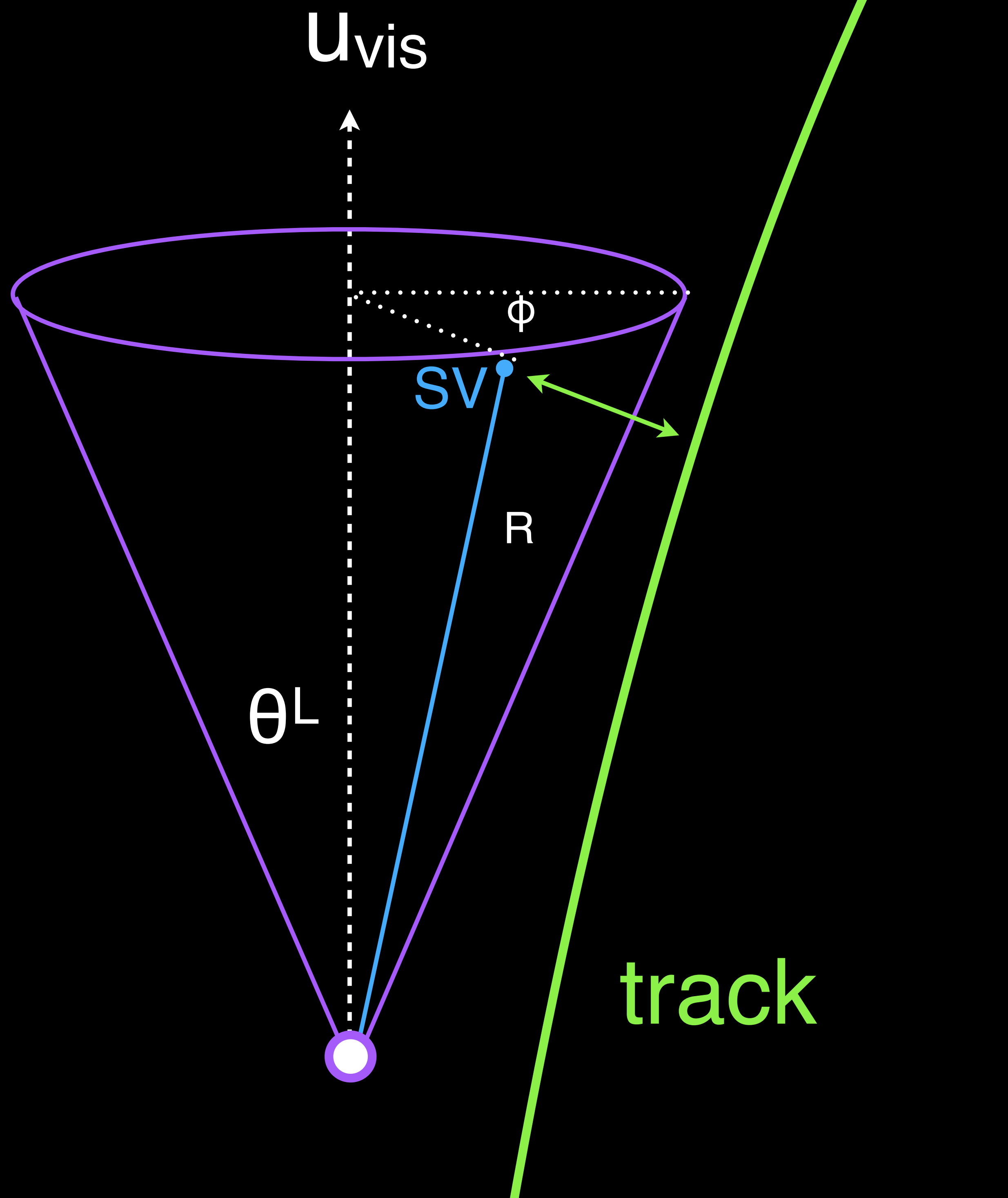
$$P_{total}(p_T) \propto P_{iso}(p_T) \otimes \exp(-\frac{p_T^2}{2s^2}) + a\Gamma(p_T, k, \theta)$$

smear, Gamma fraction and shape are
fitted as first order functions of $M_{\tau\tau}$

Tracking constraint

Taus have non-negligible lifetime: $c\tau = 87\mu\mathrm{m}$

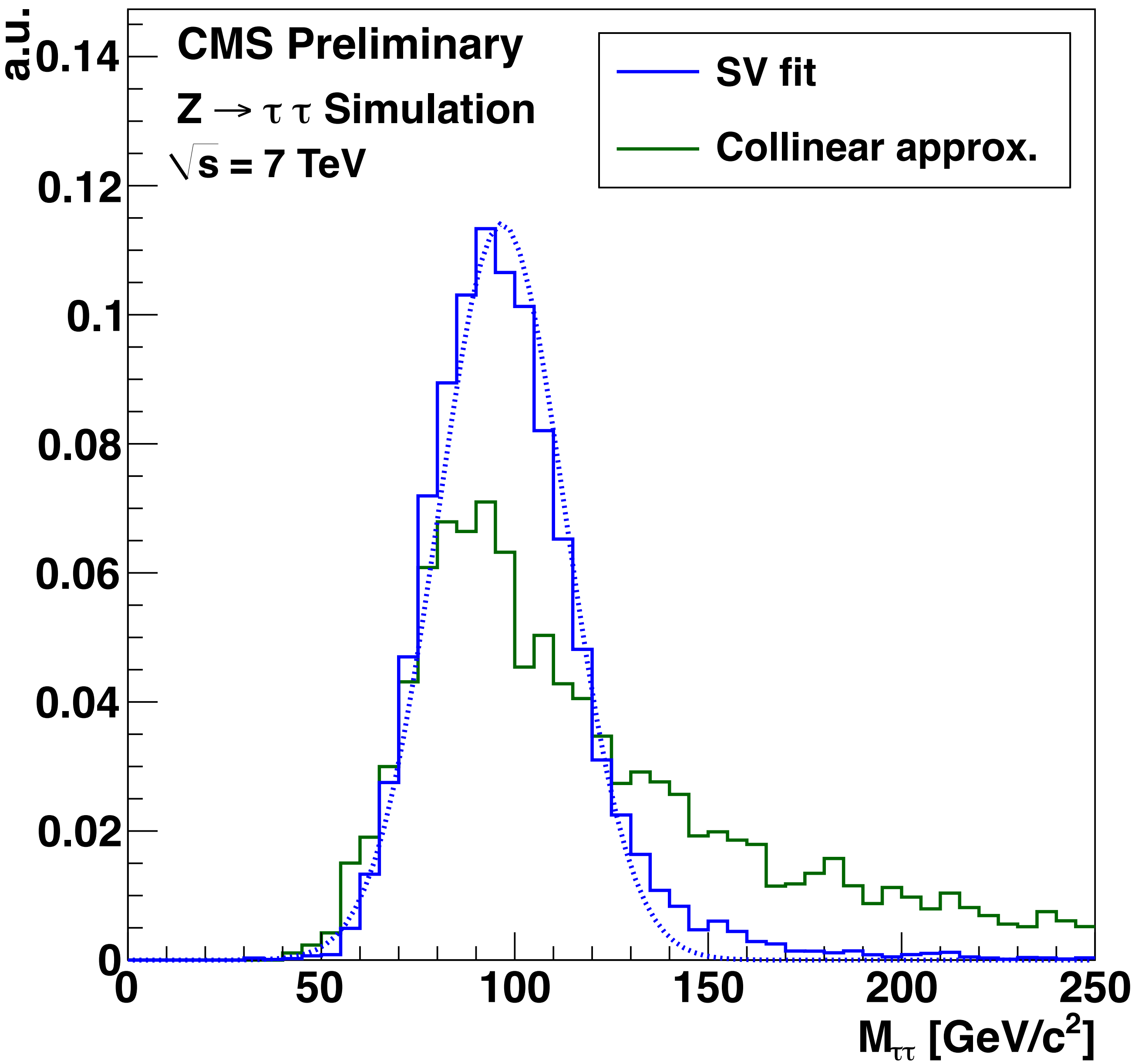
- Use tracker information
- Constrain fitted SV to lie on track with error
- Constrain flight time *R*
- Not currently used. Working to understand alignment issues.

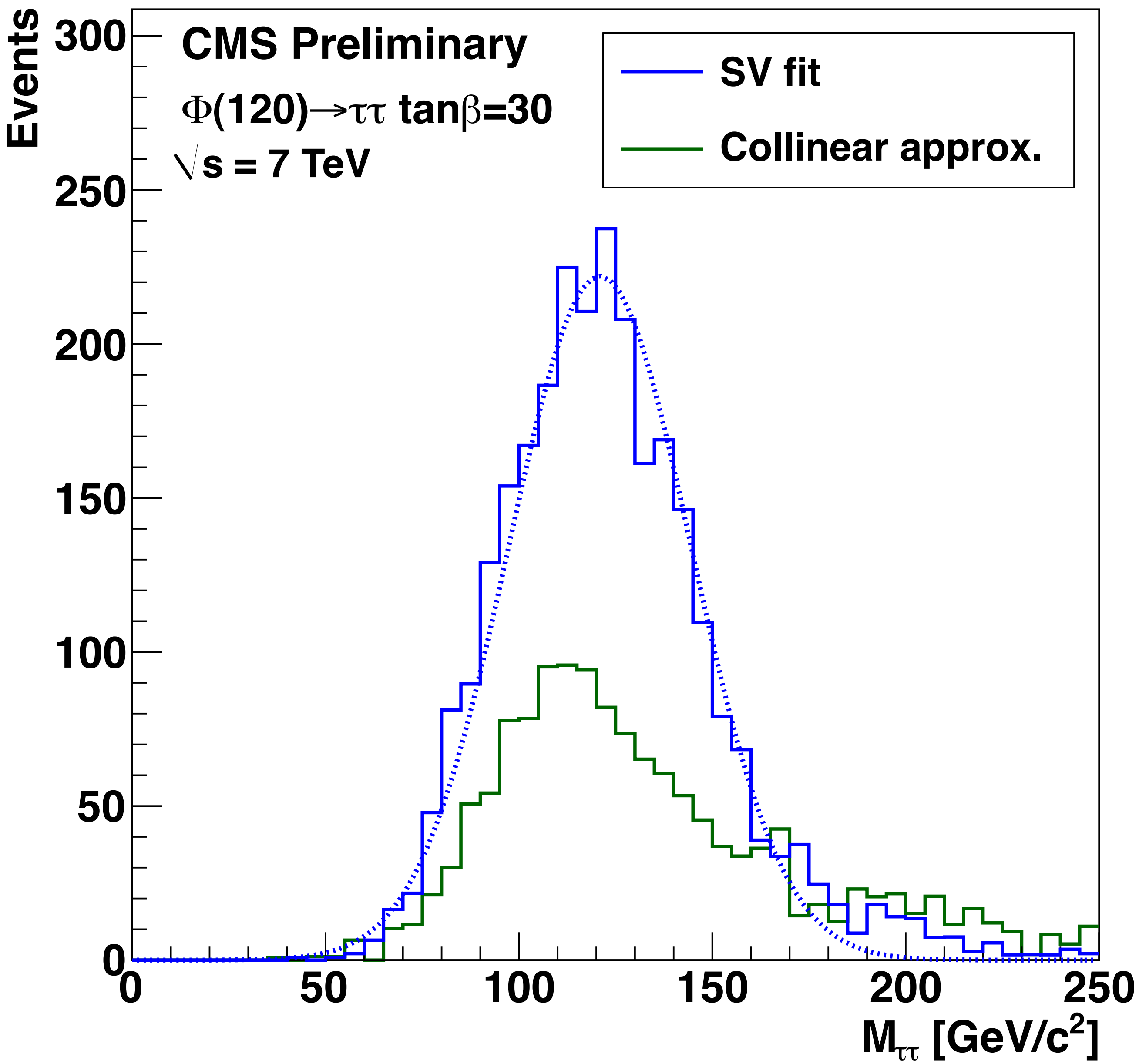


SVfit Performance

SV fit has better resolution than collinear approximation (left)

SV fit has twice the acceptance as the collinear solution (right)





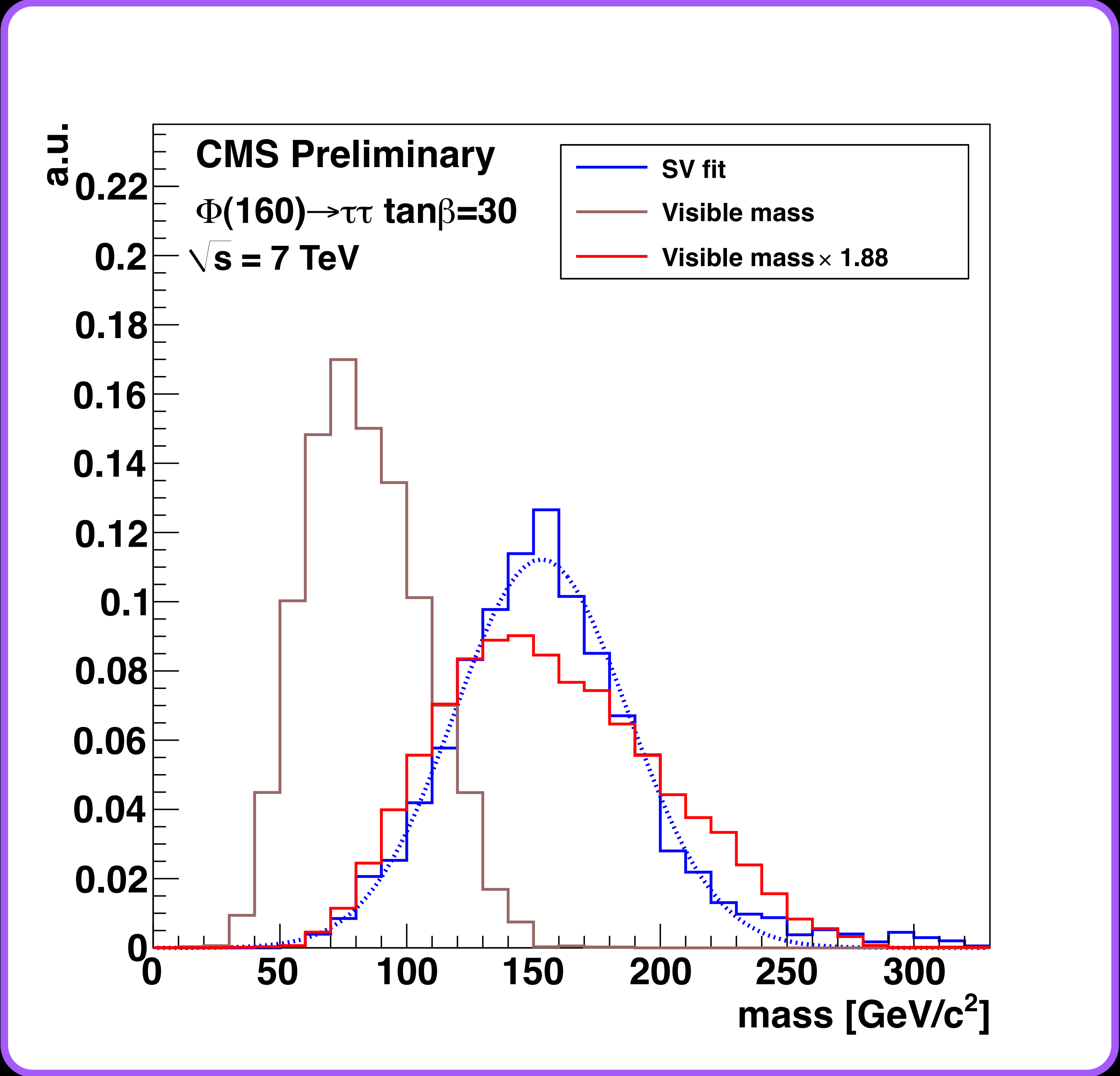
normalized to unity

normalized to luminosity

SVfit Performance

SVfit peaks at di-tau mass

better relative resolution that visible mass



Extensions to SVfit

currently under development

- Implement tracking likelihood
- Improve p_T-balance parameterization
- Use polarization information
- Integration instead of maximization (similar to DLM, MMC)
- Extend code to fit *N* leptons
  - Application to $W \rightarrow \tau\nu$, H^{++}, etc

MSSM H→ττ Search

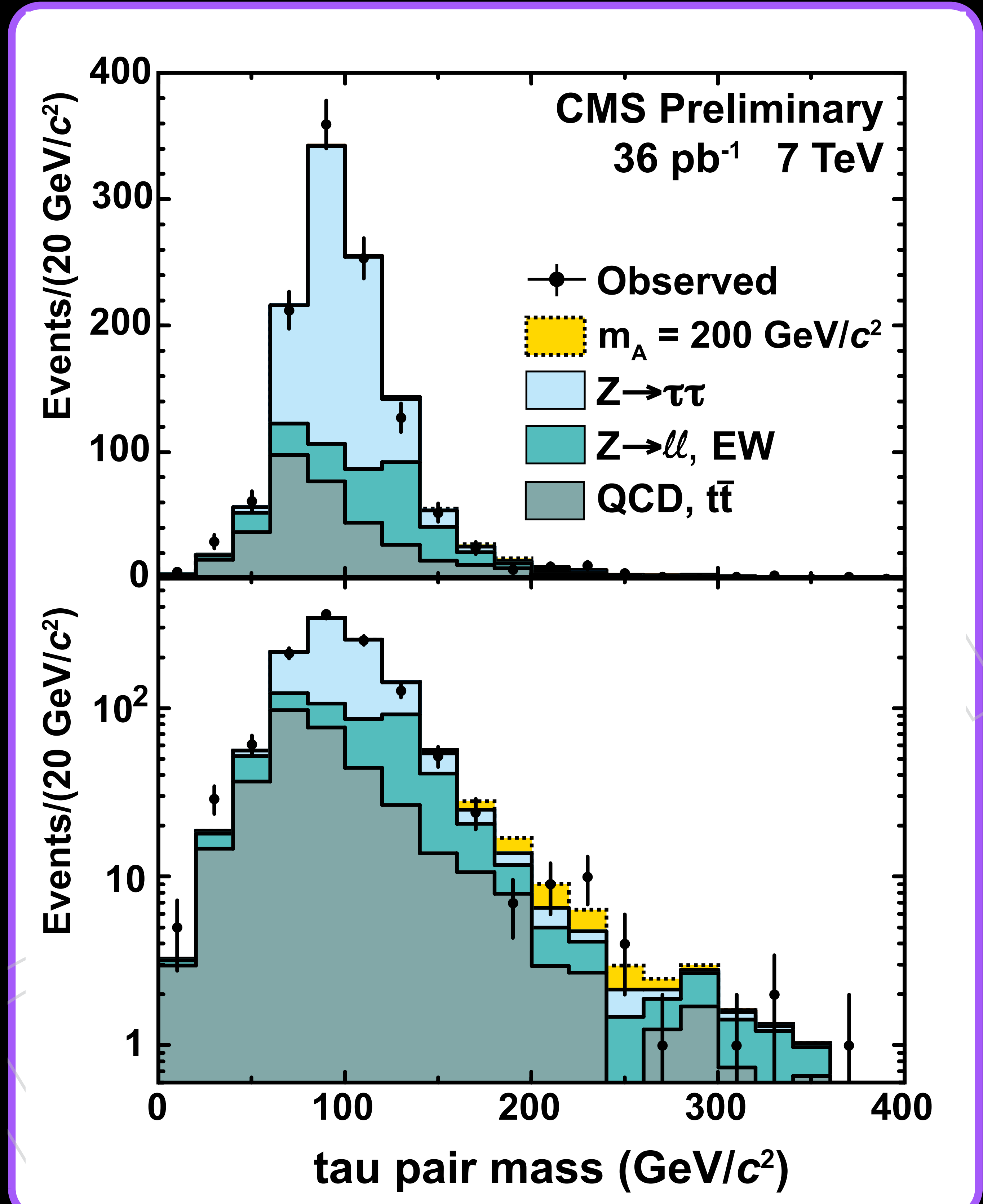
- Taus one of the best handles for MSSM
- Benefit from tanβ enhancement σ and BR
- Search performed with three channels:
  - $\mu + \tau_{had}$
  - $e + \tau_{had}$
  - $\mu + e$
- Tau ID correction factor floats in final fit, constrained by Z peak & 25% tag-probe
- All backgrounds measured from data

MSSM H→ττ Search

HIG-10-002
results

final $M_{\tau\tau}$ distribution of all channels

no bumps

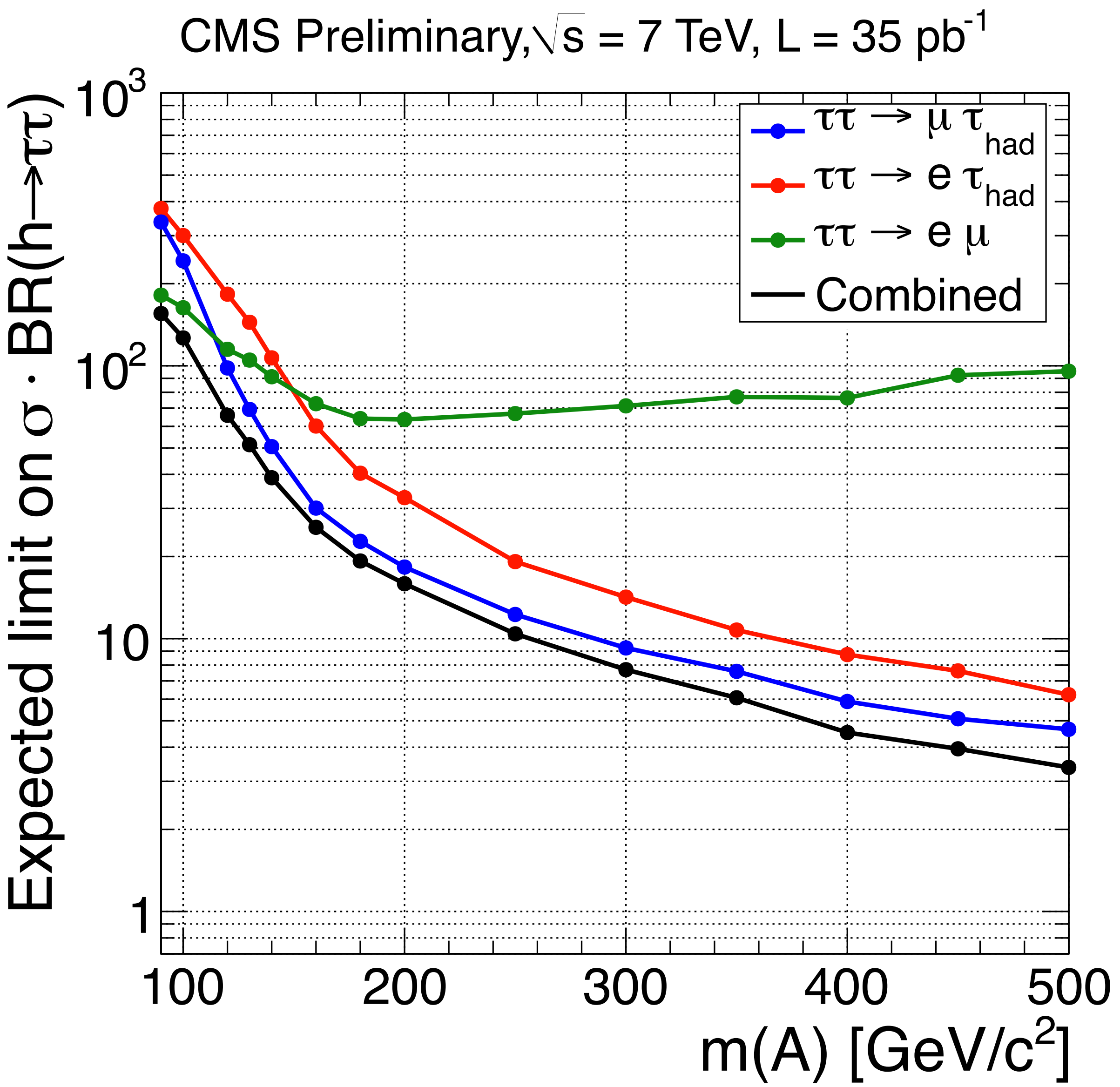


HIG-10-002

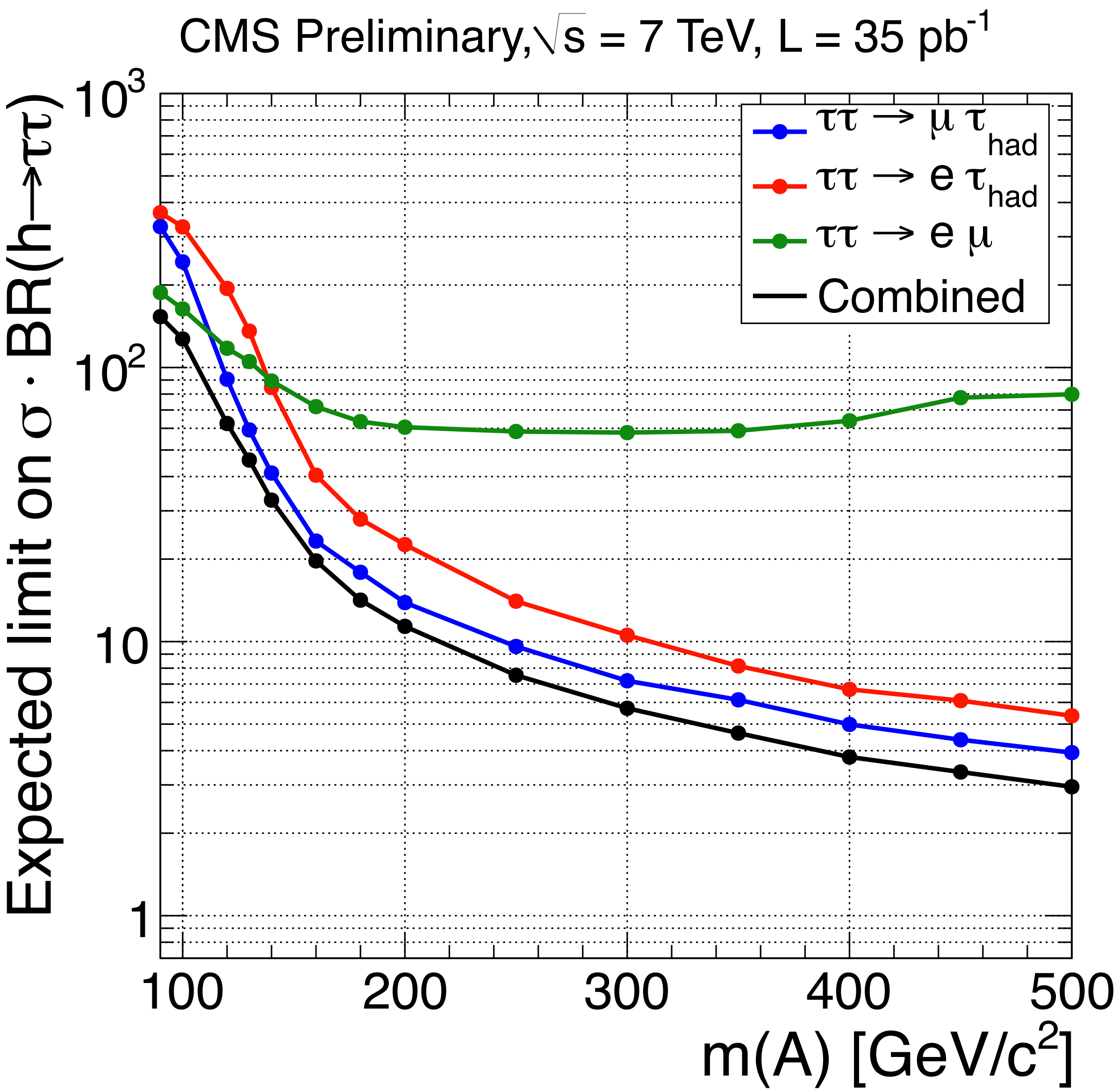
MSSM Expected Limits

Expected Bayesian 95% CL limits on $\sigma_H \times BR(H \rightarrow \tau\tau)$

Visible mass



SVfit

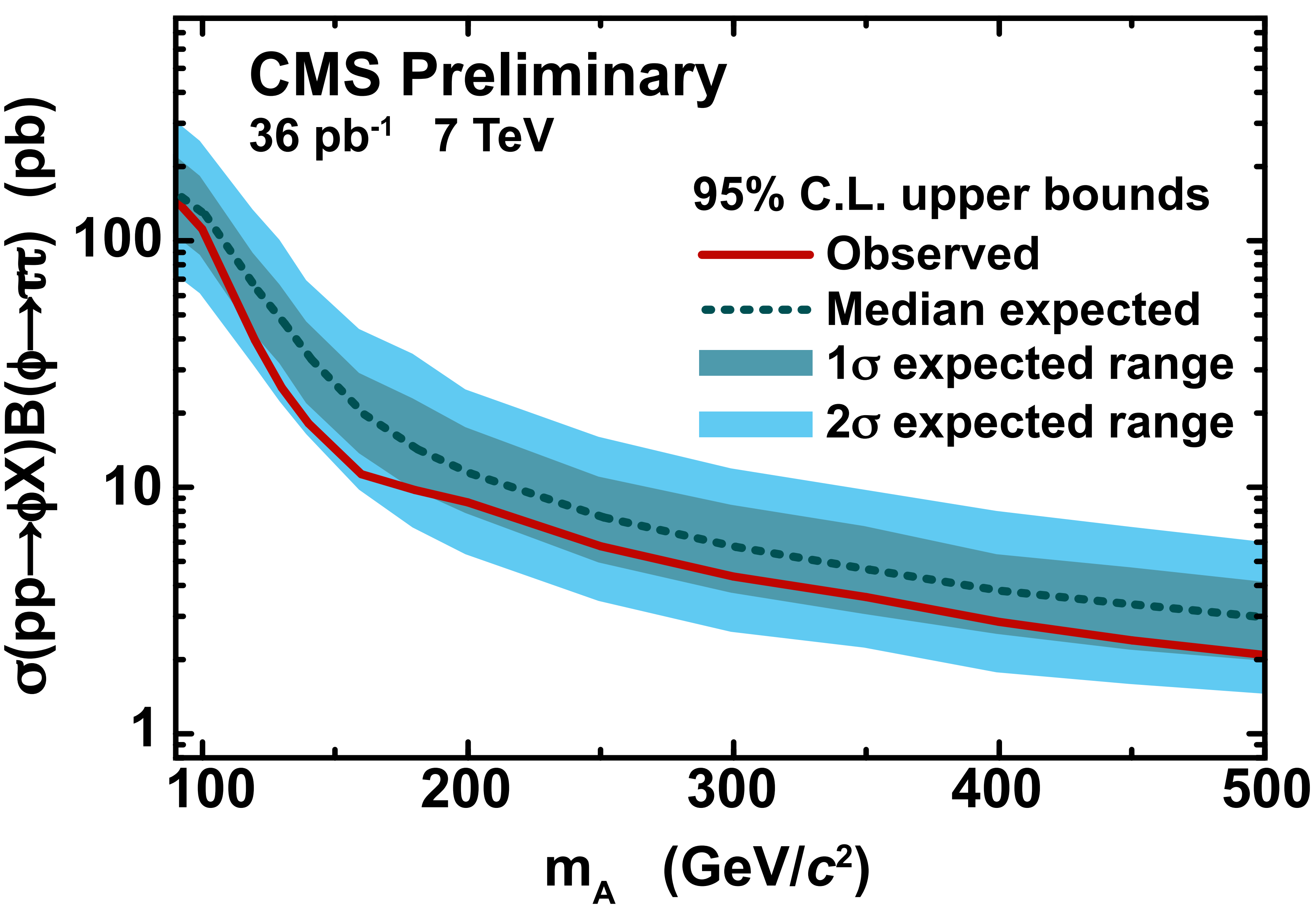


SVfit improves limit by 10%

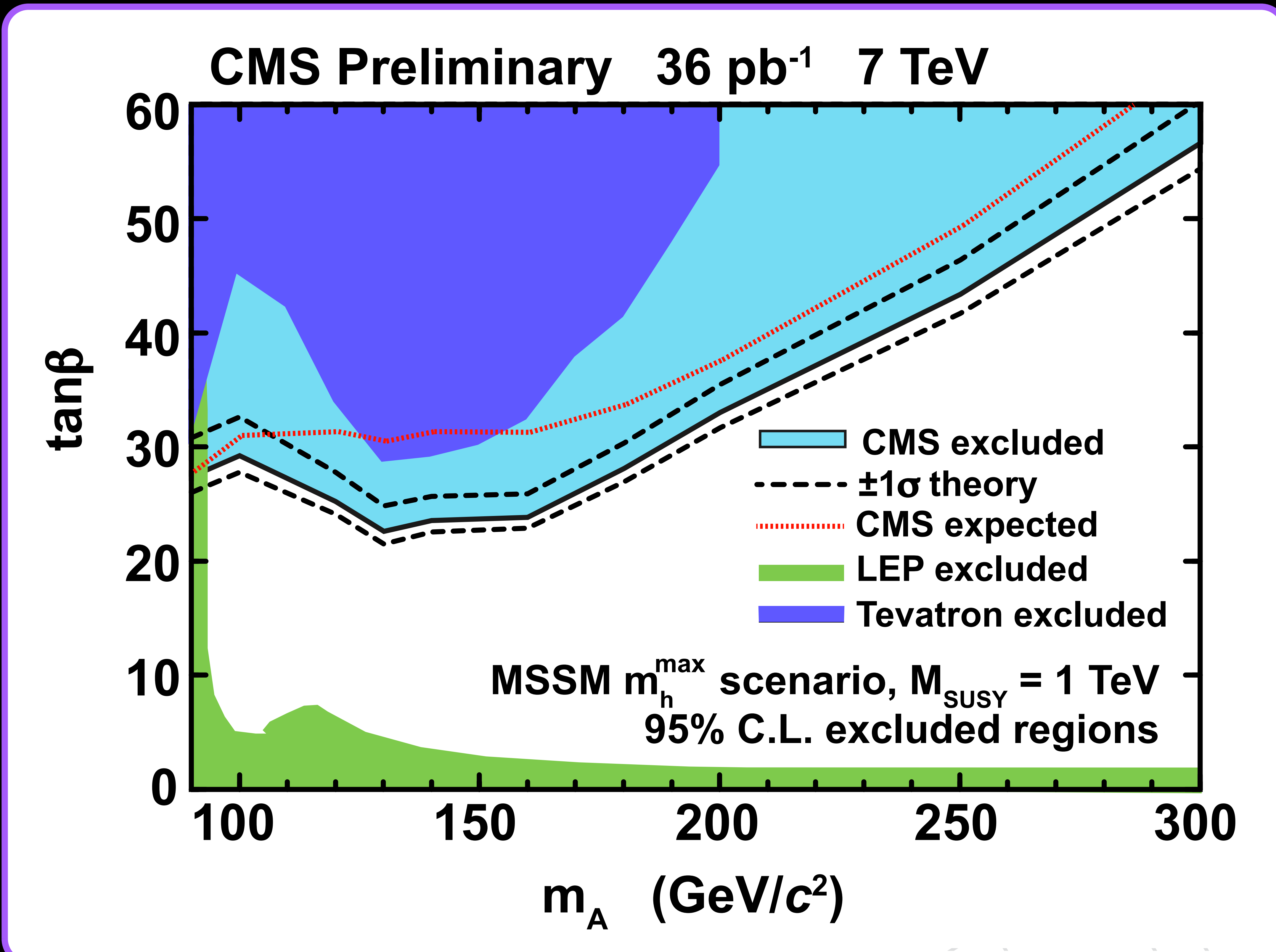
MSSM Observed Limit

HIG-10-002

observed limit on $\sigma_H \times BR(H \rightarrow \tau\tau)$



Interpretation in tanβ

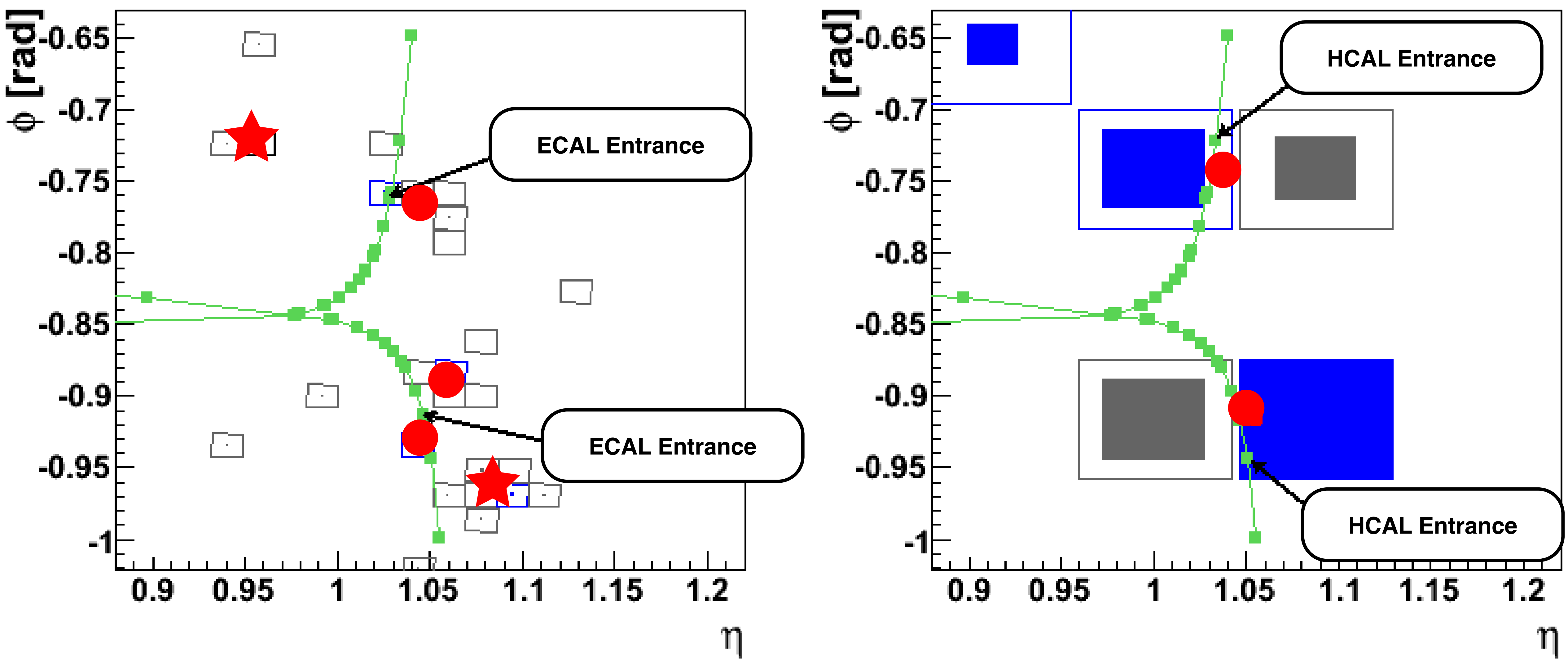


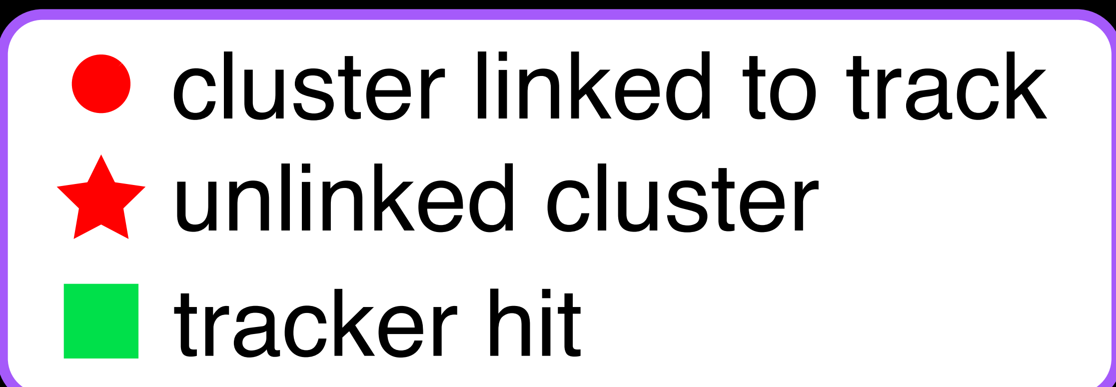
Summary

- Large improvements in τ-ID since PTDR
- Tau fake rates & **efficiency** measured in real data
- $Z \rightarrow \tau\tau$ cross section measured with high precision
- Taus well understood at CMS
- SVfit method improves $M_{\tau\tau}$ resolution
- CMS sets new limits on MSSM using taus

backup

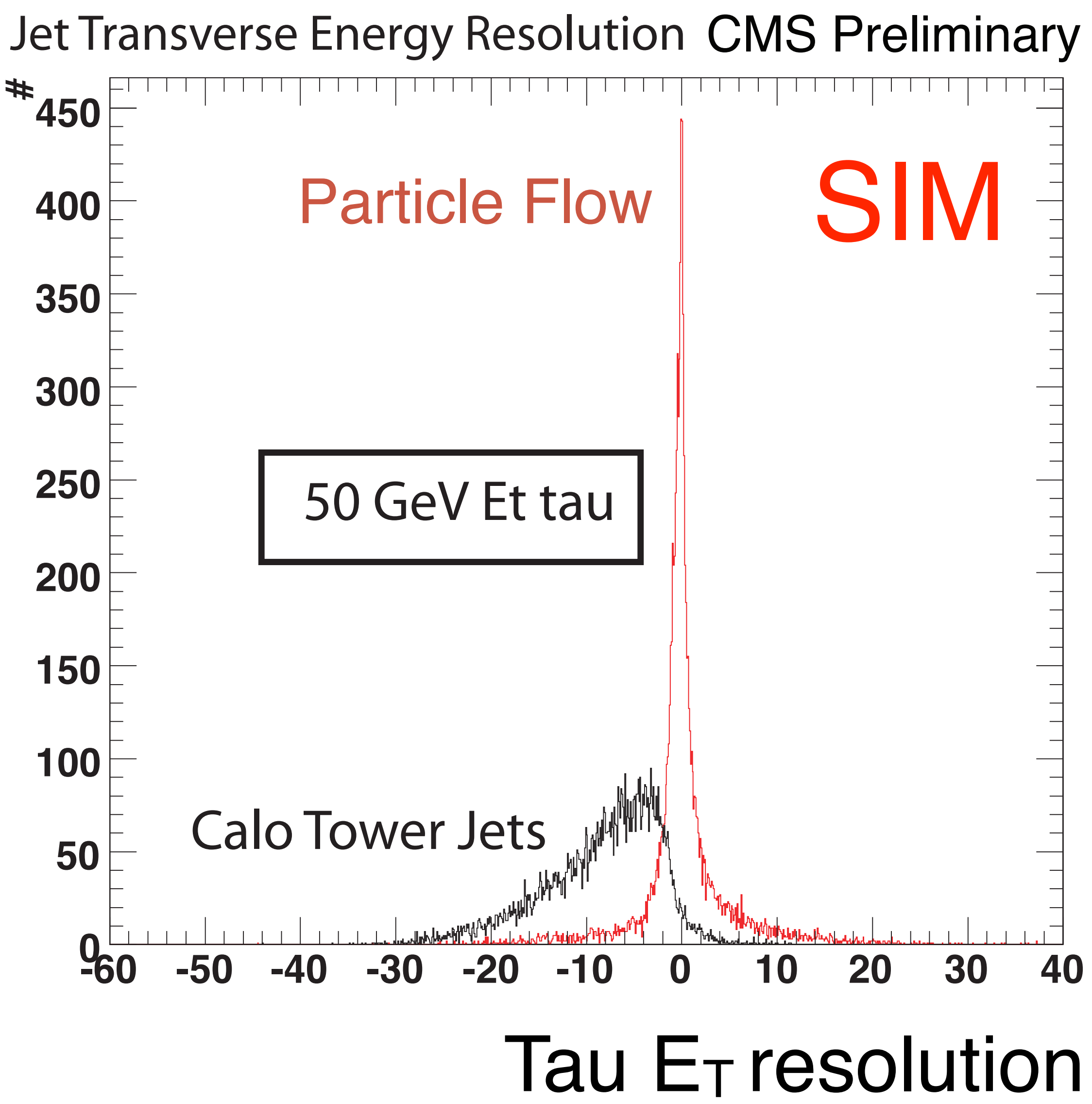
Particle Flow Algorithm

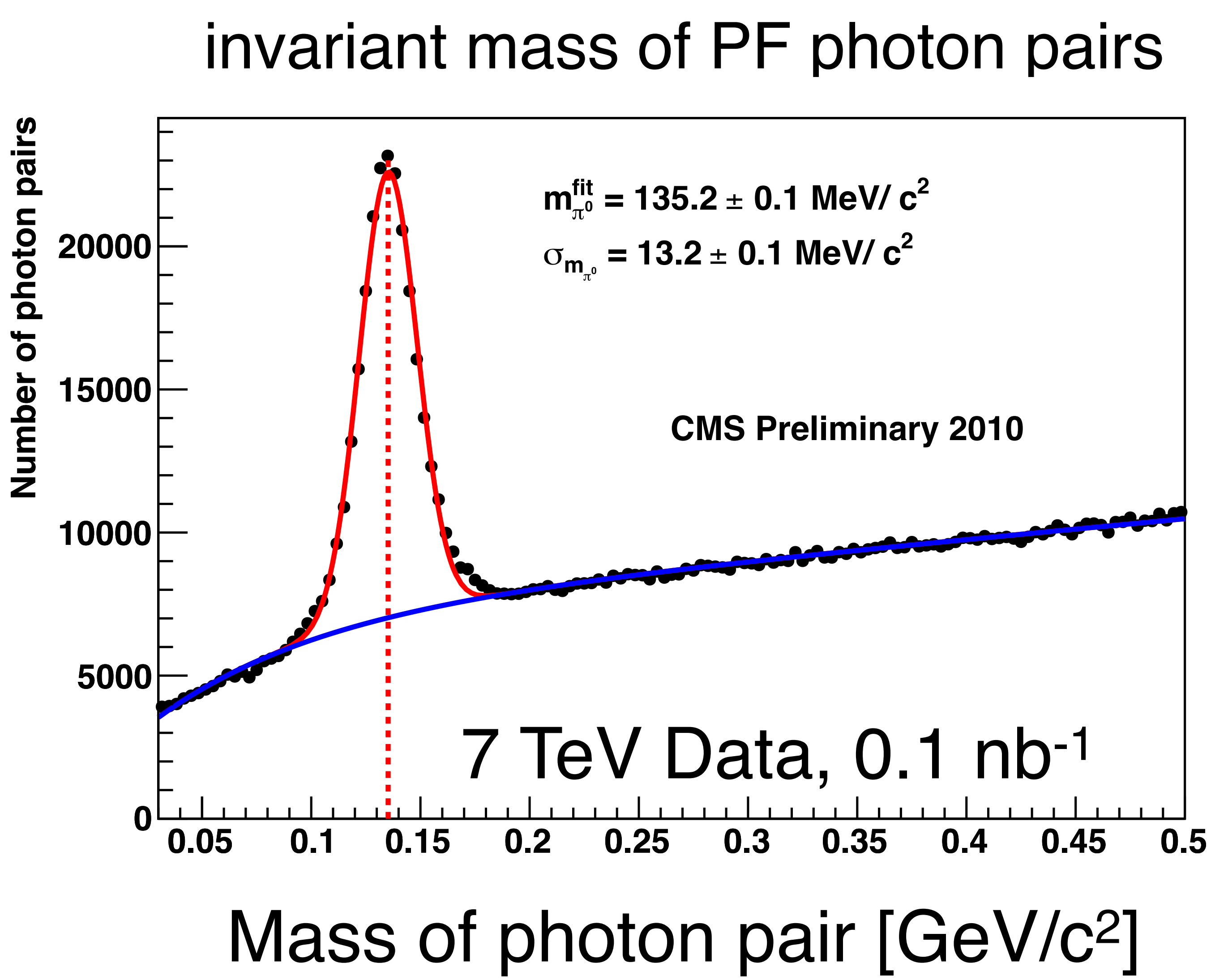




see CMS PAS PFT-10-002

Particle Flow Performance





Shrinking Cone Algorithm

reduce QCD by applying isolation requirement

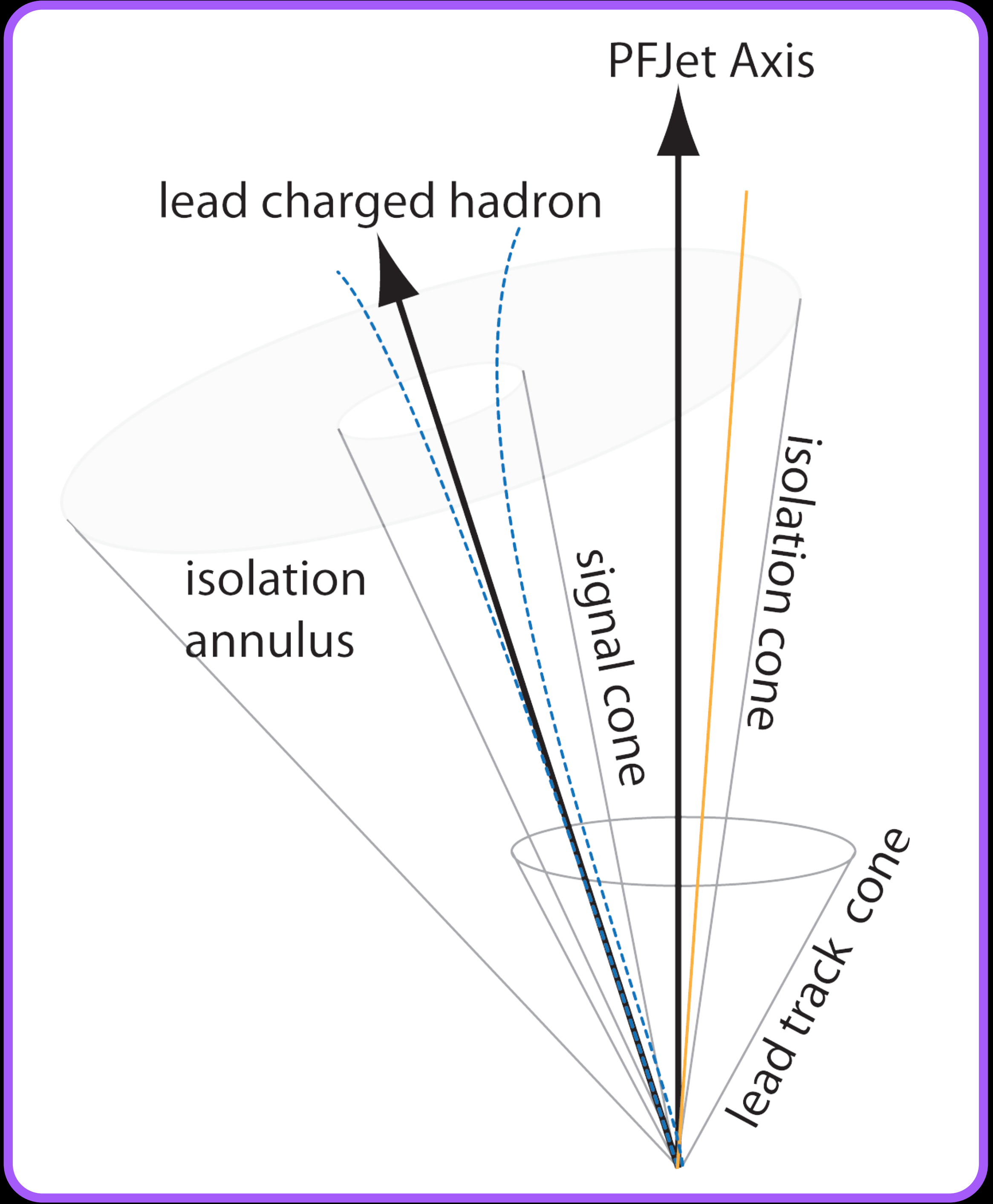
require a *leading candidate* with $p_T > 5$ within $\Delta R < 0.1$ of jet axis

signal objects are those with $\Delta R < \Delta R_{sig}$ of the lead candidate

$$\Delta R_{sig} = 5.0/E_T^{jet}$$

isolation objects are those in the region $\Delta R_{sig} < \Delta R < 0.5$ about the lead candidate

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$$



Zττ selections

Event Selection

For μ + τ_{had}, e + τ_{had} and e + μ Channels, μ + μ Channel different (Backup)

Trigger

Events triggered by single Electron/Muon Triggers
P_T thresholds 9-15 GeV, depending on instantaneous Luminosity

Lepton Selection

Electrons	Muons	had. τ Decays
P_T > 15 GeV	P_T > 15 GeV	P_T > 20 GeV
$\|\eta\| < 2.4$	$\|\eta\| < 2.1$	$\|\eta\| < 2.4$
isolated	isolated	"loose" Tau id.
		Veto against e/μ

Opposite Charge Lepton Pair

Transverse Mass

e + τ_{had}, μ + τ_{had}: M_T(l + MET) < 40 GeV
e + μ: M_T(e + MET) < 50 GeV && M_T(μ + MET) < 50 GeV

Veto Events with additional isolated Leptons

Z$\tau\tau$ yields

Event Yields

CMS Data, 36 pb^{-1} @ 7 TeV

	$\tau_\mu \tau_{\text{had}}$	$\tau_e \tau_{\text{had}}$	$\tau_e \tau_\mu$	$\tau_\mu \tau_\mu$ ($M_{\mu\mu} < 70$ GeV)
$Z \to \ell^+\ell^-$, jet fake τ_{had}	6.4 ± 2.4	15.0 ± 6.2		
$Z \to \ell^+\ell^-$	12.9 ± 3.5	109.3 ± 28.0	2.4 ± 0.3	20.1 ± 1.3
$t\bar{t}$	6.0 ± 3.0	2.6 ± 1.3	7.1 ± 1.3	0.15 ± 0.03
$W \to \ell\nu$	54.9 ± 4.8	30.6 ± 3.1		
$W \to \tau\nu$	14.7 ± 1.3	7.0 ± 0.7	1.5 ± 0.5	2.5 ± 2.5 (< 5 @95 $\%CL$)
QCD	131.6 ± 14.1	181.1 ± 22.5		
WW/WZ/ZZ	1.6 ± 0.8	0.8 ± 0.4	3.0 ± 0.4	
Total Background	228.4 ± 15.8	346.4 ± 36.7	14.0 ± 1.8	22.8 ± 2.8
Total Data	516	540	101	58

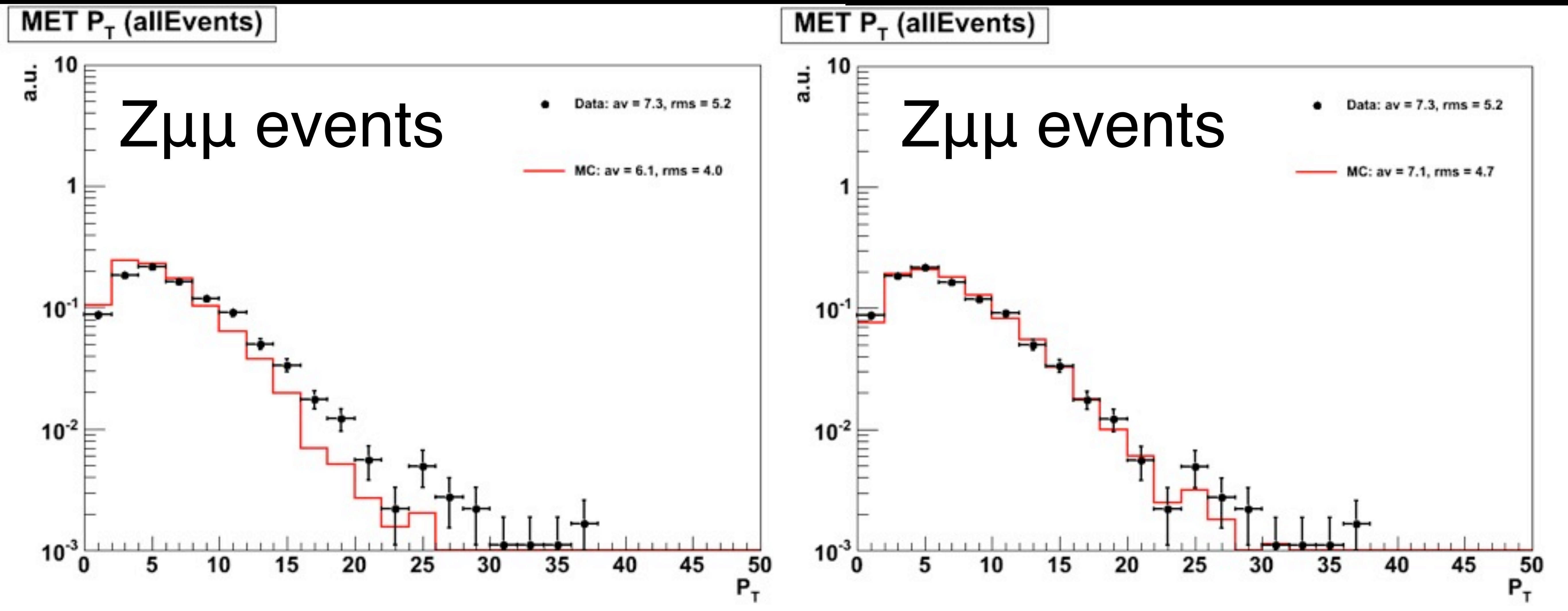
Background Estimates quoted in Table obtained from Data-driven Methods

> 600 Z → $\tau^+\tau^-$ Signal Events selected in CMS Data

Z recoil correction

MET resolution better in Monte Carlo than data

before correction after correction

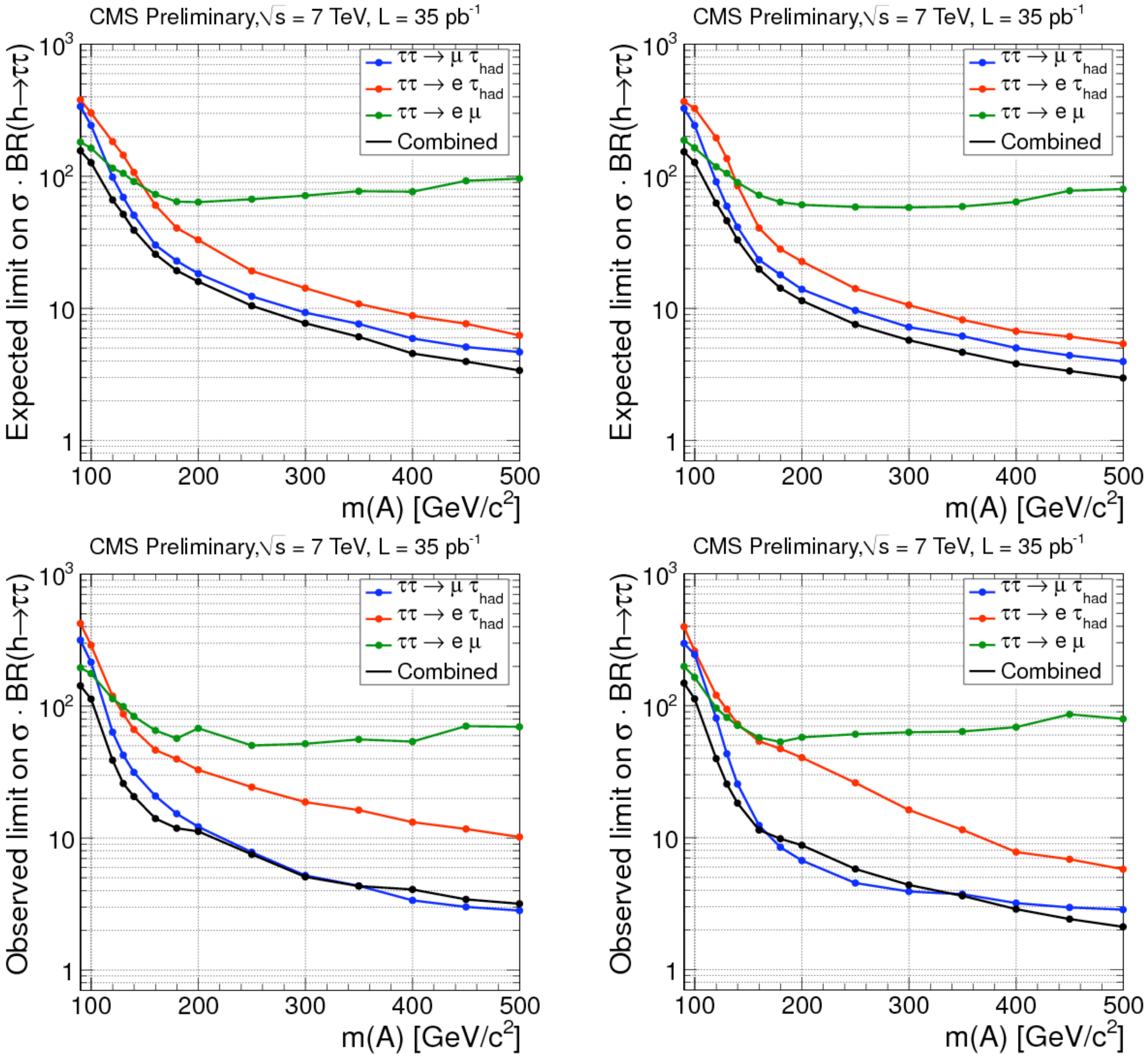


Project MET in directions parallel and perpendicular to Z and parametrize as function of Z p_T

sub title

Observed vs. expected Limits by Channel

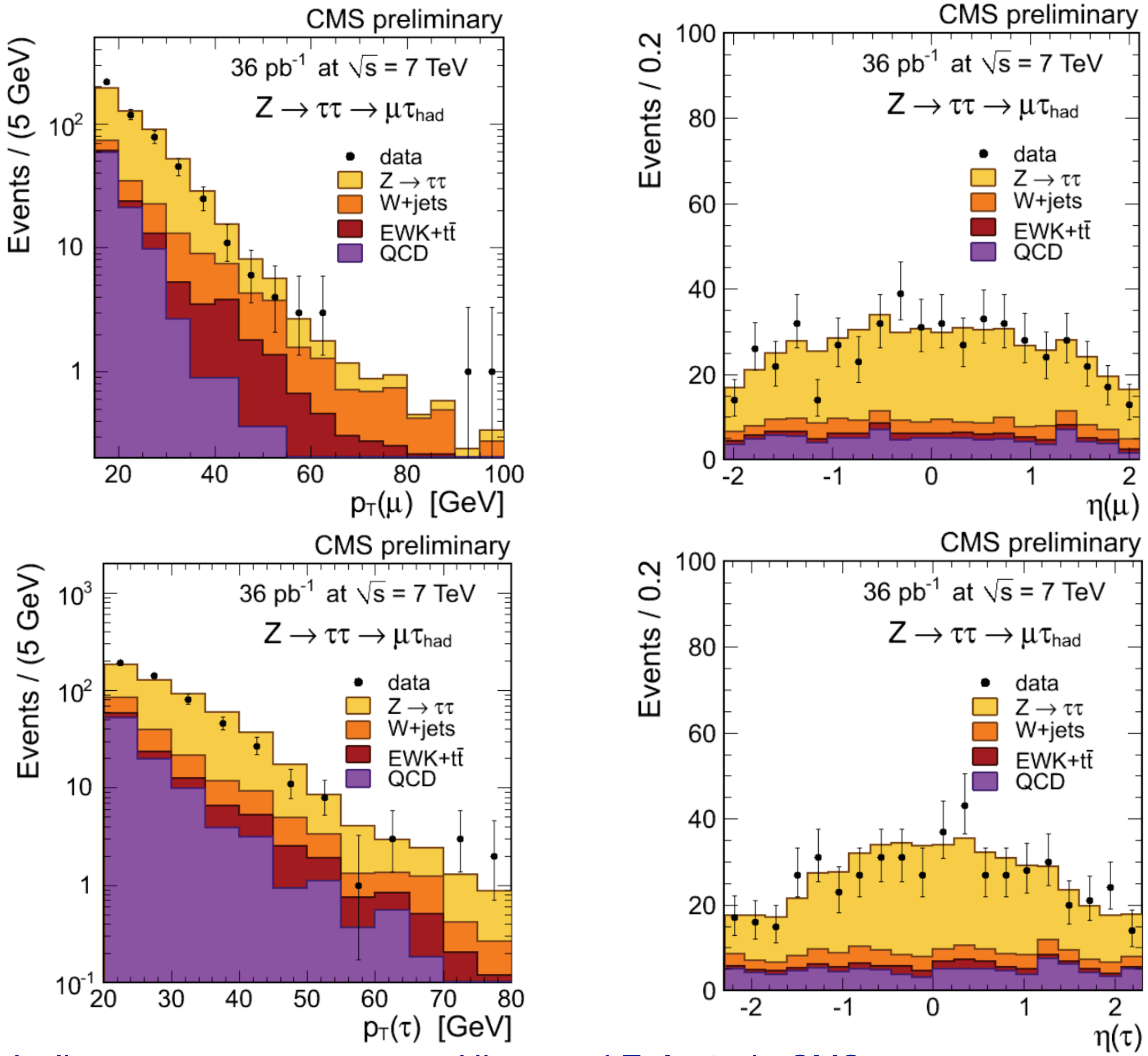
visible Mass | "full" $\tau^+\tau^-$ Mass



Higgs Search

control plots

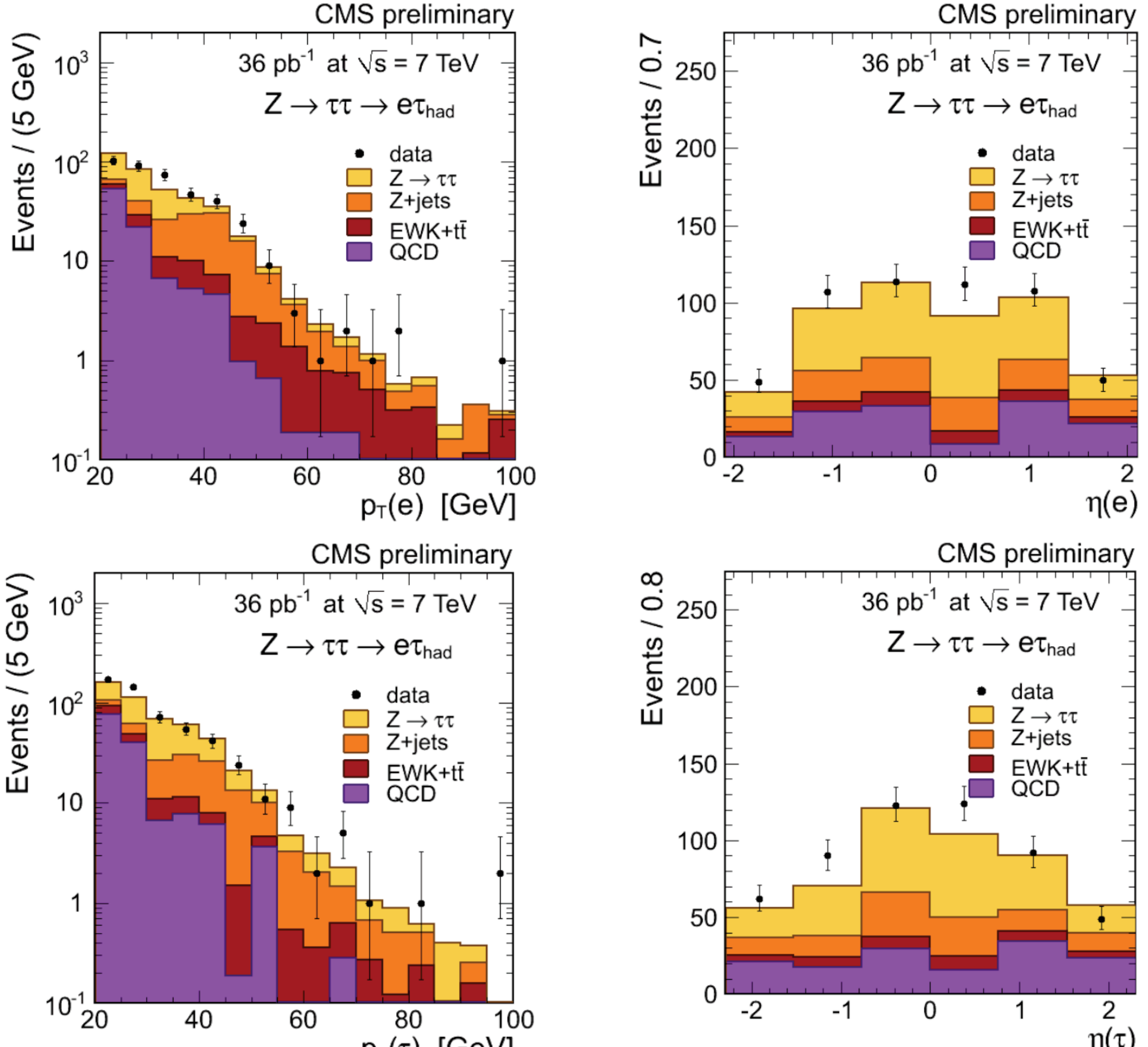
Lepton P_T and η: μ + τ_{had} Channel



Higgs Search

control plots

Lepton P_T and η: e + τ_{had} Channel



$Z \rightarrow \tau^+\tau^-$ Systematic Uncertainties

Source	$\tau_\mu \tau_{had}$	$\tau_e \tau_{had}$	$\tau_e \tau_\mu$	$\tau_\mu \tau_\mu$
trigger	0.2 %	3 %	0.2 %	0.3 %
lepton identification and isolation	1.0 %	1.1 %	1 %	1 %
τ_{had} identification	23 %		-	
efficiency of topological selections	2 %		-	
likelihood selection efficiency	-			2 %
acceptance due to τ energy scale, 3 %	3.5 %		-	
acceptance due to e energy scale, 2 %	-	1.6 %	1.6 %	-
acceptance due to μ momentum scale, 1 %	1 %	-	1 %	2 %
luminosity	4 %			
parton distribution functions	2 %			

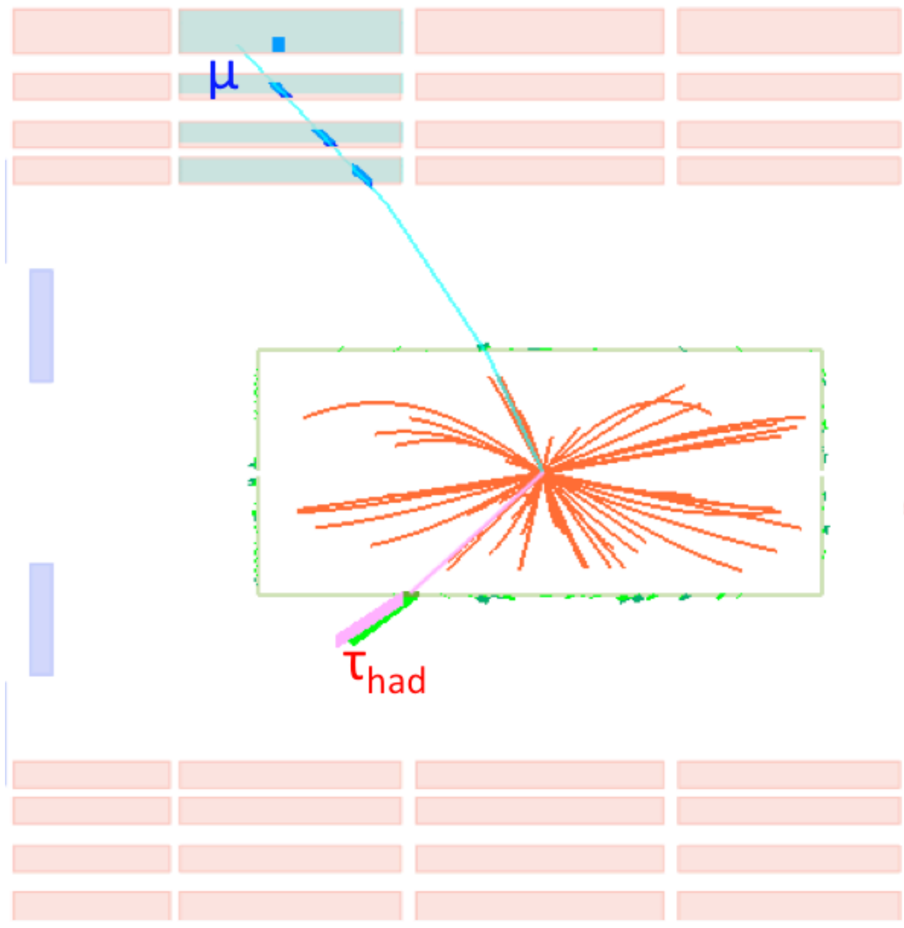
Largest Uncertainty: hadronic Tau Identification Efficiency

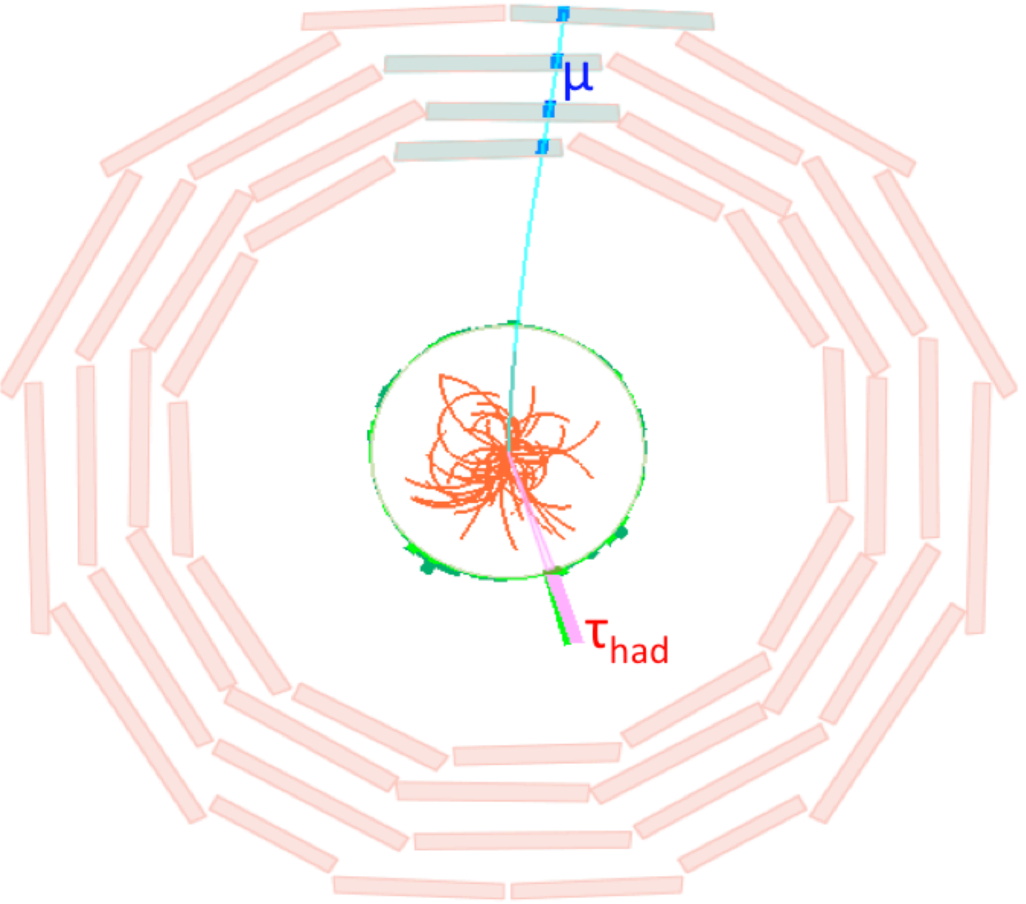
➔ τ_{had} Identification Efficiency constrained by Ratio of Event Yields in semi-leptonic/leptonic Channels

➔ Determine $Z \rightarrow \tau^+\tau^-$ Cross-section by simultaneous Fit of all four Channels

$Z \rightarrow \tau^+\tau^- \rightarrow \mu^+ + \rho^-$ Candidate

$\quad\llcorner \pi^- \pi^0$





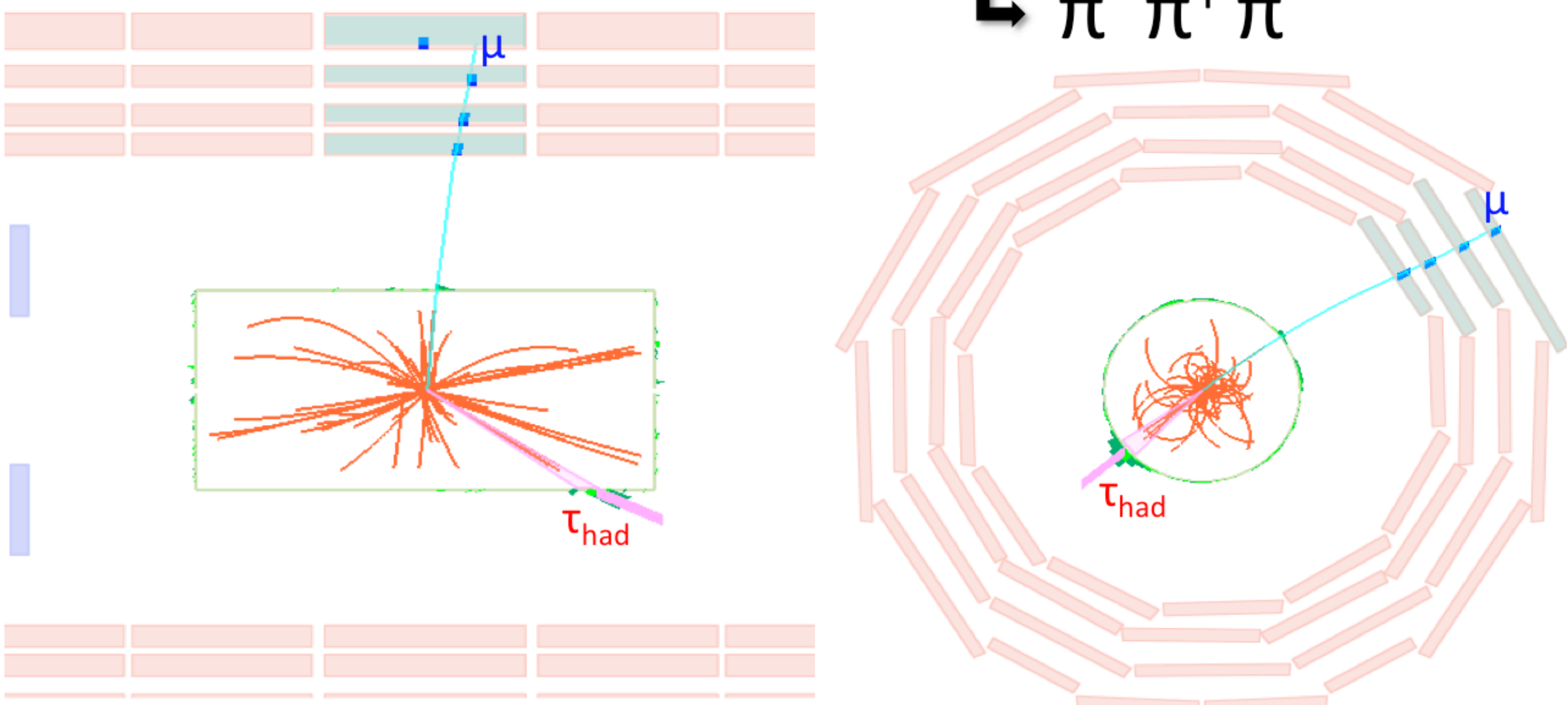
P_T^{μ} = 39.3 GeV

P_T^{had} = 28.2 GeV, lead. Track P_T = 3.3 GeV

M_{vis} = 67.0 GeV, MET = 19.9 GeV, $M_{\tau\tau}$ = 90.3 GeV

$Z \rightarrow \tau^+\tau^- \rightarrow \mu^+ + a_1^-$ Candidate

$\quad \lfloor \pi^- \pi^+ \pi$

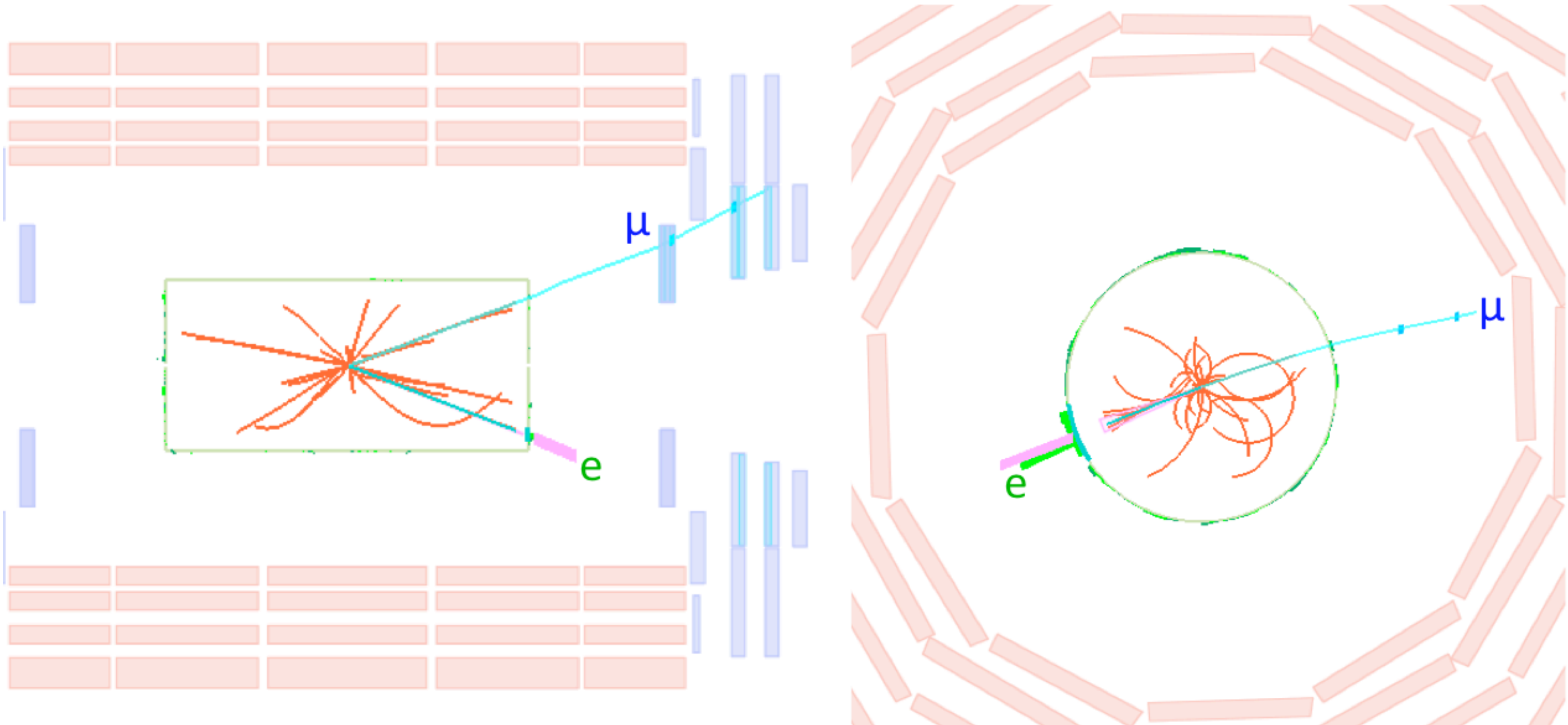


P_T^{μ} = 20.5 GeV

P_T^{had} = 35.5 GeV, lead. Track P_T = 18.5 GeV

M_{vis} = 62.7 GeV, MET = 6.2 GeV, $M_{\tau\tau}$ = 98.3 GeV

$Z \rightarrow \tau^+\tau^- \rightarrow e^- + \mu^+$ Candidate



P_t^e = 29.9 GeV

P_T^μ = 16.3 GeV

M_{vis} = 44.2 GeV, MET = 17.4 GeV, $M_{\tau\tau}$ = 91.4 GeV