Data Driven Methods to Estimate the Vector Boson + Jets Background to New Phenomena Searches involving Missing Energy at CMS

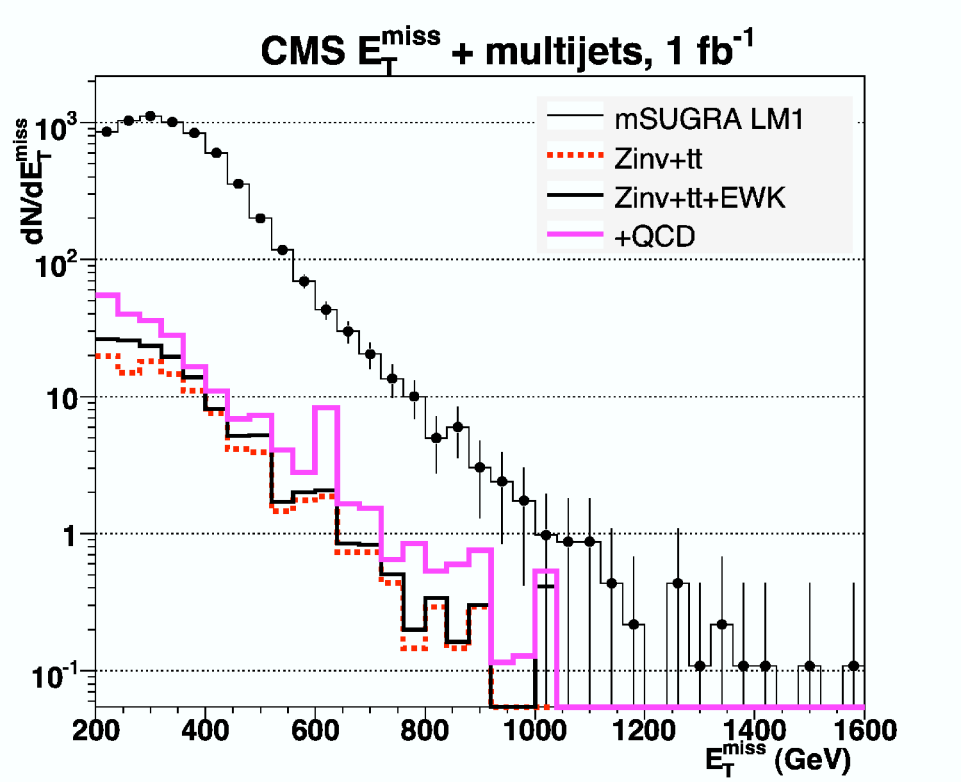
LHCDM@MCTP: The LHC and Dark Matter, 6-10 Jan 2009,

Ann Arbor, MI

James Lamb (UC Santa Barbara)

on behalf of the CMS Collaboration

Discovery of New Physics



MET+jets search, CMS PTDR, 2006

- Discovery of New Physics could take the form of an observed excess of events with high Missing Transverse Energy (MET) and several energetic jets
  - The challenge: convincing you (and ourselves) that the supposed excess doesn't come from Standard Model processes

Introduction

- Searches involving MET are good candidates for first observation of new physics
  - Supersymmetry (SUSY) has been the focus of many studies that are used as examples in this talk
  - MET-based searches are of interest to a multitude of new physics models
- Outline of the talk:
  - Introduction
  - Establishing the vector boson signal in early data
  - Selected data-driven methods studied by CMS
    - $Z \rightarrow \mu^+\mu^-$ standard candle
    - Measurements of high-p_T photon and W cross section
    - Extrapolation using forward events as a background-dominated control region

Introduction

- A typical SUSY selection:
  - Large missing E_T (MET)
  - Multiple energetic jets
  - N leptons
    - N=0,1,2+ investigates different topologies and presents different background challenges
  - Angular or event shape variables for reducing background
    - e.g. angle between MET and jets for QCD rejection
- Main backgrounds: QCD, t-tbar, W, Z
  - In regions of phase space that are poorly sampled by Tevatron measurements (large number of jets, high boson p_T)

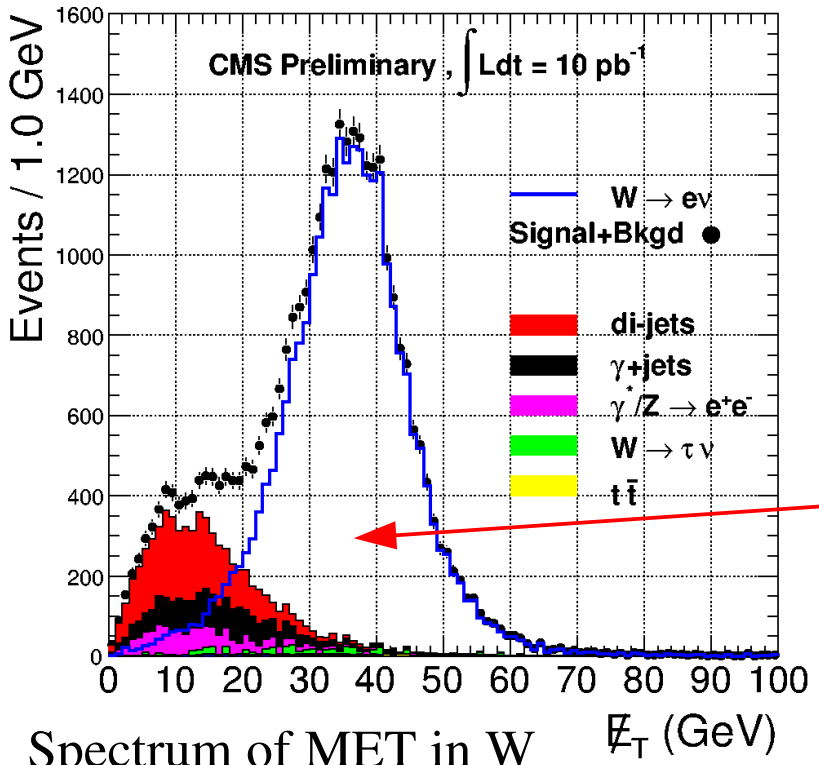
Data Driven Methods

- We would like to estimate backgrounds by taking them *directly from the data* and *minimizing dependence* on Monte Carlo ("data driven methods")
  - Look for "control regions" of phase space that are enriched in background and depleted in signal
    - Use an observable that is invariant for background events in order to extrapolate from control regions to signal region
  - Exploit ratios of production of similar processes
    - Uncertainty in what New Physics signature might look like
    - Different Standard Model processes likely to be impacted differently by new physics

Early W/Z Measurements

- Measurements of the inclusive W and Z production cross sections will be among the first results published by CMS
  - A “test bench” used to calibrate the CMS detector performance
    - Lepton selection efficiencies with $Z \rightarrow e^+e^-/\mu^+\mu^-$ “tag and probe”
    - MET resolution with W->eν/μν
  - A “playground” for physics: understand the inclusive production, then start investigating the tails
    - distribution of number of jets
    - boson p_T spectrum

Early W/Z Measurements



Spectrum of MET in W cross section measurement

[*] 10 pb^{-1} is roughly equivalent to a day of data taking at LHC design luminosity

- CMS is preparing for W and Z cross section measurements with 10pb^{-1}[*] integrated luminosity
- A nice plot: making this plot is among the first important steps in showing that MET and leptons are understood

CMS All-Hadronic Search

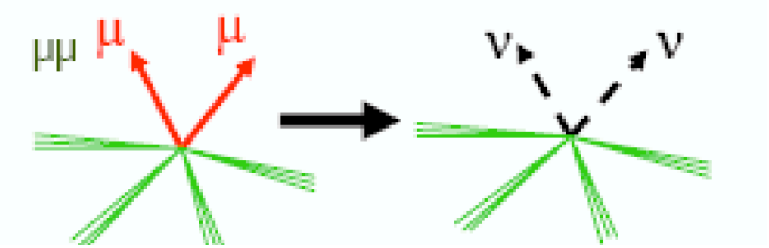
- The canonical CMS MET+Jets SUSY search (Physics TDR, 2006), relevant features:
  - MET > 200 GeV
  - >= 3 jets with 30 GeV uncorrected (~60 GeV partons)
  - Lepton veto
    - Veto events with high-p_T isolated track
    - Veto events with leading jet likely an electron (high e.m.f.)
  - Several selections derived from jet measurements (H_T, angle between MET and jets, lead jet E_T/η, etc)
    - Targeted to reduce QCD and other backgrounds

CMS All-Hadronic Search

- W and Z are backgrounds to the MET+Jets search
  - Z + jets production, with the Z decaying to neutrinos
    - An irreducible background
  - W + jets production, mostly in the channel $W \rightarrow \tau\nu$, $\tau \rightarrow$ hadrons
    - Use of the W/Z production ratio as a function of N jets
- For a W/Z event to be accepted by the MET+jets selection cuts, the boson has to have very high p_T (> 200 GeV), and pass stringent selections on jet-based observables
  - Uncertainty in Monte Carlo modeling
- Multiple approaches have been developed for estimating $Z \rightarrow \nu\nu$ based on ratios with measurable processes
  - $Z \rightarrow \mu^+\mu^-$ "standard candle"
  - methods using $W \rightarrow \mu\nu$ and photons

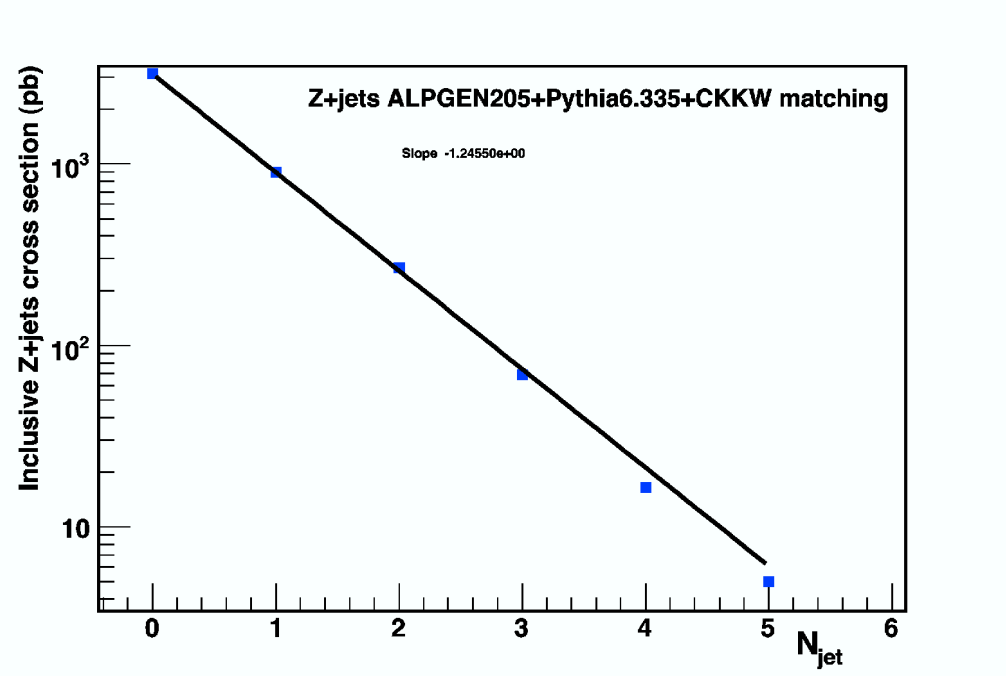
The Z Standard Candle

- If you want to know the MET spectrum of Z->νν, you can measure MET of $Z \rightarrow \mu^+\mu^-$, and scale by the relative branching ratio (~6)
  - $Z \rightarrow \mu^+\mu^-$ is easy to identify, thanks to the CMS tracker and muon system and reconstruction of the Z mass peak ("standard candle")
  - Muons are nearly invisible to the calorimeter: MET spectrum is the same
- Problem: not many events −> high statistical uncertainty

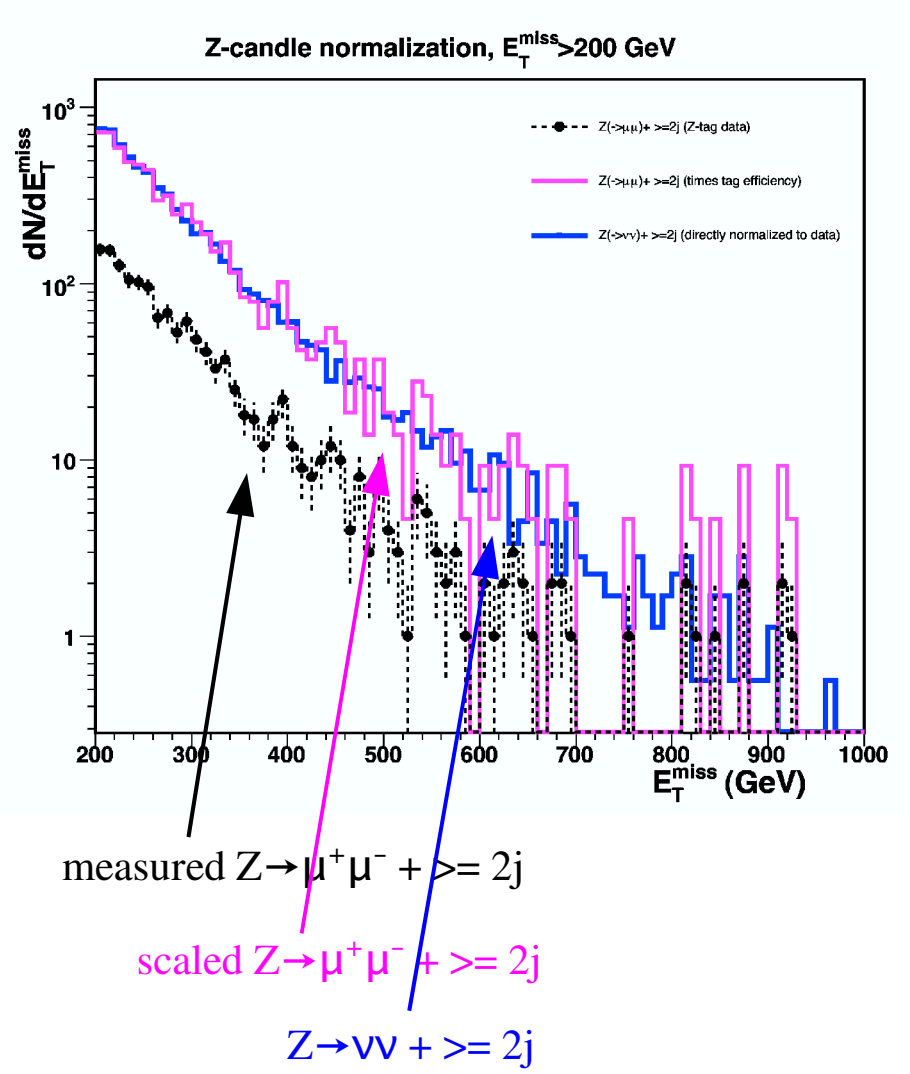


The Z Standard Candle

- Solution: extrapolate from a measurement of $Z \rightarrow \mu^+\mu^-$ with reduced cuts
  - The Z + N jets cross section is proportional to α_s^N: for each additional jet in the event the cross section falls by a factor proportional to α_s



The Z Standard Candle

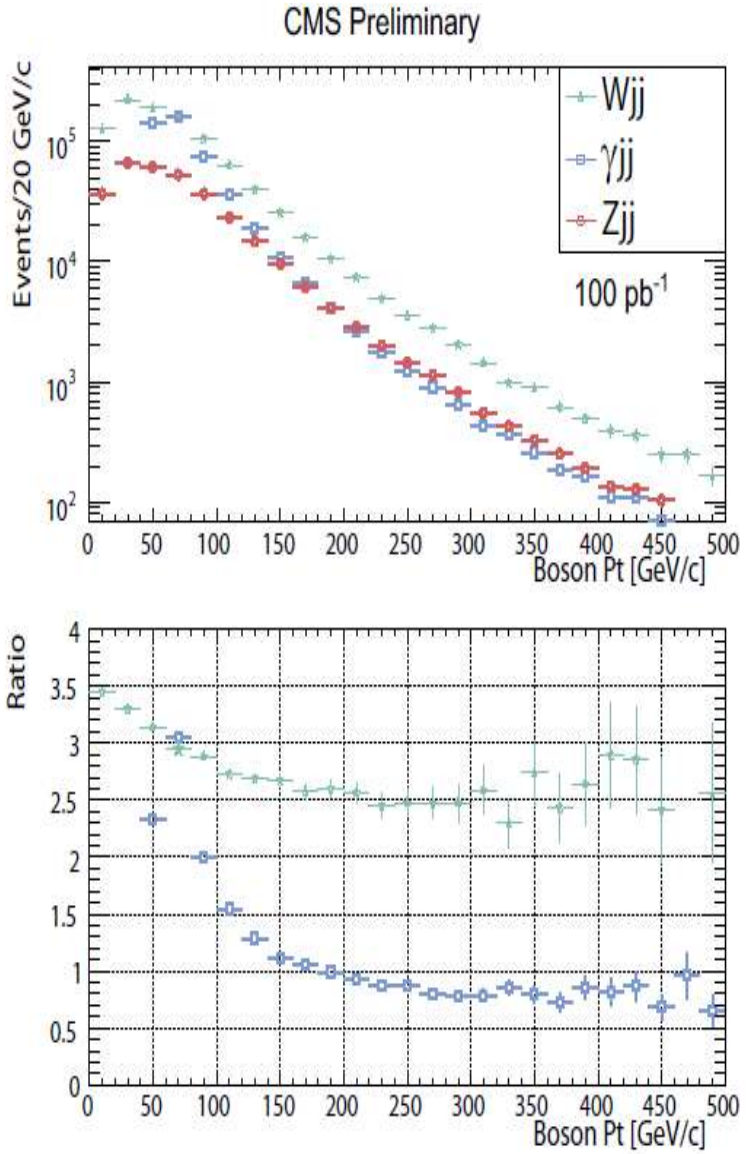


- The Monte Carlo predictions for events with >= 3 jets and Z boson p_T > 200 GeV/c can be normalized to the observed $Z \rightarrow \mu^+\mu^-$ + 2 jets data sample (where Z boson p_T > 200 GeV/c) via the measured R = dN_{events}/dN_{jets} ratio.

Estimate ~5% precision on $Z \rightarrow \nu\nu$ with 1 fb^{-1}

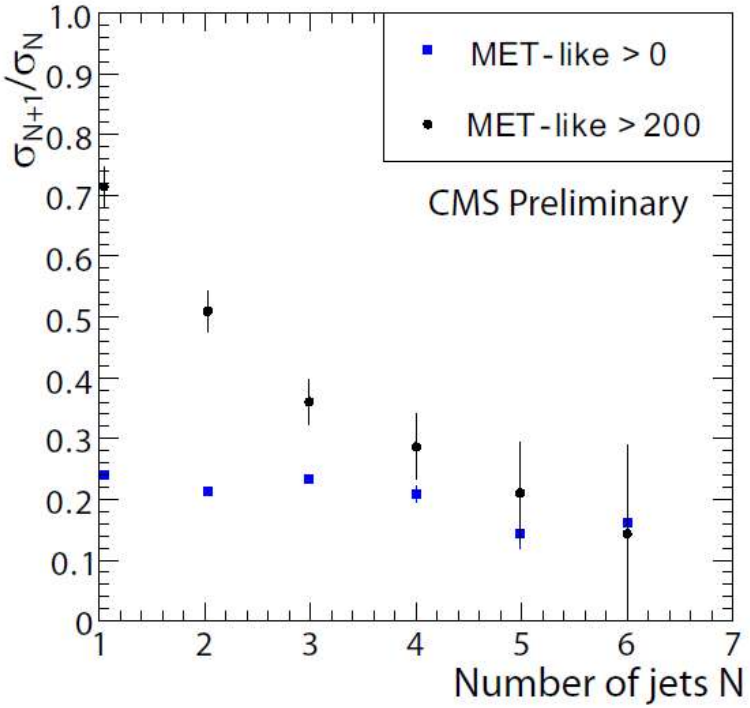
Updates of this PTDR result for smaller data samples (100 pb^{-1}) and using $Z \rightarrow e^+e^-$ are in progress.

W and Photons for Z->νν



- Photon + jets and W + jets can be used to predict Z + jets production at high boson p_T (> 150 GeV/c)
  - Relies on electroweak ratios (γ/W/Z) that may be more precisely known than Monte Carlo extrapolations in number and E_T of jets
  - Z->νν MET spectrum is simulated by removing the photon's E_T

W and Photons for Z->νν



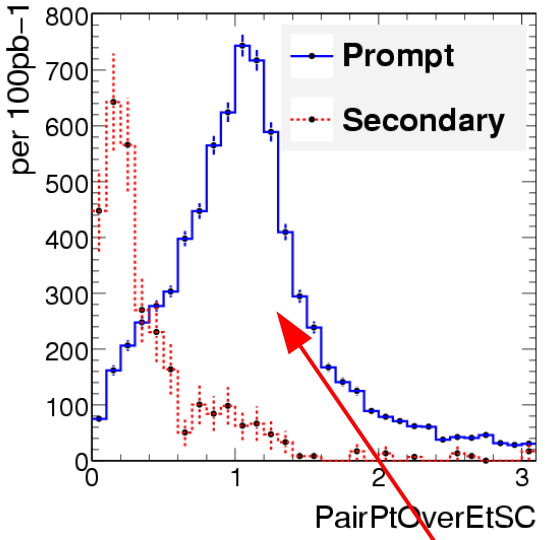
Observed N jets ratio: depends on the composition of the sample and on the experimental definition of a jet

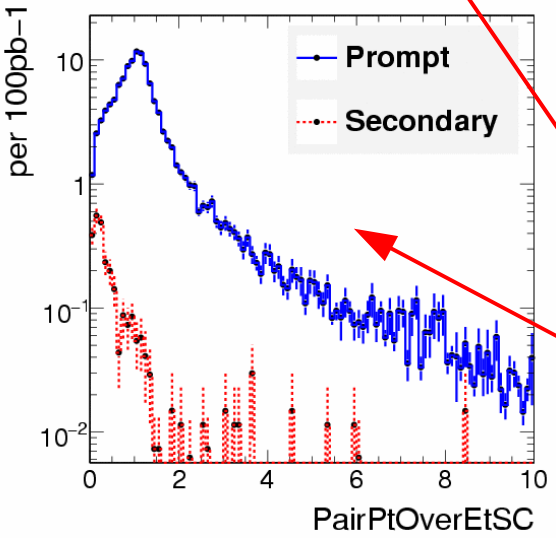
- Using photons and W is an important complement to the "standard candle" approach
  - Hadronic activity should be to good approximation independent of boson type – factors out in the ratio
- Advantage: sufficient statistics to be able to apply the full SUSY search selection

Photons Experimentally

- High-p_T photons are objects that are under excellent experimental control
  - CMS EM calorimeter provides excellent energy resolution and hermeticity out to $|\eta|<3$
  - Low background
    - QCD background via photons from decay of neutral mesons is kept very low by isolation cuts
    - Background from "lost track" electrons is small, and is kept under control by measuring electron to photon fake rate with $Z \rightarrow e^+e^-$ tag-and-probe method
  - Modeling of efficiency of photon selection cuts is checked in data ($Z \rightarrow e^+e^-$ tag-and-probe method)

Photons Experimentally



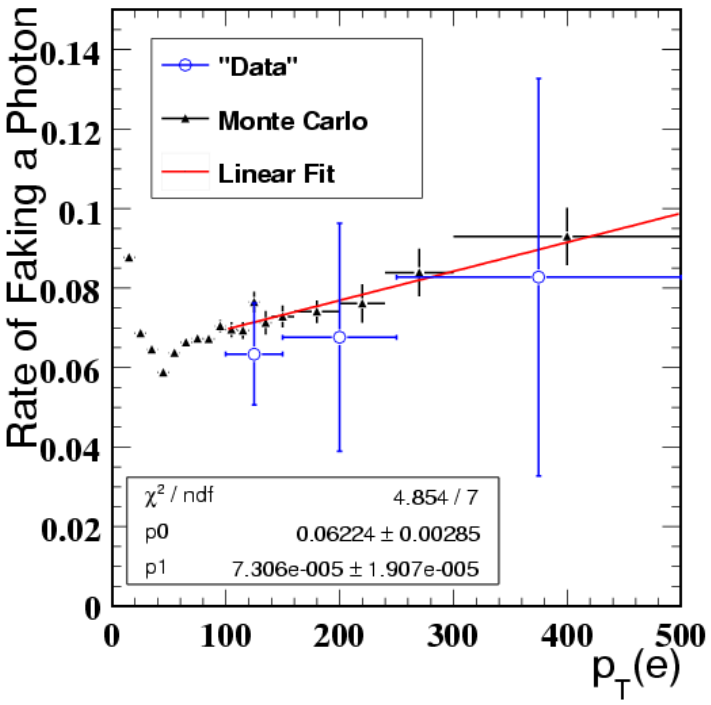


- QCD background via photons from decay of neutral mesons is kept very low by isolation cuts
  - contamination ~ 1 part in 50
- Excellent CMS tracker provides other handles to confirm the expectation of low background
  - Background photons come in collimated groups: sum of conversion track momentum doesn't add up to calorimeter energy deposit

conversions in loose photon event selection

conversions scaled to expectation in full photon event selection

Photons Experimentally

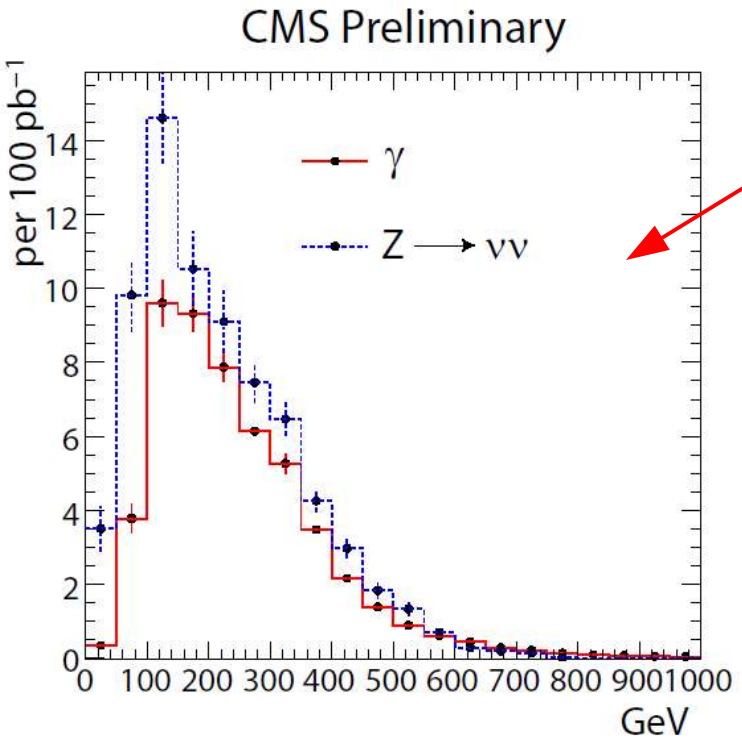


Tag-and-Probe measurement of "lost track" fake rate

- Tag-and-probe[*] method with $Z \rightarrow e^+e^-$ is a powerful tool
  - Allows to measure photon selection efficiency
    - since electrons and photons behave almost identically in the EM calorimeter
  - Provides a sample of electrons needed to estimate the "lost track" electron background

[*] Tag-and-probe: use the Z mass constraint and tight selections for one electron (tag); the other object (probe) is known to be an electron with minimal explicit selection

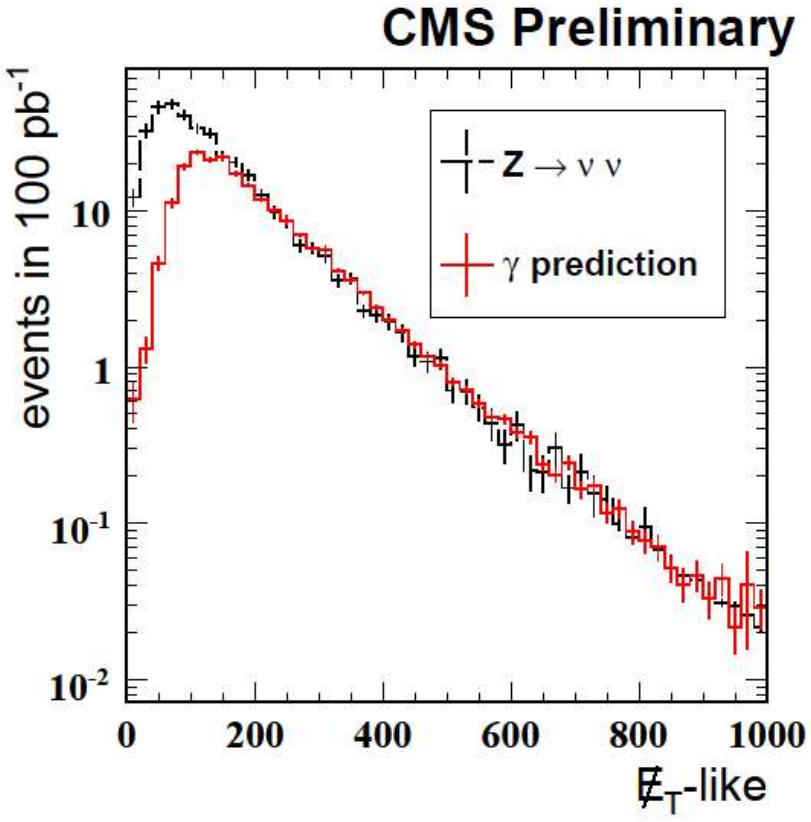
Photons for Z→νν



- ALPGEN event generator and CMS full simulation
- Experimental corrections for photon selection efficiency
- Theoretical corrections not applied

- Prediction of invisible Z MET spectrum from photon events
- Limitations:
  - Photon production is similar to Z production, but not the same
    - understanding the ratio is an area that needs more work if the power of this method is to be maximized
  - Photons are well-measured by the CMS calorimetry: no real MET means different trigger path must be used
    - Expect to be fully correctable

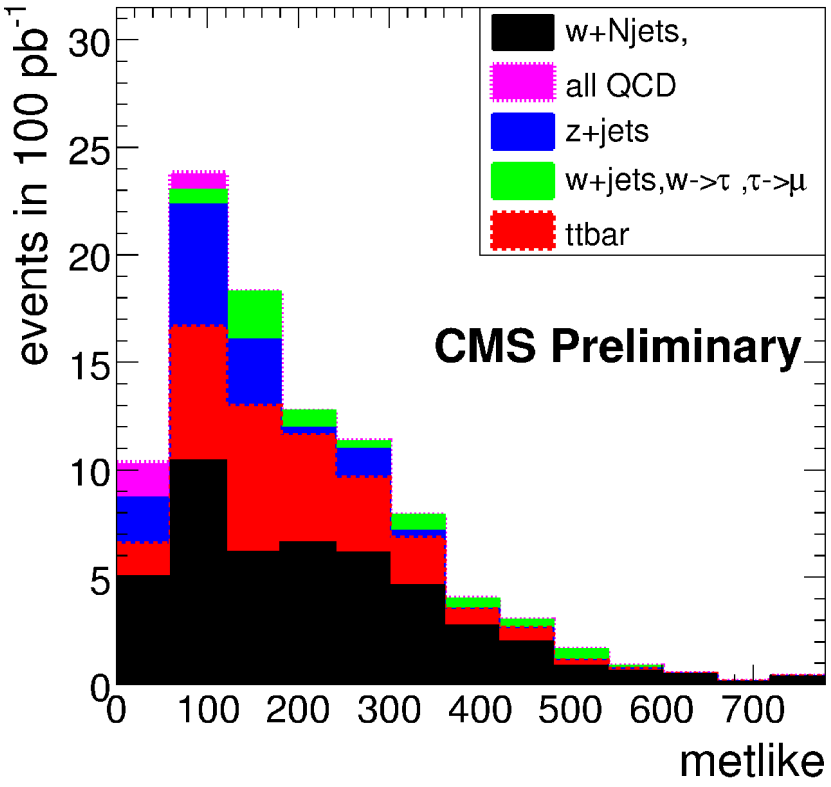
Photons for Z→νν



- Prediction of Z→νν MET spectrum from photon events
  - *With theoretical corrections applied.*
    - Corrections extracted from ALPGEN Monte Carlo

Estimate ~20% precision on Z→νν with 100pb-1
(experimental error only – not including any theory systematic)

W for Z->νν

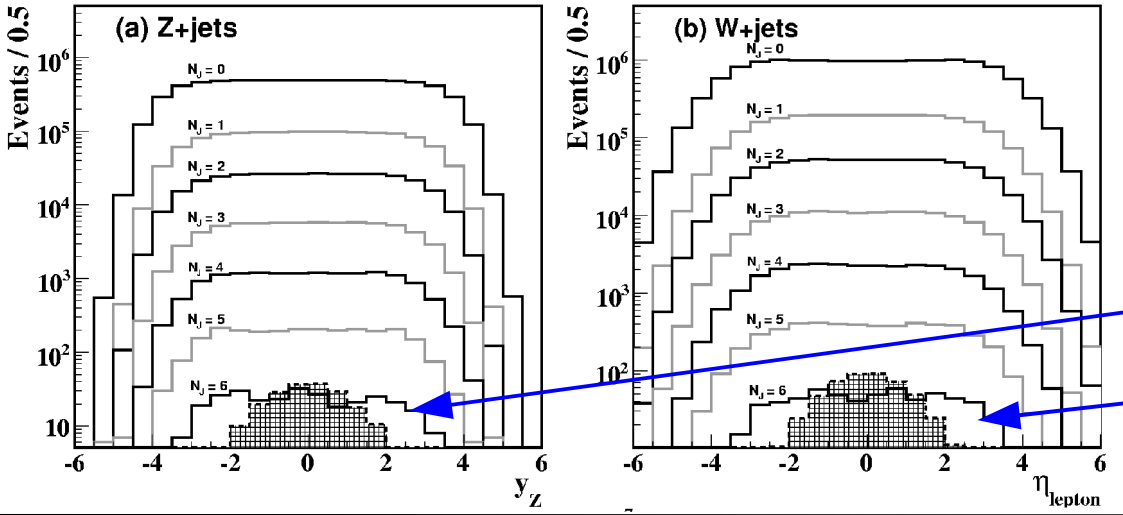


Estimate ~35% precision on Z→νν with 100pb-1

- W→μν is experimentally more challenging than photons
  - t-tbar background is significant
    - Can be estimated using b quark content (ratio of tagged/untagged events)
- W/Z ratio is better understood at the theoretical level

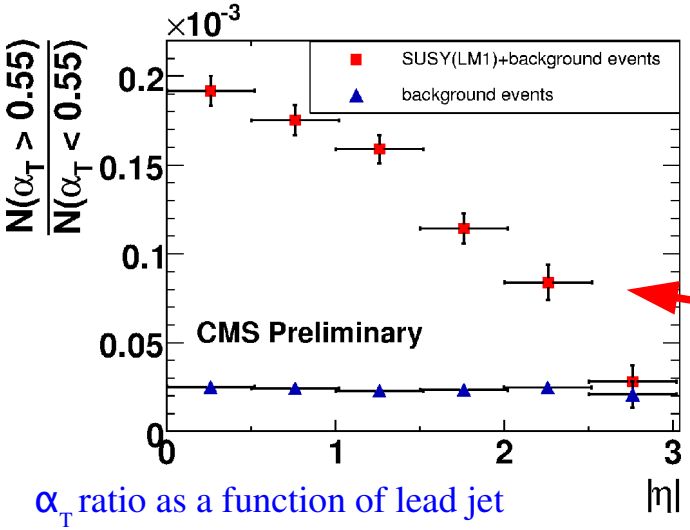
Forward Region Extrapolation

- Promising background method under study based on extrapolation from the forward region [*]
  - Heavy SUSY particles produced nearly at threshold leads to central rapidity of decay products, but low-mass SM production distributed over wide rapidity range
    - High-rapidity regions are background-dominated even when N jets or boson p_T is large
    - Rapidity distributions vary slowly as the number of jets increases



[*]Phys.Rev.D78:035012,2008

Forward Region Extrapolation



α_T ratio as a function of lead jet pseudorapidity

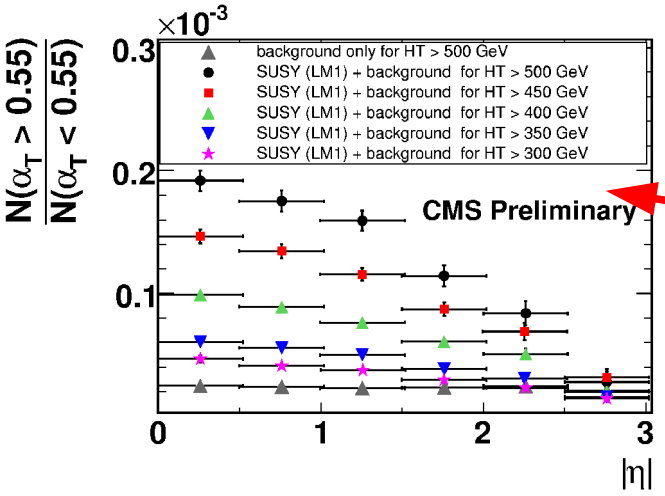


Figure 7: R_{α_T} as a function of $|\eta|$ for different HT cuts.

- A similar method is used by CMS in the search for SUSY using the α_T[*] variable in the di-jet channel
  - Ratio of high-α_T to low-α_T events is measured in the forward region (low signal contribution)
  - and applied to the number of events with low α_T in the central region (also low signal)
  - to predict the number of events with high α_T and central rapidity (where we hope the signal is)
  - nice cross check: loosen H_T cut (increase background), verify that α_T ratio is flat in η

[*] $\alpha \equiv E_T^{j2}/M_{inv}^{j1,j2}$, Phys.Rev.Lett.101:221803,2008

$\alpha_T = E_T^{j2}/M_{inv}^{T\ j1,j2}$

Conclusion

- Data driven methods to understand Standard Model backgrounds are a critical part of any new phenomena discovery at LHC
  - Look for ways to extrapolate from background dominated regions of phase space to regions where new physics might appear
- Establishing the W/Z signal with as little as 10 pb^{-1} of integrated luminosity is planned as one of the important first steps in discovery
- CMS is developing several dedicated methods to measure the W/Z backgrounds to searches involving MET
  - $Z \rightarrow \mu^+\mu^-$ standard candle
  - Methods using high-p_T photons and $W \rightarrow \mu\nu$
  - Extrapolation using forward events as a background-dominated control region