Search for Supersymmetry at the LHC

V. Daniel Elvira
Fermi National Accelerator Laboratory

- Part I: Motivation and Physics Objects ID

  Why we search for SUSY. Detector properties and Physics Object reconstruction . The Standard Model benchmarks.

- Part II: Data analysis in SUSY Searches

  Elements of a SUSY analysis and their integration in a search result. Concepts of data selection, background estimation, control and signal samples, event excess, mass and cross section limits

- Part III: Search for SUSY in CMS

  Recent public results of SUSY searches in CMS (mostly) based on a 36 pb^{-1} data sample collected during 2010. Comparison with results from ATLAS.

Bibliography

CMS Physics Results

https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResults

- Plots and Results
- Journal Publications
- Physics Analysis Summaries – public documents

ATLAS Physics Results

https://twiki.cern.ch/twiki/bin/view/AtlasPublic

Standard Model at CMS:
The foundations of discovery

V. Daniel Elvira
Fermi National Accelerator Laboratory

Standard Model Measurements

Precise measurements of Standard Model (SM) "candles" essential to establish solid ground for searches

New physics signals appear as an excess of events with respect to the SM predictions

It is important to measure accurately cross sections for:

- Jets
- W/Z+jets
- Top

This constitutes the background for SUSY searches

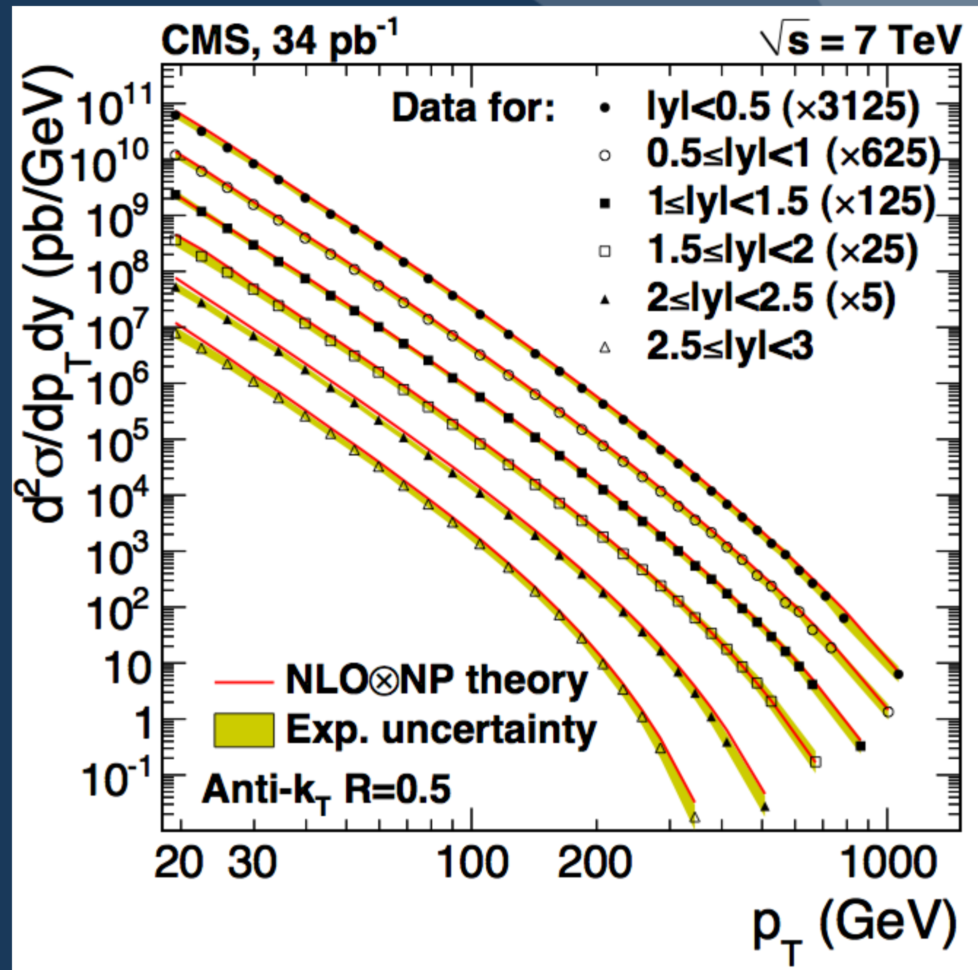
No understanding of background means no discovery

Standard Model: Jets

Inclusive jet cross-sections measured with 10-20% accuracy in most of range:

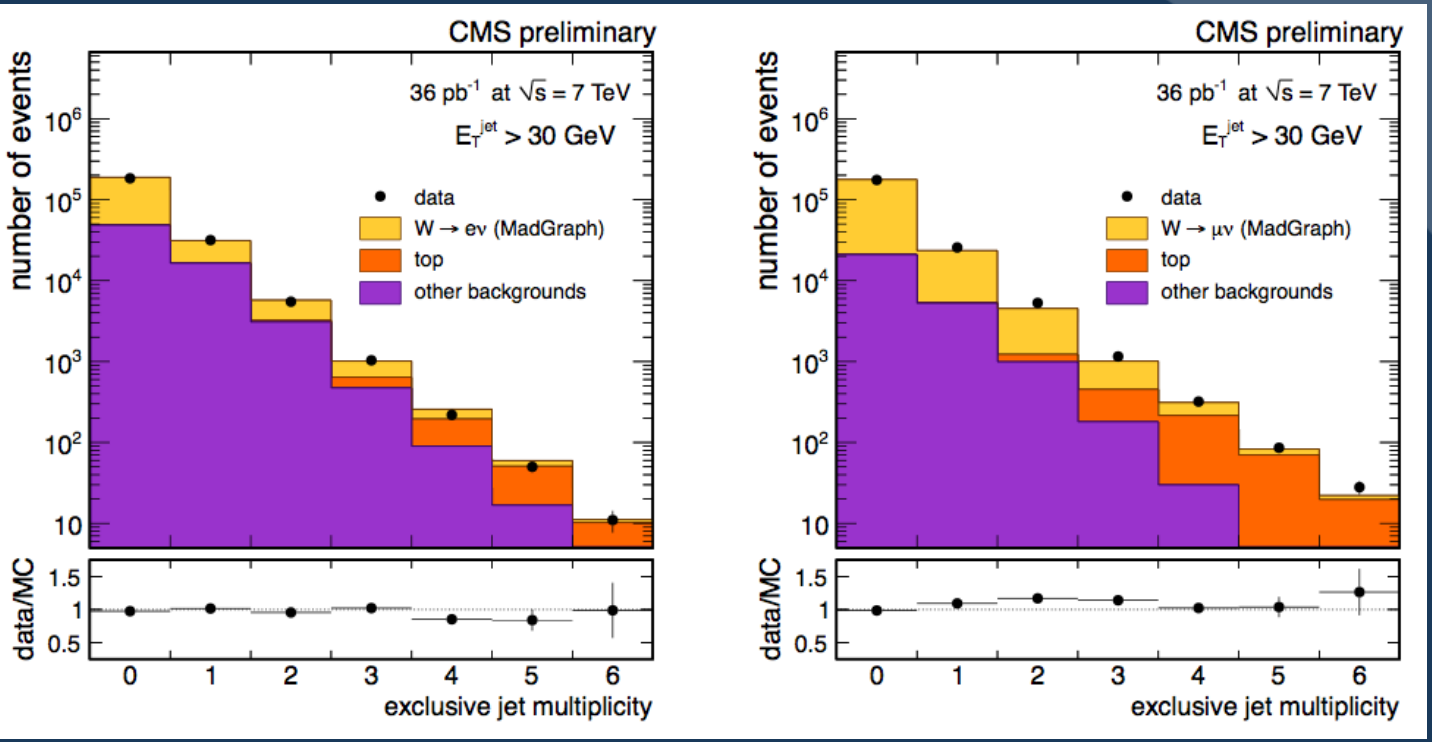
p_T=[18,1100] GeV
$|\eta|<3$

Multijet events are background in most SUSY searches



CMS-PAS-QCD-10-011

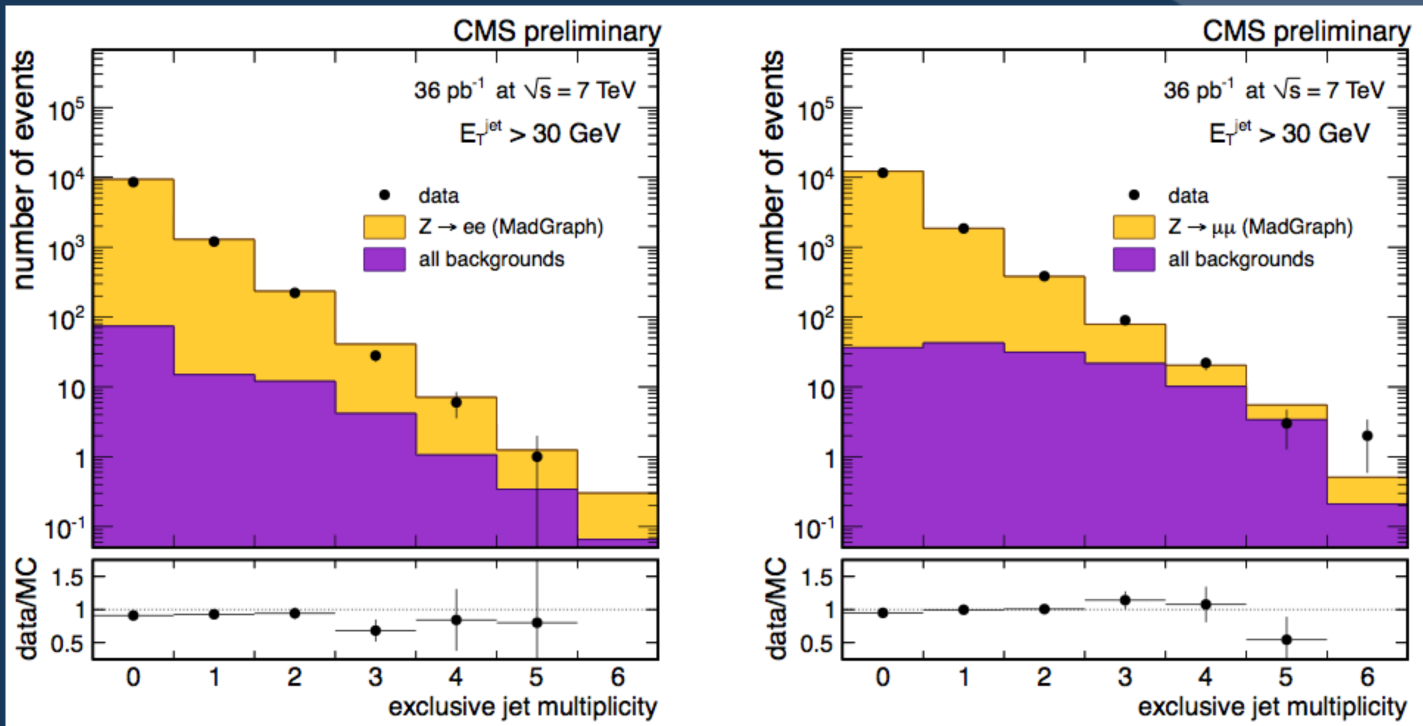
Standard Model: W/Z+jets



Jet multiplicity predicted with excellent accuracy in W candidate events

CMS-PAS-EWK-10-012

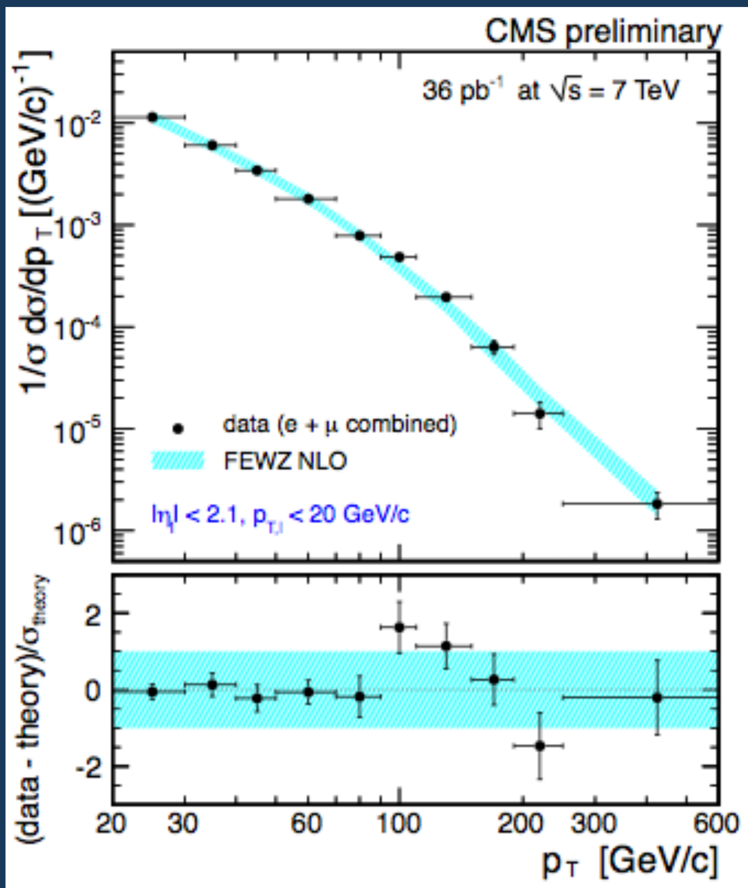
Standard Model: W/Z+jets



Jet multiplicity predicted with excellent accuracy in Z candidate events

CMS-PAS-EWK-10-010

Standard Model: W/Z+jets

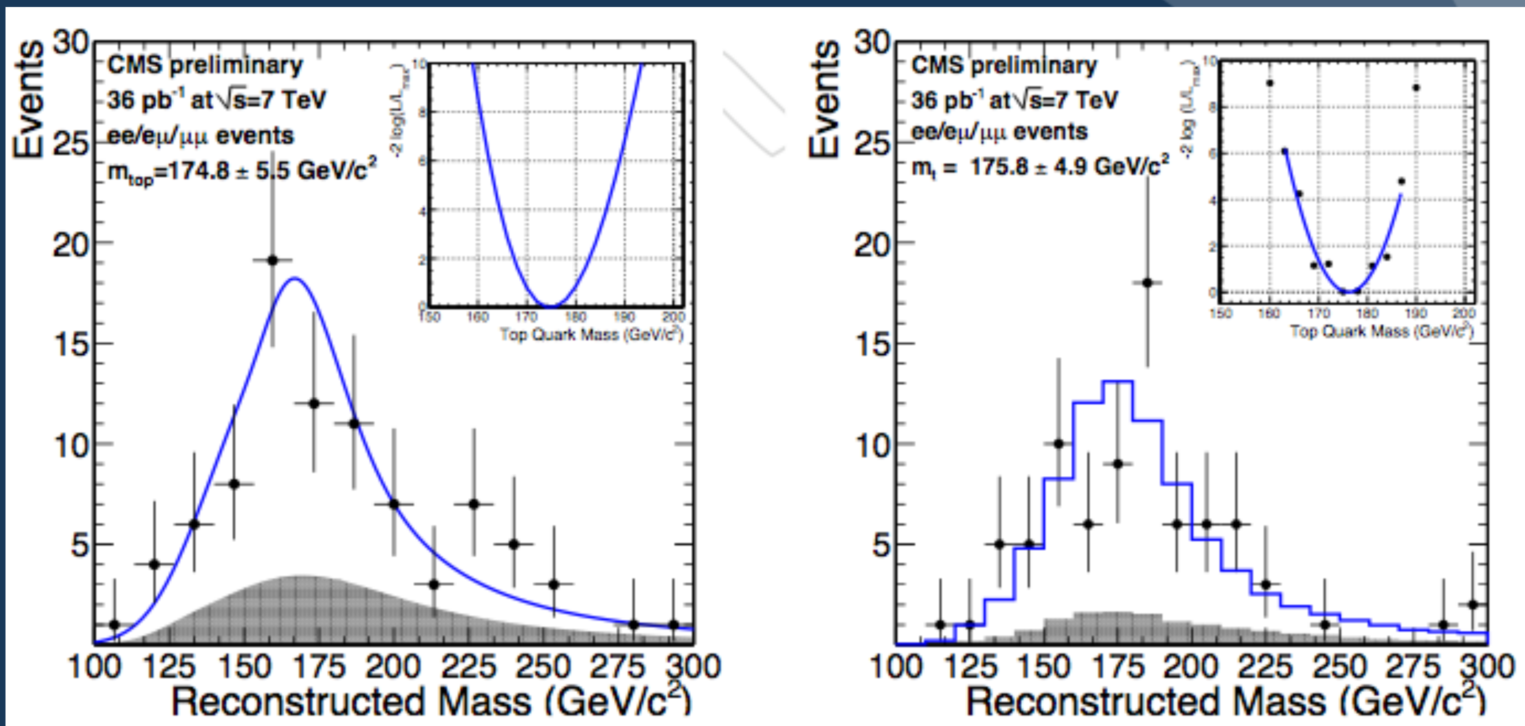


W/Z p_T cross sections measured accurately from W/Z+jets

Another background in SUSY searches

CMS-PAS-EWK-10-010

Standard Model: Top

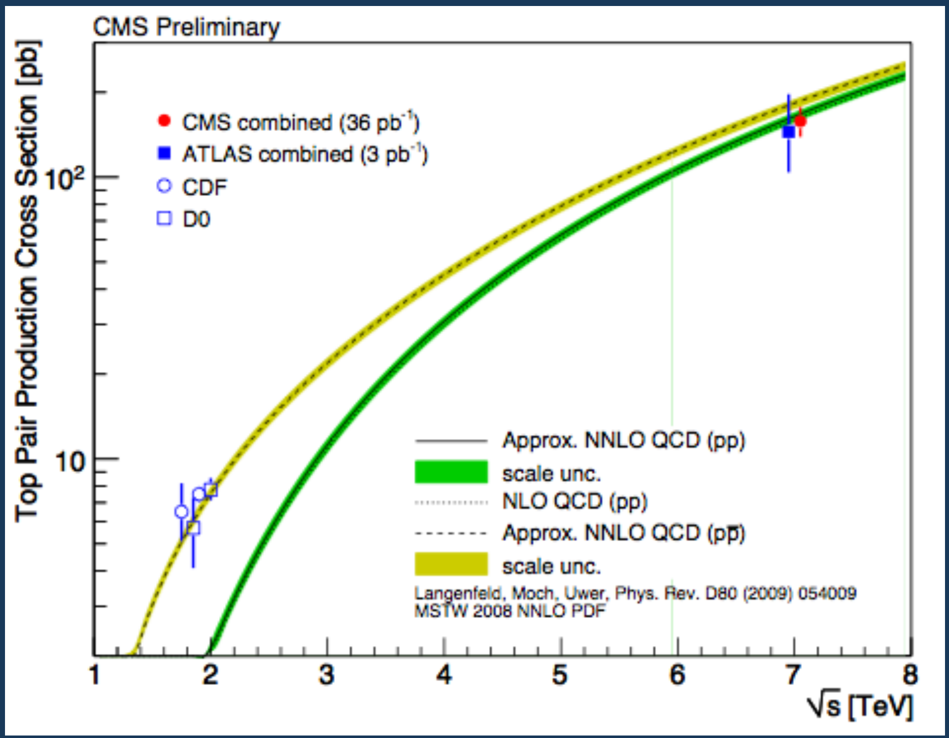


Top mass measure with good precision: 4.6 GeV

Still lower than the 3.8 GeV (D0) and 1.1 GeV (world) uncertainties

CMS-PAS-TOP-10-006

Standard Model: Top



σ = 158±10±15±6 pb^{-1}

Good accuracy

Start to test higher than NLO calculations

ttbar is background in SUSY searches

CMS-PAS-TOP-11-001

A (SUSY) Search Analysis:

How do we build the components and put everything together?

V. Daniel Elvira

Fermi National Accelerator Laboratory

Search Deconstruction

The components of a search analysis:

- Theoretical models motivate the search, but they are not essential for a discovery – until you care about its nature

  (A statistically significant deviation of the data from the Standard Model predictions is a signature of new physics)

- Sensitive variables, used to observe the data – event counting is the simplest way
- Background predictions, # of events from SM processes is subtracted from observed data
- Interpretation
  - Statistically significant excess of events – discovery

Search Deconstruction

The components of a search analysis:

- Theoretical models motivate the search, but they are not essential for a discovery – until you care about its nature

  (Any statistically significant deviation of the data from the Standard Model predictions is a signature of new physics)

- Sensitive variables, used to observe the data – event counting is the simplest way

- Background predictions, # of events from SM processes is subtracted from observed data, in case of event counting

- Interpretation

  - Statistically significant excess of events – discovery (and glory)

Search Deconstruction

The components of a search analysis:

- Theoretical models motivate the search, but they are not essential for a discovery – until you care about its nature

  (Any statistically significant deviation of the data from the Standard Model predictions is a signature of new physics)

- Sensitive variables, used to observe the data – event counting is the simplest way
- Background predictions, # of events from SM processes is subtracted from observed data, in case of event counting
- Interpretation
  - No excess does not mean failure !

Search Deconstruction

The components of a search analysis:

- Theoretical models motivate the search, but they are not essential for a discovery – until you care about its nature

  (Any statistically significant deviation of the data from the Standard Model predictions is a signature of new physics)

- Sensitive variables, used to observe the data – event counting is the simplest way
- Background predictions, # of events from SM processes is subtracted from observed data, in case of event counting
- Interpretation
  - Observation consistent with SM prediction means that new physics is not present at the mass scale we are probing – limit on mass or x-section follows

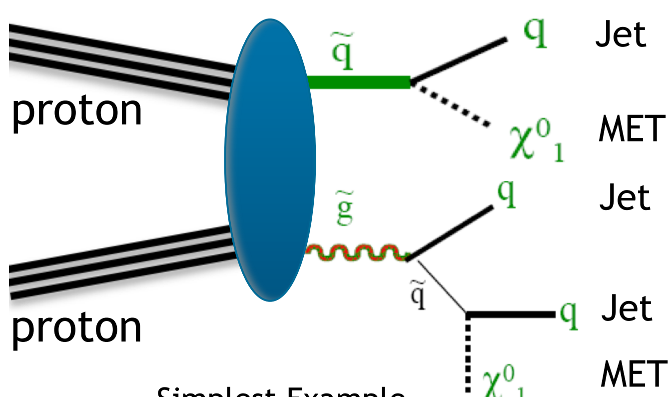
Search by examples:

The MHT Search for jets and missing transverse momentum in the all-hadronic channel

Physics Signals

A generic search for jets and MET in the all hadronic channel is motivated by R-parity conserving SUSY

- Strong production of $\tilde{g}\tilde{g}, \tilde{g}\tilde{q}, \tilde{q}\tilde{q}$
- Largest cross section, most sensitive channel - *if backgrounds are well understood*



Simplest Example

SUSY particles eventually decay to LSP (stable, neutral)

Experimental signature: Jets + Missing Transverse Momentum

In the example, LPS is χ^0_1 (neutralino)

Model independent analysis means:

Concept

- Inclusive sample selection
- High efficiency for a broad range of models associated with final state

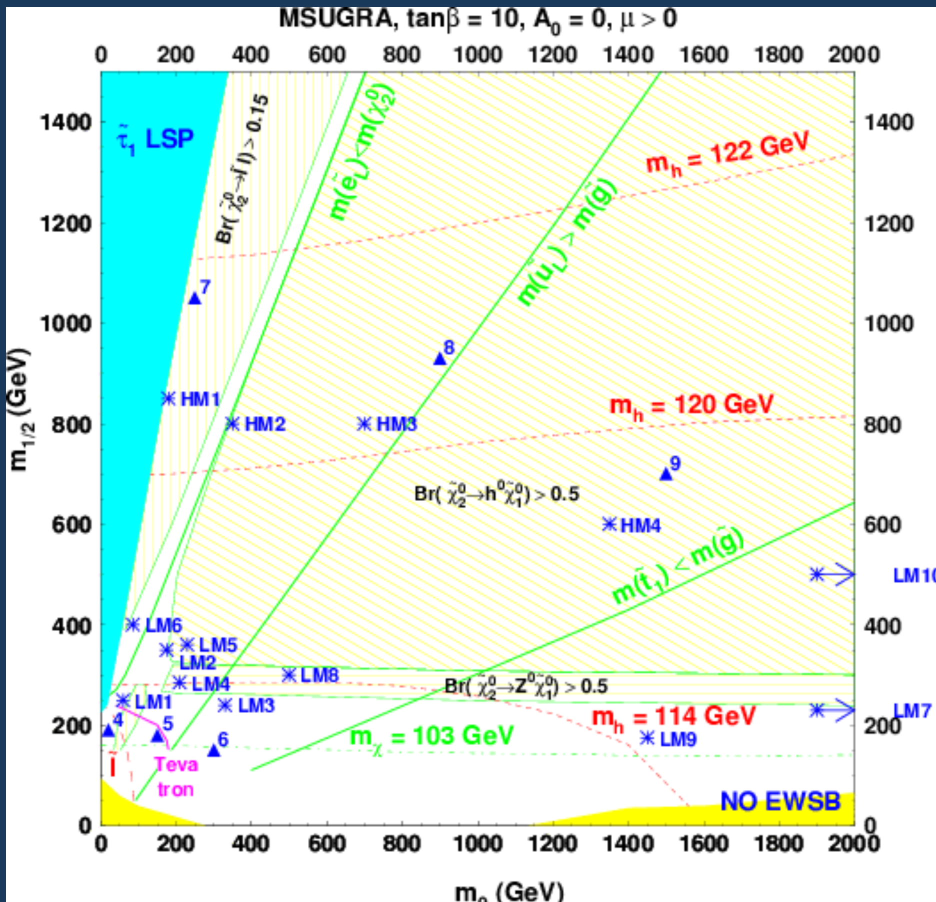
CMSSM Framework Parameters

The Constrained MSSM (CMSSM) framework includes mSUGRA

- Depends on a few independent parameters defined at the M_{GUT} scale
  - ✓ sleptons/squarks/Higgs have the same common scalar mass m_0
  - ✓ gauginos unify at the common mass $m_{1/2}$
  - ✓ Universal trilinear coupling (higgs-sfermion-sfermion) A_0
  - ✓ Ratio of the two higgs doublets VEVs is $\tan\beta$
  - ✓ Sign of higgs/higgsino mass parameter μ, $\mathrm{sgn}(\mu)$
- RGEs used to evolve parameters, compute couplings/masses at EWK scale
- LSP is often the neutralino

Different parameter values correspond to different production cross section for SUSY particles, flavor content, masses and mass hierarchy, length of the decay chain

CMSSM Benchmark Points (CMS)



- Low Mass points (LM1 to LM10), above TeV reach, target early LHC searches
- High Mass points (HM1 to HM4) defined for ultimate CMS reach

CMS Physics TDR, Vol.II, CERN/LHCC 06-021

Point	m_0	$m_{1/2}$	$\tan\beta$	sgn(μ)	A_0
LM1	60	250	10	+	0
LM2	185	350	35	+	0
LM3	330	240	20	+	0
LM4	210	285	10	+	0
LM5	230	360	10	+	0
LM6	85	400	10	+	0
LM7	3000	230	10	+	0
LM8	500	300	10	+	-300
LM9	1450	175	50	+	0
LM10	3000	500	10	+	0
HM1	180	850	10	+	0
HM2	350	800	35	+	0
HM3	700	800	10	+	0
HM4	1350	600	10	+	0

LM1(LMB): m_0=60 (400) GeV, $m_{1/2}$=250 (200) GeV, A_0=0, tan β=10 (50), sign(μ) > 0
The squark and gluino masses (LM1) are 559 GeV and 611 GeV respectively

CMSSM Benchmark Points

Experiments use benchmark points as aid for comparative assessment

Define a grid of points in parameter space for setting exclusion limits

(In CMS, $m_{1/2}$ & m_0 were scanned in 10 GeV steps for tan β=3, 10, 50 using LO generators and NLO k-factors using PROSPINO. Events are then passed through detector simulation)

ATLAS Benchmark Points

	m_0 (GeV)	m_1/2 (GeV)	A0 (GeV)	tan(beta)	σ(NLO) (pb)	Comment
SU1	70	350	0	10	10.9	Soft leptons, taus
SU2	3550	300	0	10	7.2	gluino/gaugino production, heavy flavor decays
SU3	100	300	-300	6	27.7	Generic point
SU4	200	160	-400	10	402.2	Low mass point near Tevatron bound
SU6	320	375	0	50	6.1	Tau rich

The Simplified Models

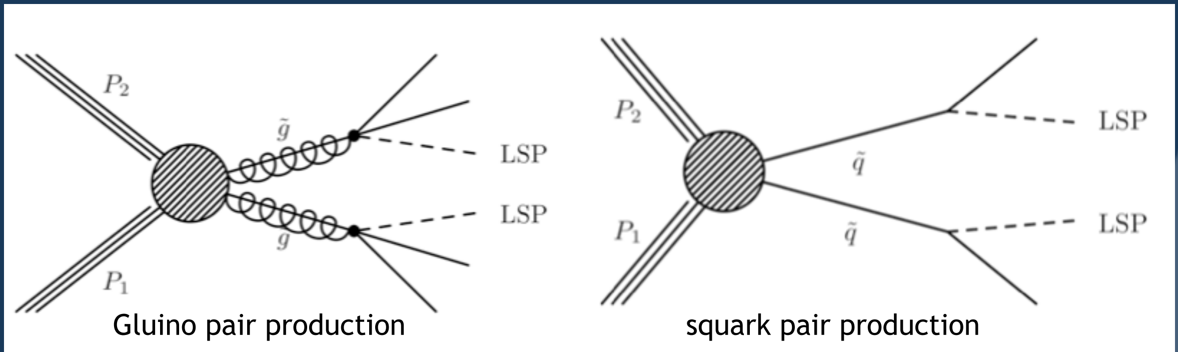
Squark & gluino strong production expected to dominate

- Final state kinematics determined mostly by pdfs and phase space factors associated with 2/3-body decays
- Cross sections depend little on the details of the SUSY model

Simplified Models (SMS)

- Characterize data in terms of small number of basic parameters (~2 x-sections, ~3 masses, ~3 branching ratios)
- Group large sectors of parameter space into a few simplified models with similar final state topologies
- Experimental data then translated to more detailed frameworks using SMS

Alwall, Schuster, Toro: Phys. Rev. D79, 075020 (2009) arXiv:0810.3921[hep-ph]



The Simplified Models are generated with PYTHIA for a range of masses of the particles involved (no fixed gluino/LSP mass as in CMSSM) and passed through detector simulation

Physics Background

Background events are events that mimic the signal — Concept

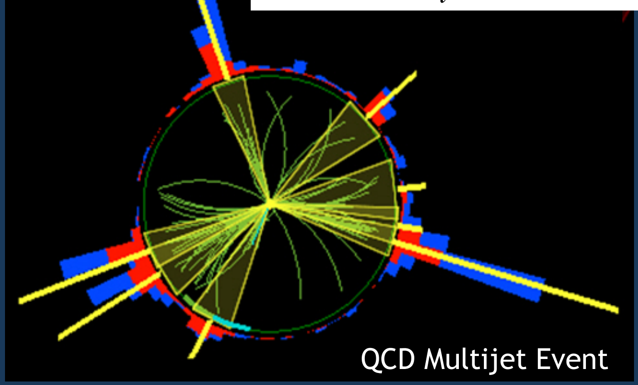
- Reducible: same final state but one or more objects are fake due to detector acceptance, response, efficiency
- Irreducible: indistinguishable from signal events, all objects are real

QCD background:

- Multijets come from QCD Standard Model production
- Large MET created by extreme detector response mis-measurement

$$\vec{p}_T^{\,event} = \sum_i \vec{p}_{Ti}^{\,jet} = 0$$

In the case of an ideal detector (perfect response)



Physics Background

Background events are events that mimic the signal

Concept

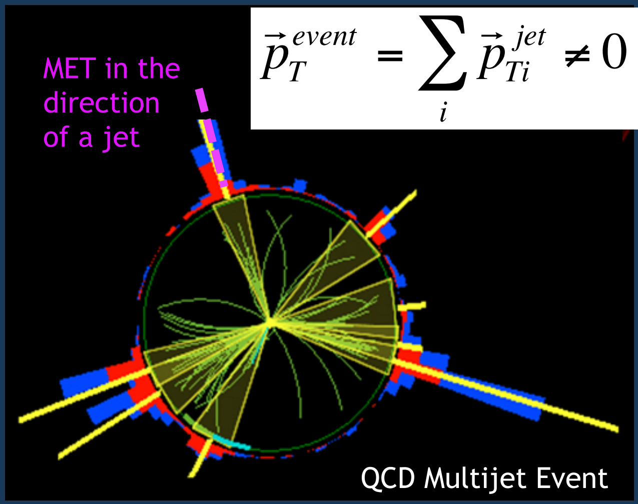
- **Reducible:** same final state but one or more objects are fake due to detector acceptance, response, efficiency
- **Irreducible:** indistinguishable from signal events, all objects are real

QCD background:

- Multijets come from QCD Standard Model production
- Large MET created by extreme detector response mis-measurement

$$\vec{p}_T^{\,event} = \sum_i \vec{p}_{Ti}^{\,jet} \neq 0$$

Detector response <1
Fake MET



Physics Background

Concept

Background events are events that mimic the signal

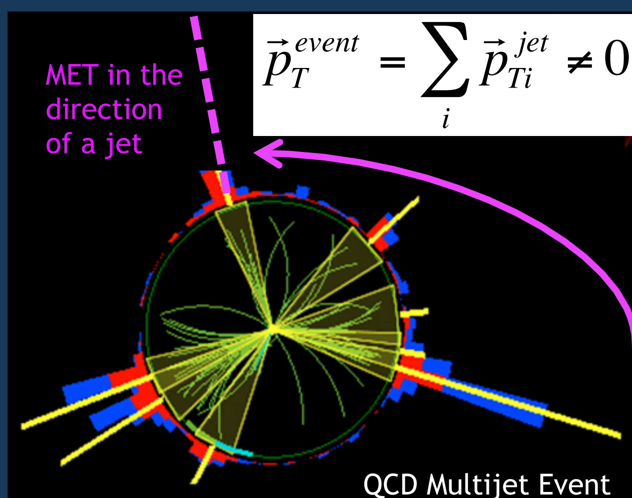
- **Reducible:** same final state but one or more objects are fake due to detector acceptance, response, efficiency
- **Irreducible:** indistinguishable from signal events, all objects are real

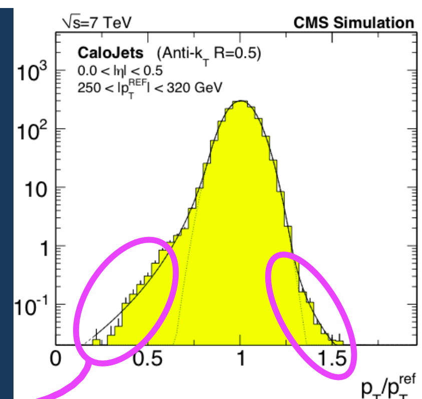
QCD background:

- Multijets come from QCD Standard Model production
- Large MET created by extreme detector response mis-measurement

$$\vec{p}_T^{\,event} = \sum_i \vec{p}_{Ti}^{\,jet} \neq 0$$

Detector response <1
Fake MET





Extreme mis-measurement
⇓
Large fake MET consistent with SUSY signals

(events in the "tails")

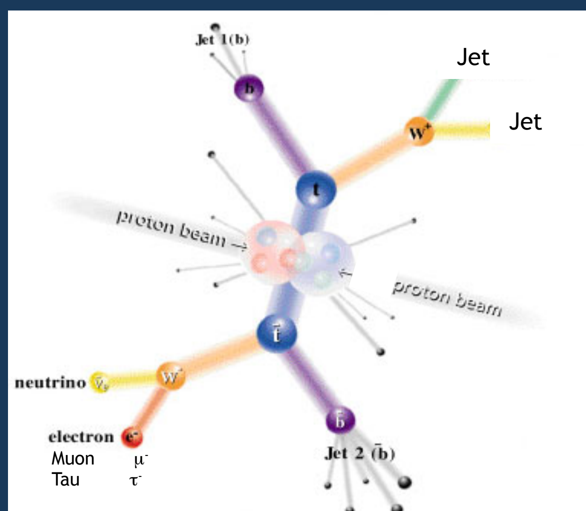
Physics Background

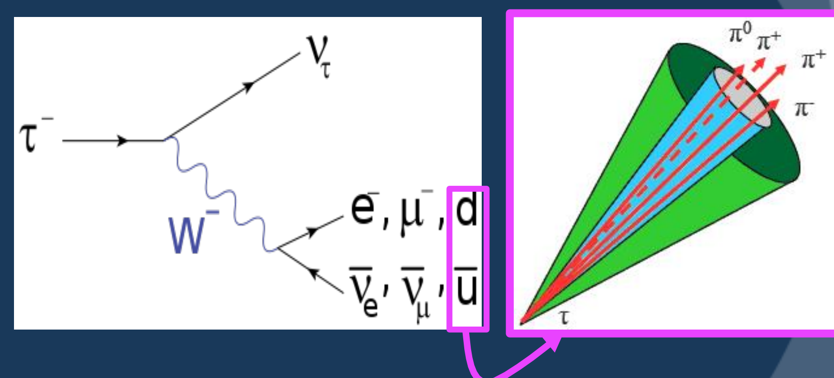
Electroweak (EWK) background:

- W+jets and top production

$$t \to W(l\nu / jets)b \equiv multijet + MET$$

If W decays to $\tau\nu$ and τ decays hadronically (irreducible background)





Physics Background

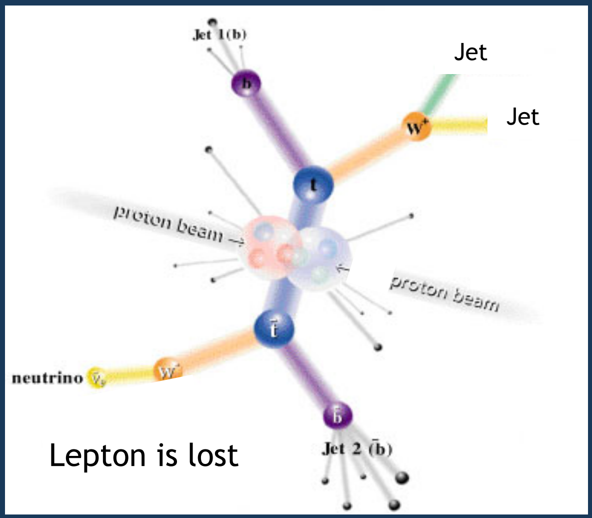
Electroweak (EWK) background:

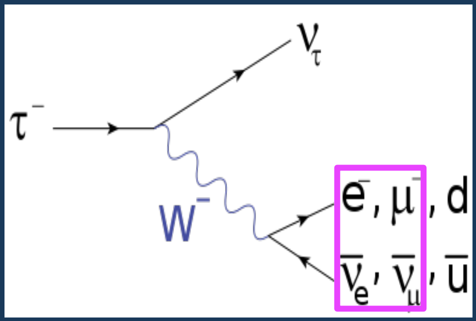
- W+jets and top production

$$t \rightarrow W(l\nu / jets)b \equiv multijet + MET$$

If W decays to $\tau\nu$ and τ decays hadronically (irreducible background)

If W decays hadronically or leptonically and e/μ is “lost” (not detected or reconstructed)





Physics Background

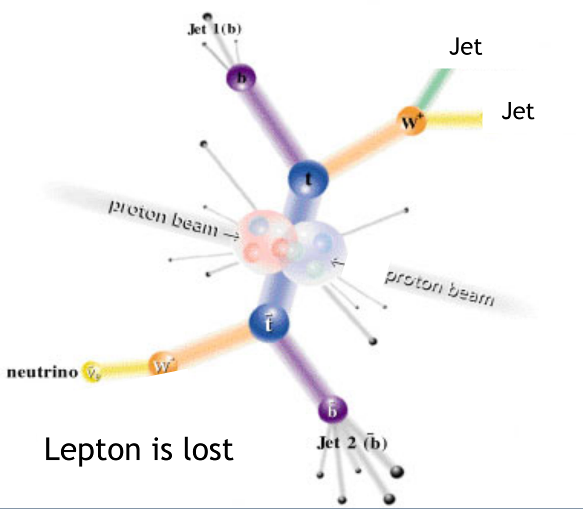
Electroweak (EWK) background:

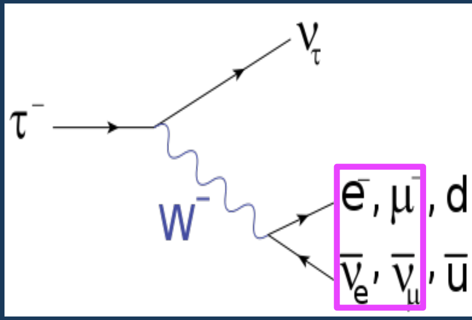
- W+jets and top production

$$t \rightarrow W(l\nu / jets)b \equiv multijet + MET$$

If W decays to $\tau\nu$ and τ decays hadronically (irreducible background)

If W decays hadronically or leptonically and e/μ is "lost" (not detected or reconstructed)





- Z+jets with Z decaying to neutrinos $Z(\nu\bar{\nu}) + jets \equiv multijet + MET$

This background is irreducible: real jets and real MET

Sample Selection

Analysis Strategy:

- Inclusive, model independent search with loose cuts to avoid kinematic bias
- Maximize signal acceptance at the cost of relatively large but well understood, accurately predicted, backgrounds
- HT and MHT are the search sensitive variables

An alternate strategy is to minimize backgrounds at the cost of signal acceptance (example analyses will be discussed later)

Baseline Event Selection:

- Online (trigger) requirement of HT> 100, 140, 150 GeV (no JEC applied)
- At least 3 jets with p_T>50 GeV, $|\eta|<2.5$ ⟵ central production
- HT>300 GeV, MHT>150 GeV [calculated from jets with with p_T>50 GeV, $|\eta|<2.5$]
- $\Delta\phi$(MET, jet[1,2,3])>[0.5,0.5,0.3] ⟵ suppress QCD background
- Isolate electron and muon veto ⟵ reduce W/top background

- Baseline + MHT>250 GeV (generic DM candidate - good bkgd rejection)
- Baseline + HT>500 GeV (heavy particle - long cascade, high multiplicity)

Object ID & Event Cleaning

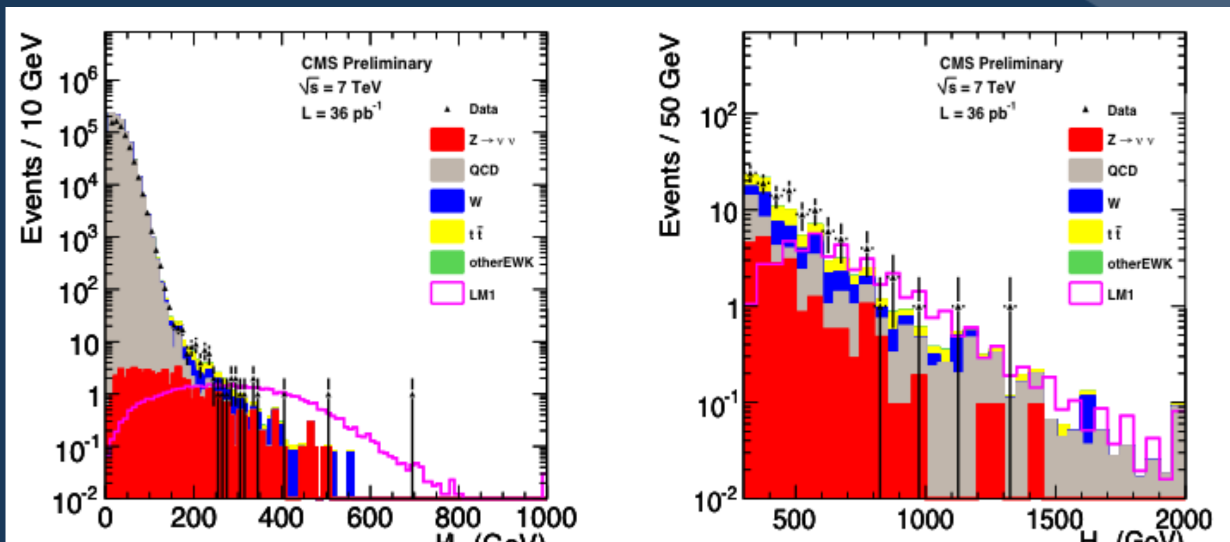
The generic all-hadronic analysis is based on PF physics objects:

- PF jets ID
  - ✓ Anti-kT (D=0.5)
  - ✓ JEC: η-dependence, p_T-dependence from MC truth corrected by data/MC ratio
  - ✓ All visible particles are clustered by PF algorithm
- PF muon and electron ID
  - ✓ p_T>10 GeV, $|\eta|$<2.4 (muons), 2.5 (electrons)
  - ✓ One good quality track matched to primary vertex :d_0<200 mm, d_z<1 cm
  - ✓ Lepton isolation defined as $\left[\sum_{\mathrm{trk}}^{\Delta R=0.3} p_T^{\mathrm{charged\ hadron}} + \sum_{\mathrm{ecal}}^{\Delta R=0.3} p_T^{\mathrm{neutral\ hadron}} + \sum_{\mathrm{hcal}}^{\Delta R=0.3} p_T^{\mathrm{photons}}\right] / p_T$ <0.2

Event cleaning:

- ➢ Require at least one good vertex reconstructed
- ➢ Remove beam related, beamhalo, background events
- ➢ Apply Hcal/Ecal noise filters
- ➢ Reject events where substantial energy was lost in the 1% of Ecal towers masked for reconstruction: check parallel trigger readout path (TP saturation veto), or enforce the energy in neighboring crystals to be < 10 GeV

MHT & HT Distributions



CMS-PAS-SUS-10-005

Observed data & MC background prediction

On left (right), baseline selection applied except for MHT (HT) cuts

LM1 benchmark for illustration

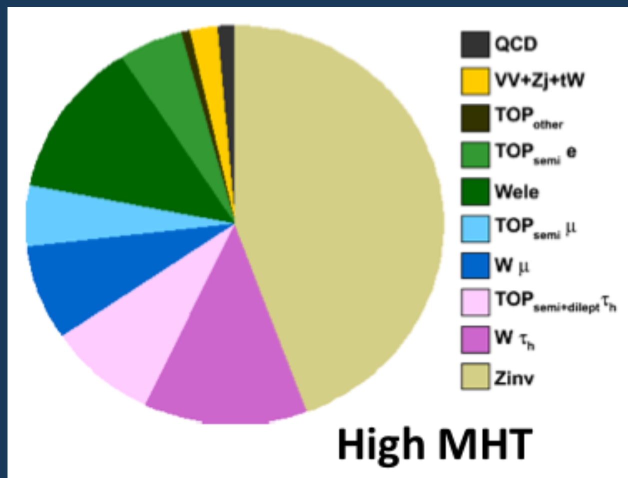
Physics generators not accurate enough
(QCD multijets, W/Z+jets)

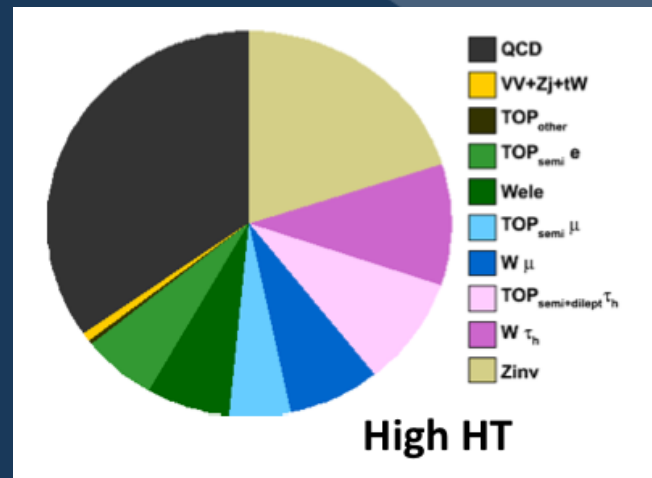
Background predictions extracted from data

	Baseline no $\Delta\phi$ cuts no e/μ veto	Baseline no e/μ veto	Baseline selection	High-$\not\!H_T$ selection	High-H_T selection
Data	482	180	111	15	40
Sum SM MC	406	149	93	14	29
QCD (PYTHIA6)	222.0	27.0	24.6	0.2	9.9
$Z \rightarrow \nu\bar{\nu}$ (MG, $\sigma = 5769$ pb)	26.7	21.1	21.1	6.3	5.7
W (MG, $\sigma = 5760$ pb)	93.9	57.8	23.5	4.7	7.6
$t\bar{t}$ (MG, $\sigma = 165$ pb)	57.5	40.1	21.9	2.6	5.7
WW+WZ+ZZ+tW +Wγ+Zγ+Z/γ^*	6.1	3.4	2.1	0.2	0.2
LM1 (PYTHIA6, $\sigma = 4.9$ pb)	71.2	60.4	45.0	31.3	33.8

Background Predictions

MC prediction of background composition





Z($\nu\nu$) dominates, almost ½
QCD very small

QCD+Z($\nu\nu$) about ½
QCD the largest

Data based predictions of backgrounds are the backbone of the analysis

Data Driven Method for background predictions

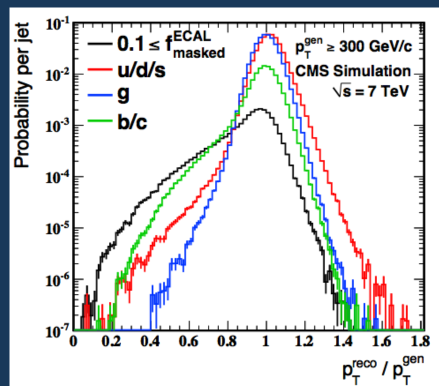
Concept

- Use "control data samples" or "control regions in data"
- Control sample/region: signal depleted sample/region from which to infer the bkgd in the signal region by use of event properties, physics laws, etc
- Signal: area of phase space where the signal is enhanced = search region (good s/b)

Background Predictions Methods

- QCD background
  - Rebalance & Smear (R+S): "unfold" data to particle level (R) and re-smear with measured jet resolutions (S).
  - Factorization: extrapolate two-variable correlation to search region
- W/top background
  - Lost lepton: use inverted lepton veto in a μ+jets control sample
  - Hadronic tau: replace muon by tau response template in a μ+jets control sample
- $Z(\nu\nu)$ background
  - From γ+jets: remove photon and scale by $Z(\nu\nu)$+jets/γ+jets ratio. High stats but non-trivial theory correspondence
  - From W+jets: remove lepton and scale by $Z(\nu\nu)$+jets/$W(l\nu)$+jets ratio. Less stats but easier theory correspondence
  - From Z+jets: remove leptons and scale by $Z(\nu\nu)$+jets/$Z(\mu\mu)$+jets ratio. Straight forward correction but limited yield

QCD Background: smearing effect

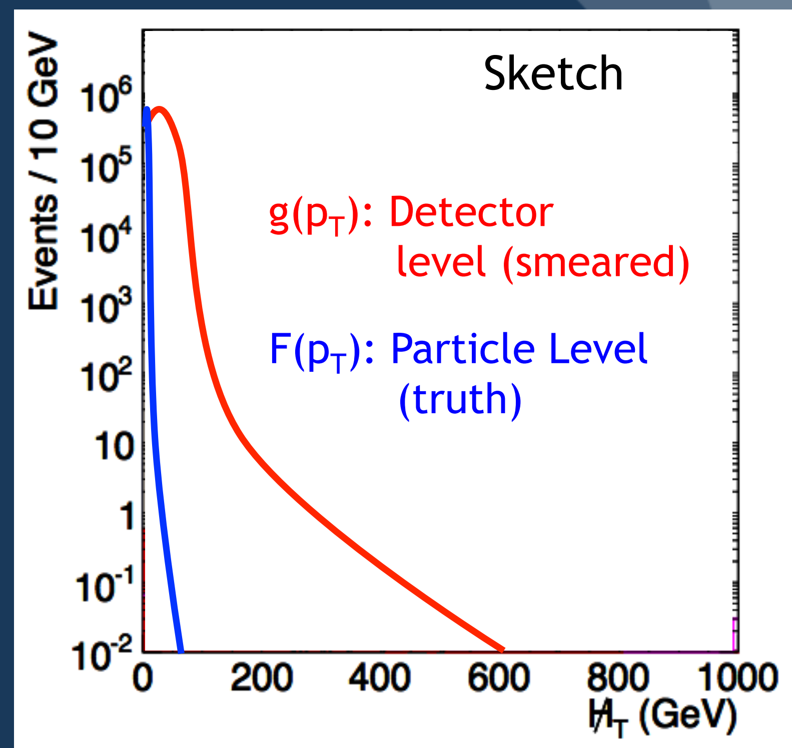
+ 

CMS-PAS-SUS-10-005

=

True distribution "smeared" due to the finite detector energy resolution

$$g^{smeared}(p_T^{meas}) = \int_0^\infty F^{true}(p_T^{true}) R(p_T^{meas}, p_T^{true}) \, dp_T^{true}$$



Jets that fluctuate to high/low response create spurious MHT tail

QCD Background: R+S concept

- Rebalance

  Jet particle level p_T restored from detector level inclusive multi-jet data sample by maximum likelihood using:

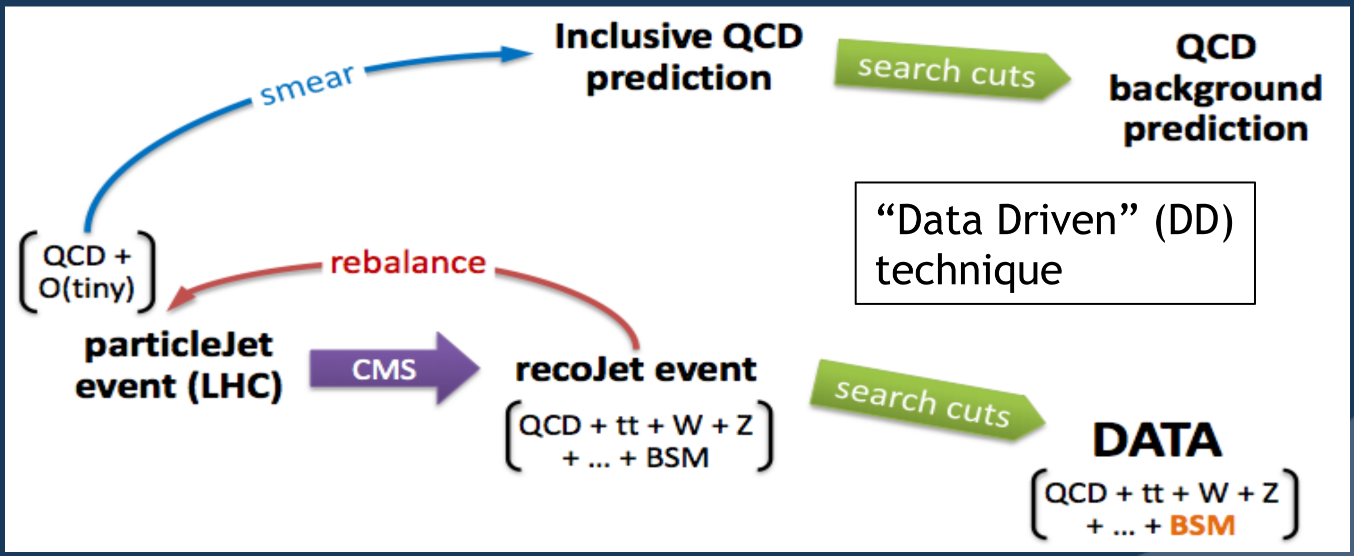
  ✓ Measured jet p_T response probability density functions

  ✓ Transverse momentum conservation $\sum_{i=1}^{n} \vec{p}_{T,i}^{\,true} + \vec{p}_{T,soft}^{\,true} = 0$

  ✓ Events with real MET are turned to QCD multi-jet events automatically

- Smear

  Rebalanced distribution is smeared by the measured jet pT resolution functions including the tails



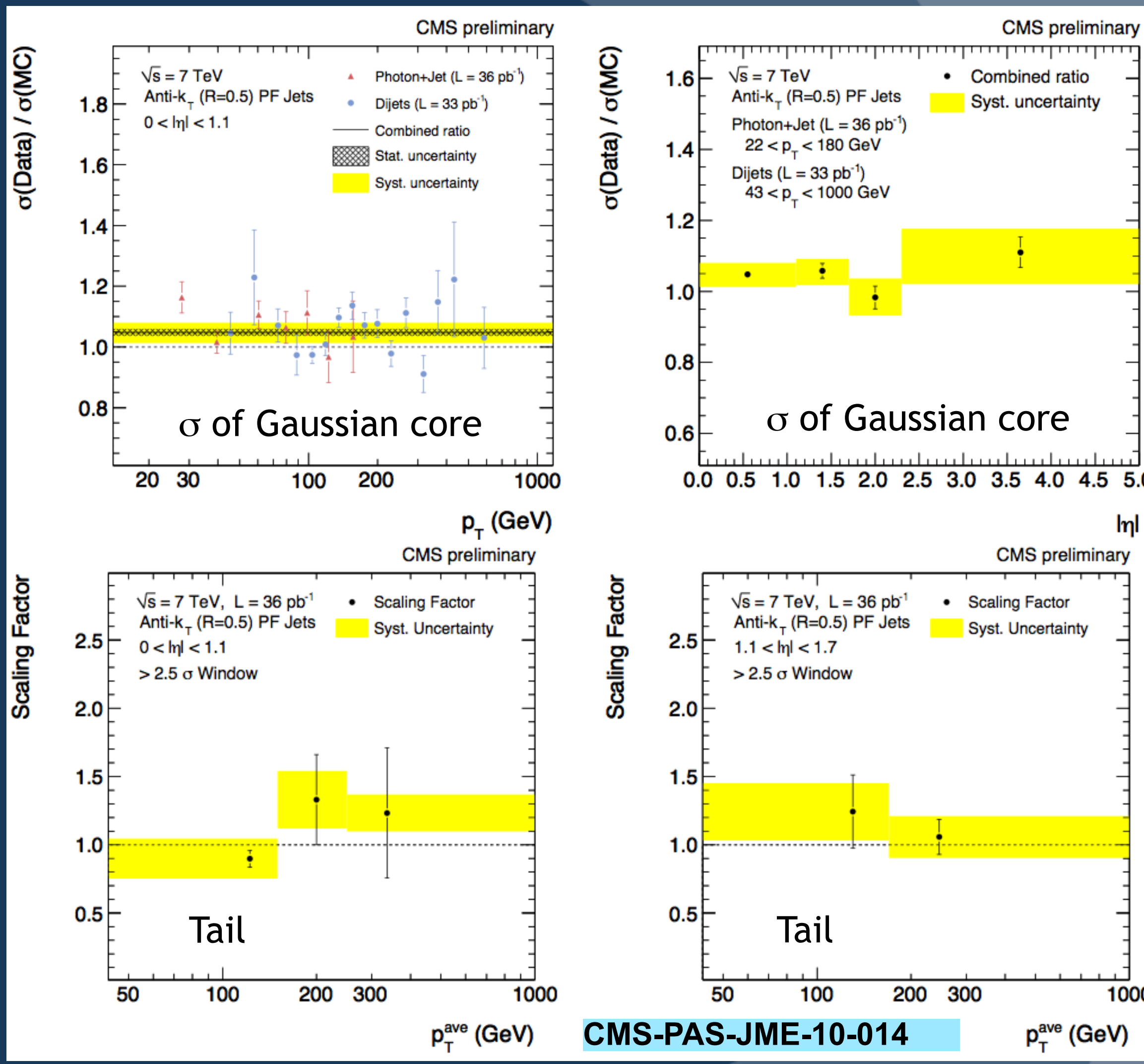
QCD Background: R+S ingredients

Jet p_T resolution functions are the main ingredient to R+S

Measured from data using object p_T balance (see lecture on physics objects)

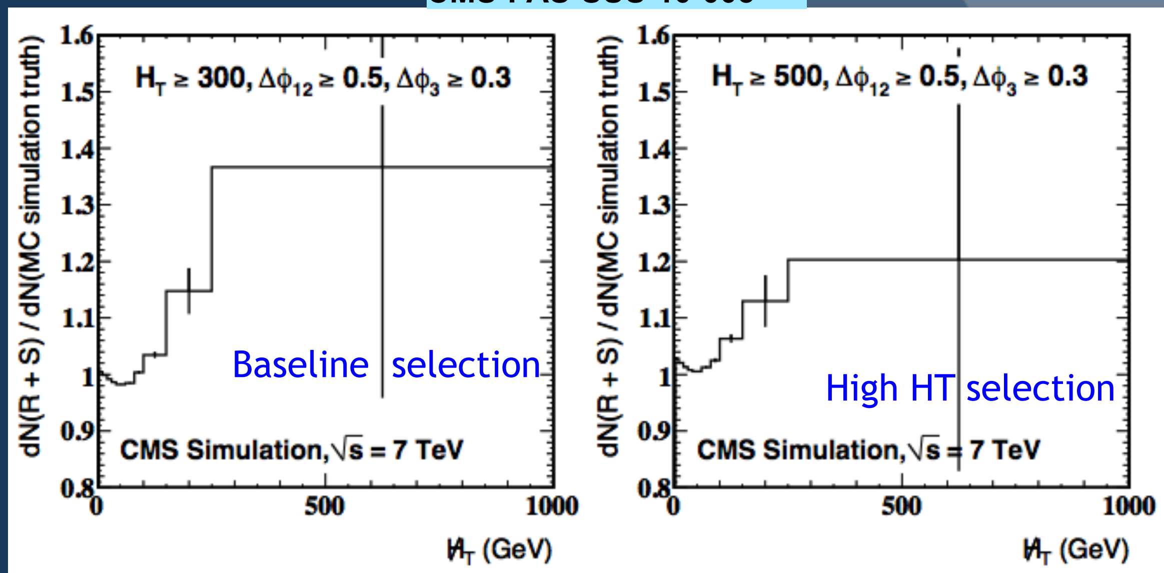
For the Gaussian core and tails the data/MC ratio was measured

MC truth resolution functions * (data/MC) were used in R+S



QCD Background: R+S closure

CMS-PAS-SUS-10-005



MC closure test of the method:

- Ratio of MC (R+S) predicted MHT (treated as data) to MC detector level MHT

Closure Test

Concept

- Using MC: evaluates the validity and accuracy of a method by comparing the "measured prediction" with the "truth" information (e.g. above)
- Using data: idem by comparing the measured prediction to the straight detector level distribution in a control region

  (e.g. R+S distribution compared to observed MHT in a signal depleted region)

QCD Background: R+S results

QCD background prediction:

- Uncertainty components

0.16 ± 0.10 *High MHT*
16.0 ± 7.9 *High HT*

(0.2 and 9.9 in MC)

CMS-PAS-SUS-10-005

QCD background		Baseline selection	high-$\not{H}_T$ selection	high-H_T selection
Nominal prediction		39.4	0.18	19.0
GenJet smearing closure	(box)	+14%	+30%	+7%
Rebalancing bias	(box)	+10%	+10%	+10%
Soft component estimator	(box)	+3%	+19%	+4%
Resolution core	(asymmetric)	+14% −25%	+0% −52%	+15% −21%
Resolution tail	(asymmetric)	+43% −33%	+56% −78%	+48% −34%
Flavour trend	(symmetric)	±1%	±12%	±0.3%
Control sample trigger	(box)	−5%	−5%	−5%
Search trigger	(symmetric)	±1%	±1%	0%
Lepton veto	(box)	±5%	±0.05%	±0.2%
Pile-up effects	(box)	±2%	±10%	±2%
Seed sample statistics	(symmetric)	±2.3%	±23%	±3.3%
Total uncertainty		51%	64%	49%
Bias-corrected prediction		29.7 ± 15.2	0.16 ± 0.10	16.0 ± 7.9

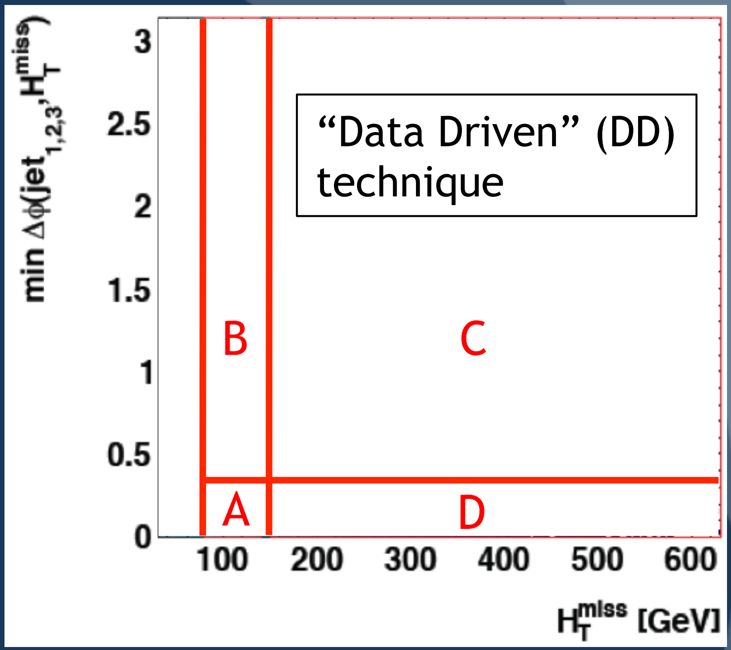
QCD Background: factorization

- A, B, D are background dominated regions
- C is the signal region

min $\Delta\phi$ (jet,MHT)>0.3, MHT>150 GeV

If variables uncorrelated:

$$N_C = N_B / N_A * N_D$$



If variables are correlated and $r(MHT)=N_B/N_A$ is understood :

$$N_C = r(MHT) * N_D$$

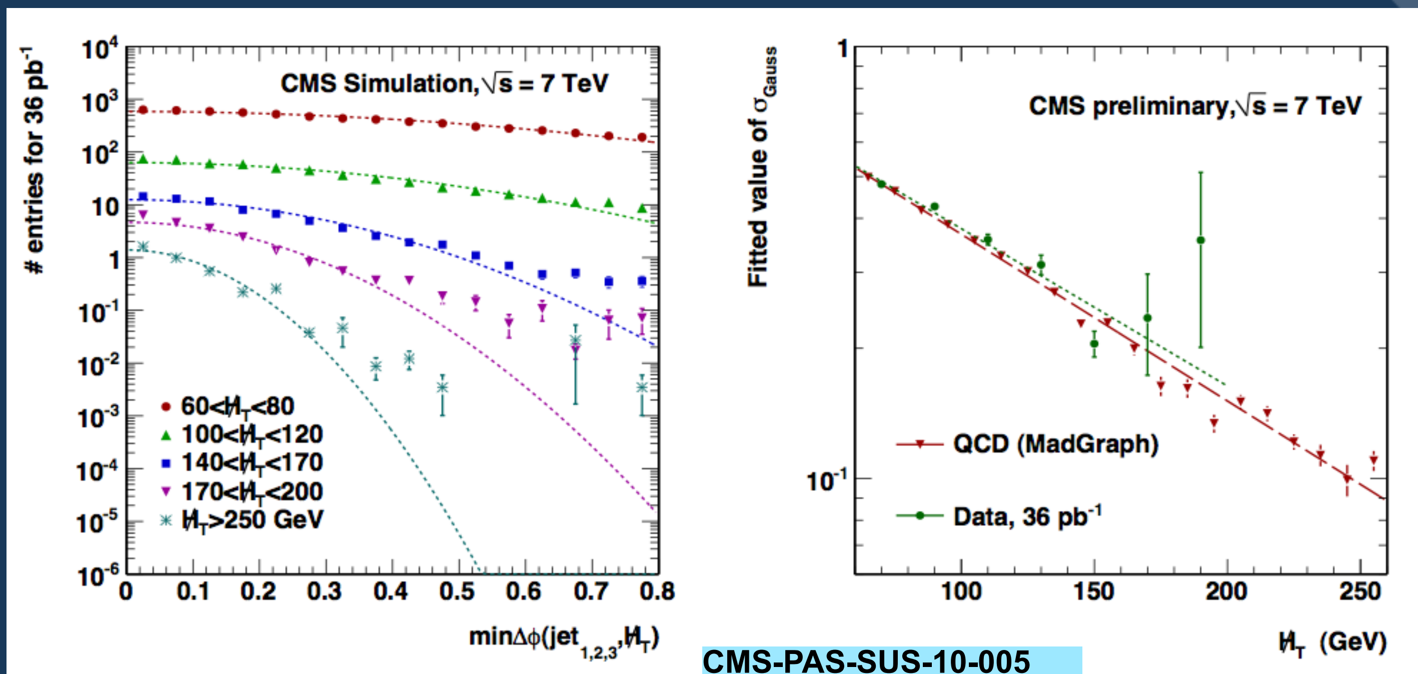
with r(MHT) extrapolated to the signal region

QCD Background: factorization

r(MHT) dependence determined empirically

- Gaussian fit to min ΔΦ(jet,MHT): $r(\not{H}_T) = \frac{1}{\mathrm{erf}\left(\frac{\Delta\phi^{cut}_{min}}{\sqrt{2}\cdot\sigma_{Gauss}(\not{H}_T)}\right)} - 1$ + C

  C taken from MC
- Exponential fit: $r(MHT) = a + \exp(-b/MHT) + c$



min ΔΦ resolution better as MHT increases (more likely single mis-measured jet), non-Gaussian tails more prominent → C constant added

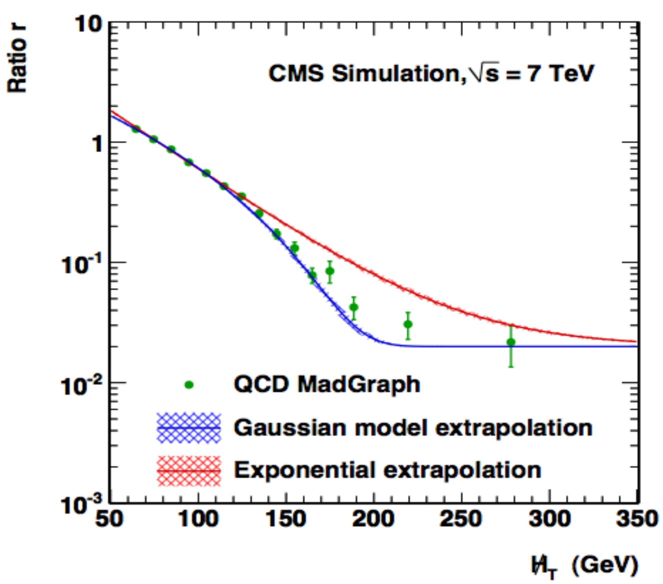
QCD Background: factorization

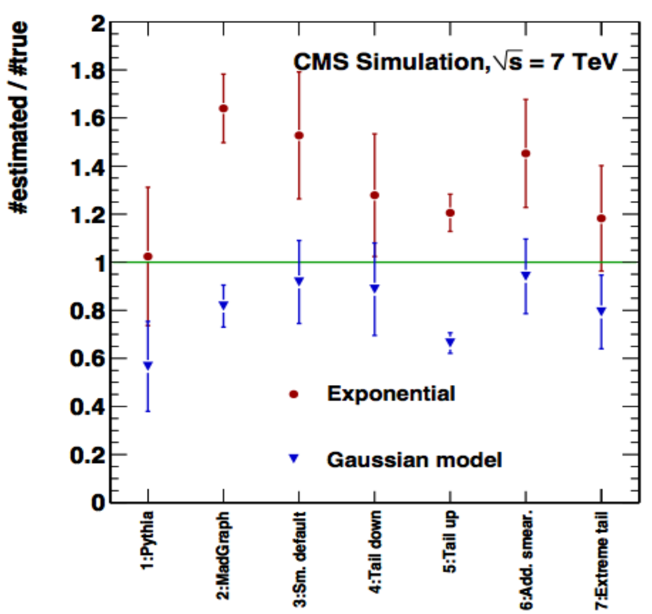
Expo and Gauss models bracket the true # of QCD events

- Gaussian underestimates, exponential overestimates

Left: MC closure test

Right: uncertainties





CMS-PAS-SUS-10-005

Method	Baseline selection	(stat.)	(syst.)	High-$H\!\!\!/_T$ selection			High-H_T selection		
Gaussian model prediction	19.0	±1.6	$^{+7.2}_{-6.5}$	0.3	±0.1	$^{+0.4}_{-0.2}$	13.0	±1.3	$^{+4.9}_{-4.4}$
Exponential prediction	31.4	±2.4	$^{+7.0}_{-6.9}$	0.5	±0.1	$^{+0.2}_{-0.2}$	21.6	±2.0	$^{+4.8}_{-4.8}$
Combined	25.2	±2.4	$^{+13.2}_{-13.1}$	0.4	±0.1	$^{+0.3}_{-0.3}$	17.3	±2.0	$^{+9.1}_{-9.2}$

W/top Background: lost lepton

Lepton veto not fully efficient rejecting W/top background. Lepton is "lost" and the event not rejected if:

- ✓ Not reconstructed
- ✓ Not Isolated
- ✓ Out of detector acceptance

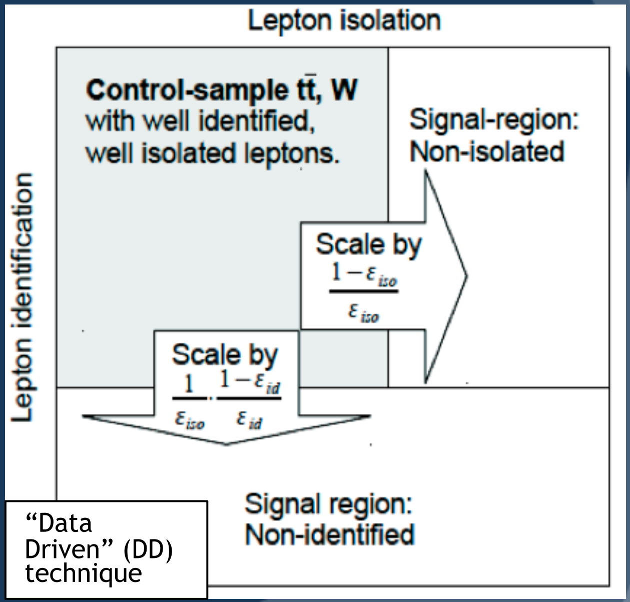
Pythia prediction for events with lost leptons passing lepton veto

36 pb^{-1}	ttbar		W+jets	
Baseline selection	**electron**	**muon**	**electron**	**muon**
Not reconstructed	1.5	0.4	0.4	0.1
Not isolated	3.2	3.8	0.6	0.6
Out of acceptance	5.5	4.8	2.1	1.9
total	10.2	9.0	3.1	2.6

Invert lepton veto technique on μ+jets control sample

(97% of events are ttbar or W+jets)

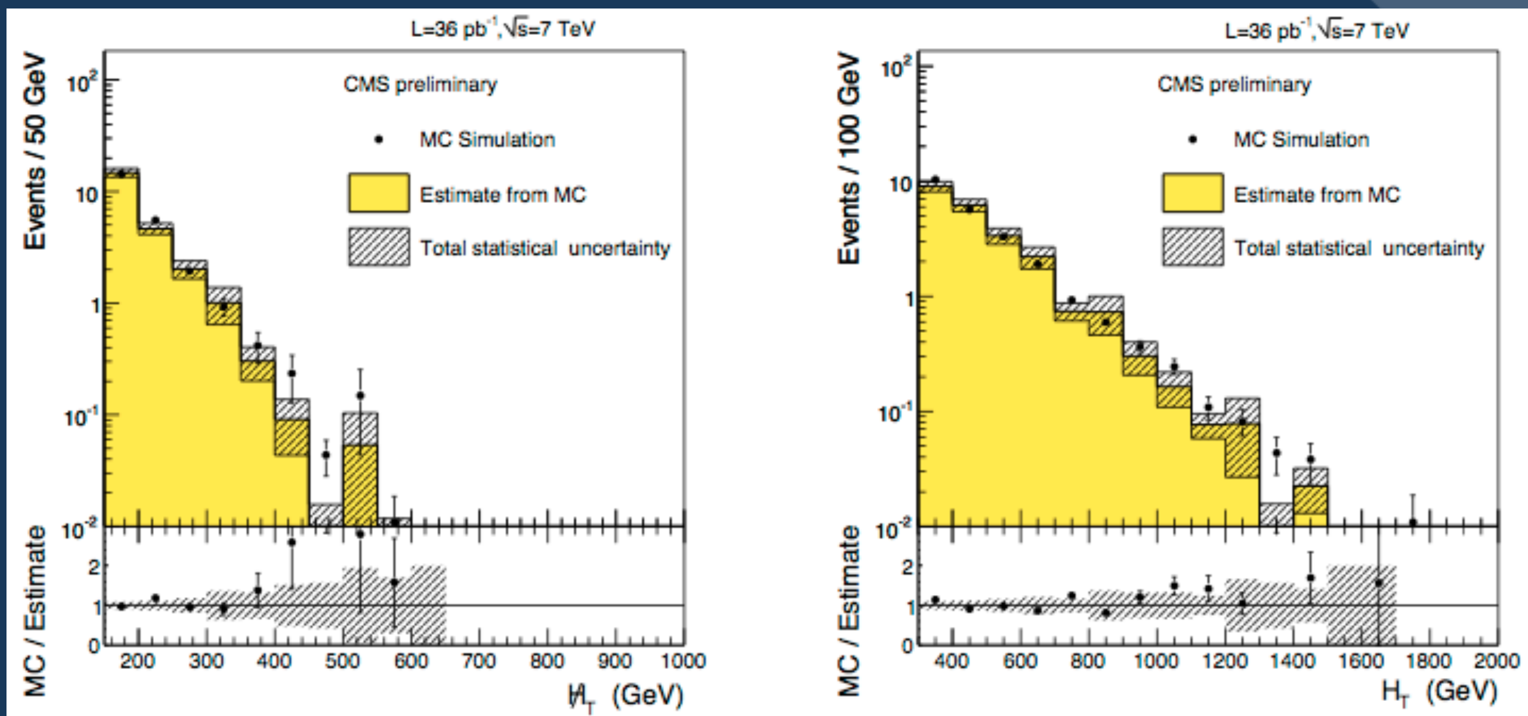
- ✓ All cuts but require one iso muon
- ✓ Events scaled by → $$\frac{1}{\varepsilon_{iso}} \frac{1-\varepsilon_{id}}{\varepsilon_{id}}$$
- ✓ ε_{iso} parameterized in p_T, $\Delta R(l,jet)$ from Z using tag and probe
- ✓ ε_{id} parameterized in p_T, η also using tag and probe
- ✓ Residual corrections (<10%) applied for differences between Z and W/top kinematics



W/top Background: lost lepton

Closure test using ttbar and W+jets MC

- ✓ Simulation (truth) and estimate (prediction) agree within stat errors
- ✓ Systematic uncertainties ~10% not included



W/top Background: lost lepton

For baseline selection, 33 events predicted, 40% more than Pythia & Madgraph

- ✓ Statistical error dominates
- ✓ Different kinematics in control and signal region, background contamination in control sample (QCD, Z, di-boson)

Method	Baseline selection	(stat.)	(syst.)	High-$\not\!\!H_T$ selection			High-H_T selection		
Estimate from data	33.0	± 5.5	$^{+6.0}_{-5.7}$	4.8	± 1.8	$^{+0.8}_{-0.6}$	10.9	± 3.0	$^{+1.7}_{-1.7}$
Estimate (PYTHIA)	22.9	± 1.3	$^{+2.7}_{-2.6}$	3.2	± 0.4	$^{+0.5}_{-0.5}$	7.2	± 0.7	$^{+1.1}_{-1.1}$
MC Truth (PYTHIA)	23.6	± 1.0		3.6	± 0.3		7.8	± 0.5	
Estimate (MADGRAPH)	20.4	± 1.5	$^{+2.6}_{-2.5}$	2.4	± 0.3	$^{+0.3}_{-0.3}$	4.8	± 0.4	$^{+0.6}_{-0.5}$
MC Truth (MADGRAPH)	21.4	± 0.7		3.0	± 0.3		5.9	± 0.4	

	Relative size		# events	
Statistics of control-sample	−17%	+17%	−5.5	+5.5
Iso- & id- efficiencies (statistical)	−13%	+14%	−4.1	+4.7
Kinematic differences $t\bar{t}$, W, Z-samples	−10%	+10%	−3.3	+3.3
SM background in control-region	−3%	+0%	−1.0	+0
MC use for acceptance calculation	−5%	+5%	−1.7	+1.7
Total Uncertainty (Baseline Selection)	−24%	+25%	−7.9	+8.1

W/top Background: hadronic τ

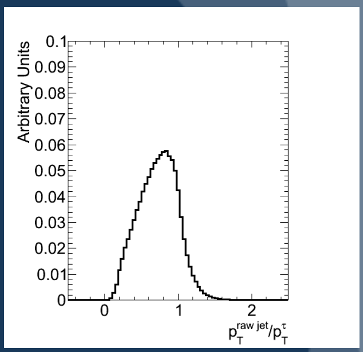
Hadronic τ method combined with lost lepton method to predict total W/top background

- ✓ Lost lepton: W/ttbar → e, μ + X
- ✓ Hadronic τ: W/ttbar → τ_{had}+ X

“Data Driven” (DD) technique

Use μ+jets control sample, p_T^{μ}> 20 GeV, $|\eta|$<2.1, muon ID & ISO

- ✓ Muons replaced by τ-jets
- ✓ τ-jet momentum obtained from simulated template of p_T^{jet}/p_T^{τ}
- ✓ Recalculate HT, MHT
- ✓ correct for muon trigger, acceptance, reco & iso efficienty



	Predicted $W/t\bar{t} \to \tau_{hadr}$
Baseline selection	22.3 ±4.0 (stat.) ±2.2 (syst.)
High-$\not{H}_T$ selection	6.7 ±2.1 (stat.) ±0.5 (syst.)
High-H_T selection	8.5 ±2.5 (stat.) ±0.7 (syst.)

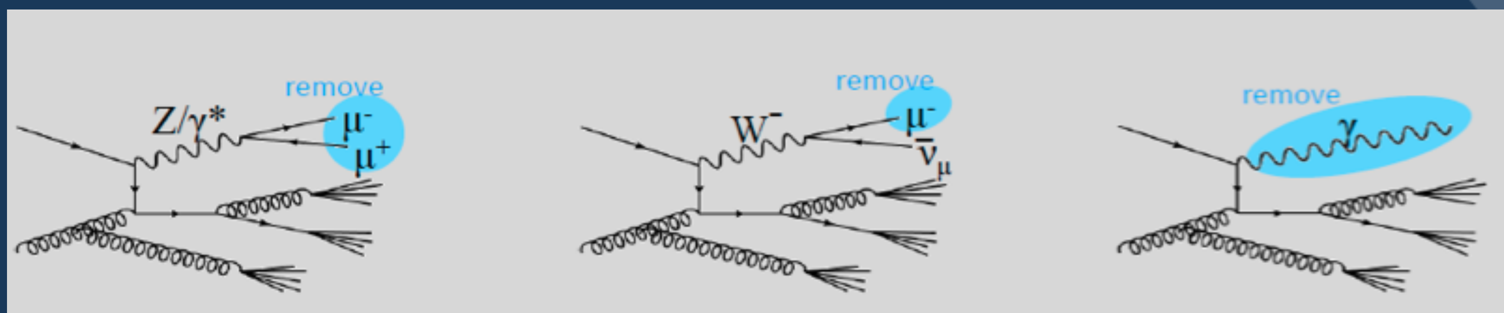
	Baseline selection	High-$\not{H}_T$ selection	High-H_T selection
τ response template	2%	2%	2%
Acceptance	+6%,-5%	+6%,-5%	+6%,-5%
Muon efficiency on data	1%	1%	1%
SM backgr. subtraction	5%	5%	5%

Z(νν) Background

Z(νν) background is ~1/2 and ~1/6 of the total in the high MHT and high HT searches respectively

Three independent data driven methods are used based on Boson substitution with MHT

Z(ll)+jets | W(lν)+jets | γ+jets



Z(ll)+jets
- Same kinematics
- Trivial Br correction

  $Br(Z \to \mu\bar{\mu})/Br(Z \to \nu\bar{\nu}) = 1/6$
- Lower stats than γ/W+jets

W(lν)+jets
- Similar kinematics
- Large backgrounds
- More stats than Z(νν) and 2.5 more than Z(μμ)

γ+jets
- Similar kinematics as Z+jets at high p_T and MHT
- Large and complex theory corrections
- High statistics

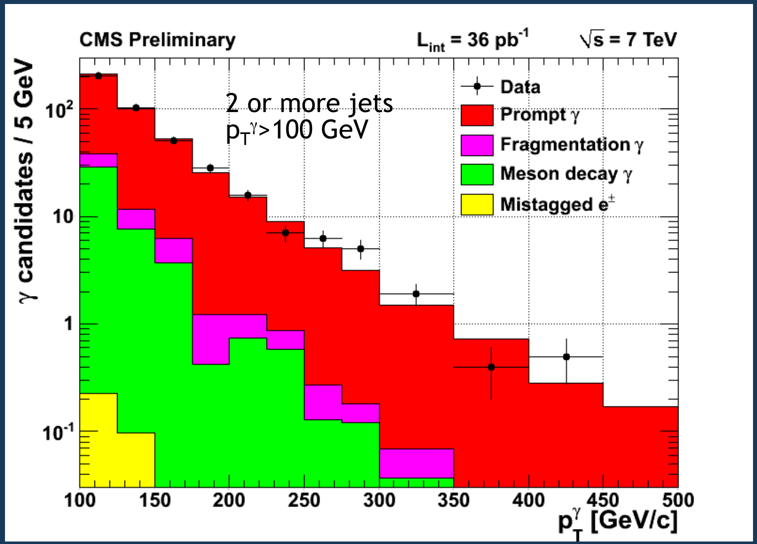
γ+jets prediction is used for the limit, Z/W+jets are cross checks

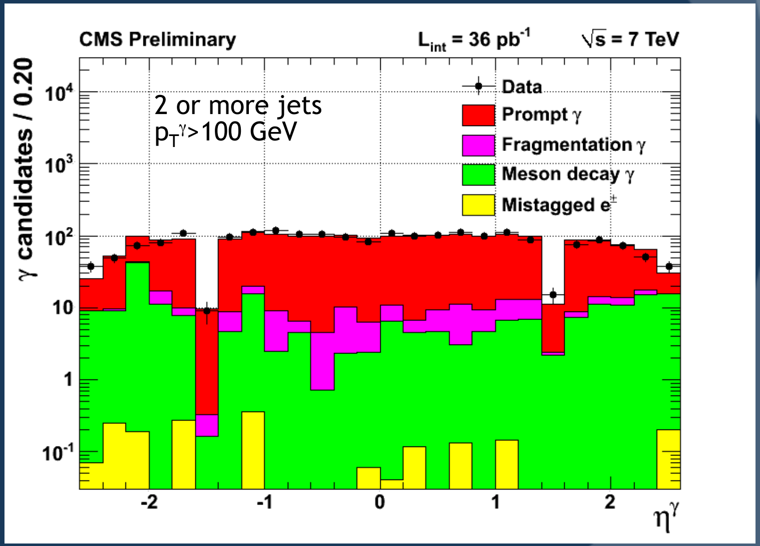
Z(νν) Background: γ+jets sample

Single photon trigger and standard cuts to select isolated photons

Photon categories

- Direct: well isolated photon from hard scatter ← selected for estimate
- Fragmented: from parton shower, non-isolated, reconstructed inside a jet
- Decay: from π, η mesons



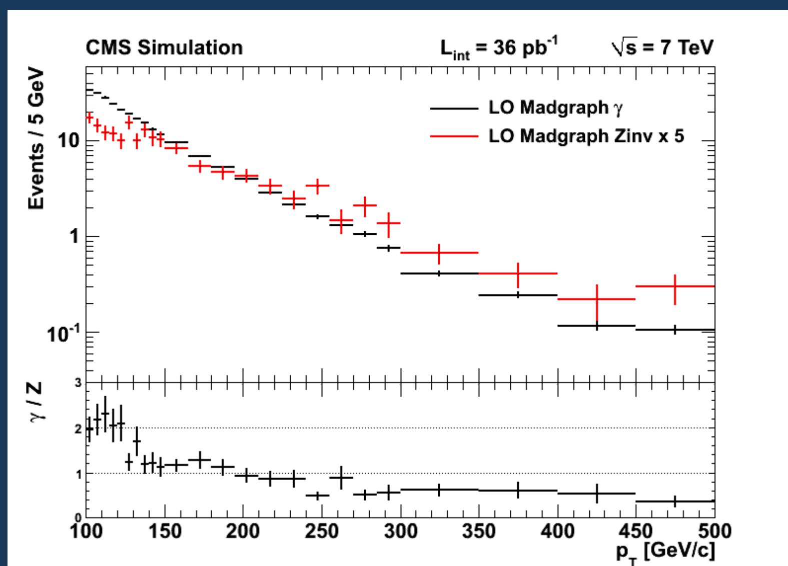


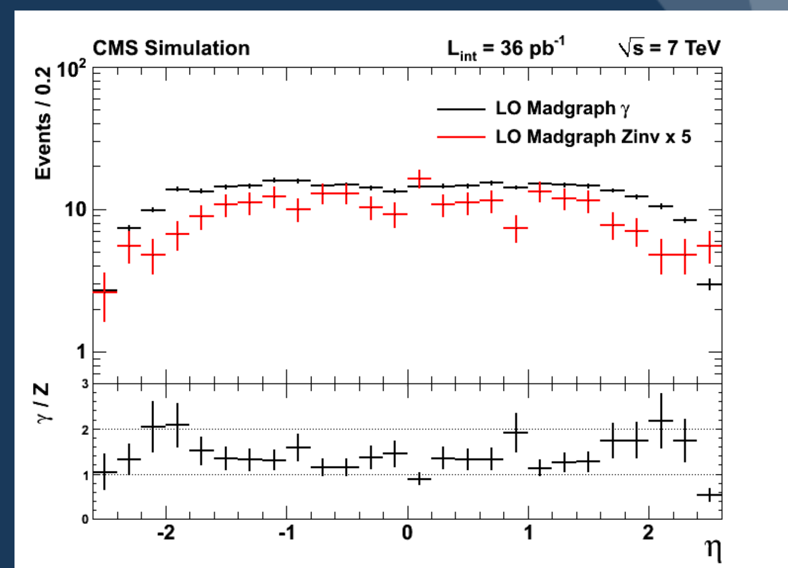
MC: Madgraph LO + detector simulation (normalized)
Excellent description of prompt photons, backgrounds

Z(νν) Background: γ+jet procedure

At p_T > 200 GeV, γ and Z spectra is similar but not the same due to the different couplings

- Background subtracted from photon sample after isolation: fragmentation photons are 5% (NLO JetPHOX), photon pairs from mesons
- LO γ+jets/LO Z+jets is computed and a correction obtained for each of the two search selections
- Detector acceptance correction folded into the γ-Z correspondence





MC: Madgraph LO + detector simulation (normalized by factor 5)

Z(νν) Background: γ+jet results

Correction factors:

		Baseline selection		High-$\not{H}_T$ selection		High-H_T selection	
Z/γ correction	±theory	0.41	±6 %	0.48	±6 %	0.44	±4 %
	±acceptance		±5 %		±5 %		±5 %
	±MC stat		±7 %		±13 %		±13 %
Fragmentation		0.95	±1 %	0.95	±1 %	0.95	±1 %
Secondary photons		0.94	±9 %	0.97	±10 %	0.90	±9 %
Photon mistag		1.00	±1 %	1.00	±1 %	1.00	±1 %
ID data/MC ratio		1.01	±2 %	1.01	±2 %	1.01	±2 %
Total correction		0.37	±14 %	0.45	±18 %	0.38	±17 %

← Uncertainty from BlackHat Collaboration

Background predictions:

	# events in γ+jets data sample	# $Z \to \nu\bar{\nu}$ events predicted	# $Z \to \nu\bar{\nu}$ events from simulation
Baseline selection	72	26.3 ± 3.2(stat.) ± 3.6(syst.)	21.2 ± 1.4
High-$\not{H}_T$ selection	16	7.1 ± 1.8(stat.) ± 1.3(syst.)	6.3 ± 0.8
High-H_T selection	22	8.4 ± 1.8(stat.) ± 1.4(syst.)	5.8 ± 0.7

This prediction is (the only) Z(νν) used in the limit calculation

Z(νν) Background: Z/W+jets

From W+jets:

- QCD background to W "signal": invert lepton isolation (veto) and normalize to signal (W) depleted region (low MET)
- ttbar background to W "signal": apply b-tag veto and estimate residual from ttbar enriched sample (tight b-tag)
- Z(ll) and W(τν) background to W "signal" is taken from MC

$$N(Z \to \nu\nu) = \frac{N_W^{obs} - N_W^{bkg}}{A_W \cdot \varepsilon_W \cdot \varepsilon_{b-veto} \cdot L} \cdot R\left(\frac{Z \to \nu\nu}{W \to l\nu}\right)$$

$$\varepsilon_W = \varepsilon_{Iso} \cdot \varepsilon_{RECO} \cdot \varepsilon_{HLT}$$

$$R\left(\frac{Z \to \nu\nu}{W \to l\nu}\right) \approx \frac{6}{10}$$

Very large uncertainty
Not included in analysis
at the moment

From Z+jets:

- Background to Z "signal" small and ignored

$$N(Z \to \nu\nu) = \frac{N_Z^{obs} - N_Z^{bkg}}{A_Z \cdot \varepsilon_Z \cdot L} \cdot R\left(\frac{Z \to \nu\nu}{Z \to ll}\right)$$

$$R\left(\frac{Z \to \nu\nu}{Z \to ll}\right) = 5.95 \pm 0.02$$

$$\varepsilon_{lepton} = \varepsilon_{Iso} \cdot \varepsilon_{RECO} \cdot \varepsilon_{trig}$$

$$\varepsilon_Z = (\varepsilon_{leptopn})^2 \cdot \varepsilon_{trig},$$

where $\varepsilon_{trig} = 1 - (1 - \varepsilon_{HLT})^2$

Baseline selection

$Z \to e^+e^-: \quad 32 \pm^{29}_{18}$

$Z \to \mu^+\mu^-: \quad 12 \pm^{16}_{8}$

$Combined: \quad 17 \pm 8$

No event survives search selection

From 2 and 1 events !
Not included in limit (Cross check)

Results

No excess of events is observed in either the high-MHT or high-HT search regions for 36 pb^{-1}

Background	Baseline selection		High-$\not{H}_T$ selection		High-H_T selection	
$Z \rightarrow \nu\bar{\nu}$ (γ+jets method)	26.3	±4.8	7.1	±2.2	8.4	±2.3
$W/t\bar{t} \rightarrow e, \mu$+X	33.0	±8.1	4.8	±1.9	10.9	±3.4
$W/t\bar{t} \rightarrow \tau_{hadr}$+X	22.3	±4.6	6.7	±2.1	8.5	±2.5
QCD (R+S method)	29.7	±15.2	0.16	±0.10	16.0	±7.9
Total background estimated from data	111.3	±18.5	18.8	±3.5	43.8	±9.2
Observed in 36 pb^{-1} of data	111		15		40	
95% C.L. limit on signal events	40.4		9.6		19.6	

At the 95% C.L. the data is consistent with no more than 9.6 (19.6) signal events for the high-MHT(HT) search regions

- If I repeat the experiment N→∞ times, 95% of the times the background will fluctuate to accommodate zero to no more than 9.6 (19.6) signal events

No Excess Means … Limits

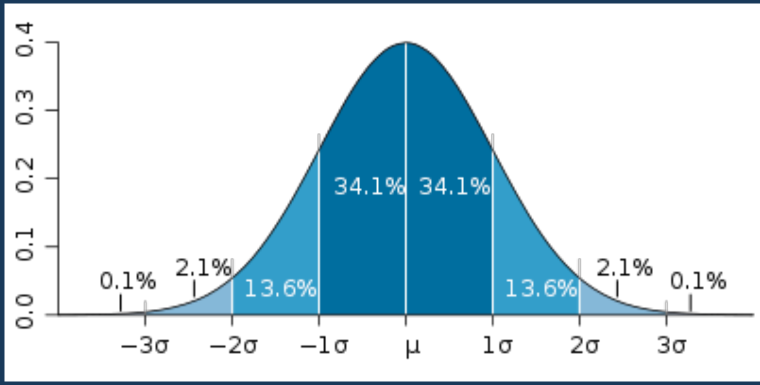
Confidence Intervals (C.I.)

A confidence interval gives an estimated range of values which is likely to include the unknown true value μ of a population parameter

$$\hat{\mu} = \langle X \rangle = \frac{1}{N}\sum_{i=1}^{n} X_i$$

The estimator of the true parameter value $\hat{\mu}$ is calculated as the mean value $\langle X \rangle$ in a given data sample

I repeat the experiment N (e.g. 100) times, each experiment generating M (e.g. 1000) values of X



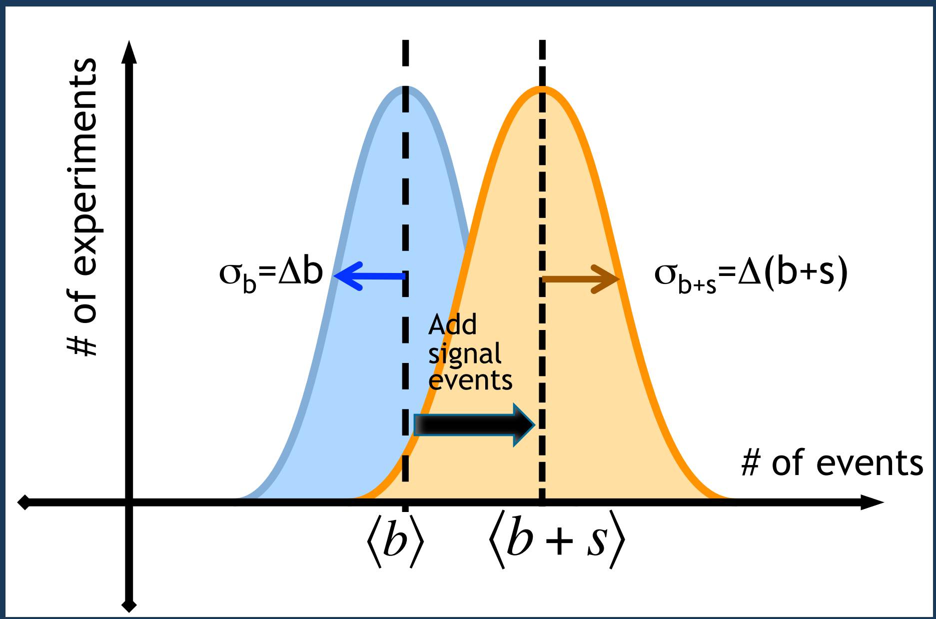
Central C.I. for Normal Distribution
1σ → 68.27%
2σ → 95.45%
3σ → 99.75%
5σ → 99.99994%

The "level" of a confidence interval (C.L. 90%, 95%, 99%, …) refers to the number of times (n/N*100 experiments) the interval will contain the true value

Expected Limit

- Generate ensemble of N experiments using the measured <b>+Δb distribution (<b> is mean of a Poisson, Δb is Gaussian)

- Question: how many signal events (s) can I add so that the b+s C.I. includes the background only prediction, <b>, 95% of the times?



Expected Limit on signal at the 95% C.L.

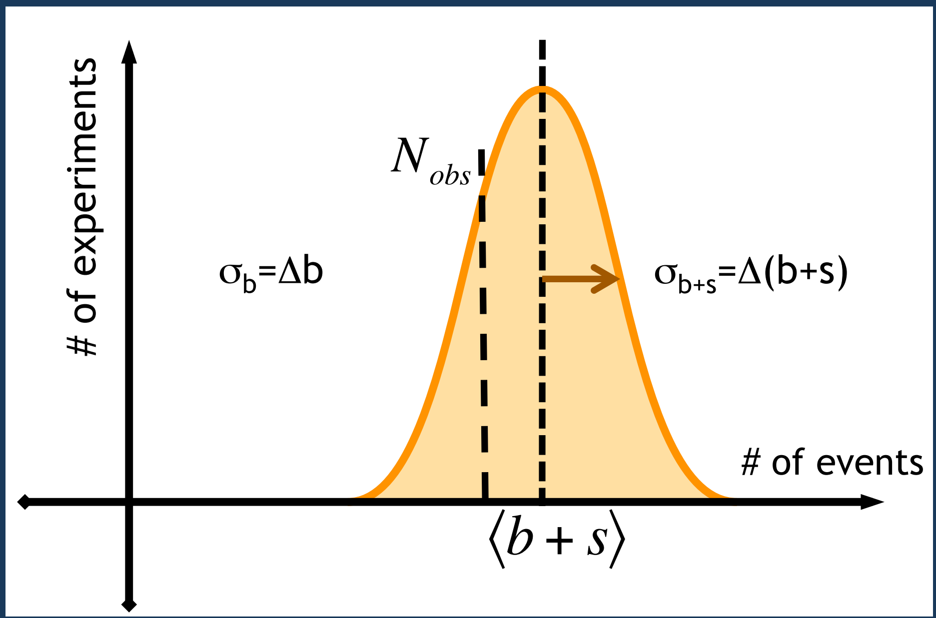
- maximum # of signal events the sample may contain consistent with <b>
- Signal events generated as explained later
- Limit translated to production x-section or masses

  (theory models and signal acceptance/efficiency)

Observed Limit

- Generate ensemble of N experiments using the measured <b>+Δb distribution (signal contamination subtracted ~3 evts.)

- Question: how many signal events (s) can I add so that the b+s C.I. includes the # of observed events, N_{obs}, 95% of the times?



Observed Limit on signal at the 95% C.L.

- maximum # of signal events the sample may contain consistent with N_{obs}
- Signal events generated as explained later
- Limit translated to production x-section or masses

  (theory models and signal acceptance/efficiency)

Comments on Limits

- Expected Limit is expressed as a band consistent with <b>±Δb
- If N_{obs} is greater than <b>, the observed limit is less than the expected
  - ✓ Small excess not "significant", most probably occurred by chance
- If N_{obs} is less than <b>, the observed limit is greater than the expected
  - ✓ Deficit means that bkgnd fluctuated low
- Zero background hypothesis is the most conservative for setting a limit
  - ✓ Lowest limit
- Zero background hypothesis is the least conservative for a discovery
  - ✓ Largest probability of wrongly accepting the signal hypothesis

Statistical Tests for Limits

CMS uses the Modified Frequentist Procedure (CL_s)

- ✓ Avoids excluding or discovering signals, that the analysis is not really sensitive to.
- ✓ Reduce dependency on uncertainty from background

CMS also uses Bayesian Framework (flat prior for the signal)

- ✓ Frequentist probability is the limit of a frequency
- ✓ Bayesian probability is a subjective degree of believe (The prior is the probability of a theory)

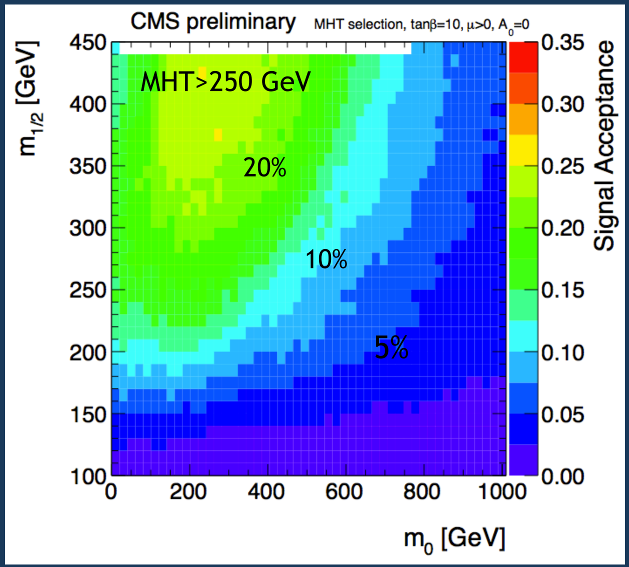
ATLAS uses Power Constraints Limits (PCL)

- ✓ Tends to give better (higher) limits for downward fluctuations in data
- ✓ ATLAS also used CL_s to allow comparison with CMS

Signal Acceptance/Efficiency

The expected number of signal events for a given model and event selection is estimated from simulated signal samples (generation + detector simulation)

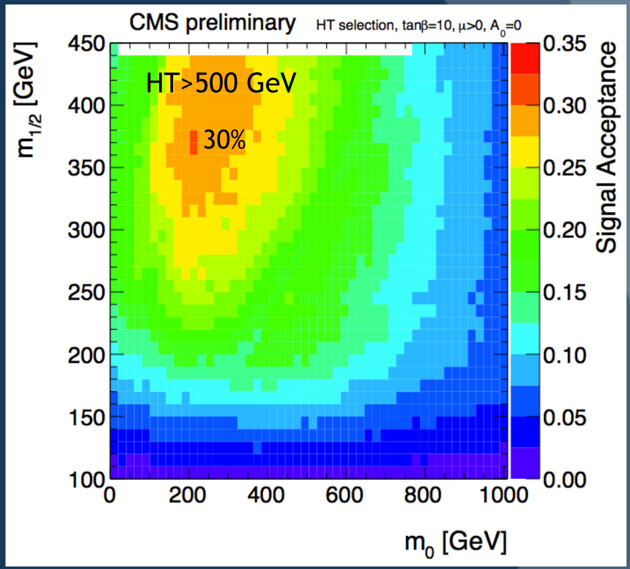
- Experimental and theoretical uncertainties from event selection, reconstruction, calibration
- Theoretical uncertainties related to event generation
- Overall luminosity uncertainty



Signal A_{cc} x E_{ff}

Acceptance (Acc): fraction of events passing the topology & kinematics requirement

Efficiency (Eff): Fraction of "accepted" events that were triggered, reconstructed, identified



Signal Uncertainties:

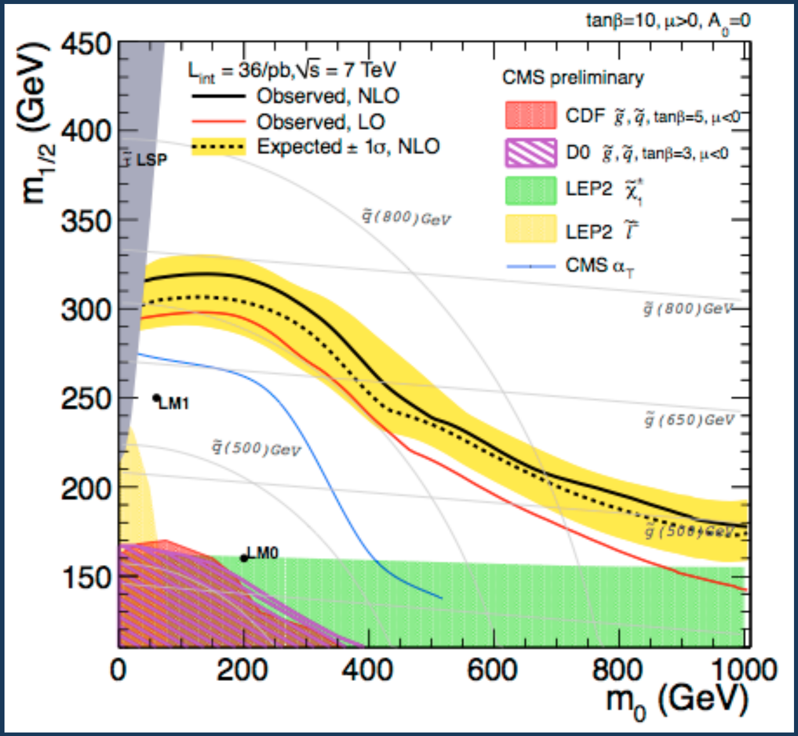
JEC and JER (8%), lepton veto/trigger efficiency (1%), dead Ecal filter inefficiency (1.5%), luminosity (4%), $\mu_{R,F}$ in NLO signal calculation (16%), PDFs (3%)

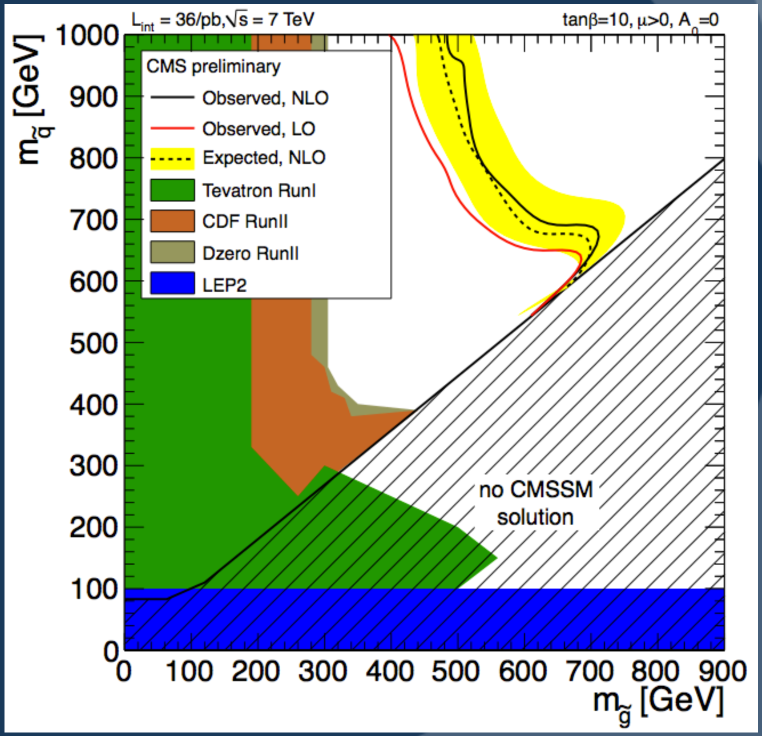
Interpretation within the CMSSM

The contours are the envelope with respect to the best sensitivity of both the HT and the MHT search selections

- For $m_0 \leq$ (>) 450 GeV the MHT (HT) selection is more powerful
- Production cross section excluded above 2-3 pb at the 95% C.L.
- Gluino masses excluded below 500 GeV for squark masses 300-1000 GeV at the 95% C.L.

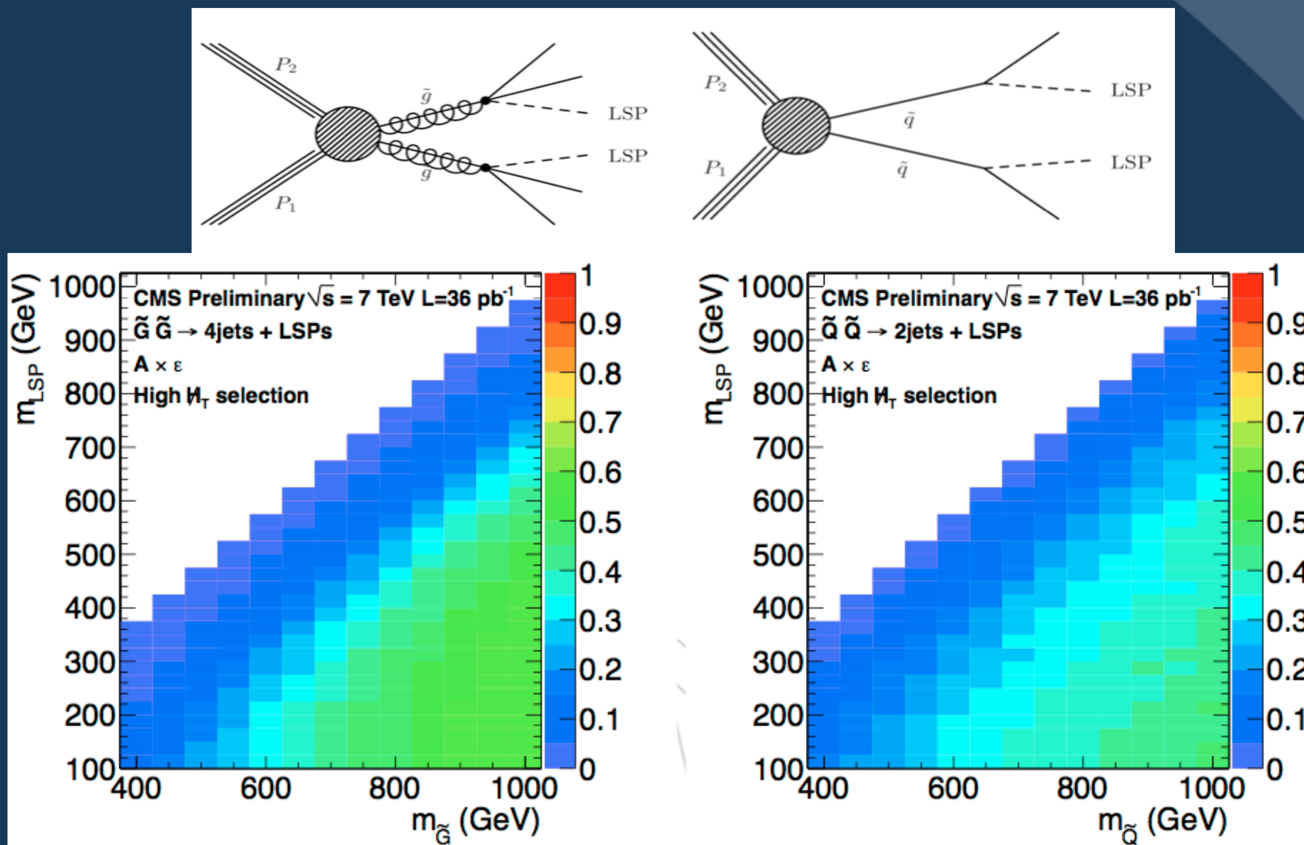
tan β=10, μ>0, A_0=0





Interpretation with Simplified Models

Signal acceptance grows at higher gluino masses and decreases in the diagonal since jets are produced with low p_T

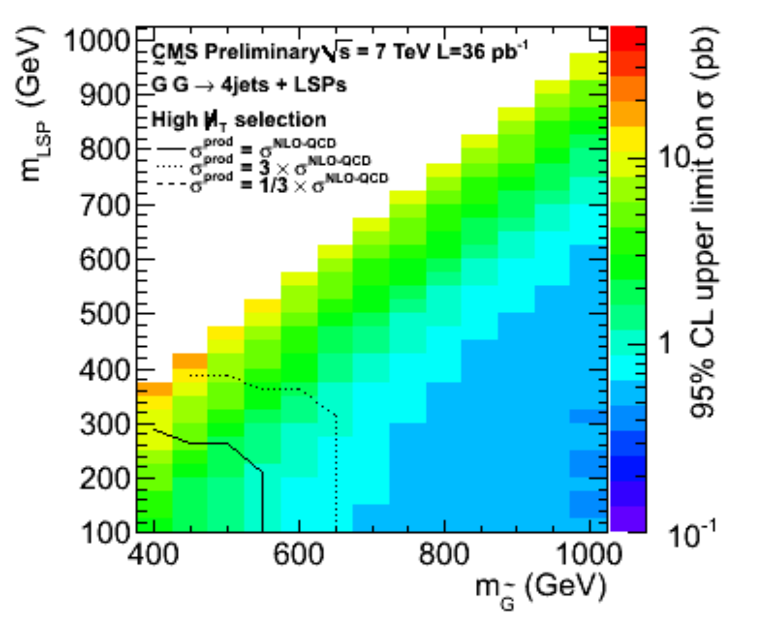


Interpretation with Simplified Models

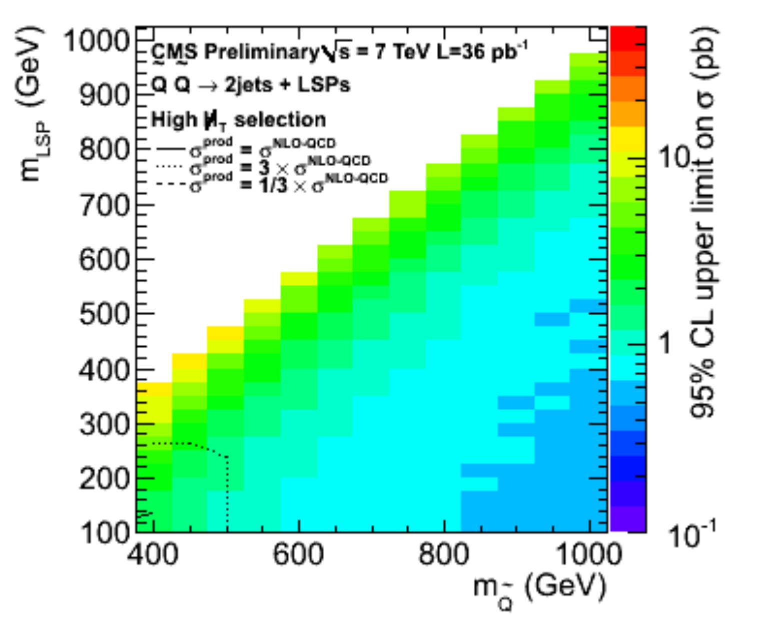
This model independent representation with the simplified model spectra allows to translate a limit to any complete model like SUSY

- Production cross section excluded above 0.5-30 pb at the 95% C.L. depending on the masses of the new particles in the decay chains

High MHT Selection



High MHT Selection

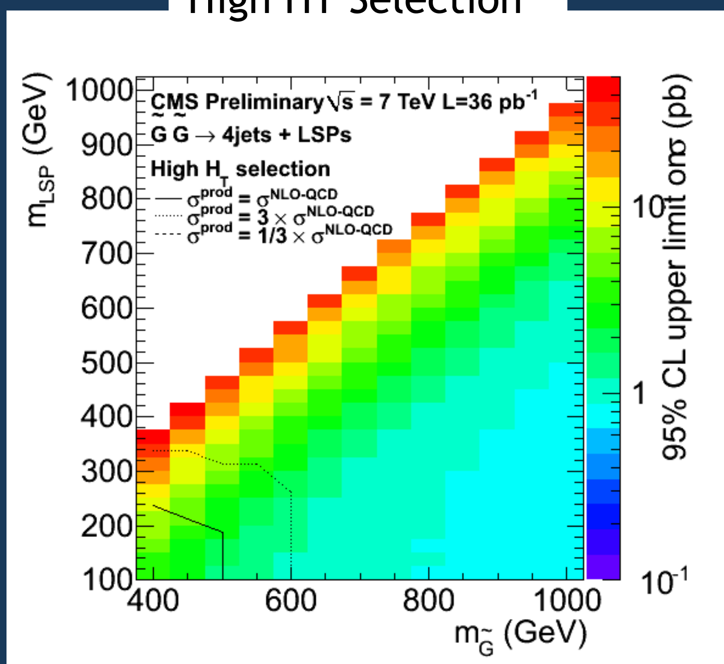


Interpretation with Simplified Models

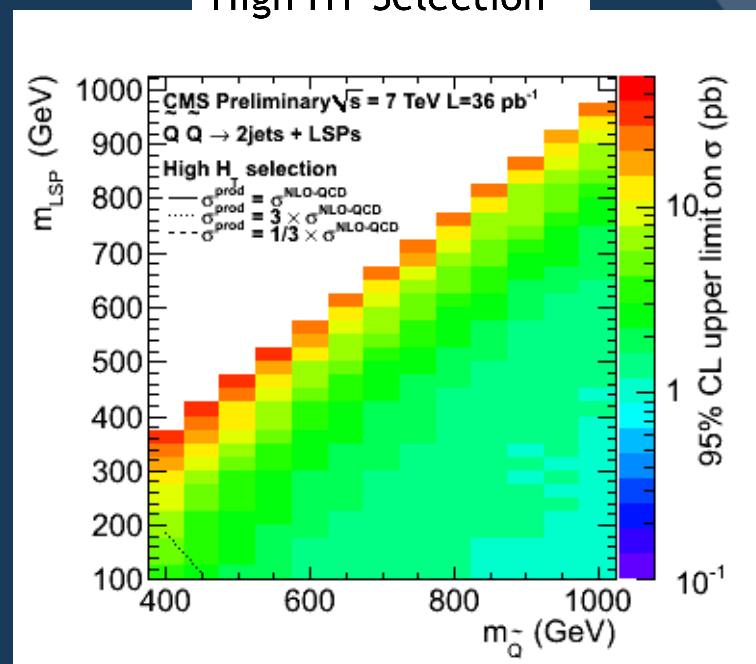
This model independent representation with the simplified model spectra allows to translate a limit to any complete model like SUSY.

- Production cross section excluded above 0.5-30 pb at the 95% C.L. depending on the masses of the new particles in the decay chains

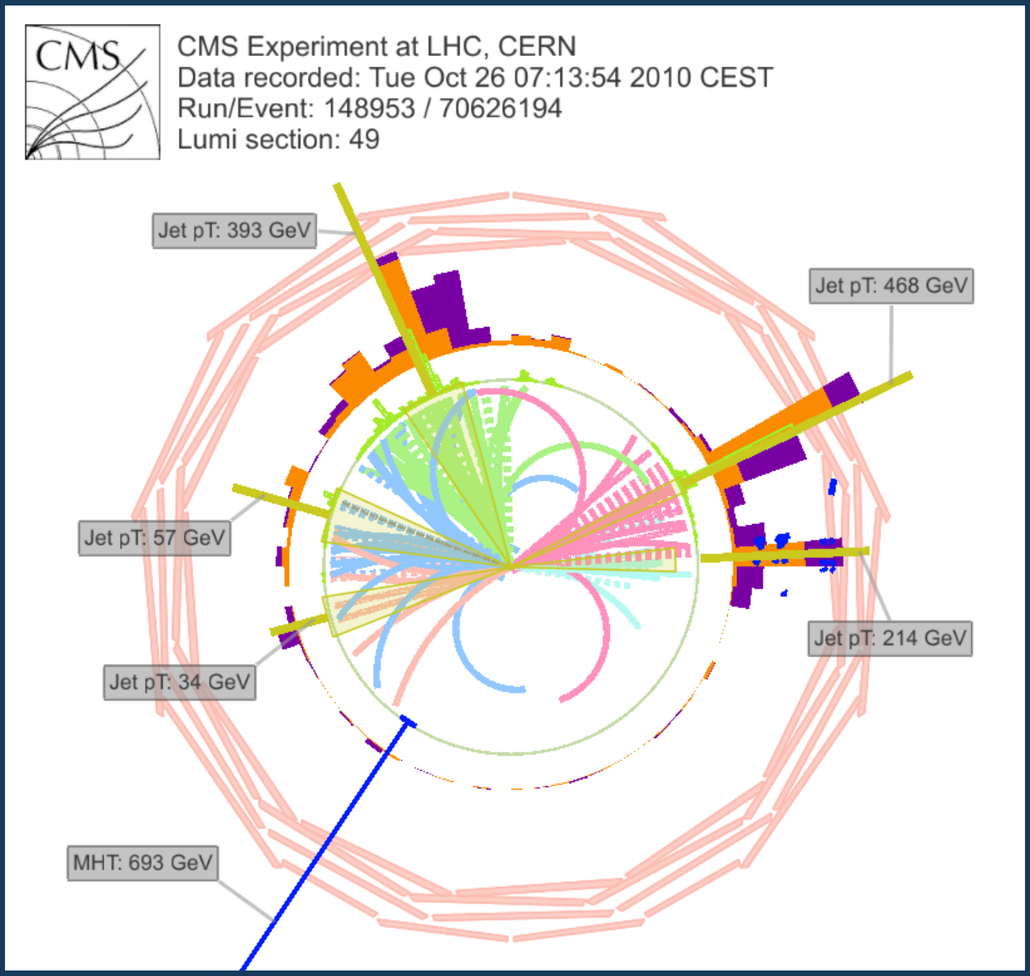
High HT Selection



High HT Selection



A Candidate Event



MHT = 693 GeV
HT= 1132 GeV
M_{eff}= MHT+HT = 1.83 TeV
No b-tagged jet
No isolated lepton
Incompatible with W or top mass