M_{T2} Analysis

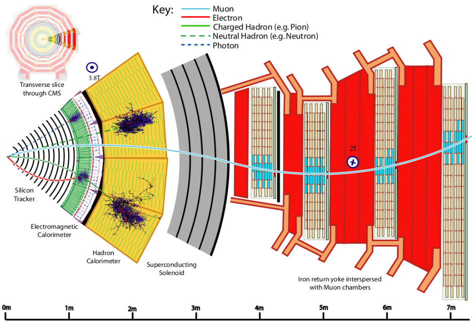
Search for Supersymmetry with the M_{T2} variable in full-hadronic final states in proton-proton collisions at $\sqrt{s} = 13$ TeV

Anne-Mazarine Lyon
anne-mazarine.lyon@cern.ch

Zürich PhD Student Seminar

9th October 2019

- One of the four experiments of CERN-LHC
- Collects the data that is produced in p-p collisions
- Different detector layers aim at measuring different properties of particles
- The information from all detectors gathered in order to reconstruct the event
- Two-level trigger system selects the interesting events



Transverse slice of the CMS detector

Introduction

Supersymmetry as a candidate for New Physics

Introduction

Supersymmetry as a candidate for New Physics

- Symmetry whose generators transform bosonic states into fermionic ones and vice versa

$$\mathcal{S}\,|\text{boson}\rangle = |\text{fermion}\rangle \qquad \mathcal{S}\,|\text{fermion}\rangle = |\text{boson}\rangle$$

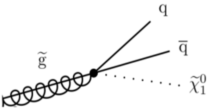
Introduction
Supersymmetry as a candidate for New Physics

- Symmetry whose generators transform bosonic states into fermionic ones and vice versa

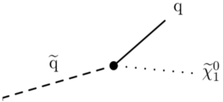
$$\mathcal{S}\,|\text{boson}\rangle = |\text{fermion}\rangle \qquad \mathcal{S}\,|\text{fermion}\rangle = |\text{boson}\rangle$$

- Strong production of SUSY
  - gluinos: fermionic superpartners of gluons
  - squarks: bosonic superpartners of quarks

- Conservation of R-parity: pair production of sparticles, decay to LSP stable (lightest SUSY particle)



gluino decay



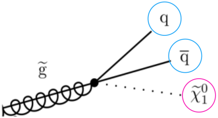
squark decay

Supersymmetry as a candidate for New Physics

- Symmetry whose generators transform bosonic states into fermionic ones and vice versa

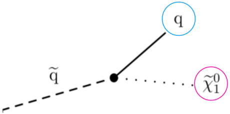
$$\mathcal{S}\,|\text{boson}\rangle = |\text{fermion}\rangle \qquad \mathcal{S}\,|\text{fermion}\rangle = |\text{boson}\rangle$$

- Strong production of SUSY
  - gluinos: fermionic superpartners of gluons
  - squarks: bosonic superpartners of quarks



gluino decay

- Conservation of R-parity: pair production of sparticles, decay to LSP stable (lightest SUSY particle)



squark decay

- Signatures:
  - Quarks ⇒ large hadronic activity
  - LSP ⇒ large missing transverse energy E_T^{miss}

Introduction
What is the M_{T2} Analysis?

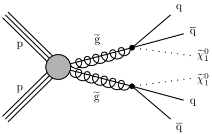
1. M_{T2} Analysis

- Search for New Physics in the all-hadronic final states, with large missing transverse energy E_T^{miss}
- Use the M_{T2} variable as a discovery variable

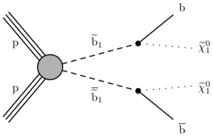
Introduction
What is the M_{T2} Analysis?

1. M_{T2} Analysis

- Search for New Physics in the all-hadronic final states, with large missing transverse energy E_T^{miss}
- Use the M_{T2} variable as a discovery variable
- Constrain mass of strong sparticles within simplified SUSY models
  - Limited set of sparticles with given production and decay modes
  - In this talk:



gluino-mediated decay

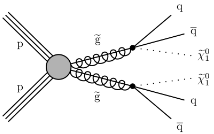


squark-mediated decay

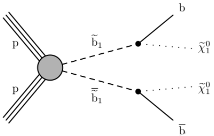
1. M_{T2} Analysis

Introduction
Methodology
Results
Interpretation
Summary

- Search for New Physics in the all-hadronic final states, with large missing transverse energy E_T^{miss}
- Use the M_{T2} variable as a discovery variable
- Constrain mass of strong sparticles within simplified SUSY models
  - Limited set of sparticles with given production and decay modes
  - In this talk:



gluino-mediated decay



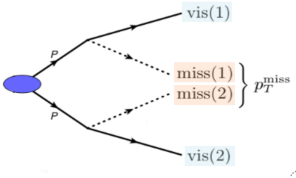
squark-mediated decay

- LHC data collected by CMS in 2016, 2017, 2018 at $\sqrt{s}$=13 TeV
- Total integrated luminosity: 137.2 fb^{-1}

2. M_{T2} Variable

2. M_{T2} Variable

- Measurement of the mass of pair-produced particles, both decaying into one visible and one invisible particle



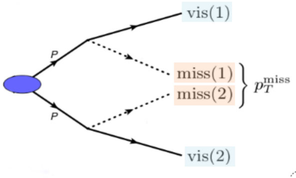
Pair production

Introduction

What is the M_{T2} Analysis?

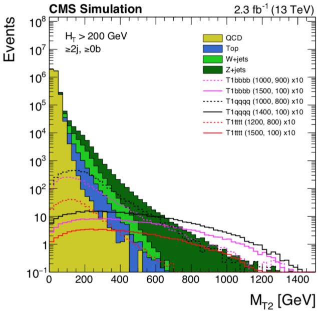
2. M_{T2} Variable

- Measurement of the mass of pair-produced particles, both decaying into one visible and one invisible particle



Pair production

- Used as a discovery variable: New Physics expected in the tail of the distribution
- Allows to get rid of QCD background



M_{T2} distribution for main backgrounds

Methodology

1. Event selection
2. Phase space binning
3. Estimation of the SM backgrounds
4. Assignment of the uncertainties
5. Signal extraction

1. Event selection

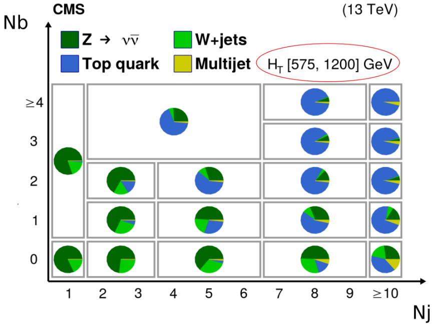
- Select events of interest, i.e with full-hadronic final states with large missing transverse energy
- Apply selection criteria in order to get rid of background
  $\Rightarrow$ Signal Region (SR)

Signal selection

At least one jet $H_T > 250$ GeV $E_T^{\text{miss}} > 250$ GeV for $H_T < 1200$ GeV > 30 GeV for $H_T > 1200$ GeV	Large hadronic activity & missing transverse energy
$\text{M}_{\text{T2}} > 200$ GeV for $H_T < 1500$ GeV > 400 GeV for $H_T > 1500$ GeV $\Delta\Phi^{\text{min}} > 0.3$ $\lvert H_T^{\text{miss}} - E_T^{\text{miss}} \rvert < 0.5\, E_T^{\text{miss}}$	QCD background rejection
Lepton vetoes	EW background rejection

2. Phase space binning

- Classify the events into regions based on the value of H_T, N_j and N_b
- Further bin these regions in $\mathrm{M_{T2}}$
- Highly binned analysis (282 bins) ⇒ large phase space coverage



Binning in N_j and N_b in the medium H_T region

3. Estimation of the SM background

Introduction
Methodology
Results
Interpretation
Summary

- In each bin, compute the counts coming from SM processes using Control Regions (CR)
- CRs are regions orthogonal to the SR: regions enriched with SM background
- Three backgrounds to be estimated:

  a. Z invisible

  b. Lost lepton

  c. QCD

a. Z invisible

a. Z invisible

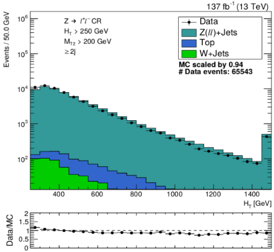
- $Z \to \nu\bar{\nu}$ events with jets, genuine E_T^{miss} carried by neutrinos

Introduction
Methodology
Results
Interpretation
Summary

a. Z invisible

- $Z \to \nu\bar{\nu}$ events with jets, genuine E_T^{miss} carried by neutrinos
- Estimated using a di-lepton Control Region

- Same selection as SR except for lepton veto
- Exactly two leptons (ee or $\mu\mu$) with opposite charge
- $Z \to \nu\bar{\nu}$ kinematics
  - invariant mass of lepton peaking at Z mass
  - Lepton p_T vectorially added to E_T^{miss}
- Enriched in Z→ ll events



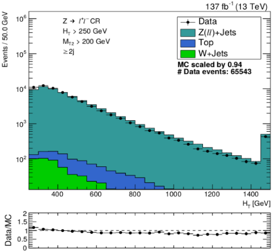
Data/MC shape comparison in the CR

Introduction
Methodology
Results
Interpretation
Summary

a. Z invisible

- $Z \to \nu\bar{\nu}$ events with jets, genuine E_T^{miss} carried by neutrinos
- Estimated using a di-lepton Control Region

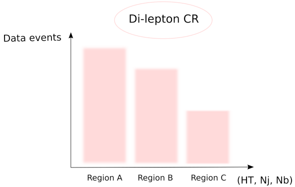
  - Same selection as SR except for lepton veto
  - Exactly two leptons (ee or $\mu\mu$) with opposite charge
  - $Z \to \nu\bar{\nu}$ kinematics
    - invariant mass of lepton peaking at Z mass
    - Lepton p_T vectorially added to E_T^{miss}
  - Enriched in Z→ ll events

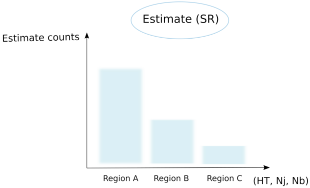


Data/MC shape comparison in the CR

- Transfer factor (TF) from CR to SR get background counts in the SR

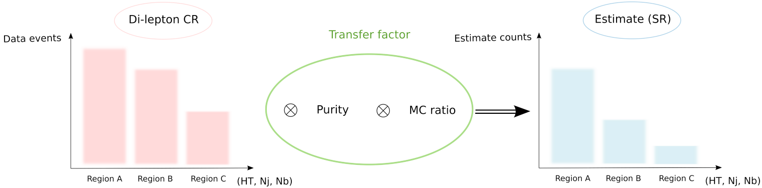
Transfer factor method



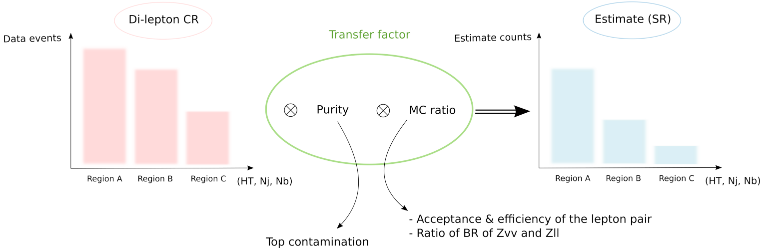


Introduction
Methodology
Results
Interpretation
Summary

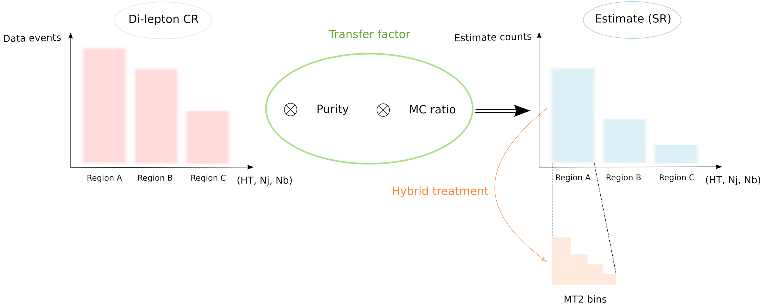
Transfer factor method



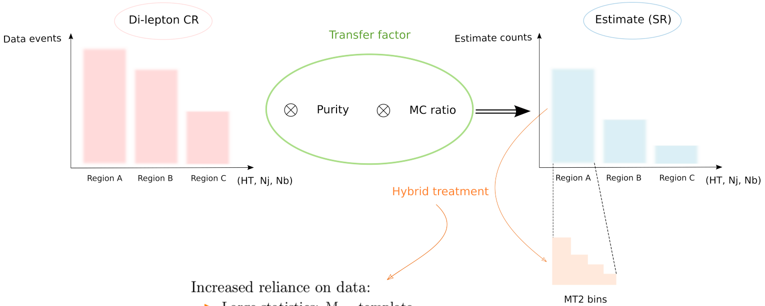
Transfer factor method



Transfer factor method

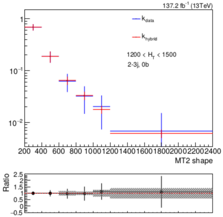


Transfer factor method



Increased reliance on data:

- Large statistics: M_{T2} template from data
- Limited statistics: M_{T2} template from MC



Normalized M_{T2} distribution in a given region

b. Lost lepton

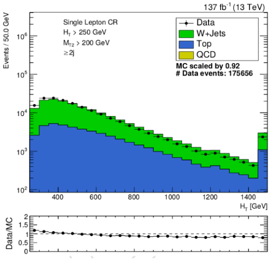
b. Lost lepton

- Lepton out of acceptance, not reconstructed, not identified or not isolated

Introduction
Methodology
Results
Interpretation
Summary

b. Lost lepton

- Lepton out of acceptance, not reconstructed, not identified or not isolated
- Estimated using a single-lepton CR
  - Same trigger and selection as SR except for lepton veto
  - Exactly one lepton (e or μ)
  - Enriched in W+jets and $t\bar{t}$ events



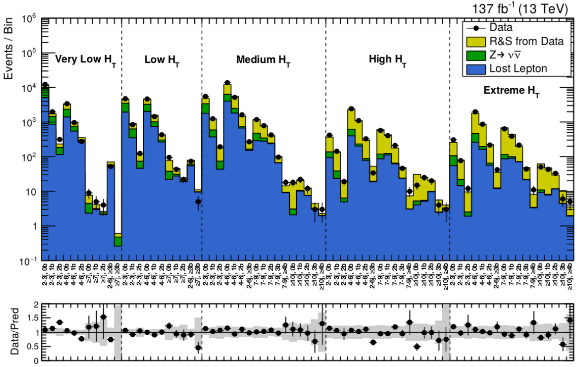
Data/MC shape comparison in the CR

- Transfer factor (TF) from CR to SR to get the background counts in the SR

c. QCD

c. QCD

- Instrumental background
- Jet mismeasurement
- Estimated using the Rebalance and Smear technique



QCD enriched Control Region

4. Assignment of the uncertainties

Z invisible

Source	Range[%]
Limited size of data control samples	5-100
Limited size of MC samples	0-50
Lepton efficiency	0-5
Jet Energy Scale	0-5
Uncertainty in $R^{\mathrm{SF/OF}}$	0-5
$\mathrm{M_{T2}}$ shape uncertainty	0-40

Lost lepton

Source	Range[%]
Limited size of data control samples	5-100
Limited size of MC samples	0-50
Lepton efficiency	0-10
τ efficiency	0-3
b tagging efficiency	0-3
$\mathrm{t\bar{t}b\bar{b}/t\bar{t}j\bar{j}}$	0-25
μ_R and μ_F variation	0-5
Jet Energy Scale	0-5
M_T(lepton, $\vec{p}_T^{\,\mathrm{miss}}$) selection efficiency	0-3
$\mathrm{M_{T2}}$ shape uncertainty (if k_{LL}) $\neq$1	0-40

Z invisible

Source	Range[%]
Limited size of data control samples	5-100
Limited size of MC samples	0-50
Lepton efficiency	0-5
Jet Energy Scale	0-5
Uncertainty in $R^{\mathrm{SF/OF}}$	0-5
$\mathrm{M_{T2}}$ shape uncertainty	0-40

Lost lepton

Source	Range[%]
Limited size of data control samples	5-100
Limited size of MC samples	0-50
Lepton efficiency	0-10
τ efficiency	0-3
b tagging efficiency	0-3
$\mathrm{t\bar{t}b\bar{b}/t\bar{t}j\bar{j}}$	0-25
μ_R and μ_F variation	0-5
Jet Energy Scale	0-5
$M_T(\text{lepton}, \vec{p}_T^{\,\text{miss}})$ selection efficiency	0-3
$\mathrm{M_{T2}}$ shape uncertainty (if k_{LL}) $\neq$1	0-40

- Statistical uncertainties dominant in extreme regions

Z invisible

Source	Range[%]
Limited size of data control samples	5-100
Limited size of MC samples	0-50
Lepton efficiency	0-5
Jet Energy Scale	0-5
Uncertainty in $R^{SF/OF}$	0-5
M_{T2} shape uncertainty	0-40

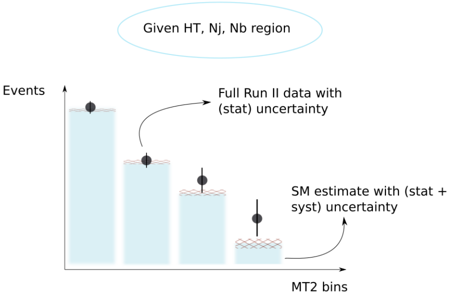
Lost lepton

Source	Range[%]
Limited size of data control samples	5-100
Limited size of MC samples	0-50
Lepton efficiency	0-10
τ efficiency	0-3
b tagging efficiency	0-3
$t\bar{t}b\bar{b}/t\bar{t}j\bar{j}$	0-25
μ_R and μ_F variation	0-5
Jet Energy Scale	0-5
M_T(lepton, $\vec{p}_T^{\,miss}$) selection efficiency	0-3
M_{T2} shape uncertainty (if k_{LL}) $\neq$1	0-40

- Statistical uncertainties dominant in extreme regions
- Systematics: main contribution comes from the M_{T2} shape

5. Signal extraction

- Compare data event yields to the background prediction
- Perform a simultaneous fit to the data in all signal regions

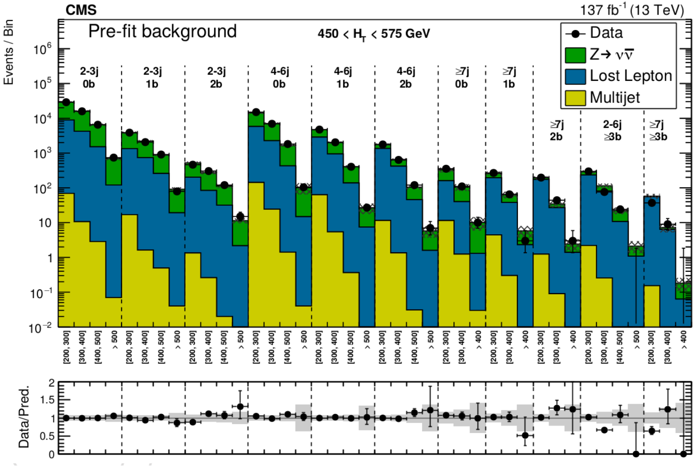


- An excess of data would be a sign of New Physics

Results

Results

- Comparison of data and prediction in a given H_T region



$(450 < H_T < 575)$ GeV

Interpretation

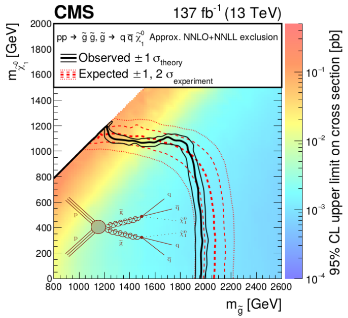
Interpretation

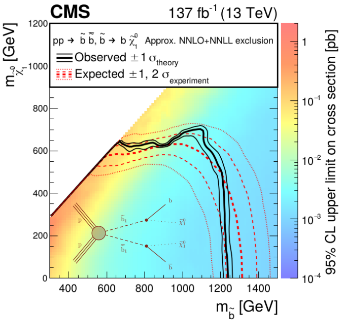
- No significant excess of data over SM background has been observed
- Largest observed significance of 1.8σ

Interpretation

- No significant excess of data over SM background has been observed
- Largest observed significance of 1.8σ

- Interpreted in the context of simplified SUSY models
- Constrain masses of gluinos, squarks and LSP within those models





- Exclusion limits improved of O(100)GeV compared to previous analysis exploiting 2016 data only

Summary

Summary

M_{T2} analysis

- Performed an inclusive search for strong SUSY
- Targeted full-hadronic with large missing energy final states
- M_{T2} used as a discovery variable
- Full Run II data studied

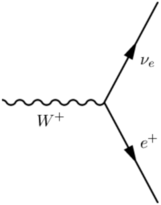
Conclusions

- No significant deviation from SM background observed
- Limits on sparticles masses improved by O(100)GeV compared to the results with 2016 data only
- Paper under review by EPJC (arXiv:1909.03460)

Backup

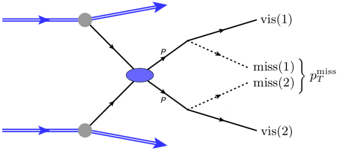
Reminder - Transverse mass

- Useful mean to measure mass of particle decaying into one visible and one invisible particle (e.g W boson)
- Kinematic endpoint at the searched mass



Generalization to $\mathrm{M_{T2}}$

- Measure of the mass of pair-produced particles, both decaying into one visible and one invisible particle
- Minimization due to the fact that one can only measure the sum of the two missing transverse momenta



$$\mathrm{M}_{\mathrm{T2}}^2 \equiv \min_{p_T^{\mathrm{miss}(1)} + p_T^{\mathrm{miss}(2)} = p_T^{\mathrm{miss}}} \left[\max \left\{ M_T^2 \left(p_T^{\mathrm{vis}(1)}, p_T^{\mathrm{miss}(1)} \right), M_T^2 \left(p_T^{\mathrm{vis}(2)}, p_T^{\mathrm{miss}(2)} \right) \right\} \right]$$

Background estimation
Z invisible

$$N^{\mathrm{SR}}_{Z\to\nu\nu} = N^{\mathrm{CR}}_{Z\to ll} \cdot P_{Z\to ll} \cdot R^{Z\to\nu\nu/Z\to ll}_{\mathrm{MC}} \cdot k_{\mathrm{hybrid}}$$

Purity

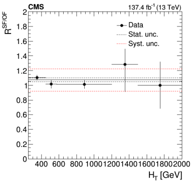
- Measure of the Top contamination in the di-lepton CR
- Estimated (in data) using two CRs: same flavour leptons (SF) and opposite flavour leptons (OF)

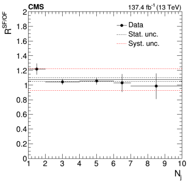
$$P_{Z\to ll} = \frac{N^{\mathrm{CR}}_{2l}\,(\mathrm{SF}) - N^{\mathrm{CR}}_{2l}\,(\mathrm{OF})\,\mathrm{R}\,(\mathrm{SF/OF})}{N^{\mathrm{CR}}_{2l}\,(\mathrm{SF})}$$

- Consistent with 1 in all topological regions

R(SF/OF)

- Estimated in a $t\bar{t}$-enriched sample (inverted di-lepton p_T and mass cuts) and applying the 2D sideband method
- Behaviour wrt kinematic variables well described by a constant $\mathrm{R(SF/OF)} = 1.06 \pm 0.15$





Backup
Z purity in the di-lepton CR - 2D sideband method

- Estimation of the Top background using 2D sideband method (evaluated in data): implementation of four different CR regions (see sketch below)
  → regions A, B and D: $t\bar{t}$-enriched regions
  → region C: Z-enriched region
- Assumption: contamination of Z in regions A, B and D negligible ($N^{A,B,D}_{Zll} \approx 0$)
- If two axes are uncorrelated, we have $\mathrm{R(SF/OF)}^{\mathrm{B/A}} = \mathrm{R(SF/OF)}^{\mathrm{C/D}}$ which implies

$$N_{\mathrm{Top}} = N^{C}_{\mathrm{Top}} = \frac{N^{B}_{\mathrm{Top}}}{N^{A}_{\mathrm{Top}}} \cdot N^{D}_{\mathrm{Top}} \qquad (1)$$

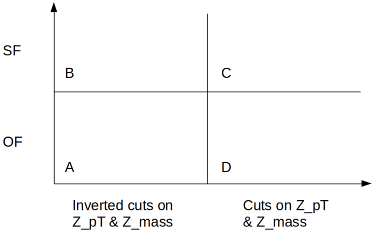


Figure: Definition of the regions of the 2D sideband method

Background estimation

Z invisible

$$N^{\text{SR}}_{Z\to\nu\nu} = N^{\text{CR}}_{Z\to ll} \cdot P_{Z\to ll} \cdot R^{Z\to\nu\nu/Z\to ll}_{\text{MC}} \cdot k_{\text{hybrid}}$$

Hybrid shape

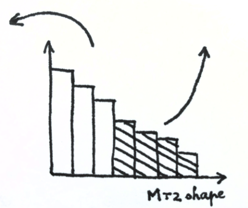
- Normalized M_{T2} shape per topological region: distribution of the estimate across the M_{T2} bins
- Measured inclusively in N_b
- Hybrid treatment: → increase reliance on data as much as statistics allows it → template normalized to unity

Limited statistics

- bins in the tail merged until the sum of the events is at least equal to 50
- k_{hybrid} built out of $Z \to \nu\nu$ MC
- $k_{\text{hybrid}} = \text{MC\_SR(bin)} \cdot \frac{\int \text{MC\_SR}}{\int \text{MC\_CR}} \cdot \frac{\int \text{data}}{\int \text{MC\_SR}}$

Large statistics

- bin by bin estimate
- k_{hybrid} built out of Z → ll data
- $k_{\text{hybrid}} = \text{data(bin)} \cdot \frac{\text{MC\_SR(bin)}}{\text{MC\_CR(bin)}}$

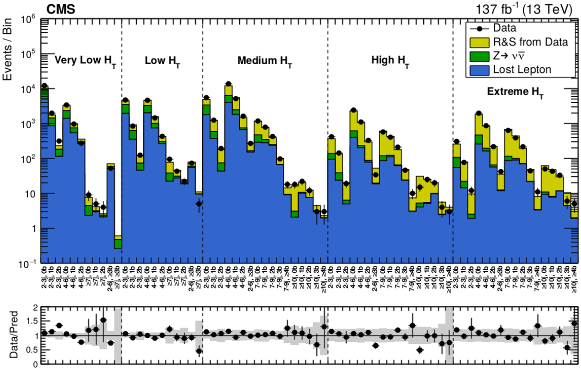


Background estimation
QCD

How to estimate this background?

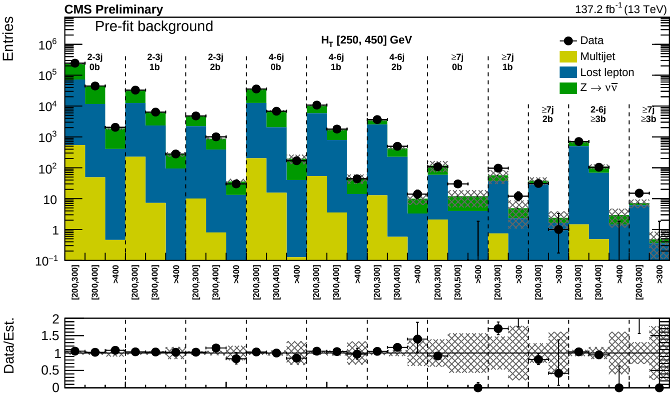
→ Two steps procedure:

1. Rebalance multijet events by adjusting the jet p_T in a way to minimize the transverse missing energy
2. Smear many times according to templates of $p_T^{\text{reco}}/p_T^{\text{gen}}$



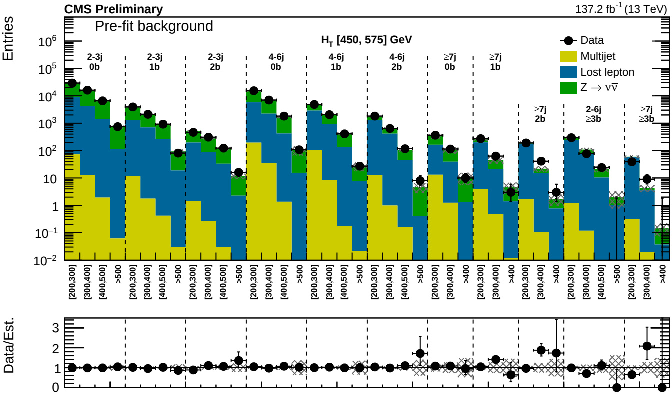
QCD enriched Control Region

Backup
Results - very low HT region



$(250 < H_T < 450)$ GeV

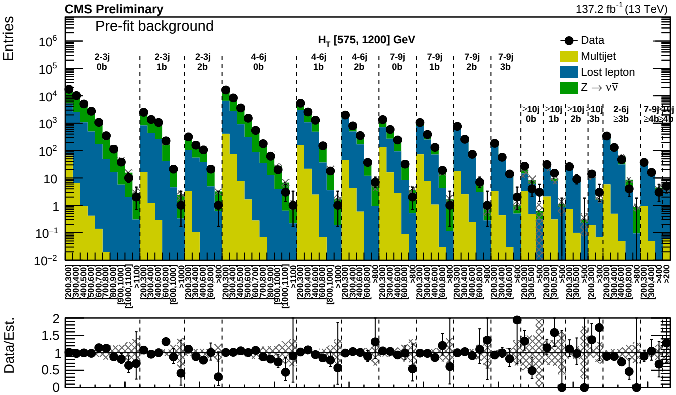
Backup
Results - low HT region



$(450 < H_T < 575)$ GeV

Backup

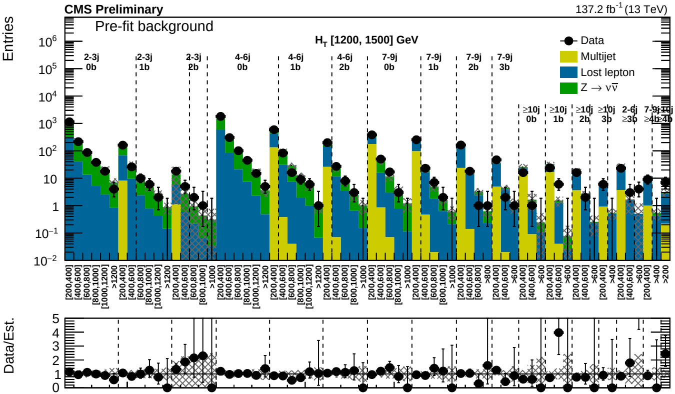
Results - medium HT region



$(575 < H_T < 1200)$ GeV

Backup

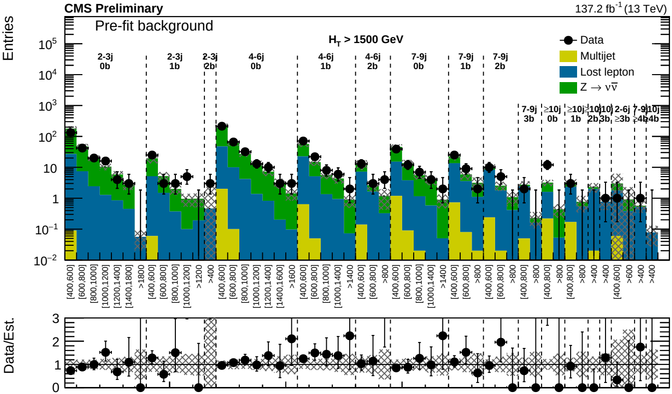
Results - high HT region



(1200 $< H_T <$ 1500) GeV

Backup

Results - extreme HT region



$H_T > 1500$ GeV

Interpretation
Residuals and significance

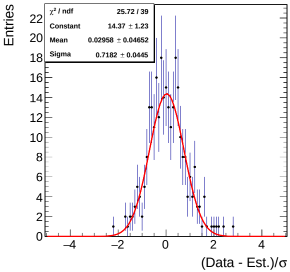
- No significant excess of data over SM background has been observed

Residuals

- Residuals computed in each bin

$$r = \frac{\text{Data} - \text{Estimate}}{\sqrt{\sigma^2_{\text{Data}} + \sigma^2_{\text{Estimate}}}}$$

- Distribution well described by a Gaussian with mean around 0



Significance

- Computation of the significance of the signals of interest
  - simultaneous fit of Sig+Bkg in all bins
  - profile likelihood ratio with asymptotic approximation
- Largest observed significance: 1.8σ (for T1 with M=2700GeV and m=1100GeV)

Samples

- Data:
  - 2016: Nano14Dec2018
  - 2017: Nano14Dec2018
  - 2018: NanoAODv4-Priv-Mo19JECs (private production)
- Background:
  - 2016: Summer16 (94X MiniAODDv3)
  - 2017: Fall17 (94X MiniAODv2)
  - 2018: Autumn18 (102X)
- Signal: 94X Fast Simulation

Primary datasets

- Signal region: JetHT, HTMHT, MET
- Control region: additionally SingleMuon, SingleElectron, DoubleMuon, DoubleEG, MuonEG

Triggers

2016:

$p_{\mathrm{T}}^{\mathrm{miss}} > 120\,\mathrm{GeV}$ and $H_{\mathrm{T}}^{\mathrm{miss}} > 120\,\mathrm{GeV}$ or

$H_{\mathrm{T}} > 300\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{\mathrm{miss}} > 110\,\mathrm{GeV}$ or

$H_{\mathrm{T}} > 900\,\mathrm{GeV}$ or jet $p_{\mathrm{T}} > 450\,\mathrm{GeV}$

2017 and 2018:

$p_{\mathrm{T}}^{\mathrm{miss}} > 120\,\mathrm{GeV}$ and $H_{\mathrm{T}}^{\mathrm{miss}} > 120\,\mathrm{GeV}$ or

$H_{\mathrm{T}} > 60\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{\mathrm{miss}} > 120\,\mathrm{GeV}$ and $H_{\mathrm{T}}^{\mathrm{miss}} > 120\,\mathrm{GeV}$ or

$H_{\mathrm{T}} > 500\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{\mathrm{miss}} > 100\,\mathrm{GeV}$ and $H_{\mathrm{T}}^{\mathrm{miss}} > 100\,\mathrm{GeV}$ or

$H_{\mathrm{T}} > 800\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{\mathrm{miss}} > 75\,\mathrm{GeV}$ and $H_{\mathrm{T}}^{\mathrm{miss}} > 75\,\mathrm{GeV}$ or

$H_{\mathrm{T}} > 1050\,\mathrm{GeV}$ or jet $p_{\mathrm{T}} > 500\,\mathrm{GeV}$

Data quality issues

- 2016+2017: L1 pre-firing inefficiency (minor effect)
- 2017: ECAL Endcap MET issue (followed recommended recipe)
- 2018: HEM failure (large effect on QCD background, veto jets or leptons in the concerned region)

Backup
Object reconstruction

Jets

- ak4 PF-CHS Jets
- p_T >30GeV, $|\eta|$<4.7
- $|\eta|$<2.4 for N_j, N_b, H_T, M_{T2}
- Jet cleaning
  - Reject event if any jet fails jet ID
  - For 1-jet region: special "monojet" ID
  - Reject event if jet with p_T >30GeV in HEM region for affected 2018 dataset

b-Jets

- DeepCSV Medium working point
- p_T >20GeV

MET

- type-1 corrected MET
- for 2017: using MET "v2" recipe

Leptons

- p_T >10GeV, $|\eta|$<2.4
- Electrons
  - POG veto ID (loose for dilep CR), miniRelIso<0.1
  - d0<0.05/0.10cm, dz<0.10/0.20cm (barrel/endcap)
- Muons
  - POG loose ID, miniRelIso<0.2
  - d0<0.2, dz<0.5

Isolated tracks (for Veto)

- PF muons/PF electrons, p_T >5GeV, |dz|<0.1cm, relTrackIso<0.2
- PF charged hadrons, p_T >10GeV, |dz|<0.1cm, relTrackIso<0.1, $|\eta|$<2.4
- M_T(cand, MET)<100GeV

Miscellaneous

MET filters: all recommended
Jet energy corrections:

- 2016: Summer16_07Aug2017_V11
- 2017: Fall17_17Nov2017_V32
- 2018: Autumn18_V3

Signal uncertainties

Source	Typical values
Integrated luminosity	2.3-5
Limited size of MC samples	1-100
Renormalization and factorization scales	5
ISR modeling	0-30
b-tagging efficiency, heavy flavours	0-40
b-tagging efficiency, light flavours	0-20
Lepton efficiency	0-20
Jet energy scale	5
Fast simulation p_T^{miss} modeling	0-5
Fast simulation pileup modeling	4.6