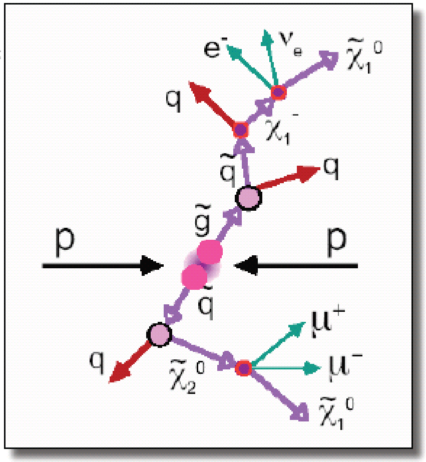
Data-driven estimations of Standard Model backgrounds to SUSY searches

Federica Legger
(Max-Planck-Institut für Physik, Munich)
on behalf of the ATLAS collaboration

Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

SUSY searches in ATLAS

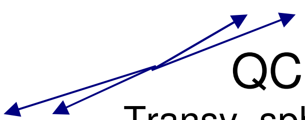


For this talk: focus on R-parity conserving, gravitino mediated (**mSUGRA**) models

- LSP is stable → large missing energy
- Sparticles produced in pairs → cascade decays
- Signature: **Multi jets + leptons + missing transverse energy ($E_{T,miss}$)**

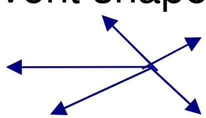
- Baseline selection cuts:
  - at least 4 jets with PT>50GeV
  - at least 1 jet with PT>100GeV
  - n leptons (e,μ) with PT > 20 GeV, n=0,1,...
  - $E_{T,miss}$ > min(100 GeV, 0.2 *Meff)
  - Transverse Sphericity > 0.2

- Effective mass $M_{eff} = \sum_{i=1}^{N} p_T^{jet,i} + \sum_{i=1}^{N} p_T^{lep,i} + E_{T,miss}$
  - Total event activity
  - correlated to mass of sparticles
- Transverse sphericity (event shape)



QCD

Transv. spher. ~ 0

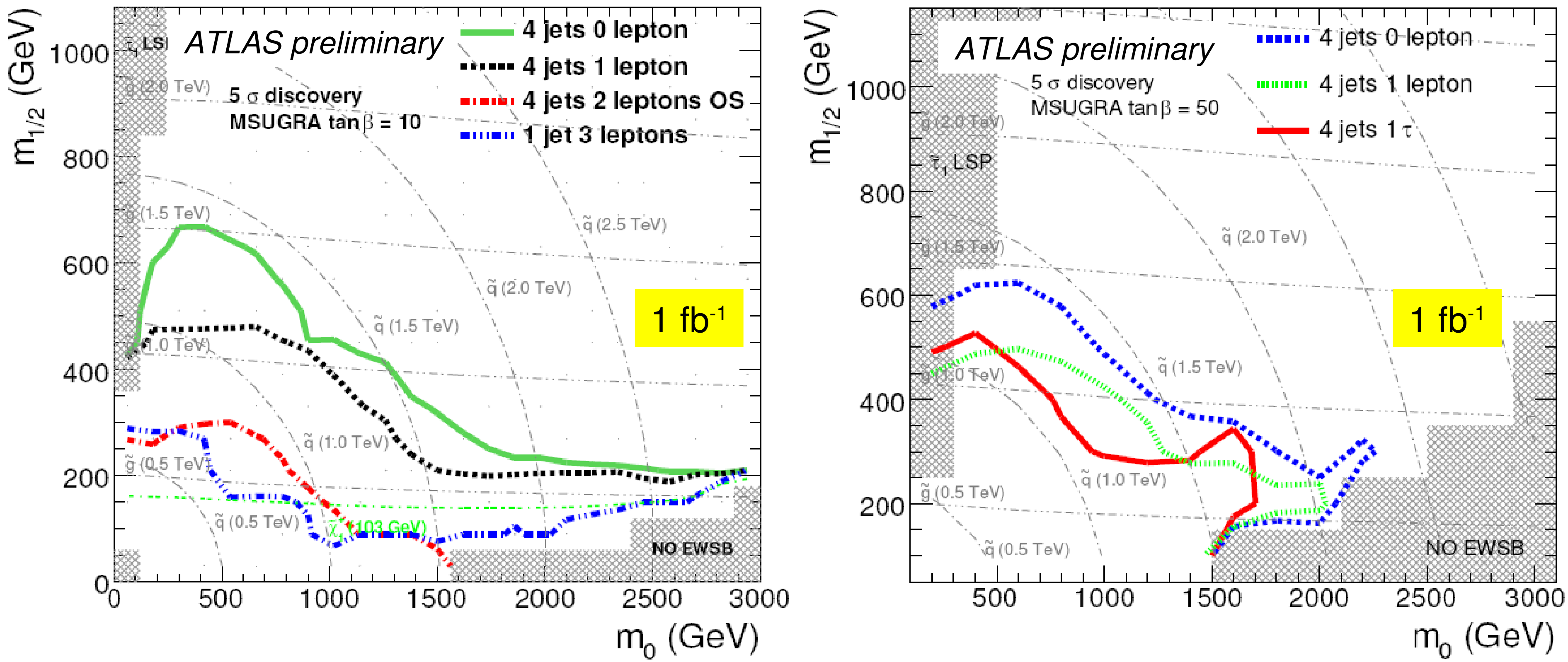


SUSY

Transv. spher. ~ 0

- Other topics:
  - **GMSB** (SUSY breaking mediated by gauge interaction, LSP is gravitino), **Split-SUSY.** Signature very analysis dependent (high pt photons, long lived sparticles)
  - Exclusive measurements

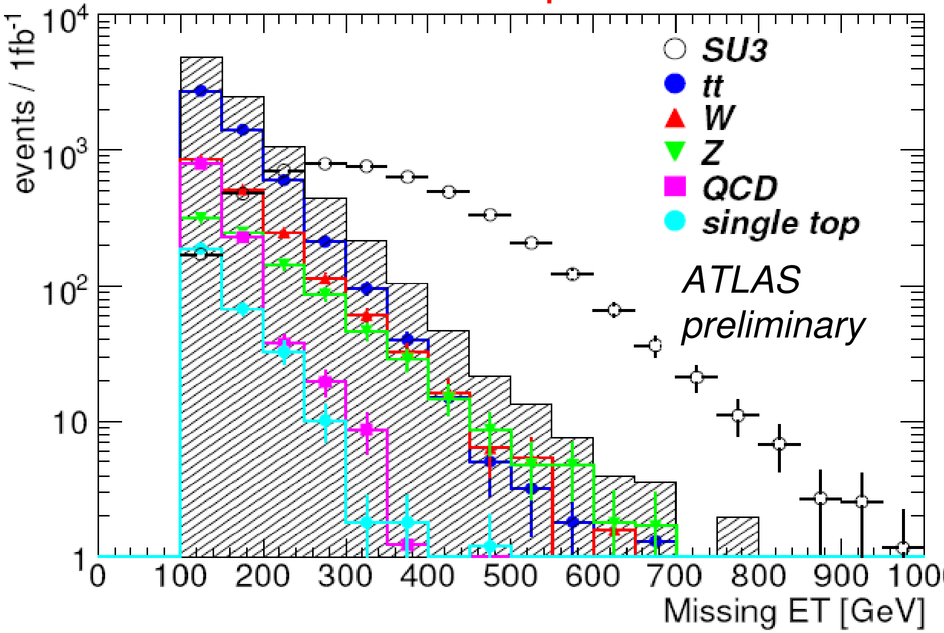
ATLAS sensitivity to SUSY



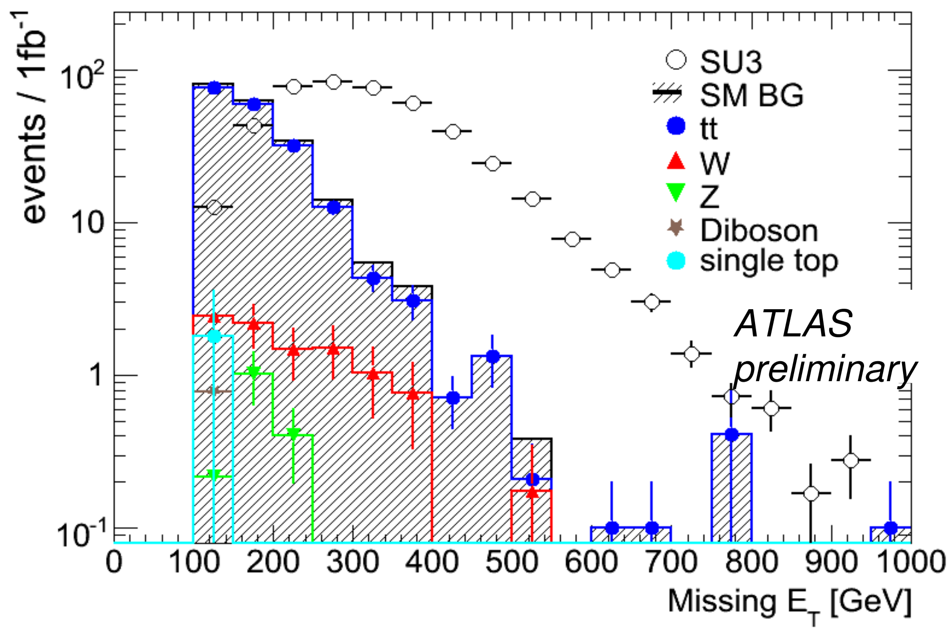
- MET + 4 jets + leptons
- Cut on effective mass optimized to get best signal significance
- Background uncertainties from data-driven methods (assuming 1 fb-1)
  - top/W/Z (20%) + QCD (50%) + 1/sqrt($N_{background}$)

SM backgrounds to SUSY searches

0-lepton mode



1-lepton mode



- Should be estimated from data because of poor knowledge of:
  - Underlying Event
  - Parton Showering
  - Cross-sections
  - Parton Distribution Functions
  - Detector Calibration (jets, $E_{T,miss}$)
  - Limited Monte Carlo statistics

Data-driven background estimation

- Estimate SM backgrounds in a **signal region** where SUSY may be present;
- SUSY may be discovered if an excess of events with respect to SM predictions is found;
- Derive prediction from a **control region**, similar to signal region but with no SUSY
  - unbiased estimation of SM background, enough statistics, low SUSY contamination

QCD	jet smearing
Semileptonic top (tau)	hadronic tau decay
Z -> νν	from Z ->ll (replacement + MC)
Top + W	transverse mass (invariant mass of $E_{T,miss}$ and lepton pt) method combined fit
Semileptonic top tt -> bbqqlν	explicit kinematic reconstruction and selection on top mass (top box method)
Dileptonic top tt -> bblνlν	HT2 (=lepton pt + 2,3,4 leading jets pt) method kinematic reconstruction top redecay

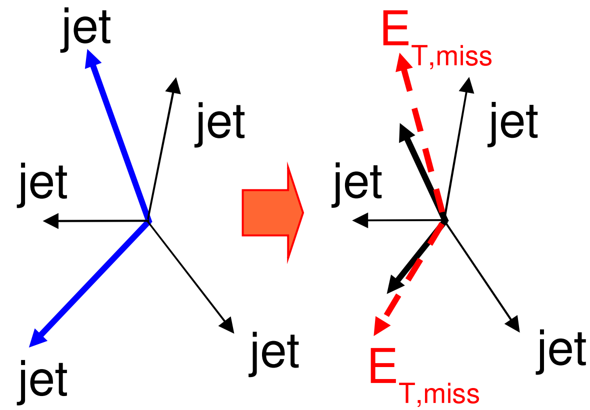
0-lepton mode

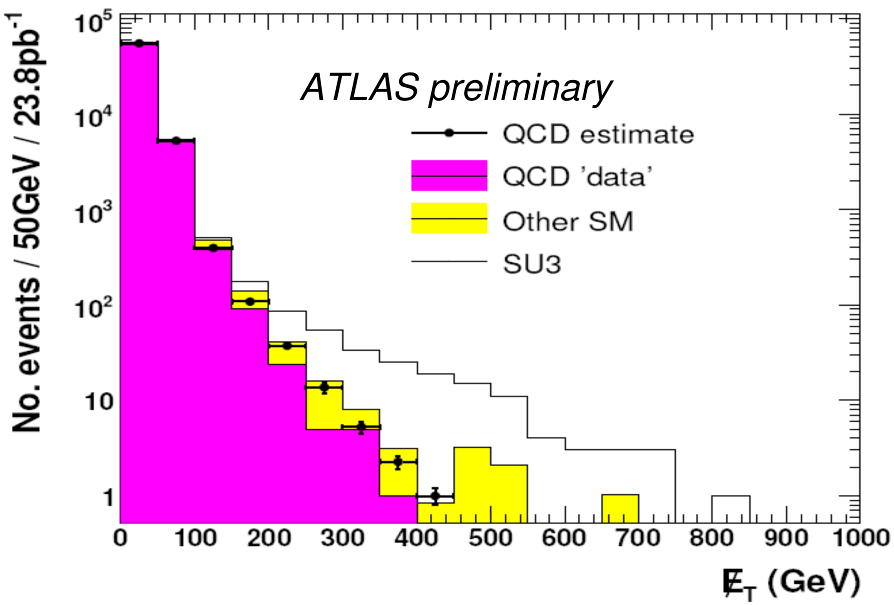
1-lepton mode

In the following, a statistic of 1 fb^{-1} is assumed

QCD background

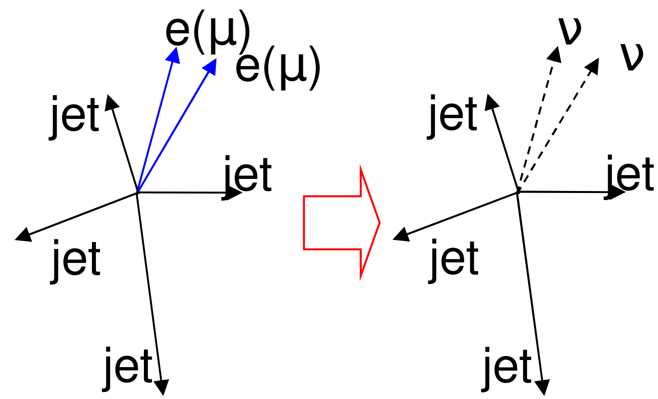
- Neutrinos emitted from semileptonic decays of b/c (**real $E_{T,miss}$**)
- Mismeasurement of jet energies (**fake $E_{T,miss}$**)
- In both cases, $E_{T,miss}$ points in one of the jet directions
- QCD background can be estimated from data from multi-jet events with no $E_{T,miss}$
  - Measure jet response function from events where $E_{T,miss}$ is (anti-)parallel to a jet
  - Apply to smear (all) jet pt in seed events with low $E_{T,miss}$
  - Normalization to QCD jet events with $E_{T,miss}$ < 50 GeV



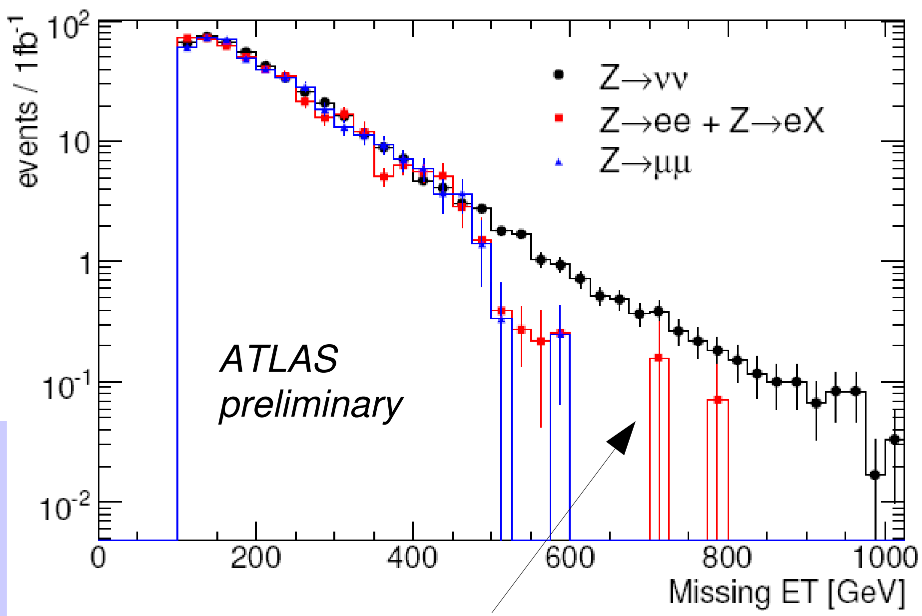


Statistic uncertainties ~1%
Systematic uncertainties ~60%
from biased event selection, statistics in non-gaussian tail and jet response function measurement
low SUSY contamination

Replacement Z -> νν



- Control sample:
  - reconstructed Z->ee or Z->μμ events
- Replace charged leptons with neutrinos
  - $E_{T,miss}$ is given by pt(ll)~pt(Z)
- Correct for lepton identification efficiency
  - from **data** with *tag and probe* method
- Correct for acceptance cuts (**MC**)
- Get Z->νν distributions (normalization and shape)
  - Use extrapolation or MC to get the shape in low stat region



BR(Z -> νν) / BR(Z->ll) ~ 6

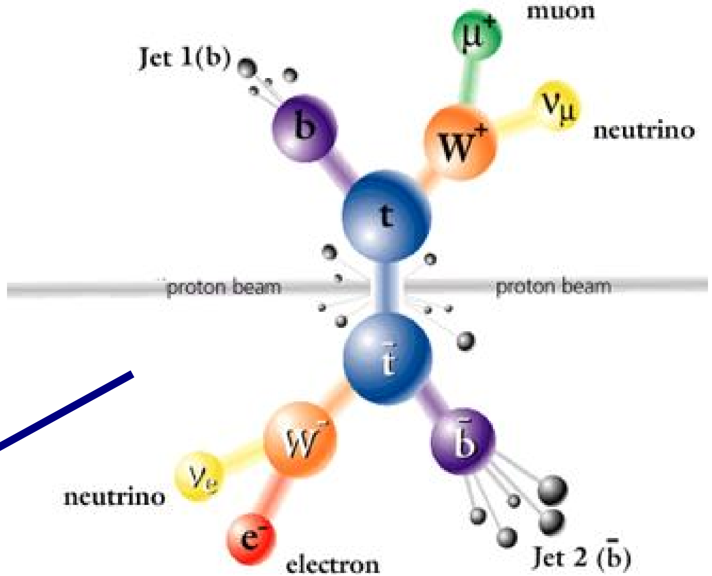
Statistic uncertainties: 13%

Systematic uncertainties: 8%

lepton ID efficiency measurement and $E_{T,miss}$ scale

low SUSY contamination

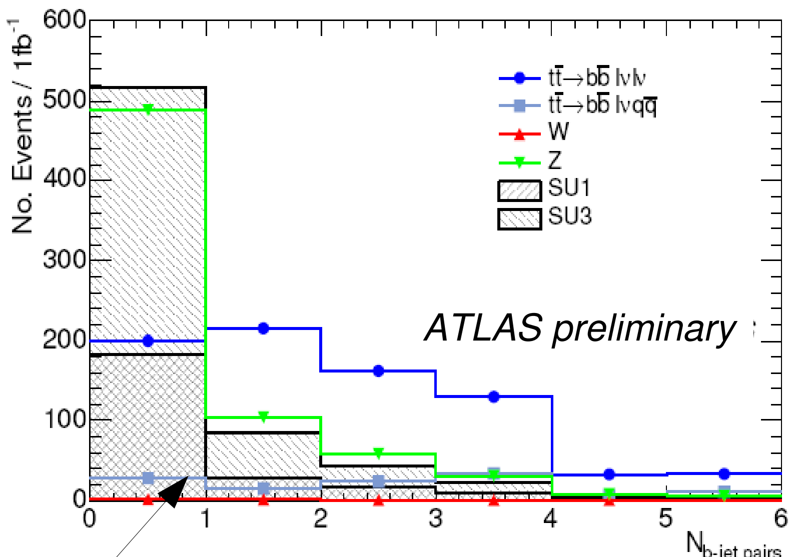
Dileptonic tt: kinematic reconstruction



- Solve system of equations for jets with pt > 20 GeV

$$
\begin{aligned}
m_W^2 &= (p_{l1} + p_{\nu 1})^2 \\
m_W^2 &= (p_{l2} + p_{\nu 2})^2 \\
m_t^2 &= (p_{l1} + p_{\nu 1} + p_{b1})^2 \\
m_t^2 &= (p_{l2} + p_{\nu 2} + p_{b2})^2 \\
E_x^{miss} &= p_{(\nu 1)x} + p_{(\nu 2)x} \\
E_y^{miss} &= p_{(\nu 1)y} + p_{(\nu 2)y}
\end{aligned}
$$

- Quartic equation: 0, 2 or 4 solutions
- no solutions: SUSY event, semi-leptonic ttbar, …
- 2 or 4 solutions: dileptonic top



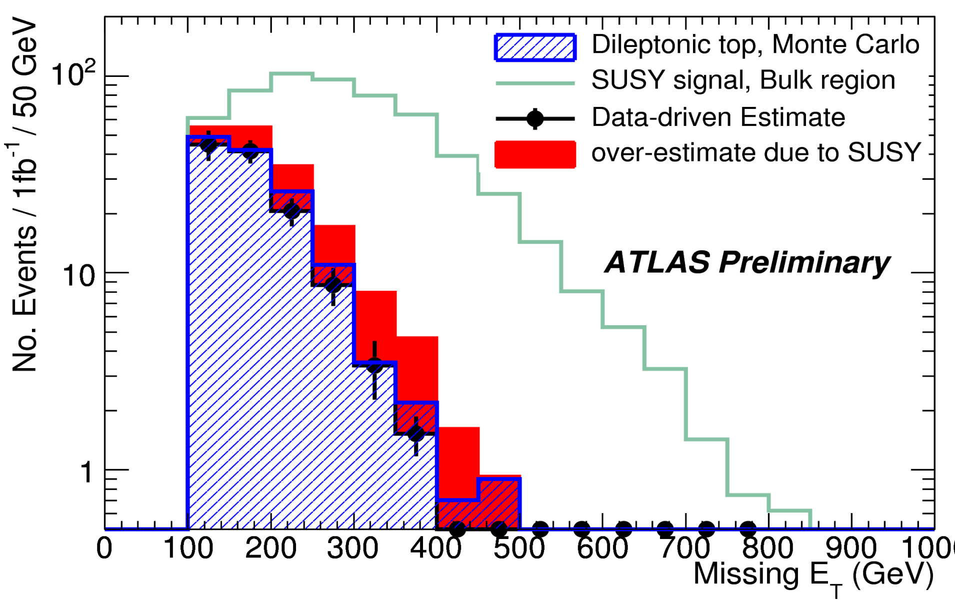
before $E_{T,miss}$ cut is applied

Dileptonic tt: kinematic reconstruction

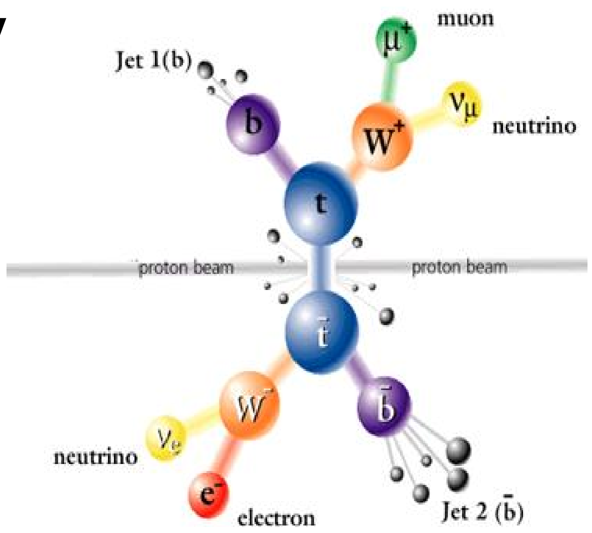
- Dileptonic top with one lepton missed because it is
  - a tau (51%)
  - Misidentified (20%)
  - Inside a jet (17%)
  - Not in acceptance (9%)
  - Both leptons are taus (3%)
- Control sample selection: 2 leptons, 3 jets, nb b-jet pairs > 0
- Normalization in low $E_{T,miss}$ region

Statistical error: 10%
Systematic uncertainties ~20%
Jet energy scale, normalization
SUSY contamination: 50%

- Contribution estimated in the control sample by
  - Replacing a lepton with a tau
  - Removing a lepton
- Recalculate event variables, then apply 1-lepton SUSY selection

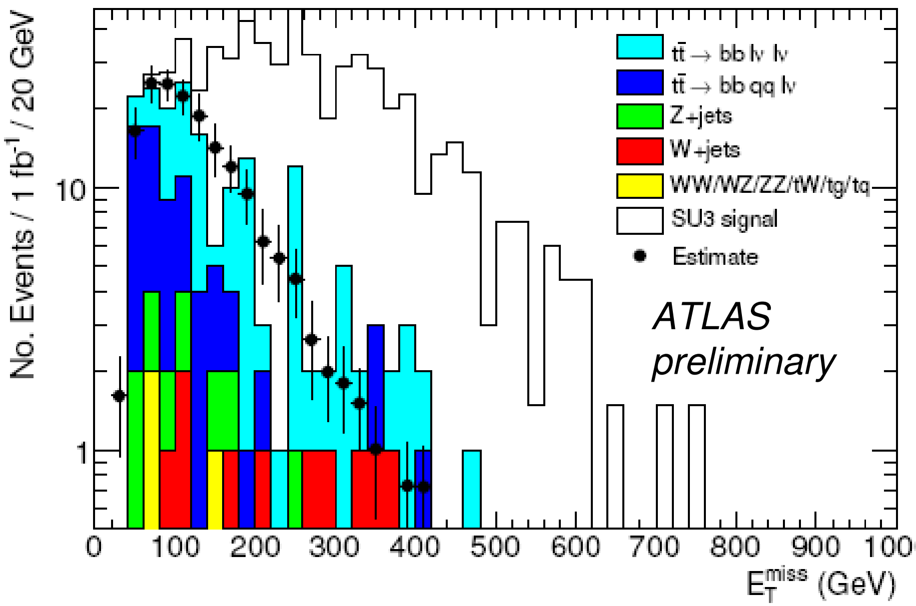


Dileptonic tt: top redecay



- Tag *seed* events (with low $E_{T,miss}$) containing 2 tops
- Reconstruct 4-momentum of tops
- Redecay/hadronize with Pythia
- Simulate decay products with fast simulation (ATLFAST)
- Remove from seed event original decay products and merge new ones
- Apply standard SUSY selection cuts on merged events
- Normalization to *data* in low $E_{T,miss}$ region

Statistic uncertainties ~30%
Systematic uncertainties ~30%
SUSY contamination ~60%



Conclusions

- Main SM backgrounds to SUSY searches are **tt, W+jets, Z+jets, QCD** events
- Several methods are being developed in ATLAS to estimate SM backgrounds
  - Complementary methods are necessary for such a crucial issue!!!

	Stat.	Syst	SUSY
QCD	1%	60%	<1%
Semileptonic top (tau)	6%	10-15%	<1%
Z -> $\nu\nu$	8-13%	10-15%	<1%
Top + W	4-8%	15%	15%
Semileptonic top	5%	22%	<1%
Dileptonic top	10%	20%	50%

0-lepton mode

1-lepton mode

Assuming 1 fb^{-1}

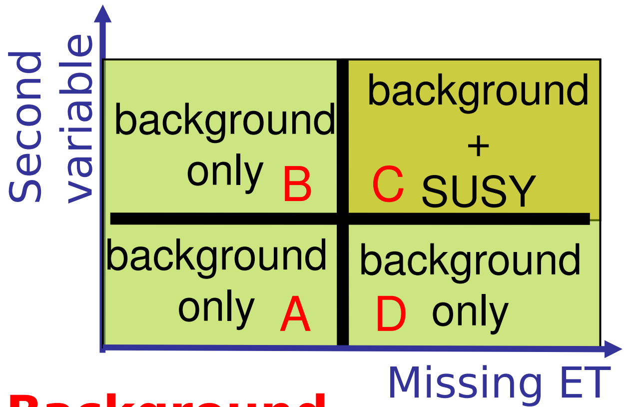
- Presence of SUSY will affect background estimates, however SUSY excess will be larger (even with 1fb-1)
- **Data-driven estimation methods are necessary to keep background under control and key to SUSY discovery**

Spare slides

0-lepton mode

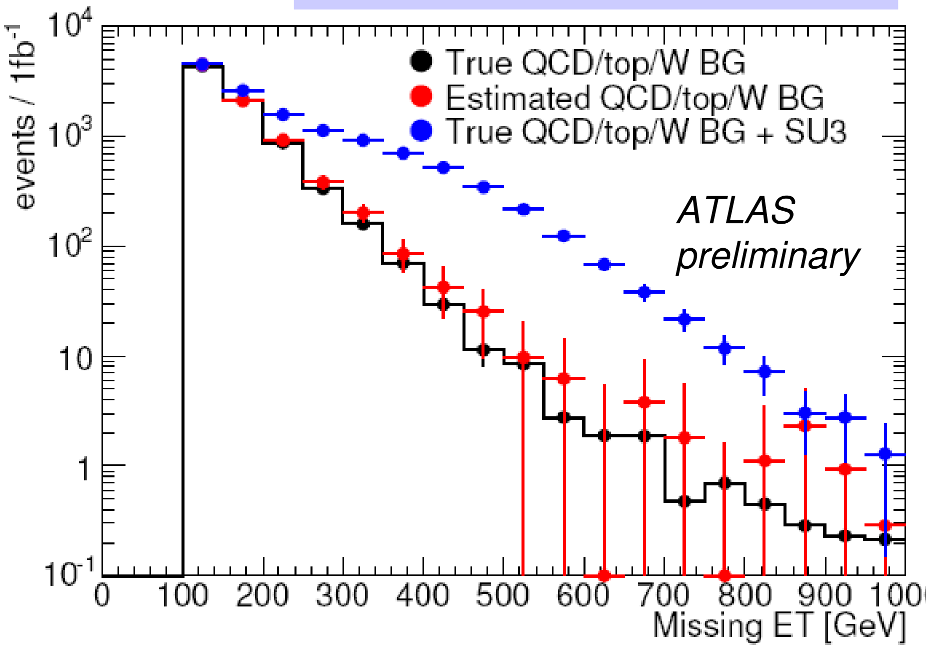
tt + W: transverse mass

- Semileptonic top can contribute to 0-lepton mode searches when the lepton is not identified
  - Tau, out of acceptance, inside jet
- Control sample
  - SUSY selection + MT < 100 GeV + 1 lepton
- The isolated lepton is then removed from the event, and all kinematic variables recalculated
- Normalization
  - 100 GeV < MET< 200 GeV
- QCD estimation also included
- SUSY contamination:
  - extract from control sample



Background in C = D x B/A

Systematic uncertainties ~15%



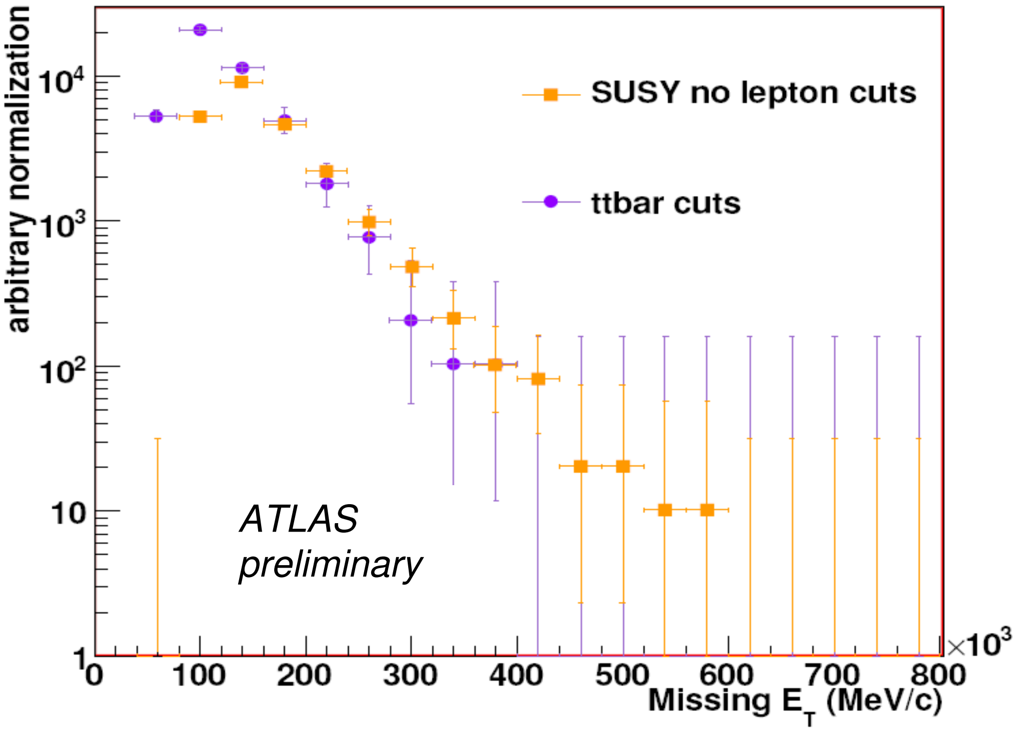
Semileptonic tt (with tau)

- Independent event reconstruction on hadronic and leptonic side
  - Hadronic top: W (dijet combination with mass closest to PDG value) + closest b-jet (in ΔR)
  - Leptonic W: tau + MET (collinear approximation)

Statistic uncertainties ~6%

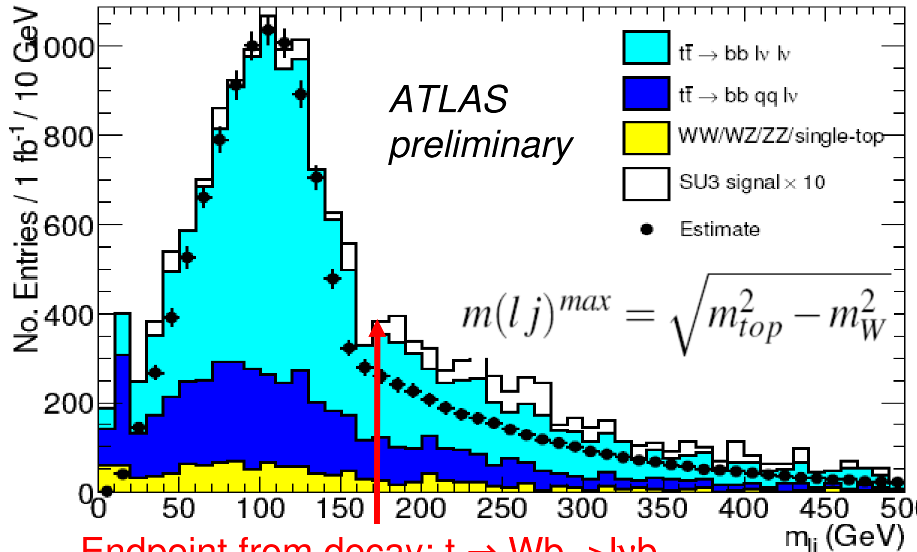
Systematic uncertainties ~15%

Systematic variation	Cross section variation (%)
Jet Energy Scale	2.5
b-tagging efficiency	7.5
light quark rejection in b-tag	1.3
τ-ID efficiency	3.4
light quark rejection in τ-ID	4.5



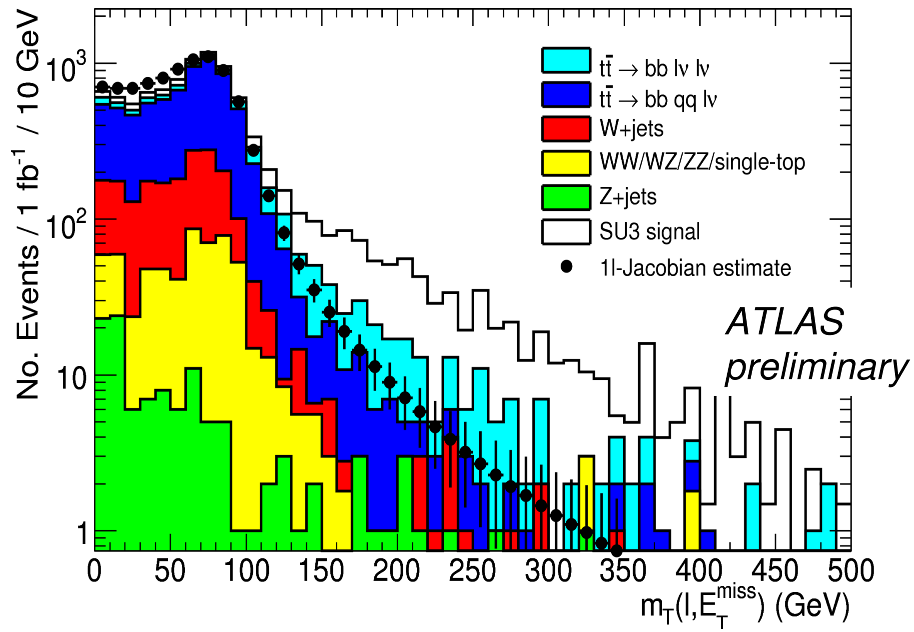
Dileptonic tt: top redecay

- Dileptonic top selection
  - J45_xE50 jet + MET trigger
  - 2 jets with pt > 20 GeV
  - 2 OS leptons pt > 20 GeV
  - MET < ½ (pt(lepton1) + pt(lepton2))
  - mass(lepton,jet) < 155 GeV
  - Solve system for p(ν)

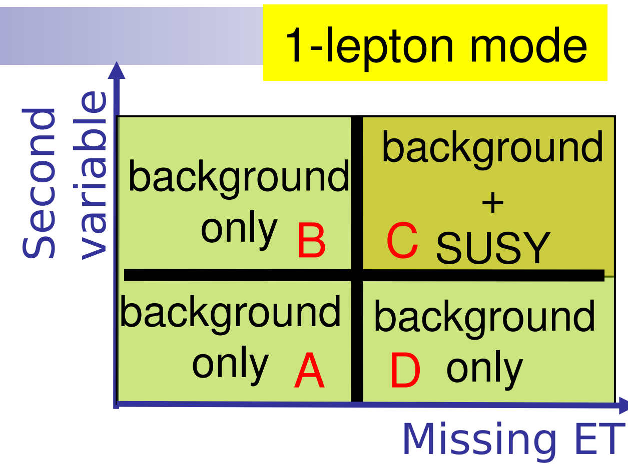


Endpoint from decay: t → Wb ->lνb (neglecting m_b)

- Semileptonic top, W, Z contribution estimated from MET distribution from events with MT < 100 GeV
  - hard MT cut (MT>150 GeV) → semi-leptonic background is sub-dominant.
  - events in Jacobian peak smeared with MC function to simulate tail of MT distribution



tt + W: transverse mass

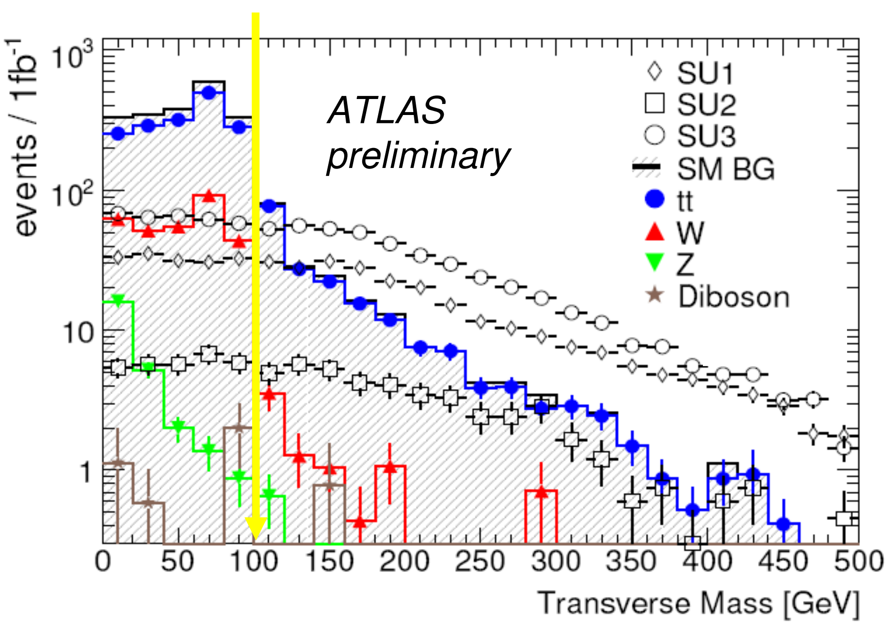


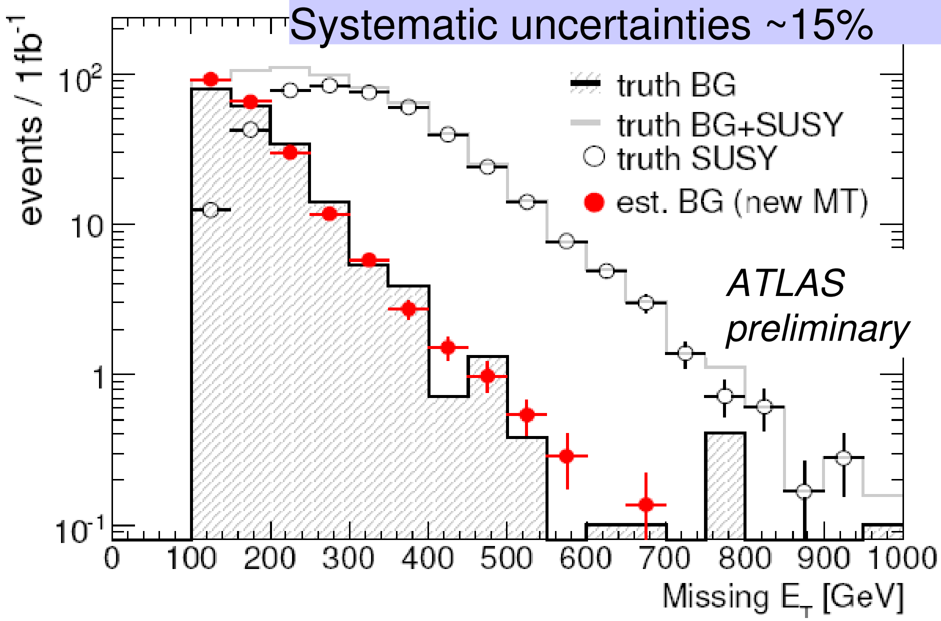
Background in C = D x B/A

- Transverse mass and MET uncorrelated
- Control sample
  - SUSY selection + MT < 100 GeV
- SUSY contamination: extract from control sample
  - assume same SUSY signal ratio in control and signal region for all SUSY samples

	Syst. error
Jet energy scale	$< 5\%$
Lepton ID efficiency	7%
MC@NLO vs ALPGEN	8%
MC parameter variation (ALPGEN)	$< 5\%$

Systematic uncertainties ~15%





1-lepton mode

Dileptonic tt with one misidentified lepton: HT2

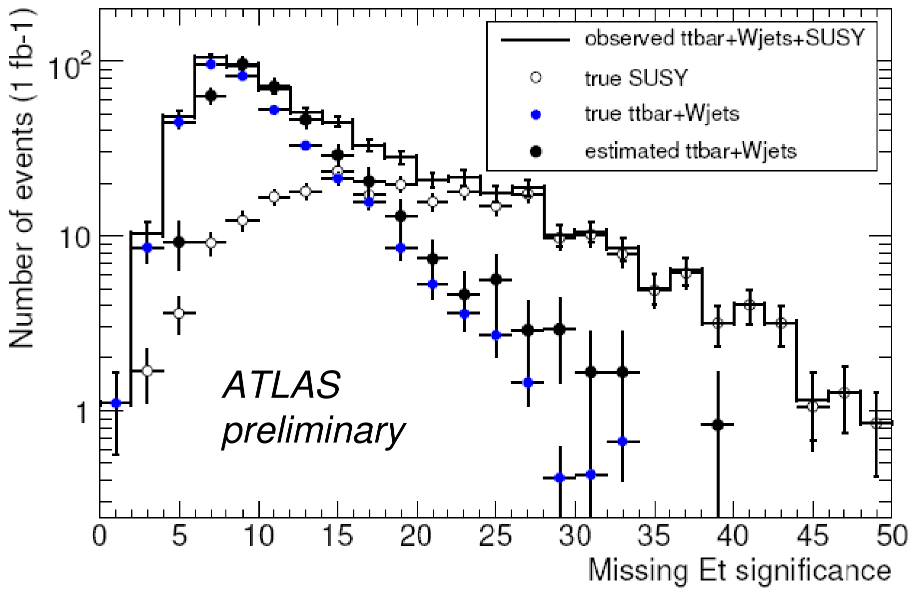
- Control sample
  - SUSY selection + HT2 < 300 GeV

$$\mathrm{HT2} = \sum_{i=2}^{4} p_T^{\mathrm{jet}i} + p_T^{\mathrm{lepton}}$$

- MET significance uncorrelated to HT2

$$\not{E}_T/[0.49 \cdot \sqrt{\sum E_T}$$

- Normalization region:
  - HT2 > 300 GeV and 8<MET significance<14 (low MET region)



Systematic uncertainties (MC) ~20%

Systematic uncertainties (detector) ~20%

1-lepton mode

Semileptonic tt: top box

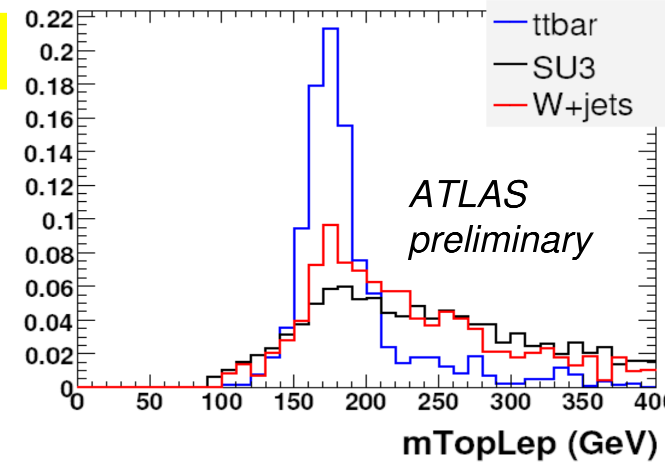
- Reconstruct leptonic W assuming neutrino from W responsible for all MET
- Reconstruct “best” (mass closest to top mass) leptonic top with one of the leading jets
- Reconstruct best hadronic W with the three remaining leading jets
- Reconstruct best hadronic top
- Top box cuts (define control sample)

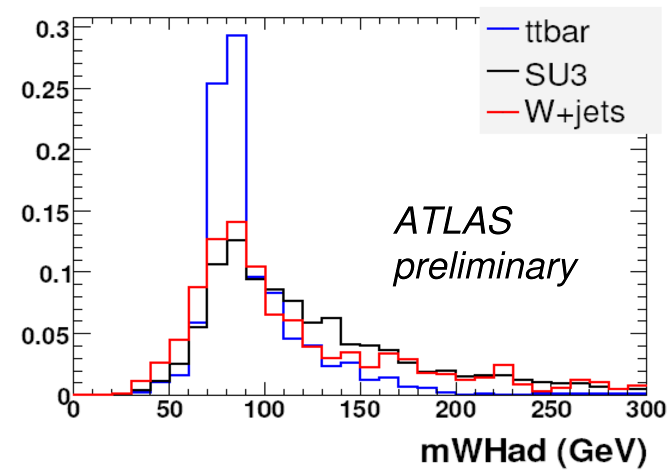
$$|M_{Top-lep} - M_{Top}| < 25 \text{ GeV}$$
$$|M_{W-had} - M_{W}| < 15 \text{ GeV}$$
$$|M_{Top-had} - M_{Top}| < 25 \text{ GeV}$$

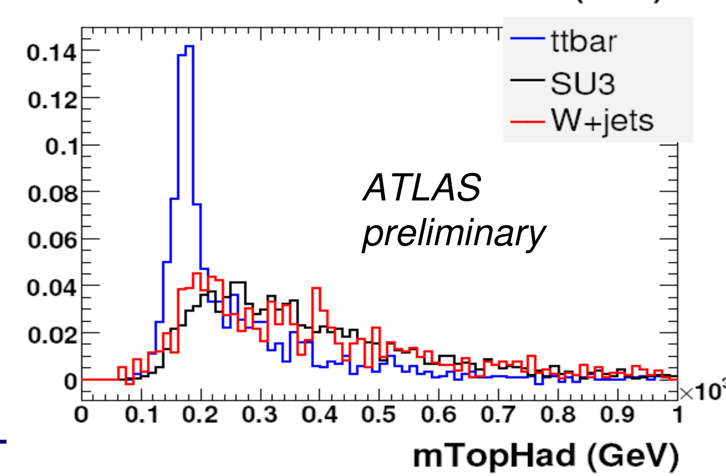
- Extrapolation to signal region using MC

Source	Contribution %
Jet energy scale	20
$\not{E}_T$ scale	2
MC Model dependence of R_{tt}	8

Systematic uncertainties ~22%





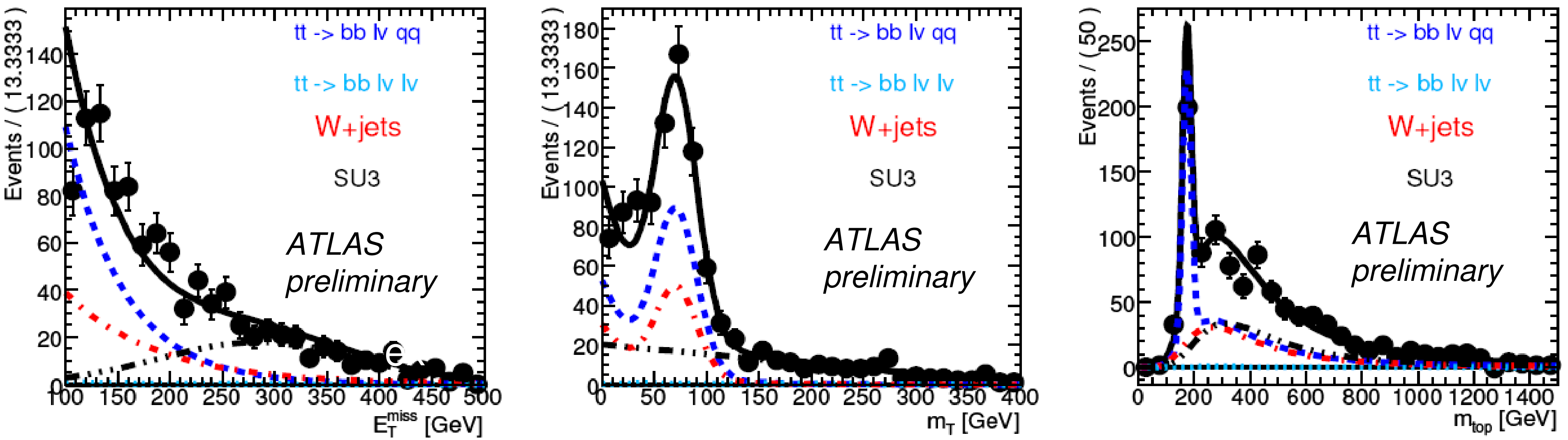


1-lepton mode

tt + W: combined fit

- Fit three observables: **MET, MT and Mtop** (invariant mass of 3 jets with largest vector PT sum)
- **Sideband:** SUSY selection + MT < 150 GeV OR MET < 200 GeV
- **Signal:** SUSY selection + MT > 150 GeV AND MET > 200 GeV
- All SUSY models (except SU4) have similar behavious in SB region in MT and MET → build a model background only vs background+SUSY
- Relax all parameters except the SUSY ansatz shape

Systematic uncertainties ~20%

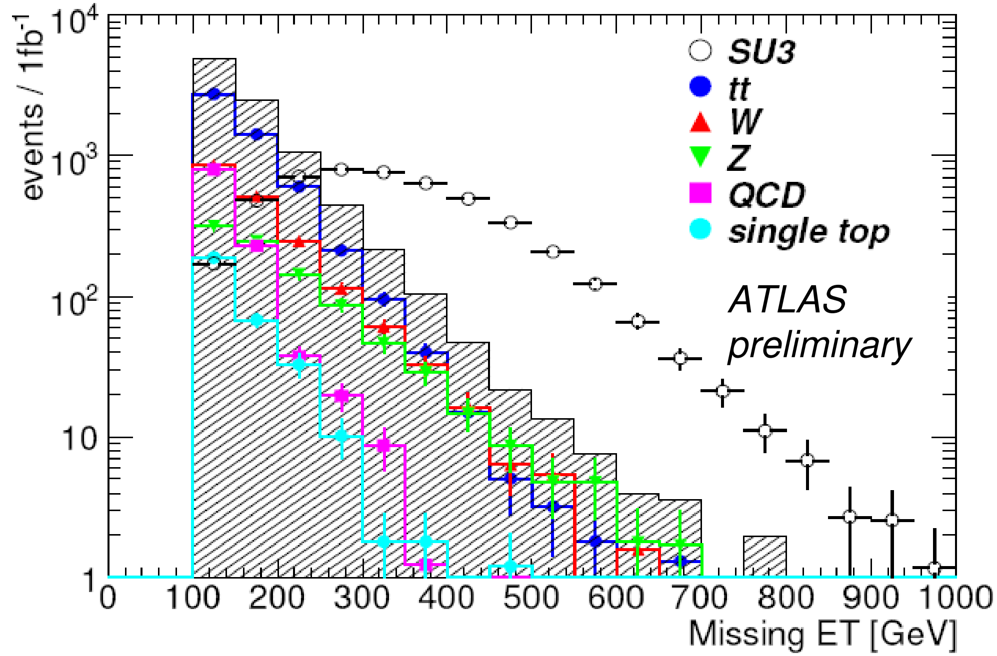


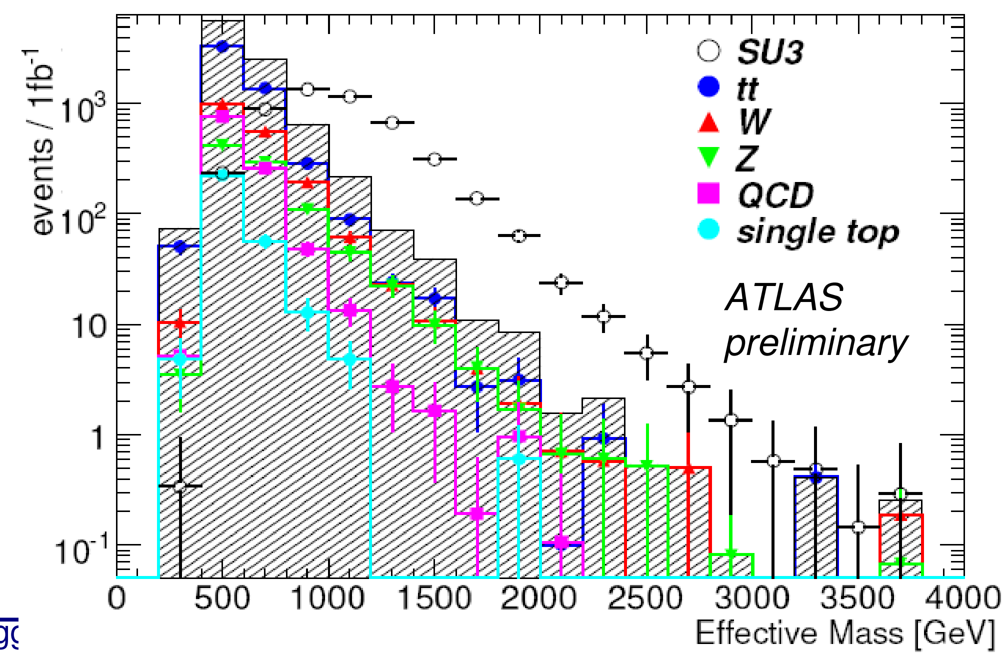
0-lepton search mode

- Selection cuts:
  - at least 4 jets with PT>50GeV
  - at least 1 jet with PT>100GeV
  - 0 lepton (e, μ) with PT > 20 GeV
  - MET > 100 GeV
  - MET > 0.2 effective mass
  - Transverse Sphericity ST > 0.2
  - Δφ(ET – jet *i*) > 0.2 (*i* = 1, 2, 3)

- Main backgrounds:
  - tt
  - W+jets
  - Z+jets
  - QCD

SM	0-l
tt	62%
W	17%
Z	10%
QCD	10%



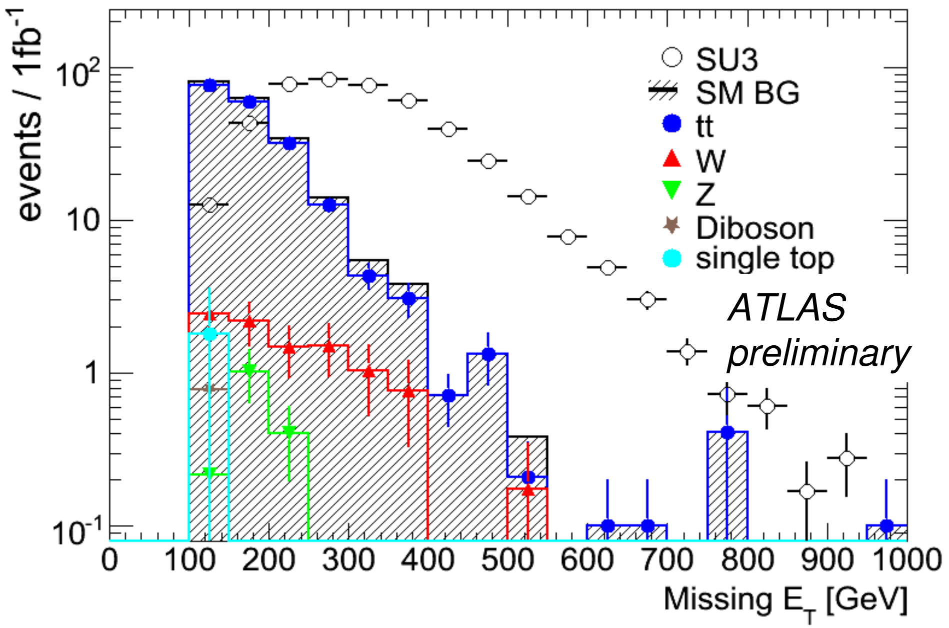


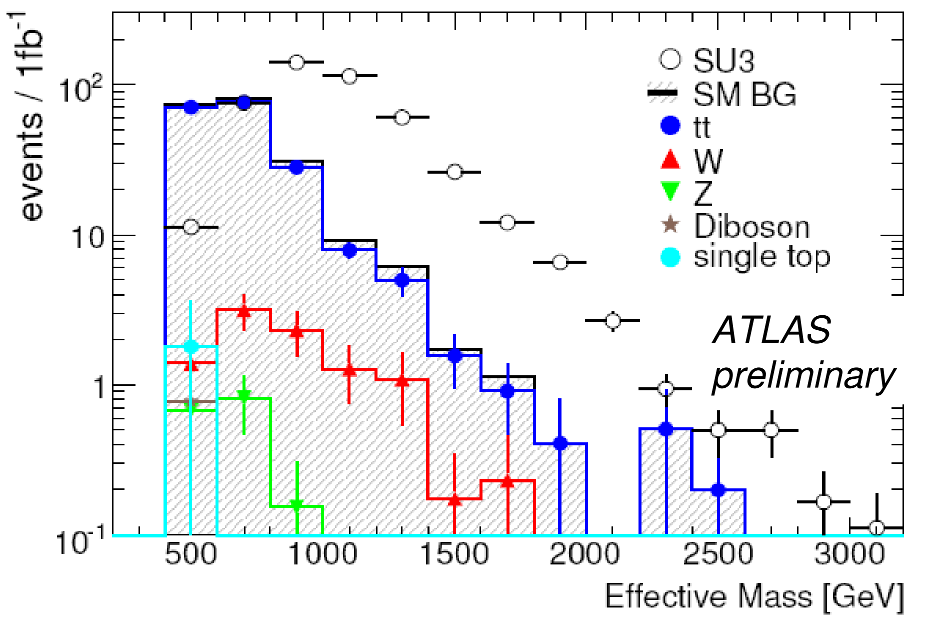
1-lepton search mode

- Selection cuts:
  - at least 4 jets with PT>50GeV
  - at least 1 jet with PT>100GeV
  - 1 lepton (e, μ) with PT > 20 GeV
  - MET > 100 GeV
  - MET > 0.2 effective mass
  - Transverse Sphericity ST > 0.2
  - transverse mass(lepton, ET) > 100GeV

- Main backgrounds:
  - tt
  - W+jets

SM	1-l
tt	91%
W	7%
Z	1%
QCD	<1%





Object definition

- **Electrons**
  - Pt > 10 GeV and |eta|<2.5
  - Veto on events with an electron in the crack (1.37<|eta|<2.5)
  - Calorimeter isolation in a cone (0.2) <10 GeV
  - Angular distance to closest jet > 0.4 (after overlap removal)
- **Muons**
  - Pt > 10 GeV and |eta|<2.5
  - Chi2 > 100
  - Calorimeter isolation in a cone (0.2) <10 GeV
  - Angular distance to closest jet > 0.4 (after overlap removal)
- **Jets**
  - Pt > 20 GeV and |eta|<2.5
- **Electron/Jet overlap removal**
  - Jets matching an electron within 0.2 cone
- **Transverse sphericity:** use all jets with |eta|<2.5 and leptons
- **Effective mass:** use 4 leading jets with |eta|<2.5 and leptons

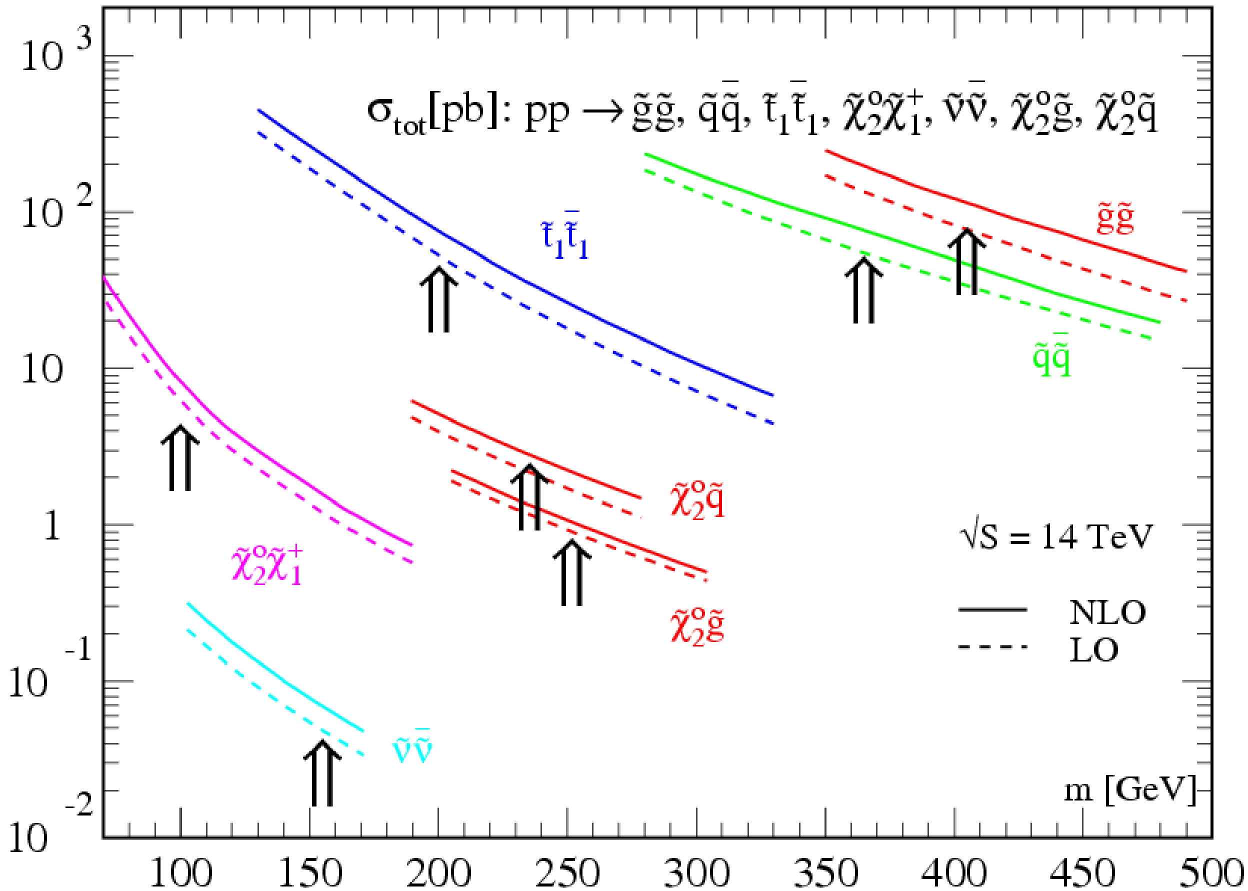
MC background estimation

- Will **ROUGHLY** be subject to the following uncertainties:
  - Underlying Event & Parton Distribution Functions 20%
  - Cross-sections 50%
    - No NLO calculations for tt
  - Parton Showering 50%
    - After accurate normalization to data has been made
  - Detector Calibration (JES, MET) 30%
  - Detector simulation 100%
  - Limited Monte Carlo statistics

Background estimation for multi-leptons analysis

- OS 2-lepton & tau searches
  - MT method
  - HT2 method
  - Top redecay
  - Top kinematic reconstruction
- SS 2-lepton searches
  - Lepton isolation

Cross sections at LHC



mSUGRA benchmark points

- We consider the following points in the mSUGRA parameter space:

SU1 $m_0 = 70$ GeV, $m_{1/2} = 350$ GeV, $A_0 = 0$, $\tan\beta = 10$, $\mu > 0$. Coannihilation region with nearly degenerate $\tilde{\chi}_1^0$ and $\tilde{\ell}$.

SU2 $m_0 = 3550$ GeV, $m_{1/2} = 300$ GeV, $A_0 = 0$, $\tan\beta = 10$, $\mu > 0$. Focus point region near boundary where $\mu^2 < 0$, so light Higgsions which annihilate efficiently.

SU3 $m_0 = 100$ GeV, $m_{1/2} = 300$ GeV, $A_0 = -300$ GeV, $\tan\beta = 6$, $\mu > 0$. Bulk region: relatively light sleptons enhance LSP annihilation.

SU4 $m_0 = 200$ GeV, $m_{1/2} = 160$ GeV, $A_0 = -400$ GeV, $\tan\beta = 10$, $\mu > 0$. Low mass point close to Tevatron bound.

SU6 $m_0 = 320$ GeV, $m_{1/2} = 375$ GeV, $A_0 = 0$, $\tan\beta = 50$, $\mu > 0$. Funnel region with $2M_{\tilde{\chi}_1^0} \approx M_A$. Since $\tan\beta \gg 1$, A is wide and τ decays dominate.

SU8.1 $m_0 = 210$ GeV, $m_{1/2} = 360$ GeV, $A_0 = 0$, $\tan\beta = 40$, $\mu > 0$. Variant of coannihilation region with $\tan\beta \gg 1$, so that only $M(\tilde{\tau}_1) - M(\tilde{\chi}_1^0)$ is small.

- For all these points, gluino mass < 1 TeV, and it's 6-8x neutralino mass. For all points except SU2, squark and gluino masses are comparable, therefore they are strongly produced and decay giving hard jets, leptons and MET.