Search for Dark Matter with top pair production in single-lepton channel

Zurich PhD Seminar 2014

September 11th - 12th

Deborah Pinna

Introduction

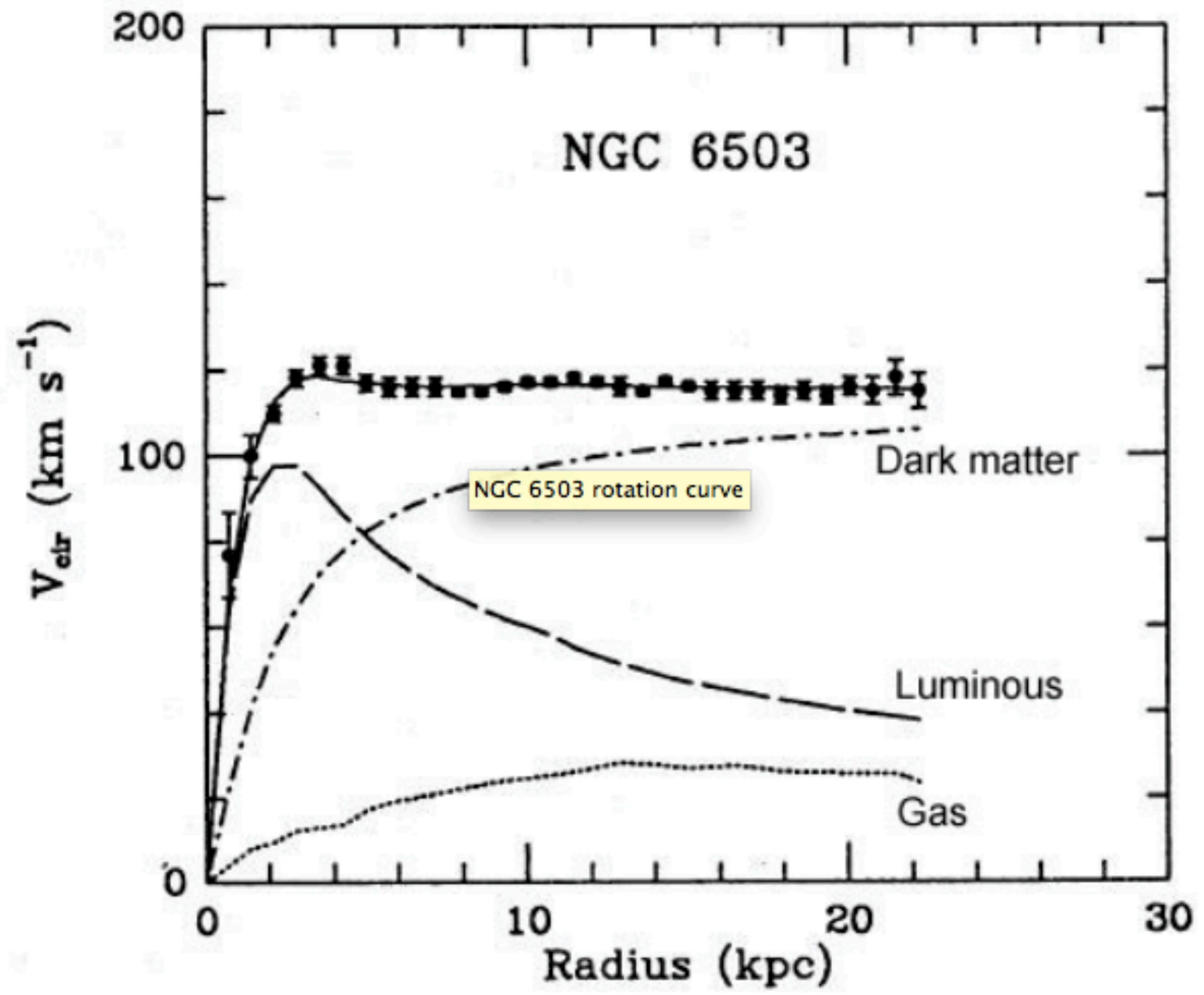
- Evidence at different observable length scales for Dark Matter (DM)

    dispersion velocity of galaxies in galactic cluster too large to be explained by luminous matter

    rotation curves on singular galaxies constant beyond luminous region

        velocity is expected to go like r $^{-1/2}$

        differences explained by non luminous mass (dark matter)



- Studies at different scales provide overwhelming evidence that Universe is composed mainly of non-luminous matter

    Dark matter abundance ~25% of the Universe, five times the amount of baryonic matter

- Evidence based on gravitational interactions, no information of what is the nature of Dark Matter

Introduction

- Dark Matter leading empirical evidence for new physics beyond the Standard Model (SM) in addition to problems like gauge hierarchy and neutrino mass

- Large variety of Dark Matter candidates motivated both by cosmology and particle physics

  the most studied is a **Weakly Interacting Massive Particle (WIMP)**
    neutral particle, mass in the range 10 GeV - TeV, weak interactions, correct relic density, may be detected in different ways

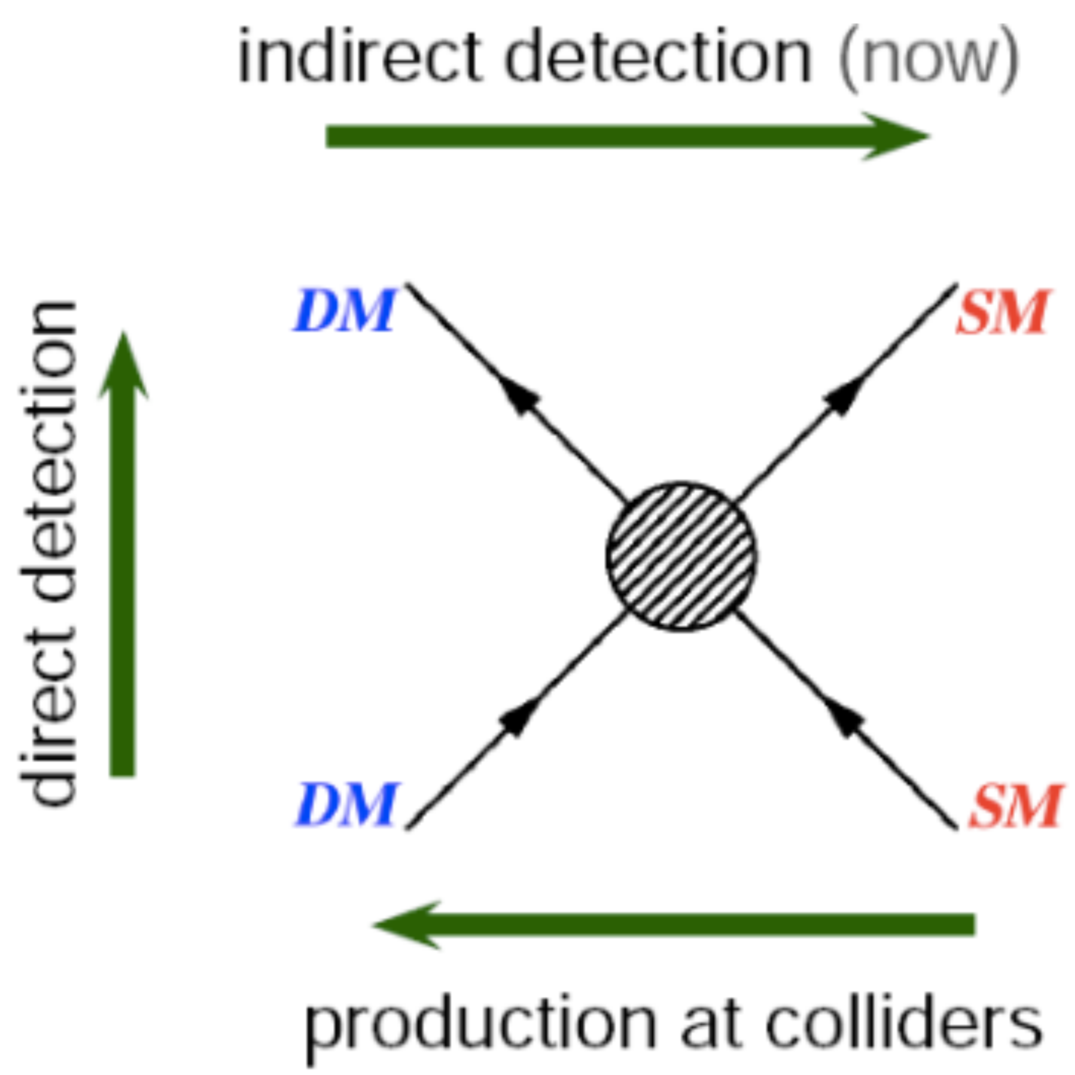
  many particle physics theories provide this candidate, important to perform model-independent DM searches
    Effective Field Theory (EFT), interaction between DM and SM particles described by effective operators

- Studies at largest and smallest observable length scales

  indirect searches: products of DM annihilations or decays

  direct searches : scattering DM-heavy nucleons

  collider searches: signature of DM production



Effective Field Theory

- EFT approach valid when the energy scale of considered SM-DM process (momentum transferred Q_{tr}) is small compared to the underlying microscopic interaction (cutoff scale M_*)

- Assuming DM is a Dirac fermion χ the effective operators are of form $\bar{q}\Gamma q\bar{\chi}\Gamma\chi$

Name	Initial state	Type	Operator
D1	qq	scalar	$\frac{m_q}{M_\star^3}\bar{\chi}\chi\bar{q}q$
D5	qq	vector	$\frac{1}{M_\star^2}\bar{\chi}\gamma^\mu\chi\bar{q}\gamma_\mu q$
D8	qq	axial-vector	$\frac{1}{M_\star^2}\bar{\chi}\gamma^\mu\gamma^5\chi\bar{q}\gamma_\mu\gamma^5 q$
D9	qq	tensor	$\frac{1}{M_\star^2}\bar{\chi}\sigma^{\mu\nu}\chi\bar{q}\sigma_{\mu\nu}q$
D11	gg	scalar	$\frac{1}{4M_\star^3}\bar{\chi}\chi\alpha_s(G^a_{\mu\nu})^2$

these operators encode the underlying microscopic interaction

the interaction strength depends on the parameter M_* (but not the kinematics)

more precise informations comes from knowing details of the UV theory

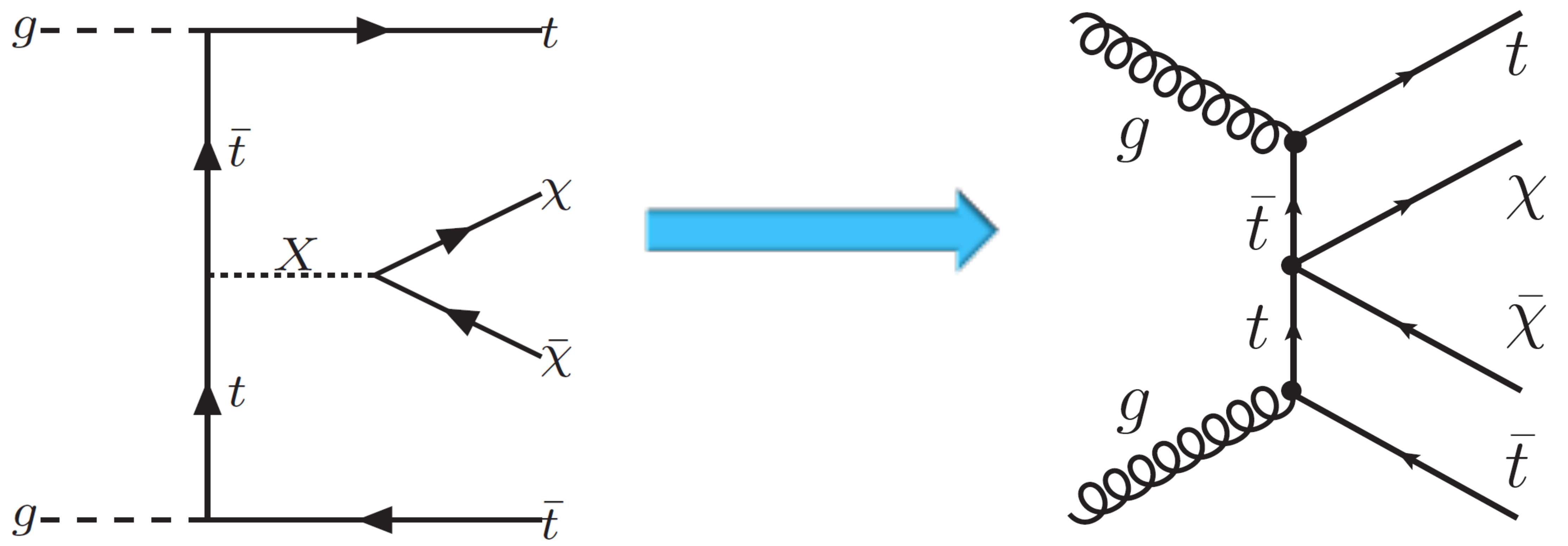
- Scalar interaction
proportional to quark mass, better constraints from events where DM couples to massive quarks

- Study of production of DM in association with top quark pair

Effective Field Theory

- Supposing a heavy mediator of mass M in the s-channel coupling to DM and SM with couplings g_1 g_2.

  Considering only the lowest order operators in the EFT approach is connected to the propagator expansion



$$\frac{g_1 g_2}{Q_{tr}^2 - M^2} = -\frac{g_1 g_2}{M^2}\left(1 + \frac{Q^2}{M^2} + \mathcal{O}\left(\frac{Q_{tr}^4}{M^4}\right)\right) \simeq -\frac{g_1 g_2}{M^2} \text{ for } Q_{tr}^2 << M^2$$

the coefficient of the effective operator should match to reproduce the UV theory, i.e. for D1 $M_* = \left(\frac{m_q M^2}{g_1 g_2}\right)^{1/3}$

Effective Field Theory

- In general the EFT field theory is valid when $Q_{tr} << M_*$

- The validity of the truncation of the propagator expansion requires

$$Q_{tr} < M$$

from the assumed UV details (heavy mediator, s-channel) $\frac{M_*^3}{m_q} > \frac{M^2}{g_1 g_2}$

from kinematics $Q_{tr} > 2m_\chi$

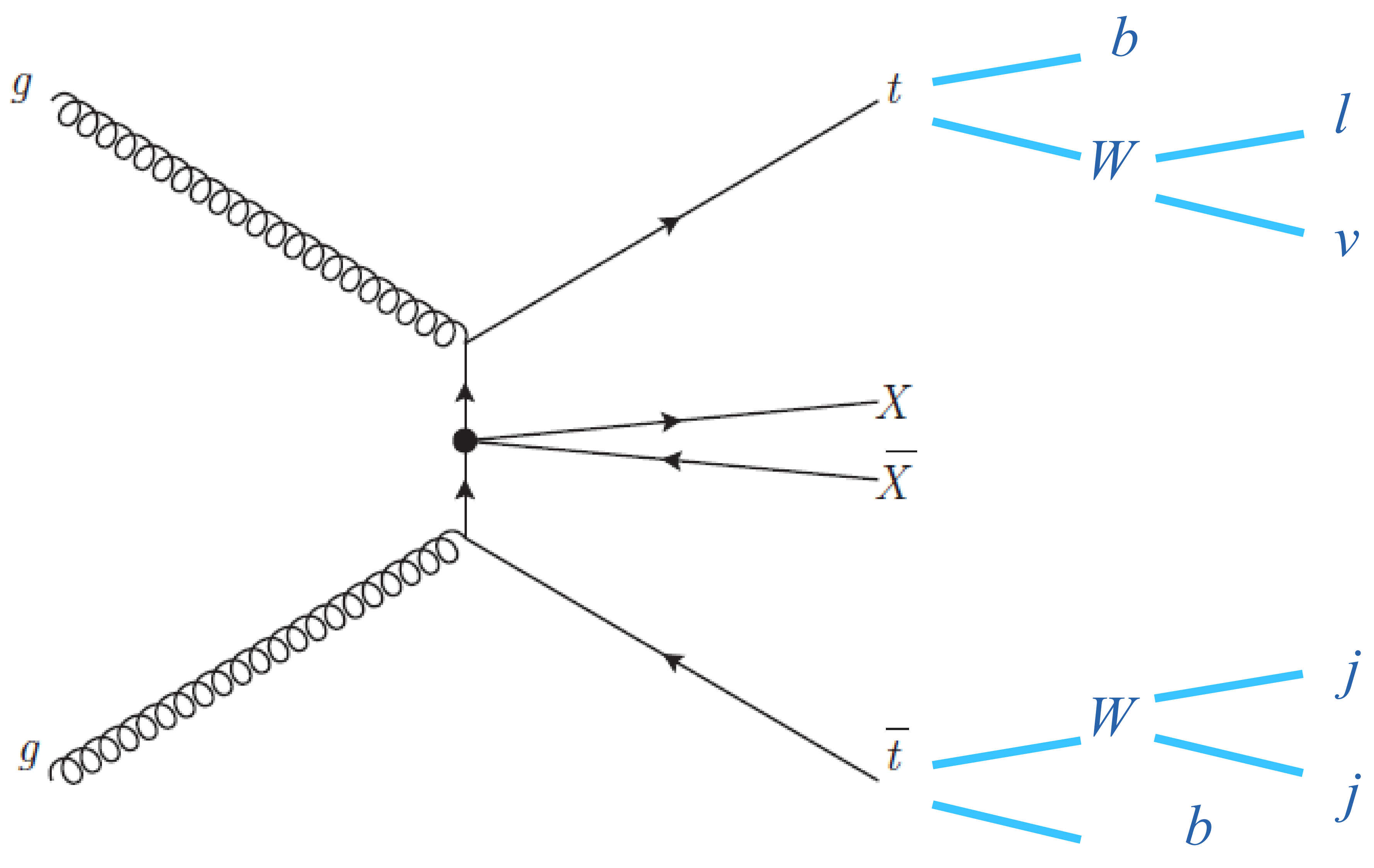
assuming most strongly coupled scenario in the perturbative regime $\sqrt{g_1 g_2} < 4\pi$

$$\sqrt{\frac{M_*^3}{m_q}} > \frac{M_\chi}{2\pi}$$

This is a very minimal requirement on M_* and it depends on the details of the UV completion

Scalar interaction: DM + top production

- Study of production of DM in association with top quark pair (semi-leptonic channel)



Signature

large unbalanced momenta in the transverse plane (MET)

one reconstructed lepton (from W decay)

Multiple jets and 2 jets from fragmentation and hadronization of b-quark from top

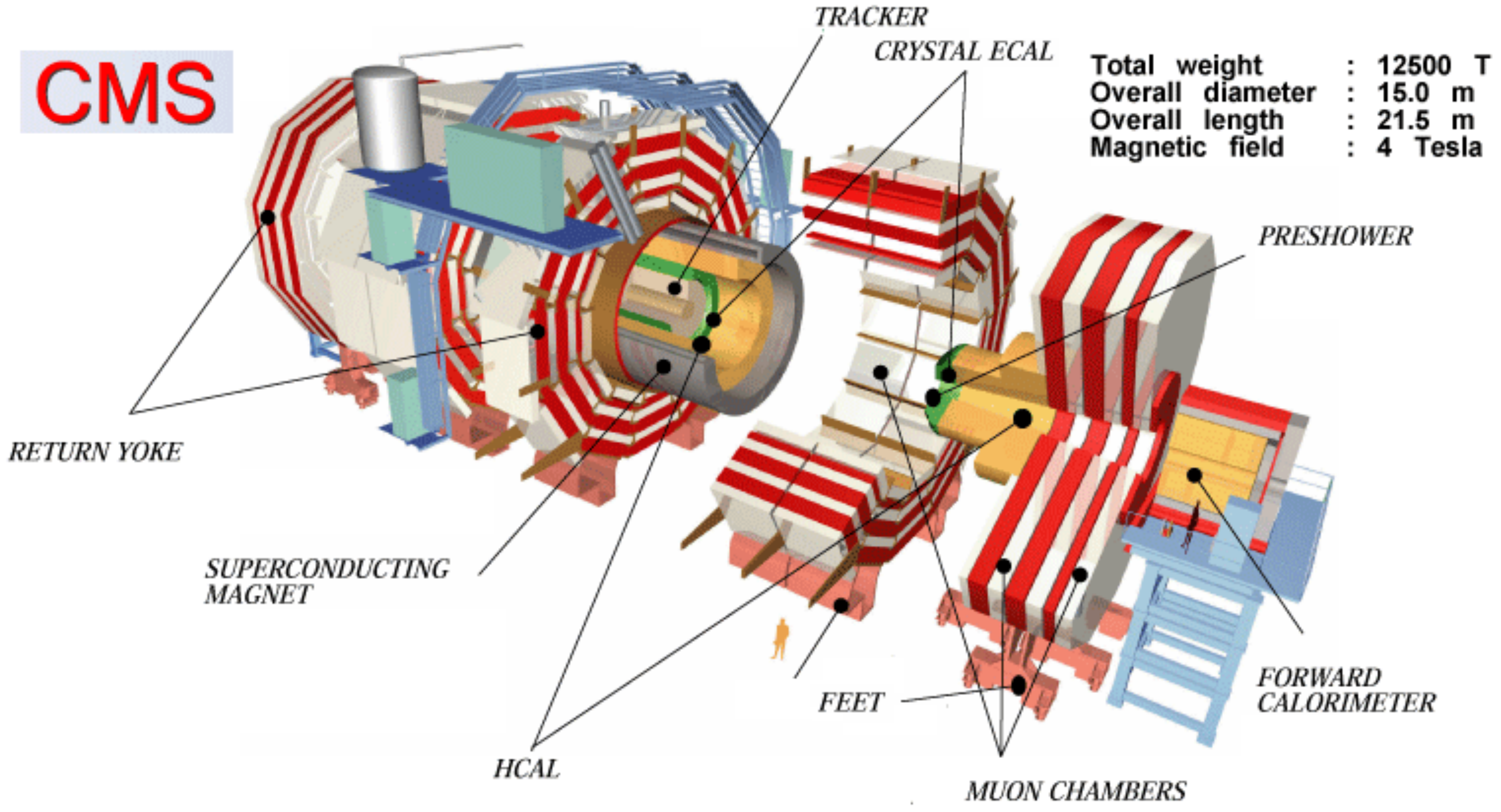
- SM processes that can mimic signatures of the process under study are referred as background processes

  tt+jets, W+jets, Z+jets (Drell-Yan), single top, WW/WZ/ZZ (dibosons)

Experimental apparatus: CMS detector

- Compact Muon Solenoid (CMS) one of four main experiments at LHC

  Is a multi-purpose experiment: aims to study unknown physics beyond the SM, while provides excellent performances for SM studies



Search performed using data collected by CMS experiment during 2012

center of mass energy 8 TeV

integrated luminosity 19.7 fb^{-1}

Di-lepton channel already investigated

CMS-PAS-B2G-13-044

Single-lepton channel presented in this talk

CMS-PAS-B2G-14-004

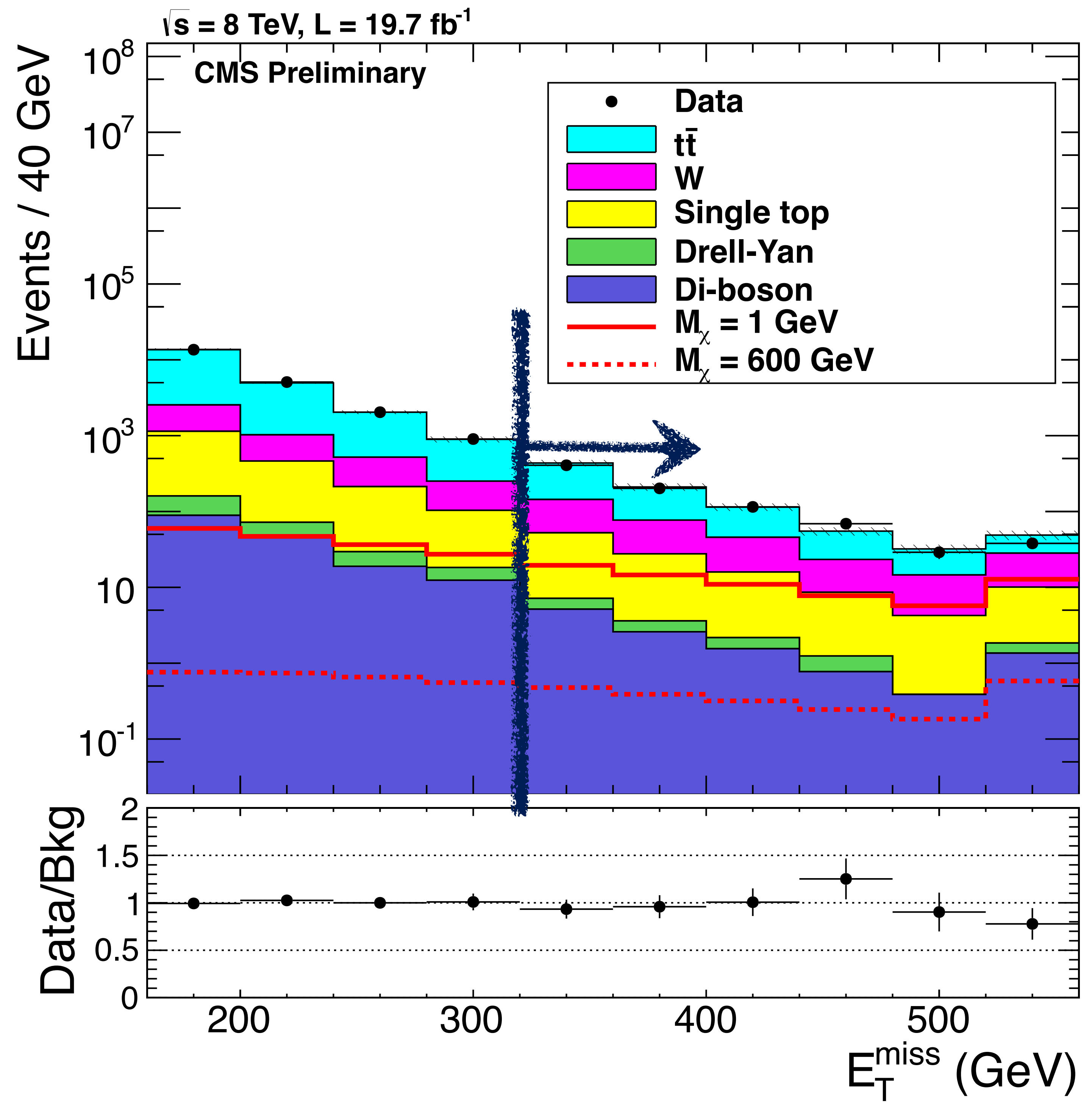
Event selection

- Pre-selection

  1 lepton
  at least 3 jets of which at least one is b-tagged (Combined Secondary Vertex medium working point)
  Missing Transverse Energy (MET) > 160 GeV

- Selection: contamination from backgrounds can be reduced using variables discriminating on the kinematics

Variable	Cut
MET	> 320 GeV
$m_T = \sqrt{2p_T^{lep} E_T^{miss}(1-\cos(\Delta\phi))}$	> 160 GeV
$\min(\Delta\Phi_{j1,MET}, \Delta\Phi_{j2,MET})$	> 1.2
M_{t2W}	> 200 GeV



signal has large MET from DM particles which escape detector

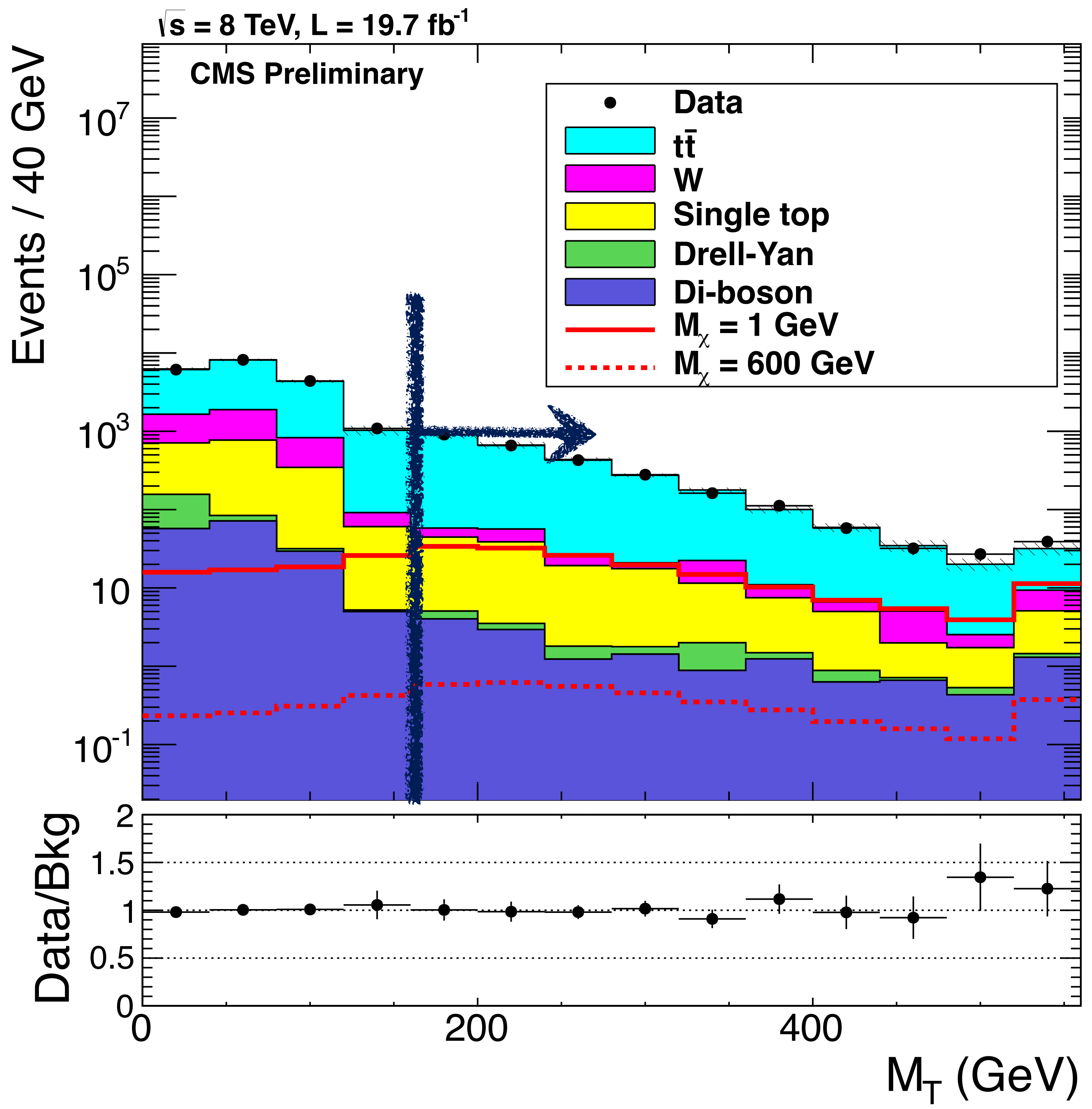
Event selection

- Pre-selection

  1 lepton
  at least 3 jets of which at least one is b-tagged
  Missing Transverse Energy (MET) > 160 GeV

- Selection: contamination from backgrounds can be reduced using variables discriminating on the kinematics

Variable	Cut
MET	> 320 GeV
$m_T = \sqrt{2p_T^{lep} E_T^{miss}(1-\cos(\Delta\phi))}$	> 160 GeV
$\min(\Delta\Phi_{j1,MET}, \Delta\Phi_{j2,MET})$	> 1.2
M_{t2W}	> 200 GeV



Most W+jets and tt+jets semi-leptonic events $M_T < M_W$. Signal events distribution peaks at higher values

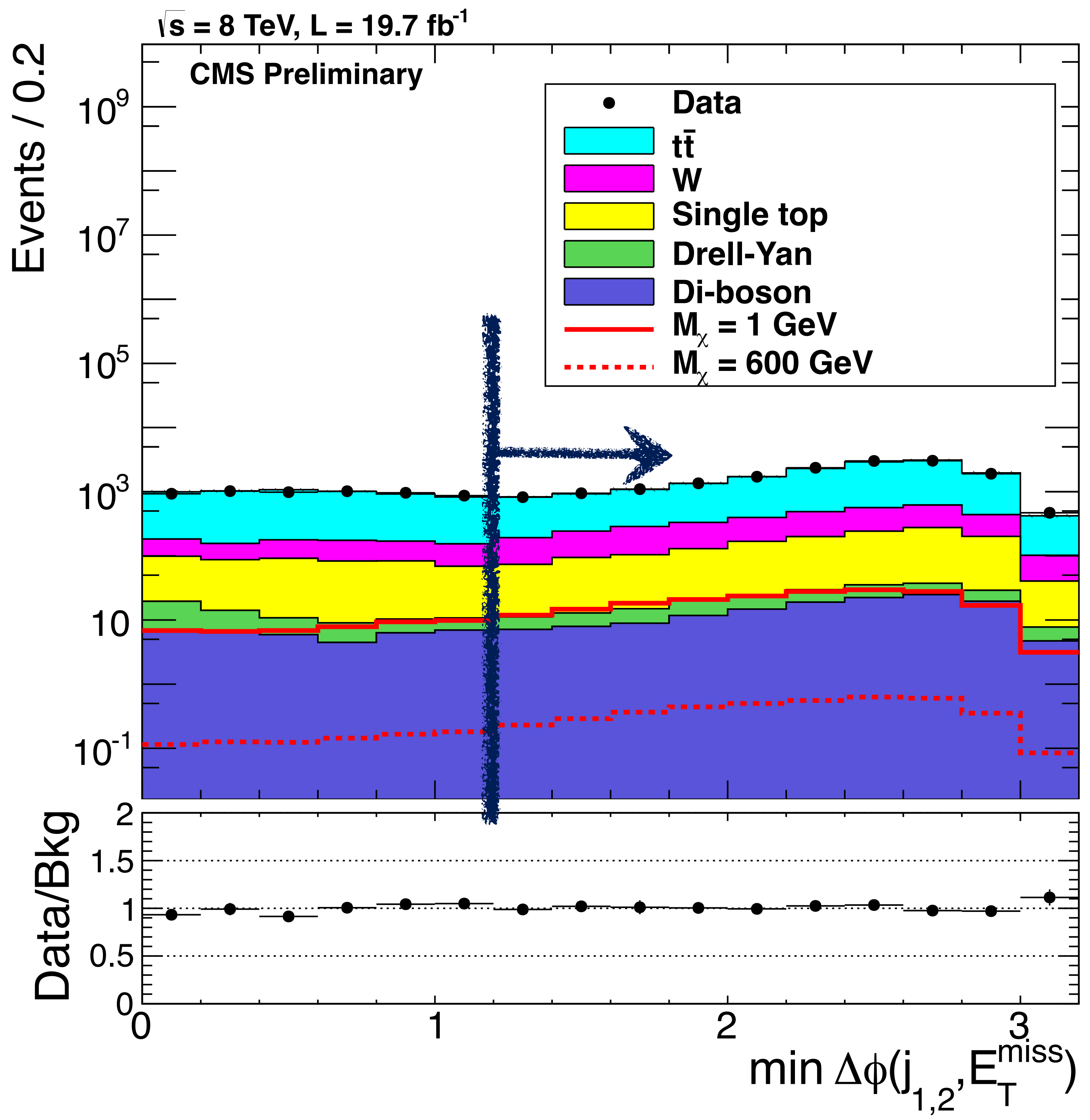
Event selection

- Pre-selection

  1 lepton
  at least 3 jets of which at least one is b-tagged
  Missing Transverse Energy (MET) > 160 GeV

- Selection: contamination from backgrounds can be reduced using variables discriminating on the kinematics

Variable	Cut
MET	> 320 GeV
$m_T = \sqrt{2 p_T^{lep} E_T^{miss} (1 - \cos(\Delta\phi))}$	> 160 GeV
$\min(\Delta\Phi_{j1,MET}, \Delta\Phi_{j2,MET})$	> 1.2
M_{t2W}	> 200 GeV



The jets and the MET tends to be more separated in Φ in signal events than in tt and in single top events

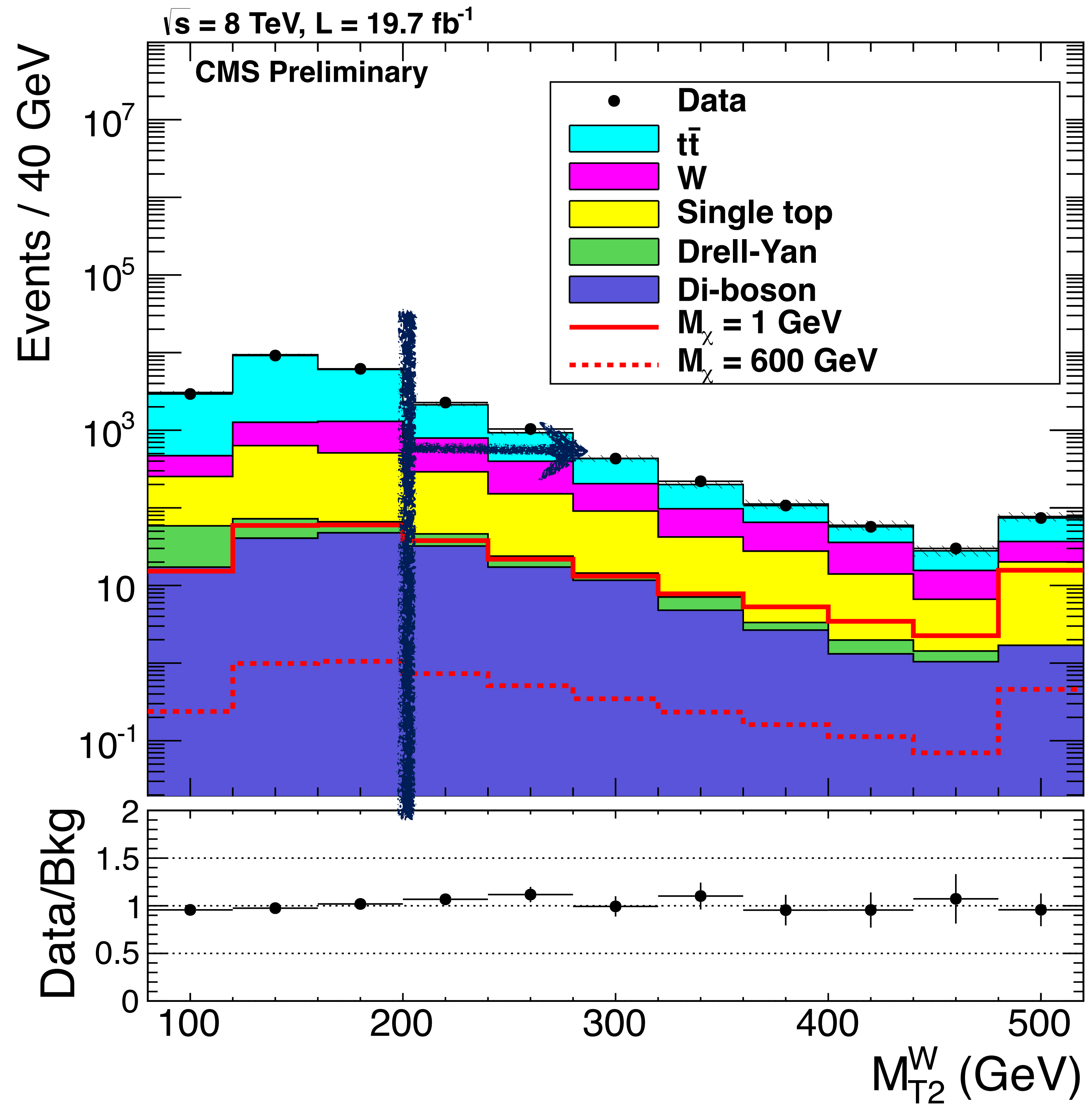
Event selection

- Pre-selection

  1 lepton
  at least 3 jets of which at least one is b-tagged
  Missing Transverse Energy (MET) > 160 GeV

- Selection: contamination from backgrounds can be reduced using variables discriminating on the kinematics

Variable	Cut
MET	> 320 GeV
$m_T = \sqrt{2 p_T^{lep} E_T^{miss} (1 - \cos(\Delta\phi))}$	> 160 GeV
$\min(\Delta\Phi_{j1,MET}, \Delta\Phi_{j2,MET})$	> 1.2
M_{t2W}	> 200 GeV



For most tt+jets di-leptonic events $M_{T2W} < M_{top}$. Signal events distribution shows higher tails

Background control

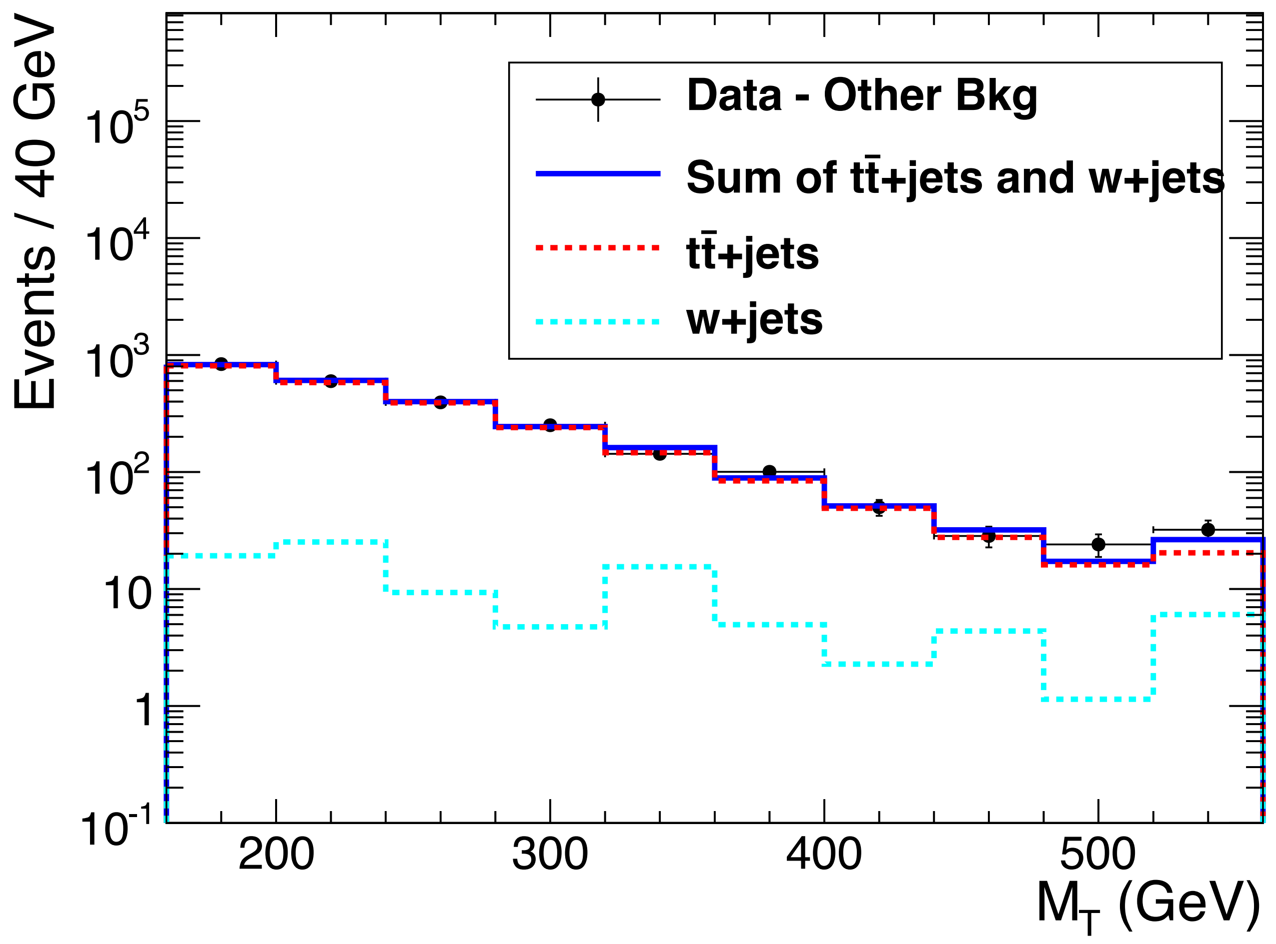
- Background estimation

dominant background from tt+jets and W+jets: from simulation with data-to-simulation scale factors (SF)
other small background: from simulation

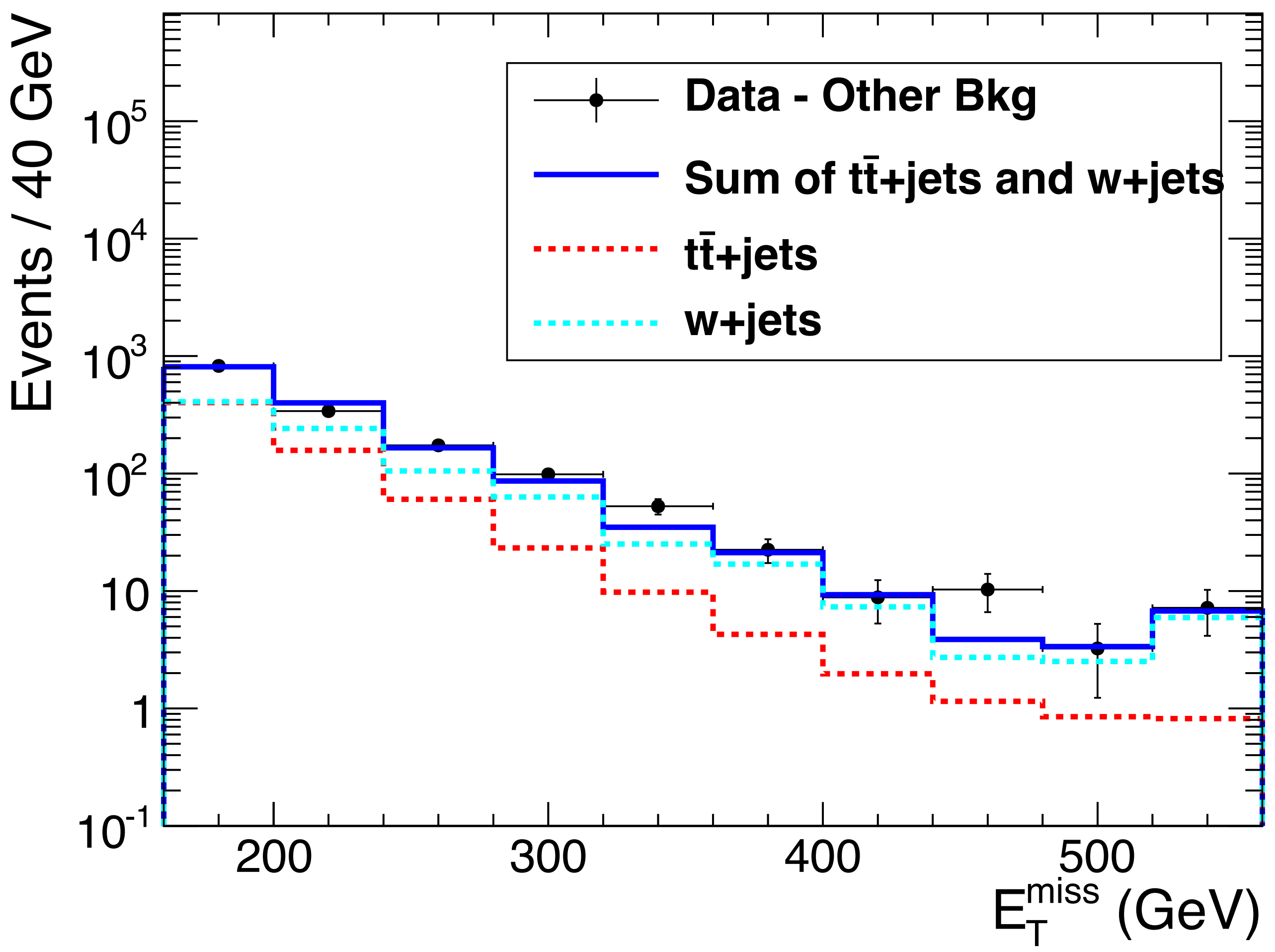
tt+jets and W+jets SFs

SFs are extracted in control regions (CRs) enriched in background composition and with negligible contribution from signal fitting simultaneously two different simulated templates distributions to data

CR1 (tt+jets enriched)



CR2 (W enriched)



CR1: pre-selection + $M_T > 160$ GeV
CR2: as CR1 but 0 b-tag

SF(tt+jets) 1.11 ± 0.02 (stat)
SF(W+jets) 1.26 ± 0.06 (stat)

Background control

- predicted background yields and uncertainties are then propagated from CR to signal region (SR)

$$\mathrm{N(bkg\ pred.\ in\ SR)} = SF(CR) \cdot \mathrm{N(bkg\ from\ MC\ in\ SR)}$$

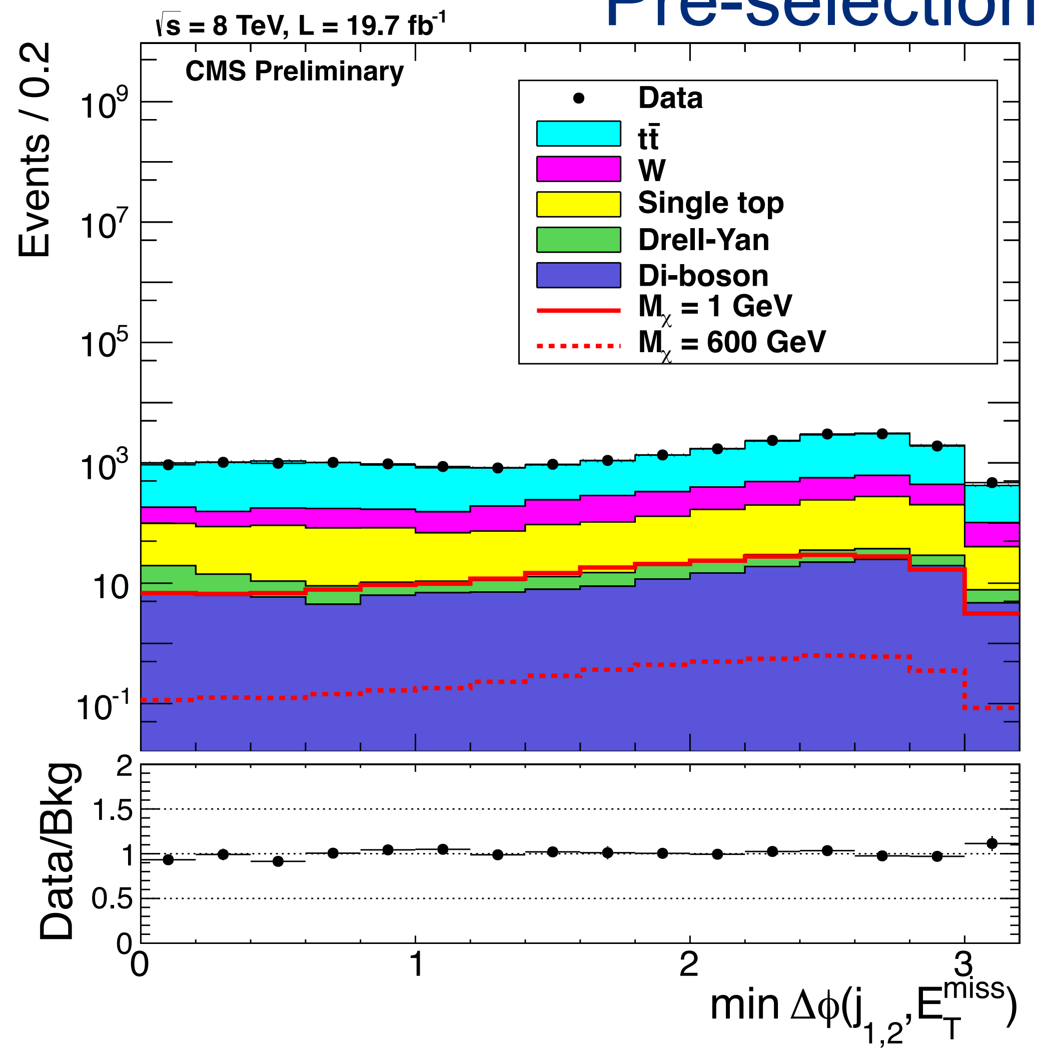
- Systematics

normalization unc. already covered by SFs
shape unc. constrained in CRs and SFs propagated in SR, differences taken as systematic

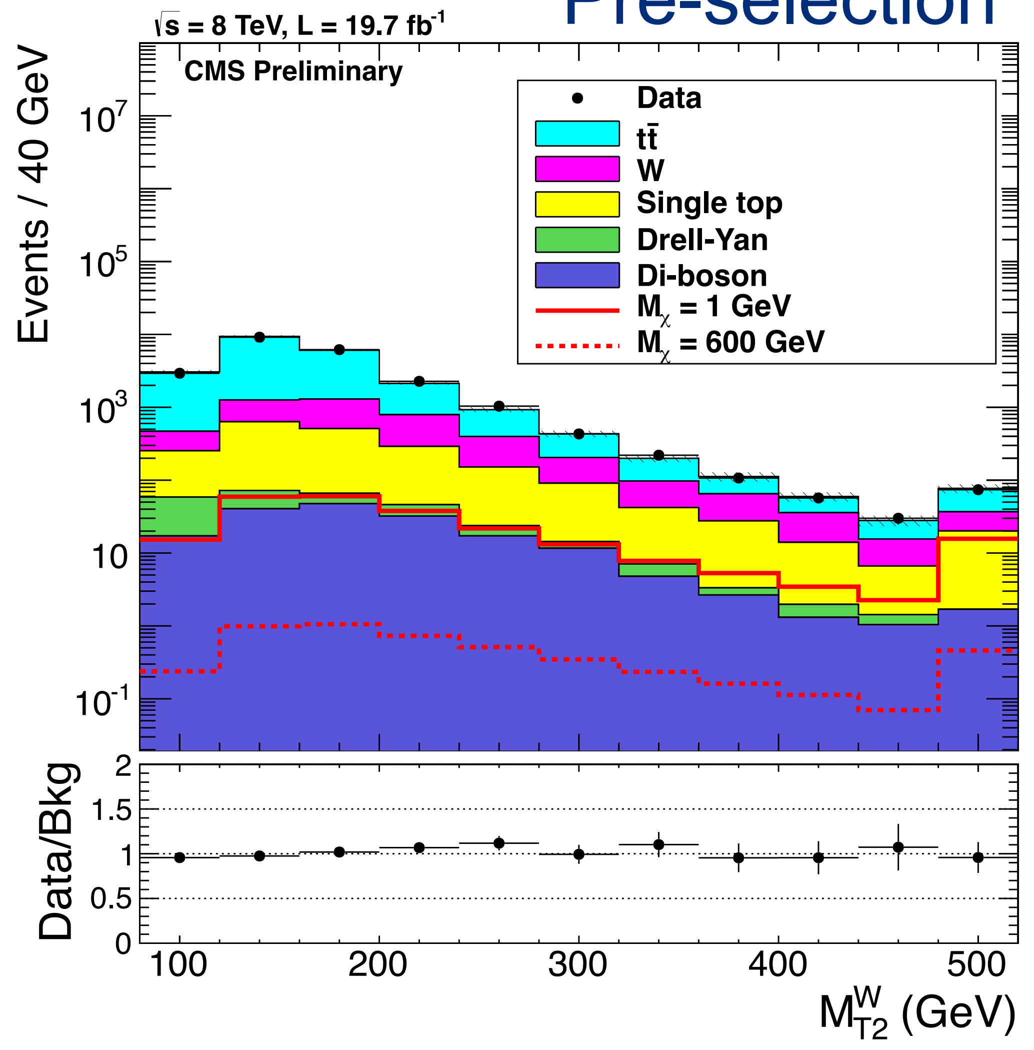
Main systematics tt+jets: top p_T reweighing (14%)
W+jets jet energy scale (11%)

- Agreement between data and simulation after SFs is used to check the validity of method

Pre-selection



Pre-selection



in the distributions a good agreement between data and simulation is observed after SFs are applied

Results: number of events

- Number of events after full selection

Source	Yield (± stat. ± syst. unc.)
Data	18
Signal	$38.3 \pm 0.7 \pm 2.1$
Total Bkg	$16.4 \pm 2.2 \pm 2.7$
$t\bar{t}$	$8.2 \pm 0.6 \pm 1.9$
W	$5.2 \pm 1.7 \pm 0.6$
Single top	$2.3 \pm 1.1 \pm 1.1$
Di-boson	$0.5 \pm 0.2 \pm 0.2$
Drell-Yan	$0.3 \pm 0.3 \pm 0.1$

- No excess of events observed in the search region

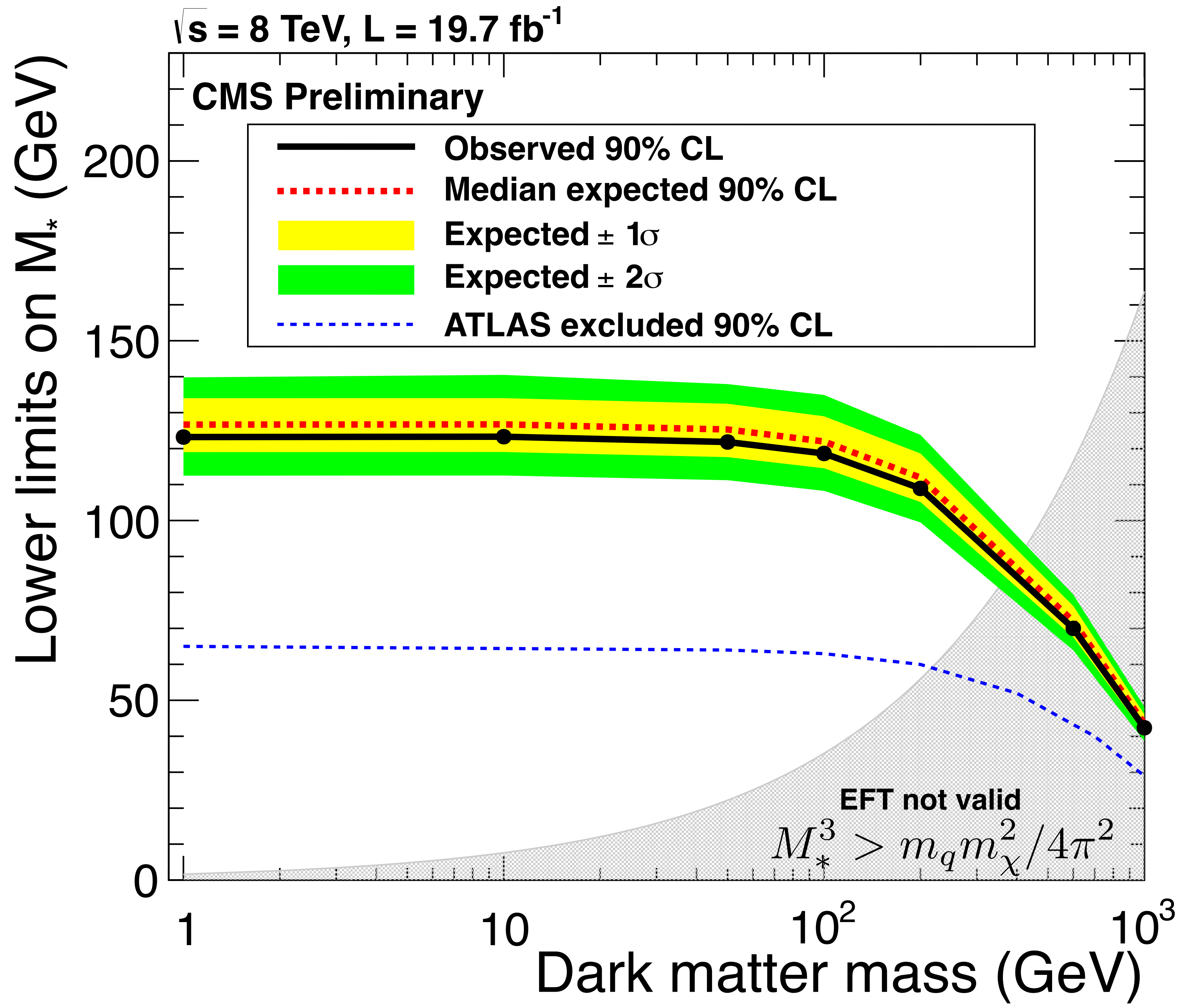
- Theories on DM particles do not predict their masses

  we can set lower limits on the interaction strength parameter M_* for different mass hypothesis

  these translate in upper limits values of production cross section for DM + top as a function of DM mass

Results: lower limits on M*

- Lower limits on interaction scale M*



- The dashed line shows the average behavior in absence of a DM particle based on the simulation
  The yellow and green bands indicate the corresponding 68% and 95% certainty of those values

- The solid line shows the observed limit and it is compared with results by ATLAS collaboration from DM + W/Z events (Phys. Rev. Lett. 112 (Jan, 2014) 041802)

- Values below the observed limit are excluded

  Assuming a dark matter particle with a mass of 100 GeV M* is excluded at 90% CL below 118 GeV

- The grey area represent only minimal requirement on M* for the EFT to be valid. There could be other areas on the plane where the EFT breaks down

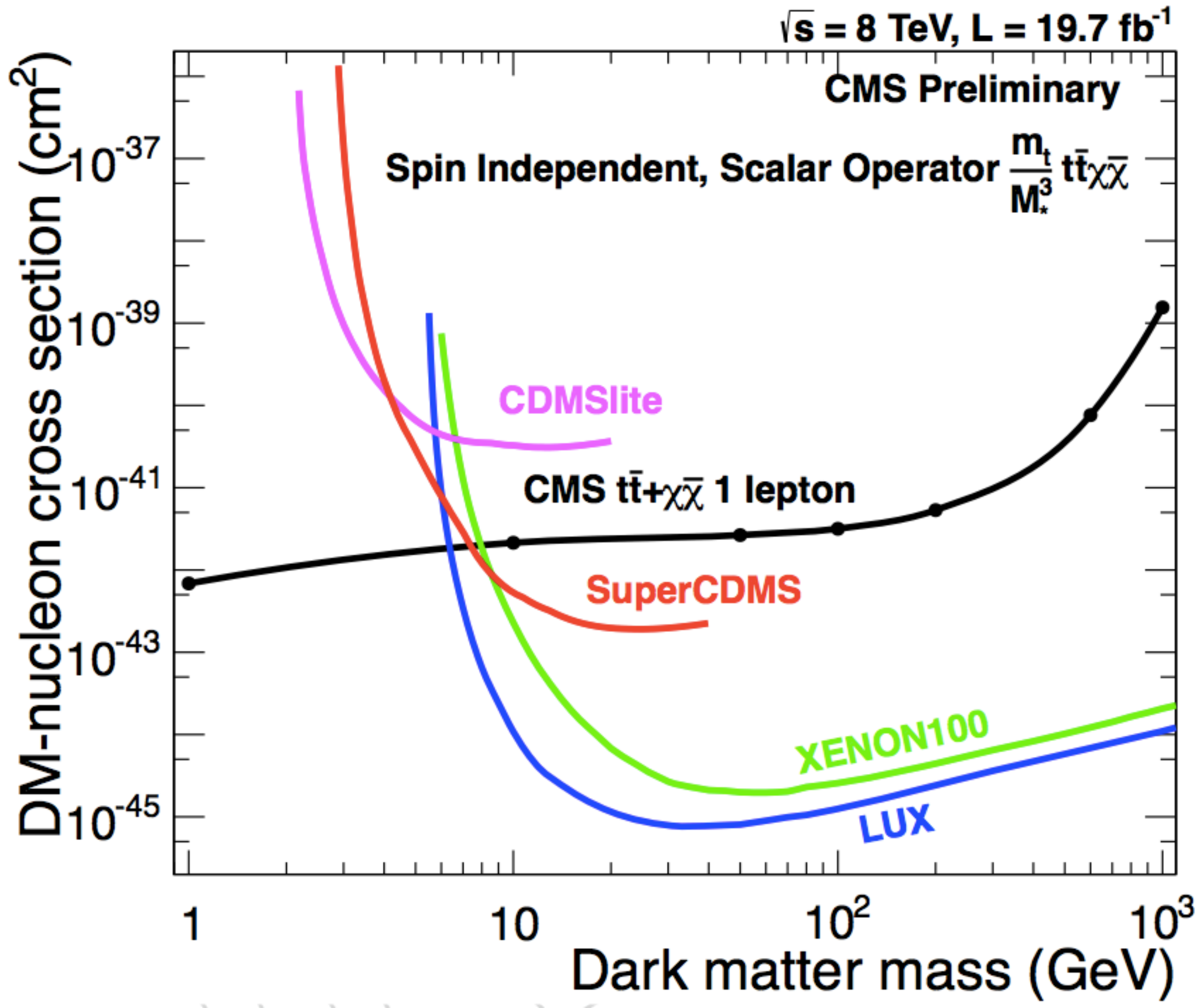
Results: upper limits on cross section

- Upper limits on DM + top events cross section

M_χ GeV	Signal efficiency (%) ($\pm$ stat. $\pm$ syst. unc.)	$\sigma^{\text{lim}}_{\text{exp}}$ (fb)	$\sigma^{\text{lim}}_{\text{obs}}$ (fb)
1	$1.01 \pm 0.02 \pm 0.05$	47^{+21}_{-13}	55
10	$1.01 \pm 0.02 \pm 0.05$	46^{+21}_{-13}	54
50	$1.20 \pm 0.02 \pm 0.06$	38^{+18}_{-11}	45
100	$1.46 \pm 0.02 \pm 0.07$	32^{+15}_{-9}	37
200	$1.73 \pm 0.02 \pm 0.08$	27^{+12}_{-8}	32
600	$2.40 \pm 0.03 \pm 0.11$	19^{+9}_{-6}	23
1000	$2.76 \pm 0.04 \pm 0.13$	17^{+8}_{-5}	20

Cross sections higher than 20 to 55 fb are excluded at 90% CL for DM particles with masses ranging from 1 GeV to 1 TeV

- Can be translated in upper limits on DM-nucleon cross section for comparison with results from direct detection experiment



Excluded dark matter-nucleon cross sections higher than $1 - 2 \times 10^{-42}$ cm^2 (10-20 fb) for DM masses from 1 to 6 GeV

Summary

- Search for DM produced in association with a top quark pair has been presented in the single-lepton channel

  better sensitivity for scalar interactions where coupling strength is proportional to quark mass

- No excess of events observed in the search region

- Lower limits on interaction scale M_* as a function of DM mass

  Assuming a dark matter particle with a mass of 100 GeV M_* is excluded at 90% CL below 118 GeV
  Regions of validity of EFT approach should be taken into account

- Upper limits on cross section as a function of DM mass

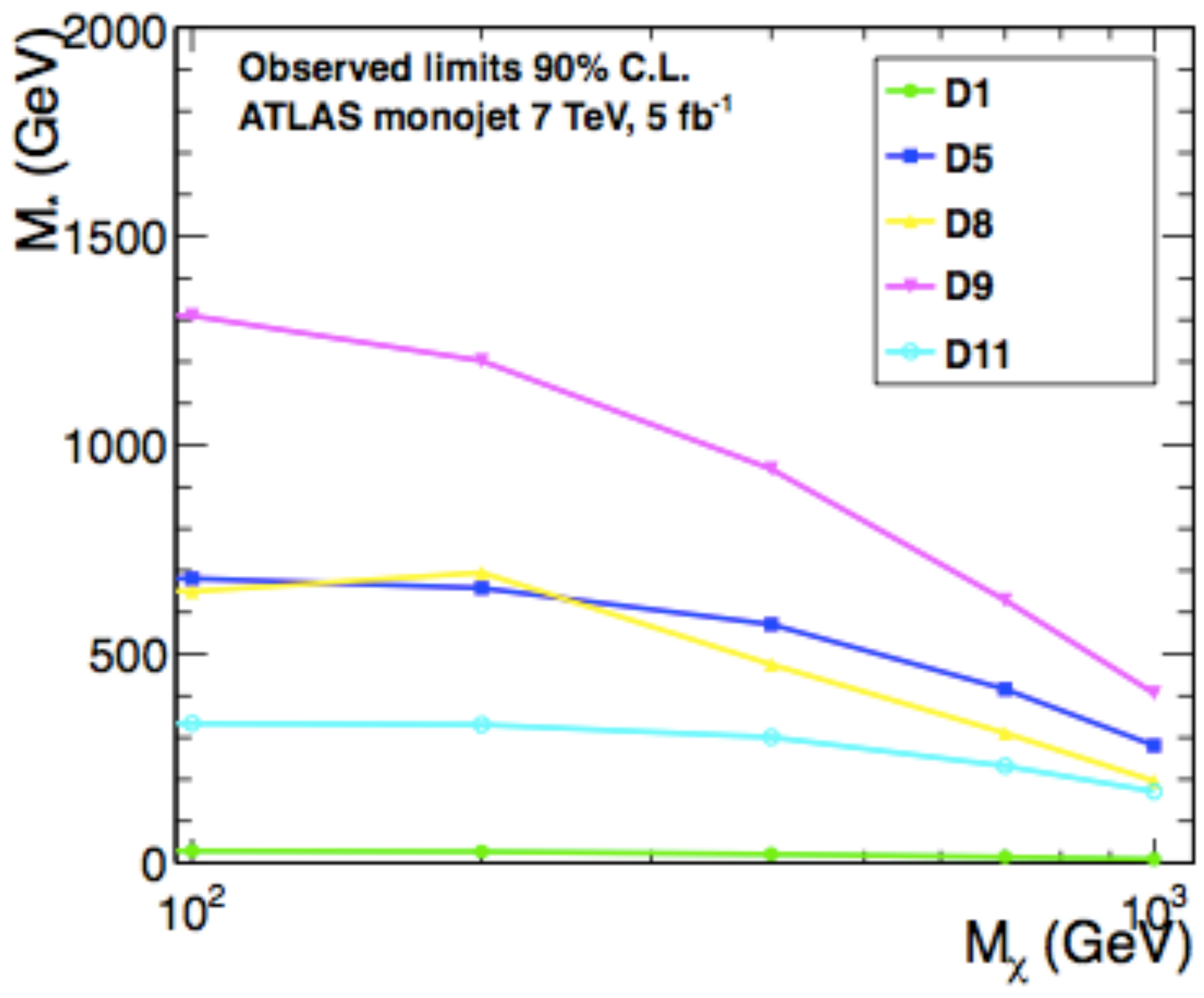
  Cross sections for DM + top events higher than 20 to 55 fb are excluded at 90% CL for dark matter masses from 1 GeV to 1 TeV

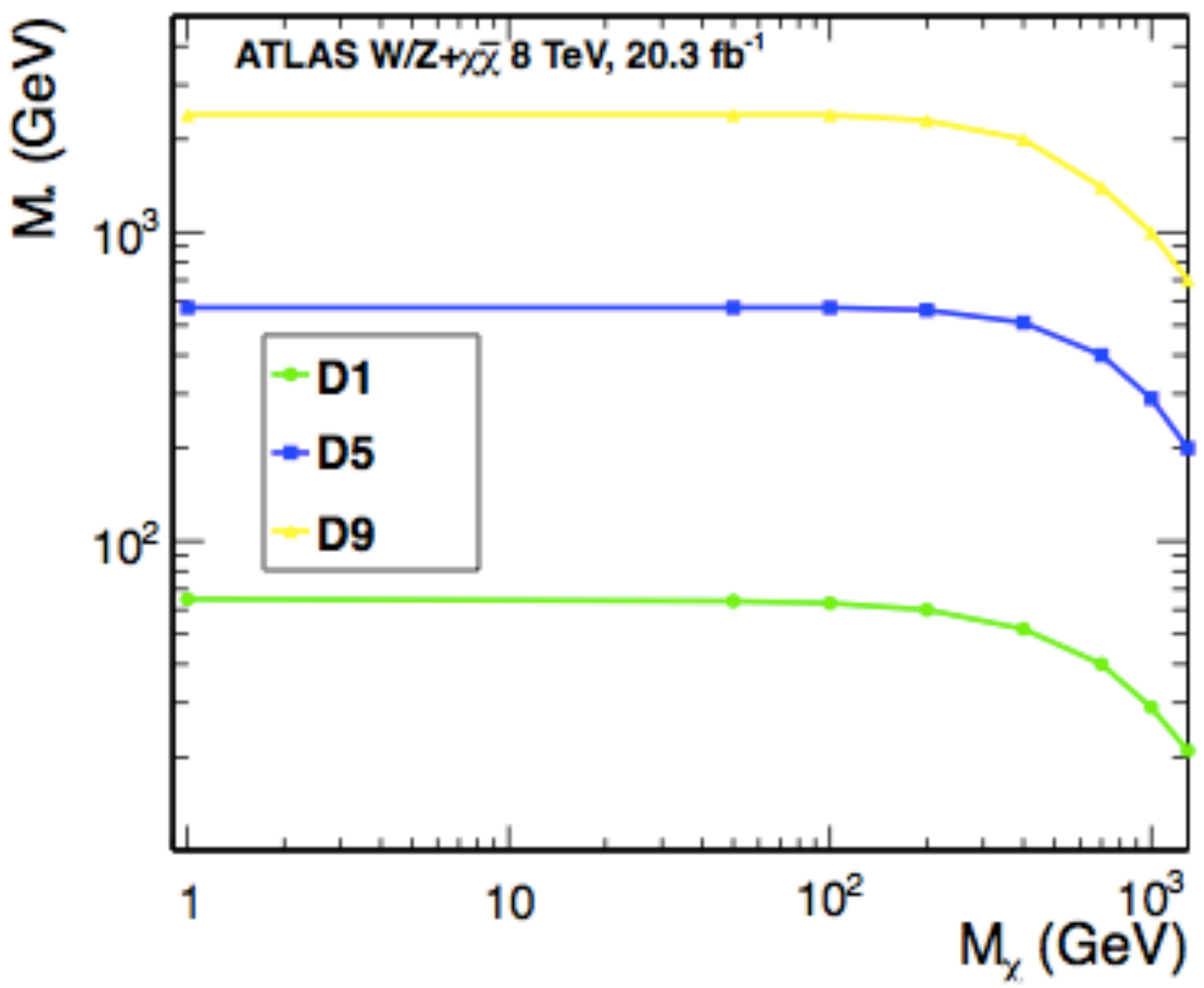
  DM-nucleon cross sections higher than $1 - 2 \times 10^{-42}$ cm^2 (10-20 fb) for DM masses from 1 to 6 GeV are excluded at 90% CL

Backup slides

Current limits

Previous limits for 5 different operators





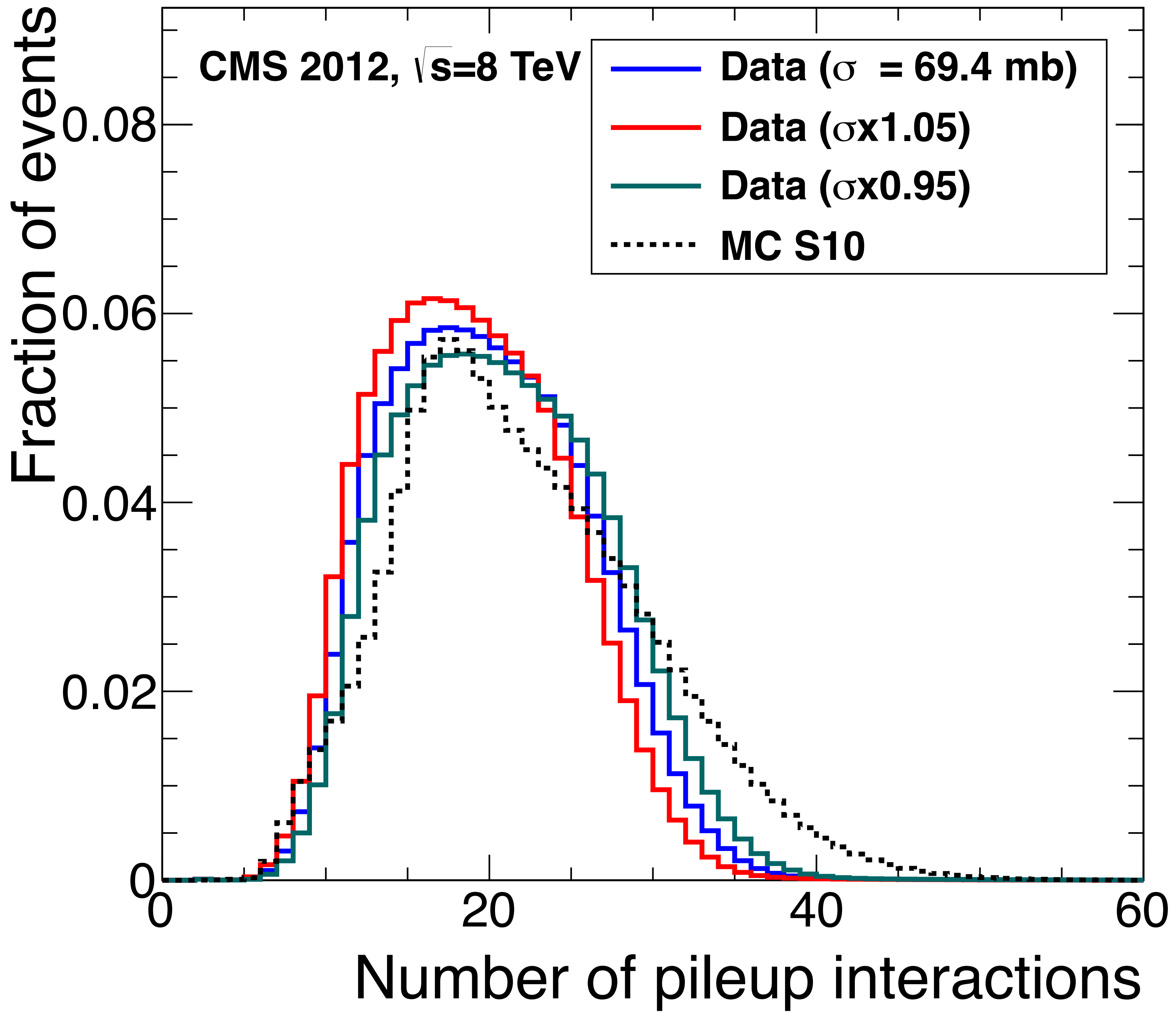
- Weakest limits on D1 operator compared to other types interactions
- Improvement in constraints on scalar interactions using ttbar final states can be achieved

Pileup reweighting

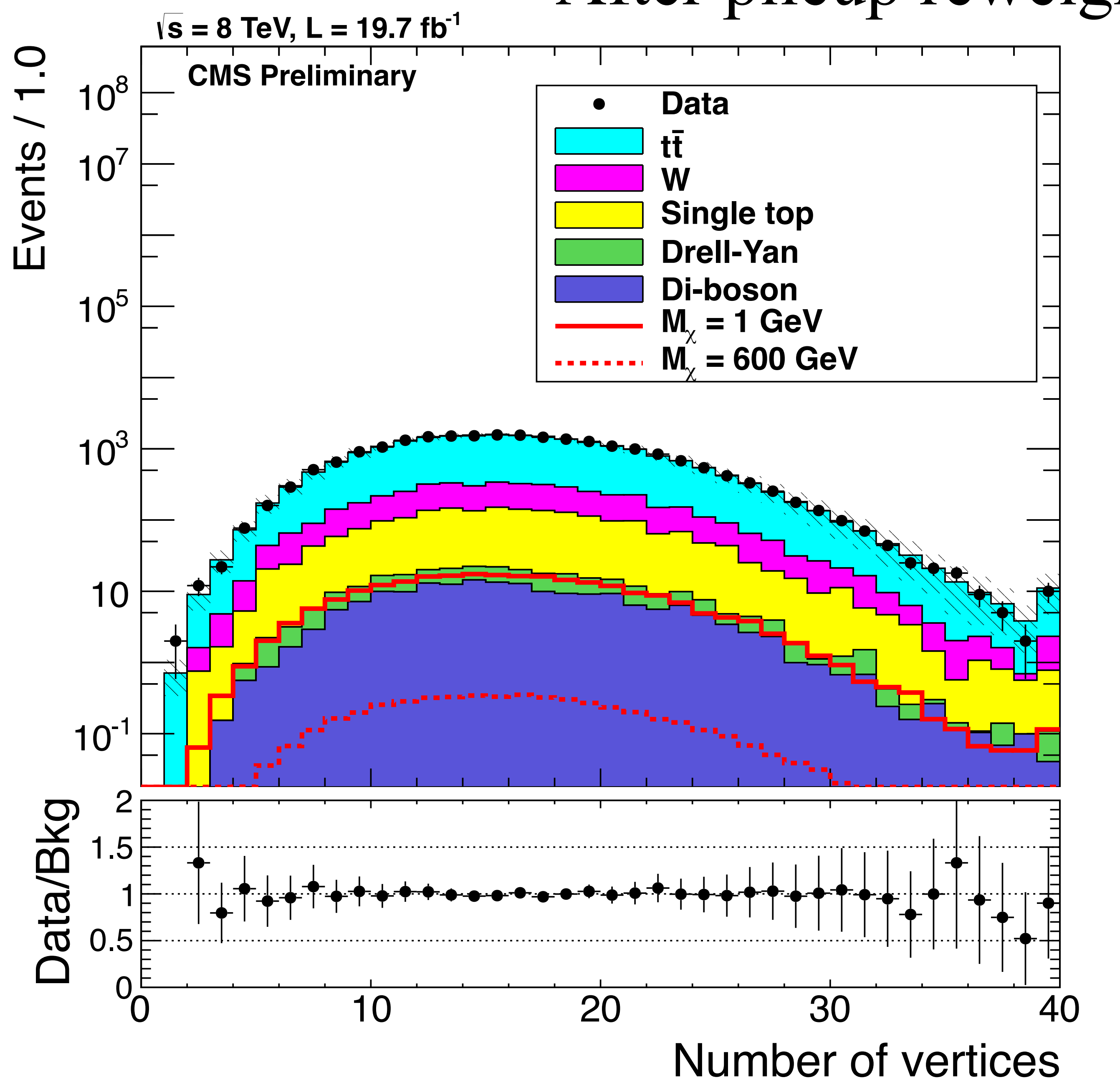
Pileup distribution different in MC w.r.t data

- MC number pileup reweighted to match data
- Data distribution re-calculated with ±5% variation on cross section to cover pileup mis-modeling syst. unc.

Good agreement data-MC after pileup reweighting



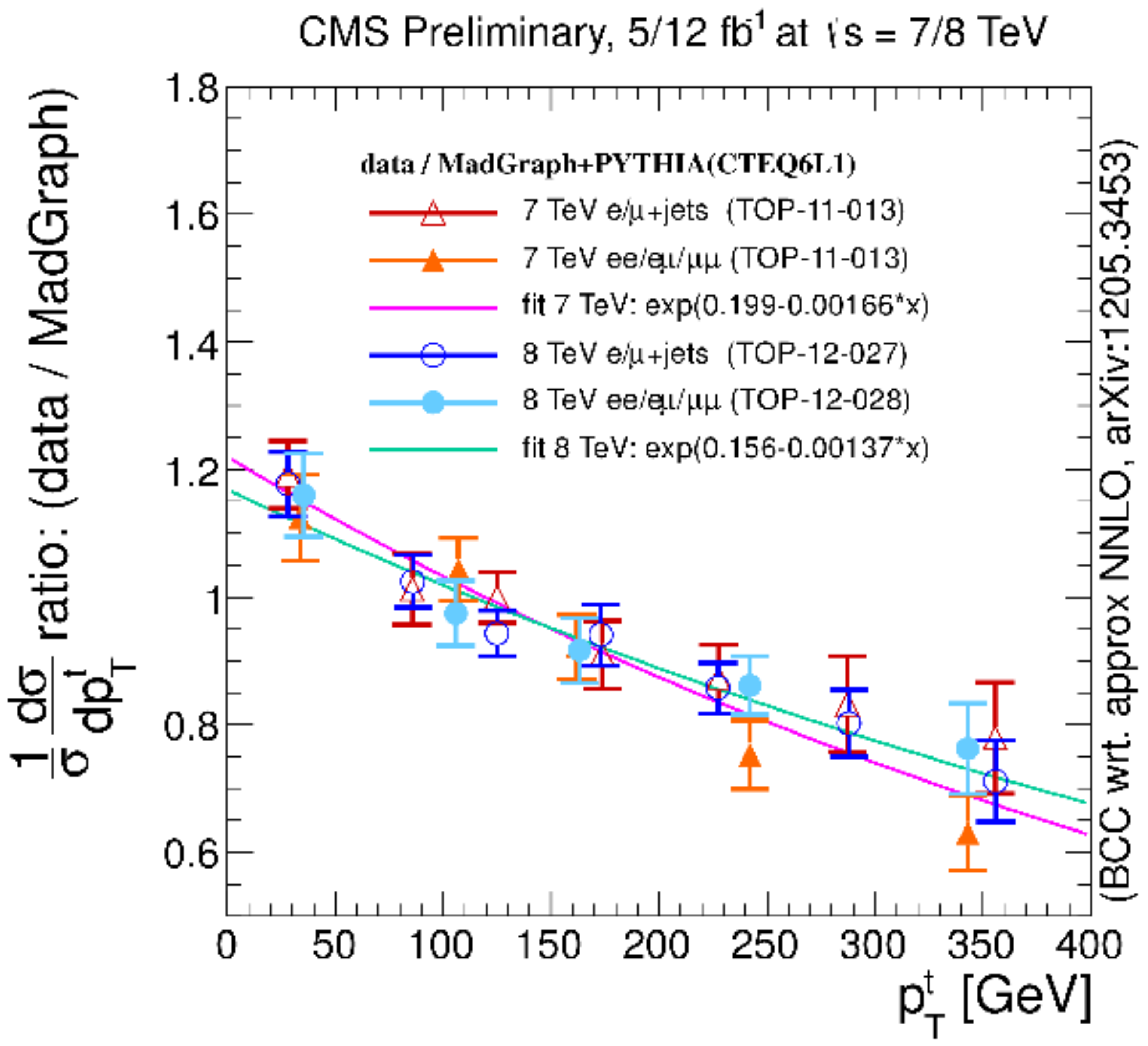
After pileup reweighting



Top p_T reweighting

p_T distribution of leptons and jets from tops softer in data w.r.t Madgraph simulation

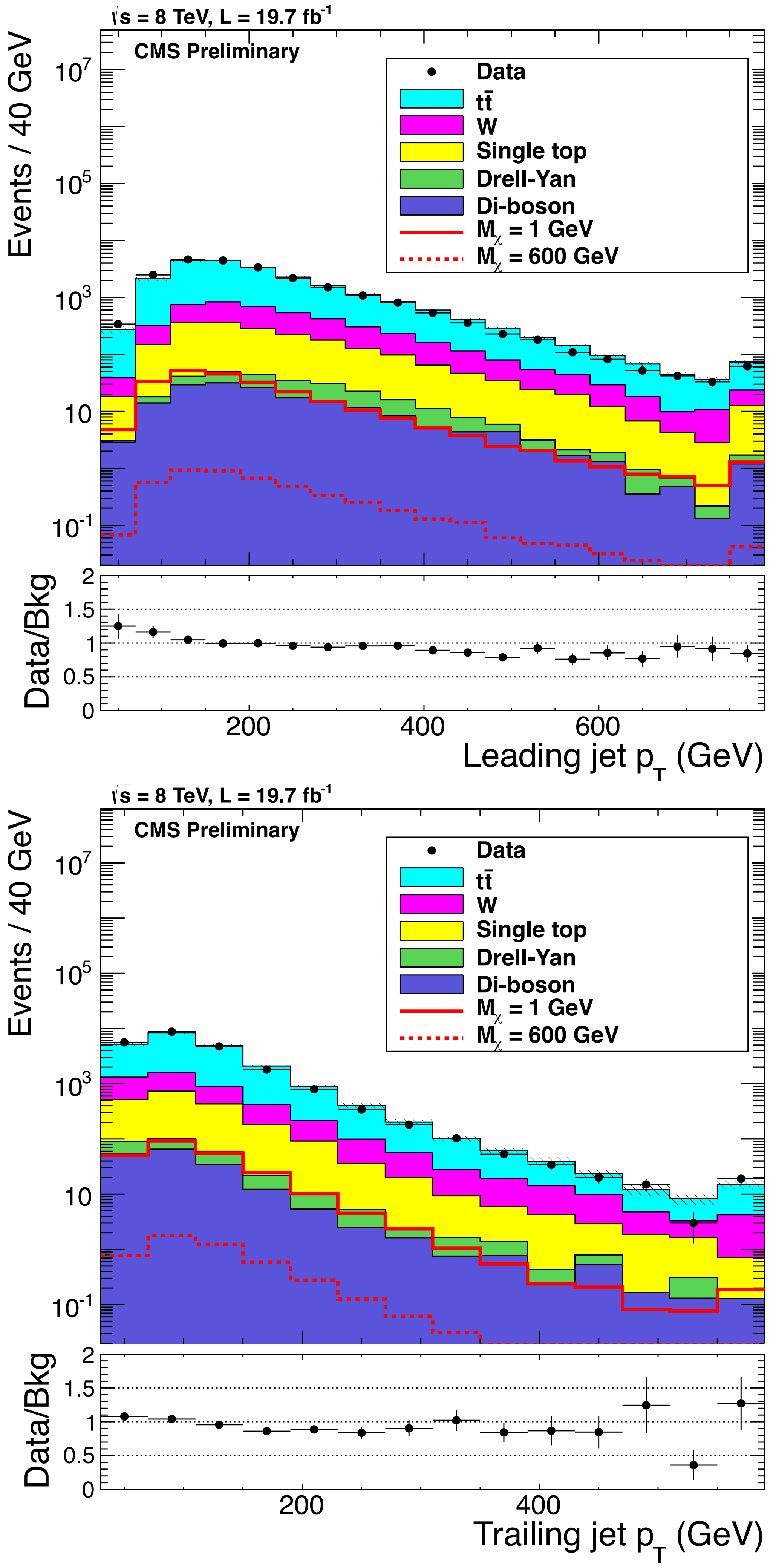
- Top differential cross section measurement provide SFs for correction
- Each event weighted by geometric mean of SFs from 2 tops (assumed flat > 400 GeV)
- Syst. unc.: no SF, SF applied twice



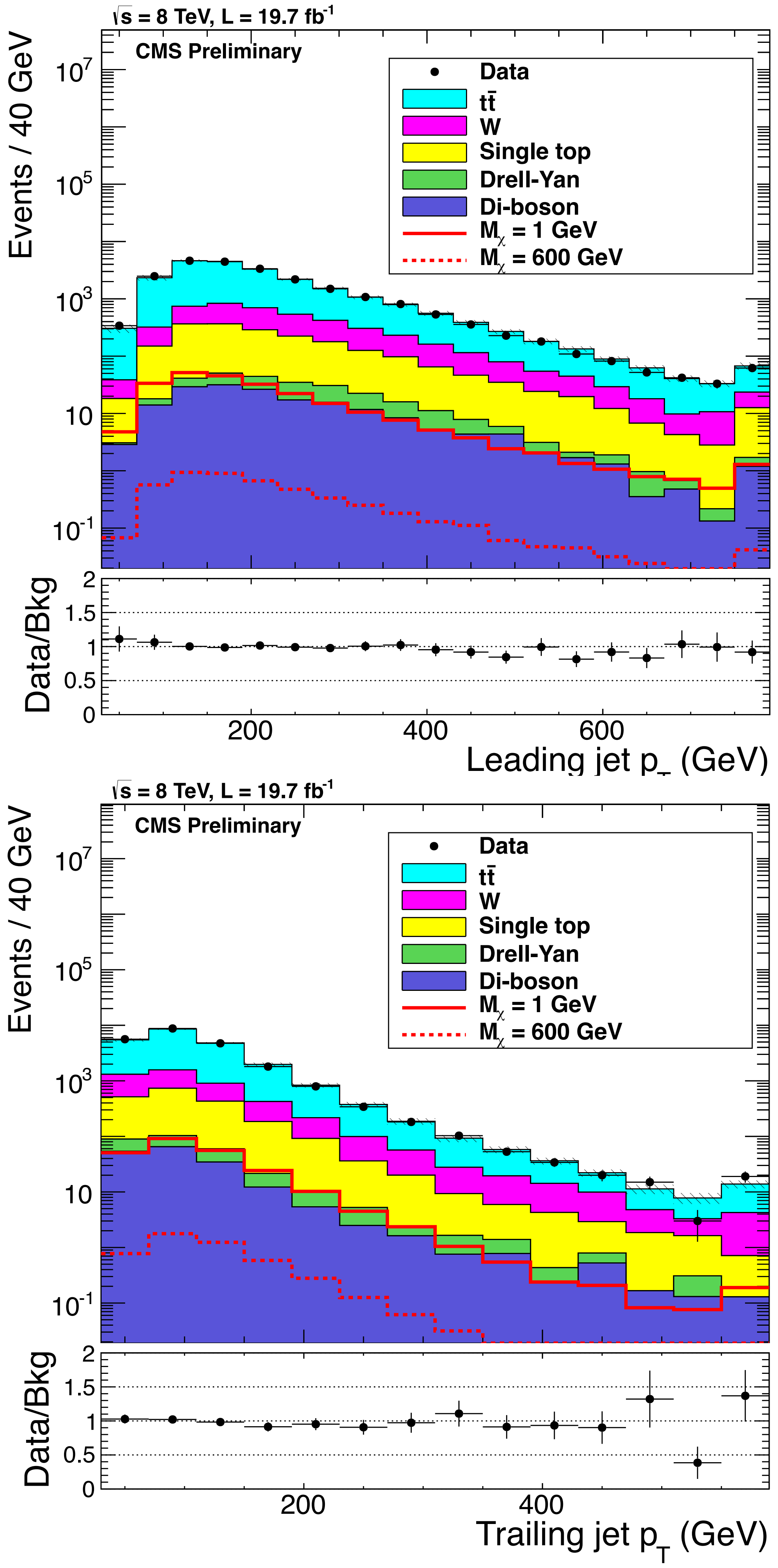
$$SF = e^{0.156 - 0.00137 \times p_T}$$

Top p_T reweighting data-simulation agreement

Left-hand: before reweighting

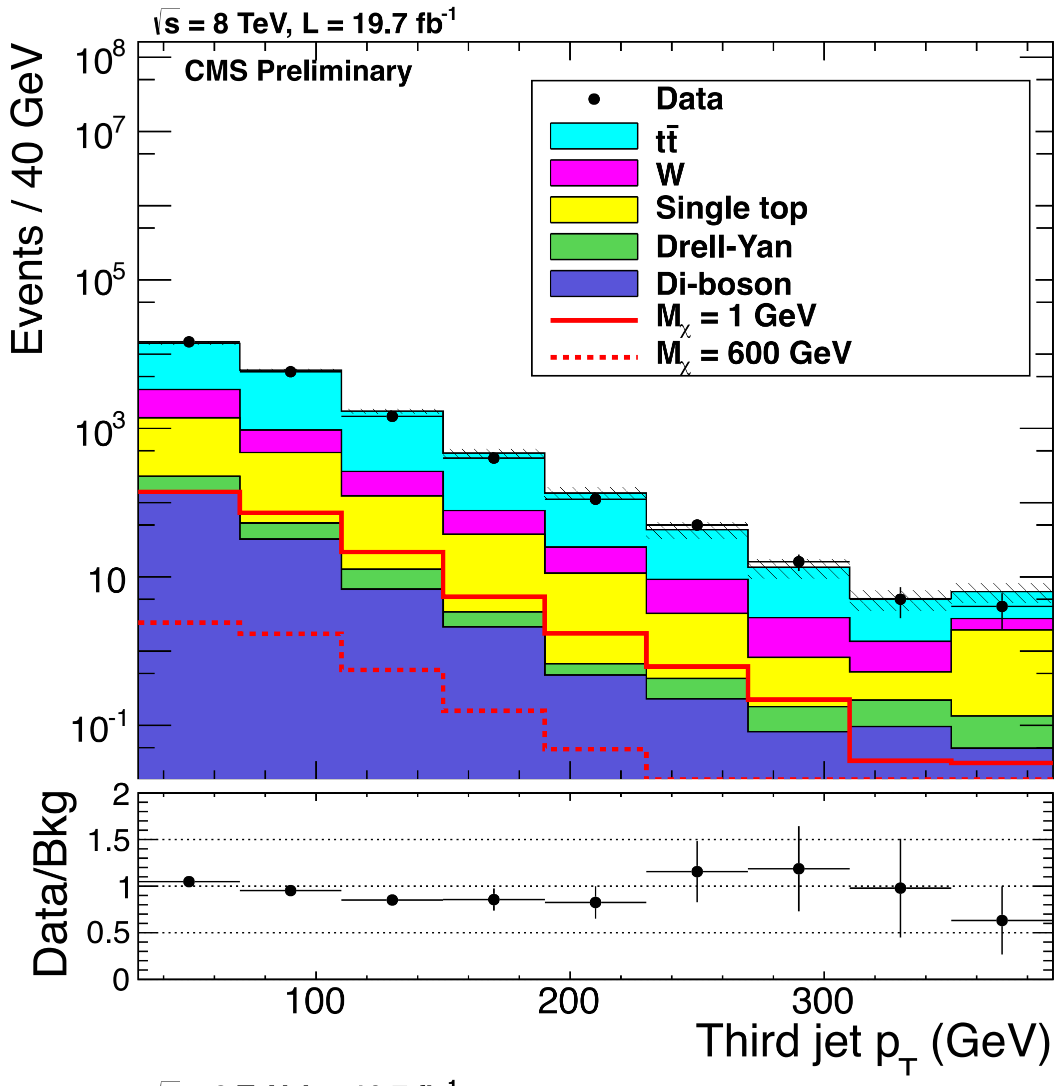


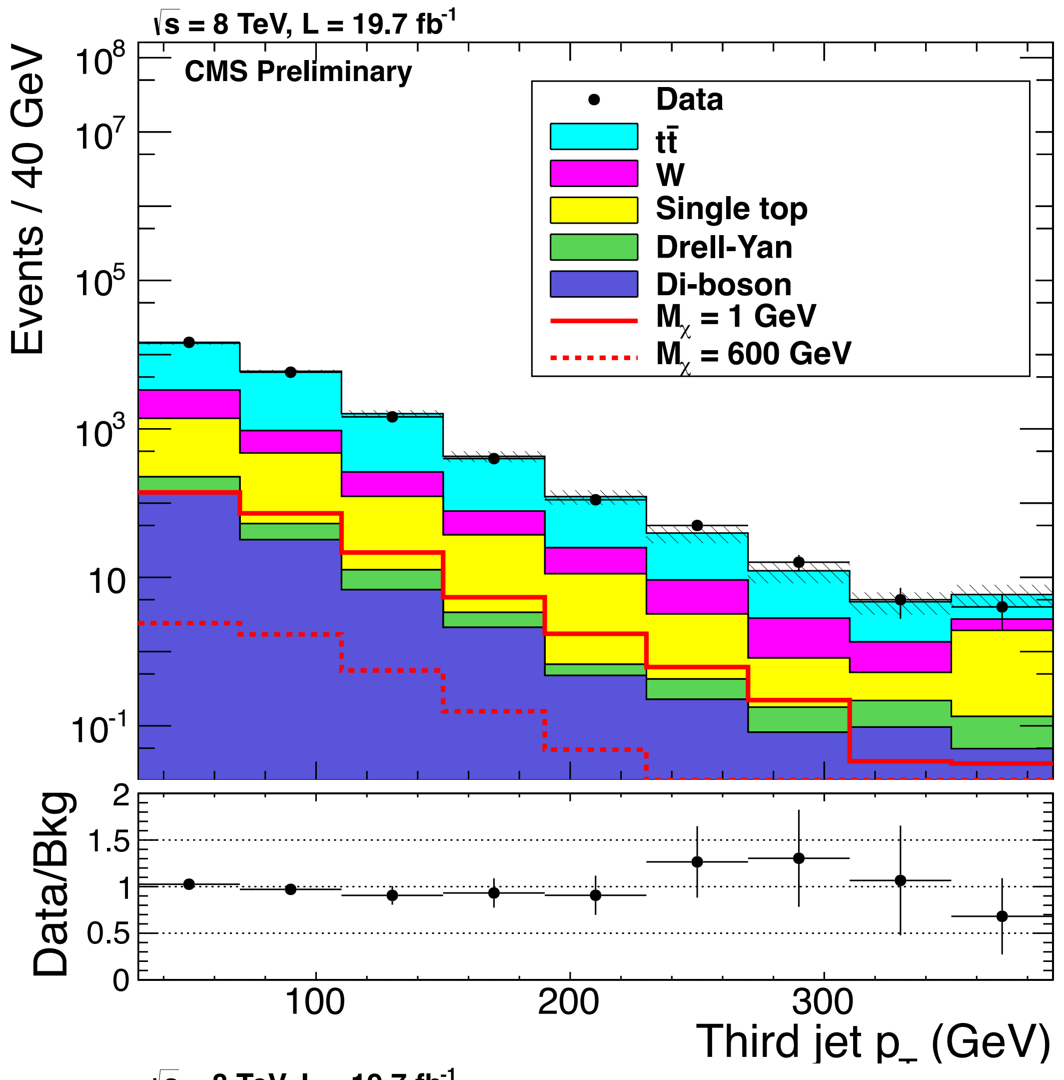
Right-hand: After



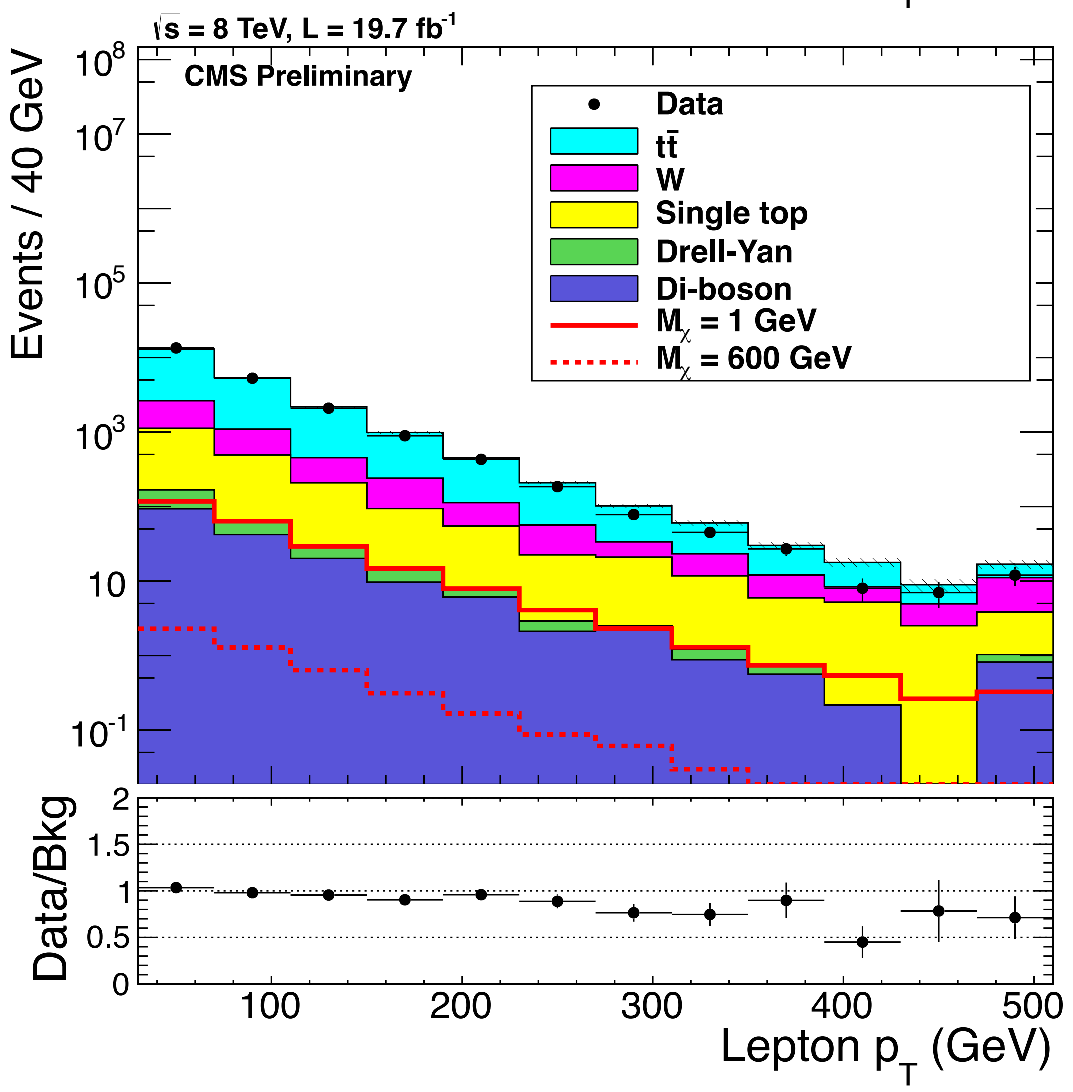
Top p_T reweighting data-simulation agreement

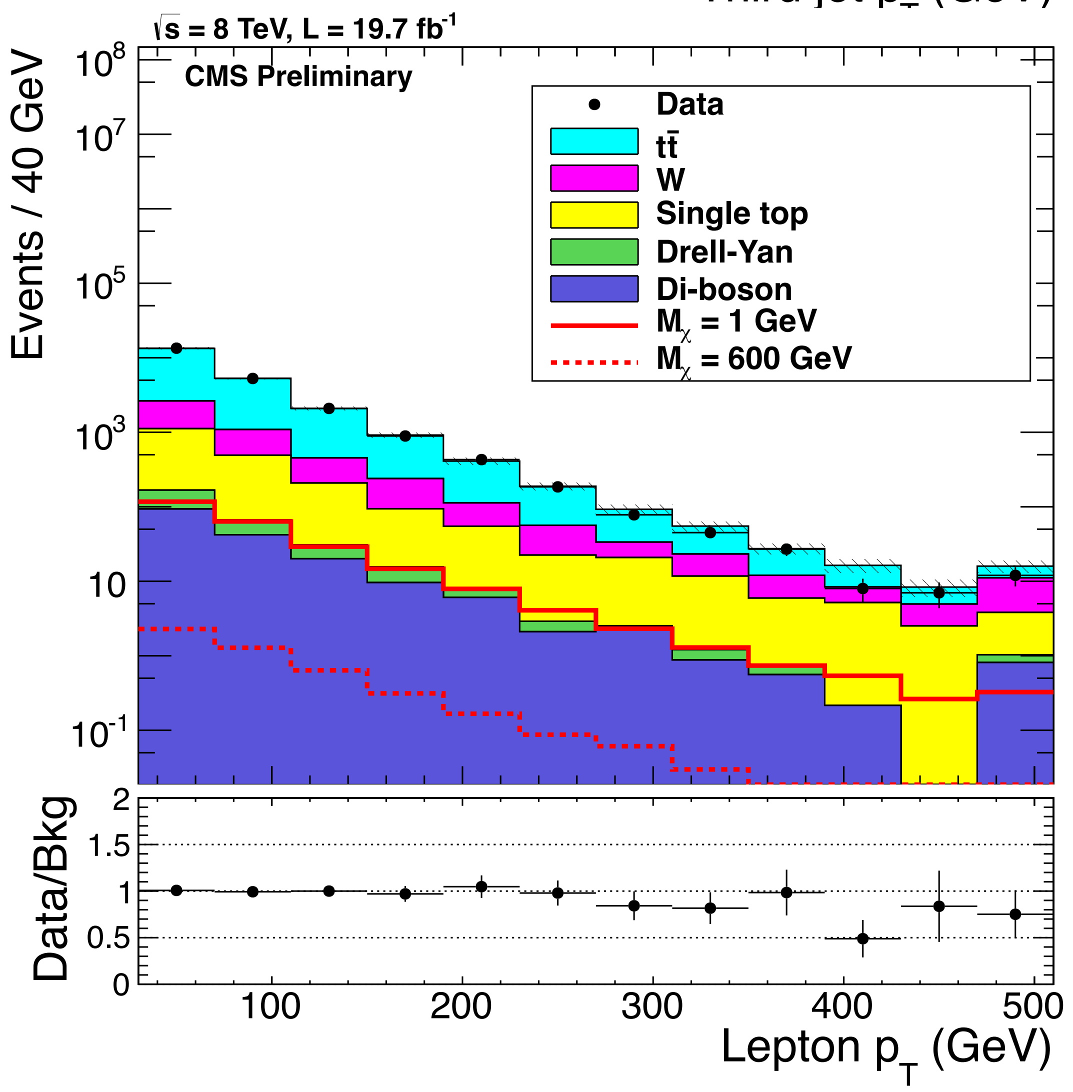
Left-hand: before reweighting



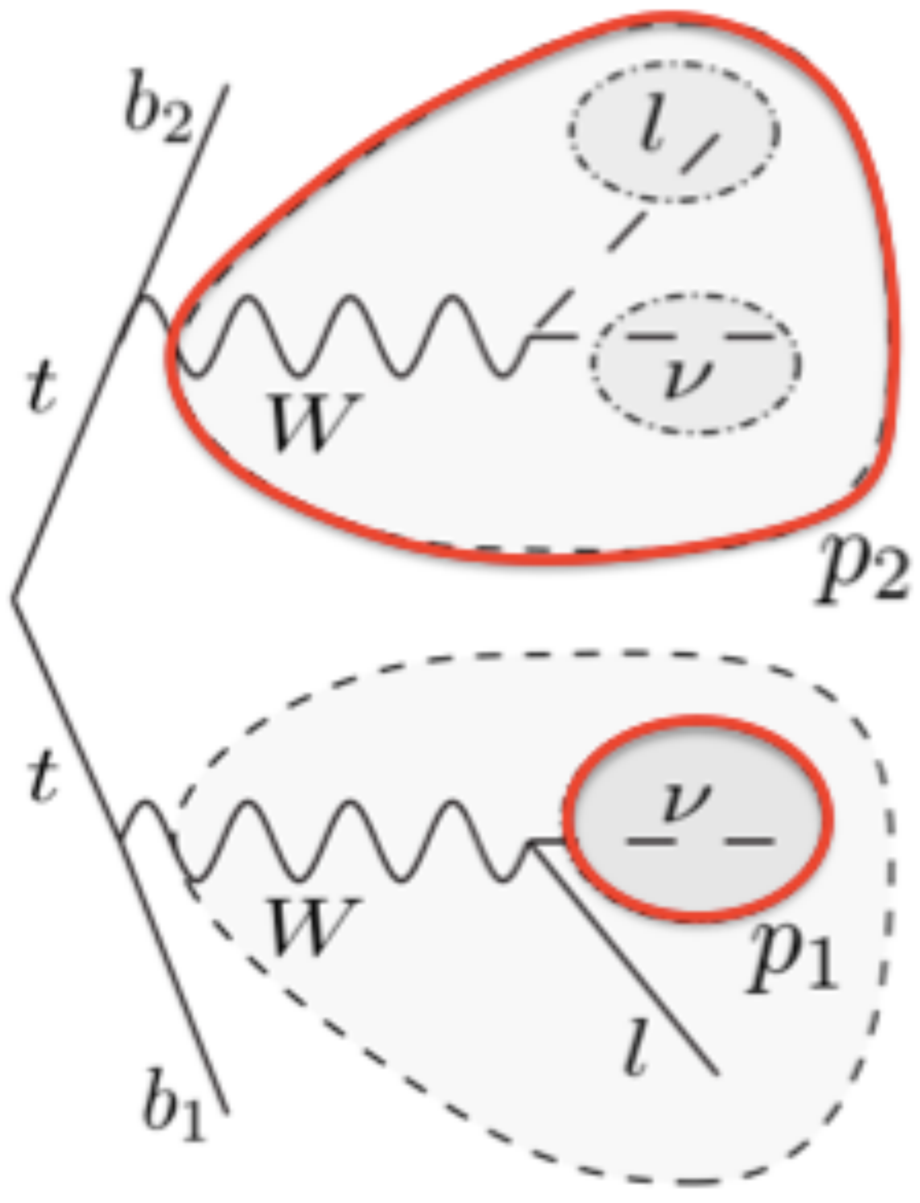


Right-hand: After





M_{T2W} as discriminating variable



Most irreducible background from tt di-leptonic

- Large MET can arise from neutrinos and missing lepton
- M_T higher than W mass because of additional missing particles

Transverse mass M_{T2} can be used to reject background event

- minimal mother particle mass compatible with assumed event topology and daughter particle mass

A variable where the intermediate W are considered on shell can be used

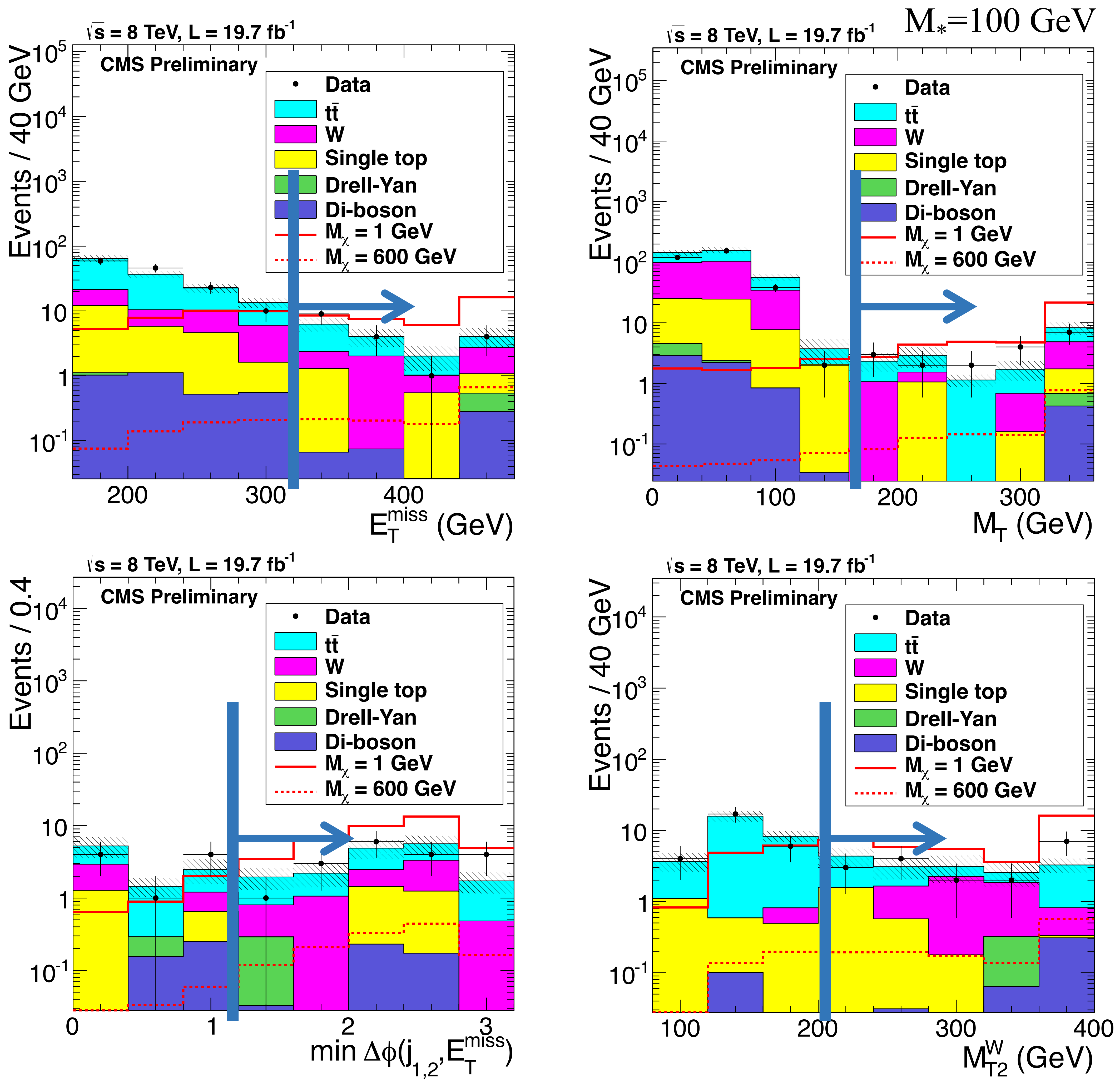
$$M_{T2}^W = \min\left\{m_y \text{ consistent with: } \left[\begin{array}{c} \vec{p}_1^T + \vec{p}_2^T = \vec{E}_T^{\text{miss}},\ p_1^2 = 0,\ (p_1 + p_\ell)^2 = p_2^2 = M_W^2, \\ (p_1 + p_\ell + p_{b_1})^2 = (p_2 + p_{b_2})^2 = m_y^2 \end{array}\right]\right\}$$

it adds other kinematical info w.r.t to other M_{T2} variables

Bai, Cheng, Gallichio, Gu
JHEP 07 (2012) 110

N-1 plot: discriminating power of variables

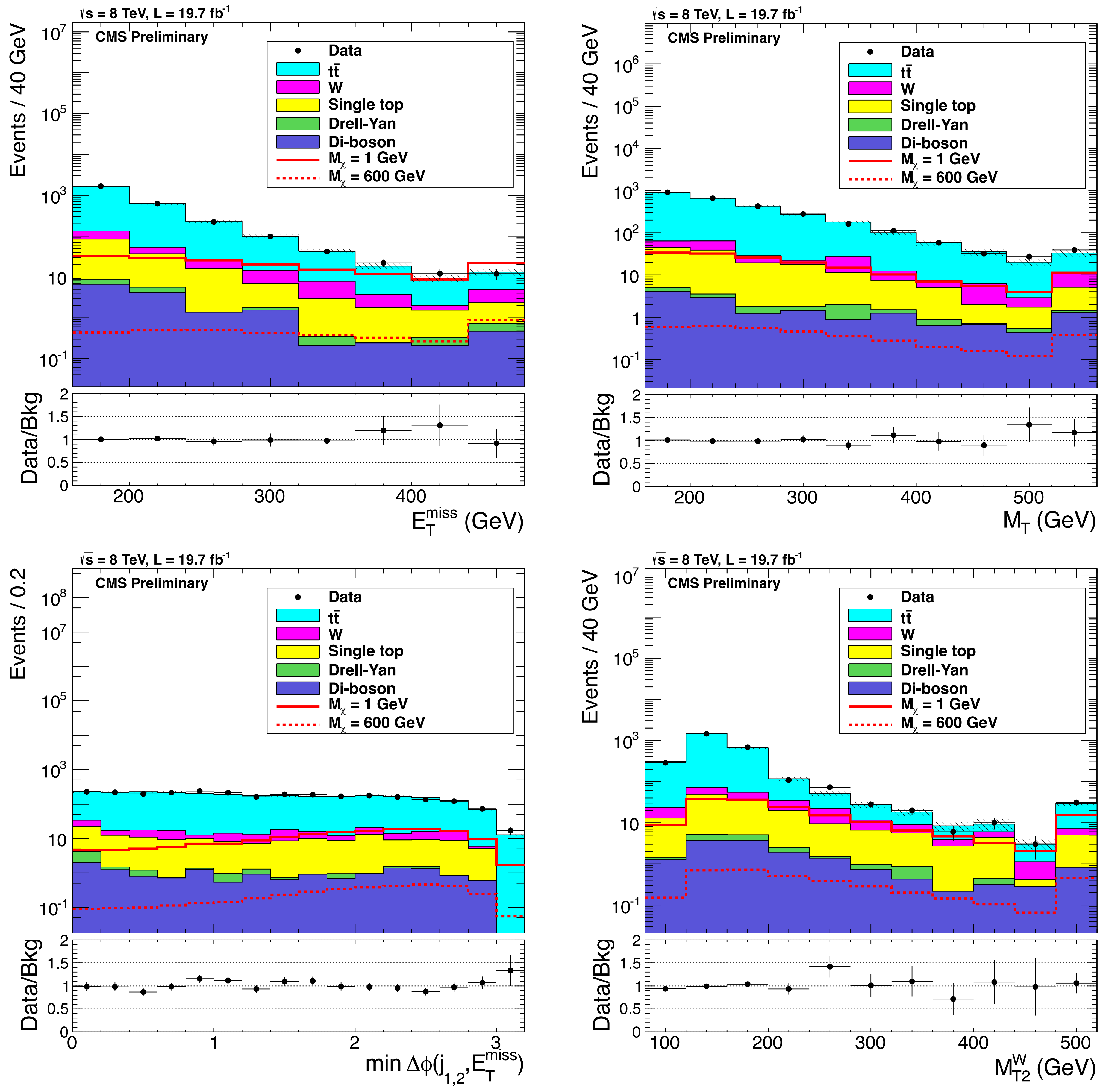
- All the cuts on the discriminating variables are applied, expect the one on the variable that it is showed



Distributions show that each variables has strong discriminating power between signal and background

Data-simulation agreement

Agreement between data and MC samples is used to check the validity of this estimate

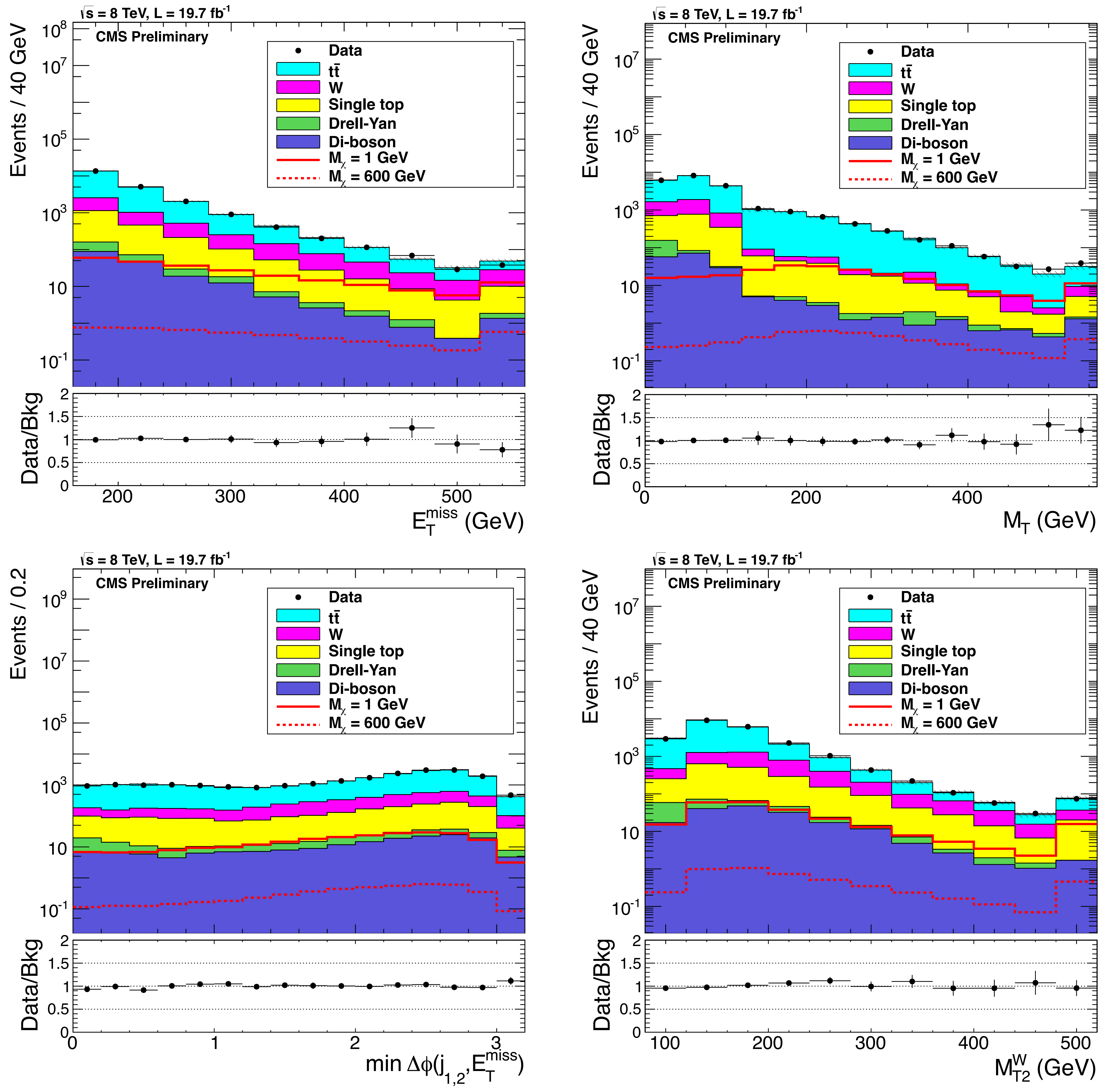


CR1 (tt enriched)

- In all distributions good agreement between data and background prediction is observed after SFs applied

Data-simulation agreement

Agreement between data and MC samples is used to check the validity of this estimate



PRE-SELECTION

- In all distributions good agreement between data and background prediction is observed after SFs applied

Backup slides

Background

Normalization uncertainties already covered by SFs for tt+jets and W+jets

- the integrated luminosity
- lepton identification
- trigger efficiency
- and the cross sections

Other systematic uncertainties constrained by refitting MC to data in CRs and propagating results in SR

Source of systematic uncertainties	Relative error on total background (%)
Normalization of other bkg	10
SF_w(stat)	1.5
$t\bar{t}$+jets jet-parton matching	8.2
$t\bar{t}$+jets Q^2	6.6
$t\bar{t}$+jets top p_T reweighting	3.9
Jet energy scale	4.0
Jet energy resolution	3.0
b-tagging correction factor (heavy flavor)	1.0
b-tagging correction factor (light flavor)	1.8
Pileup model	2.0

Signal

The total systematic uncertainty for signal efficiency is estimated to be 5-6%.

Physics objects: b jets

b jets

b-tagging algorithm: **C**ombined **S**econdary **V**ertex (CSV)

- Standard CMS b-tagging algorithm
- Used to identify jets likely to come from b quarks fragmentation-adronization
- Exploits long lifetime of b hadrons

  large impact parameter and presence of a secondary vertex as input
- Continuous output: allows selection of optimal working points

Efficiencies, mis-tag rates

for CSV > 0.90 b quark tag: 50%

c quark tag: 6%

light quark tag: 0.15 %

Systematics

Jet Energy Scale (JER)

the jet energy is determined from energy deposits in the hadronic calorimeter. Applying a scale to this value the energy of the parton generating the jet can be inferred. A systematic uncertainties is associated to this calculation

Jet Energy Resolution (JES)

the jet energy resolution reflects the ability to distinguish between two different energy values. It is generally defined as the ratio of the full-width-at-half-maximum (FWHM) of the peak divided by its centroid position. Small values of this ratio correspond to narrow peaks and then to a better resolution.

Backup slides