A Model Independent General Search in ATLAS

Simone Amoroso

[1]Albert Ludwigs Universität Freiburg

YRMR

Rome 20/01/2012

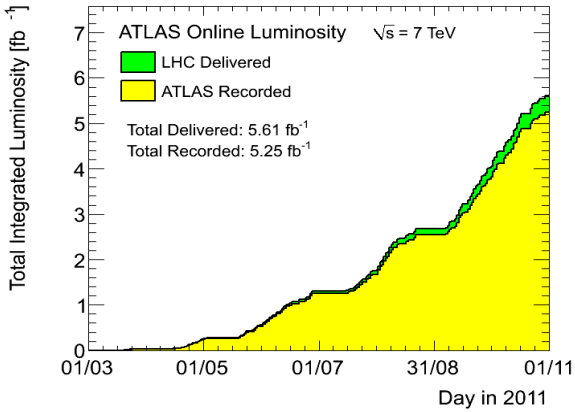
The LHC accelerator

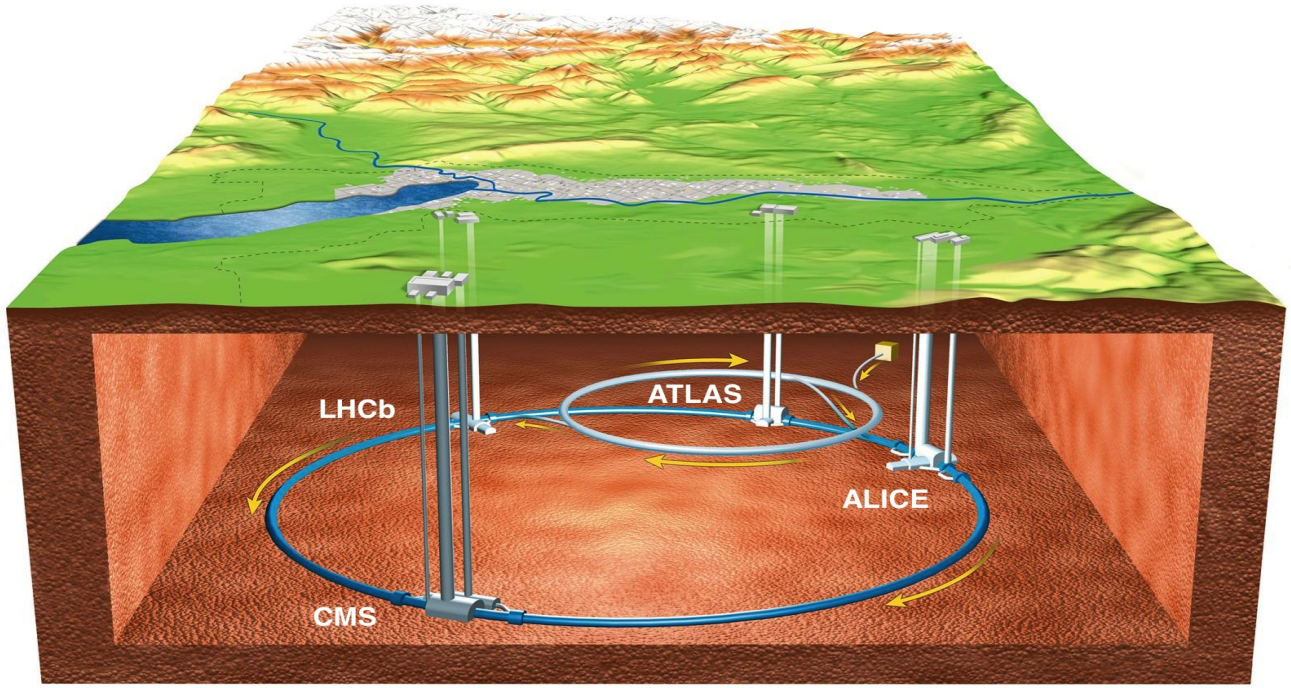
A proton-proton collider at 7 TeV center-of-mass energy

2011 LHC performance

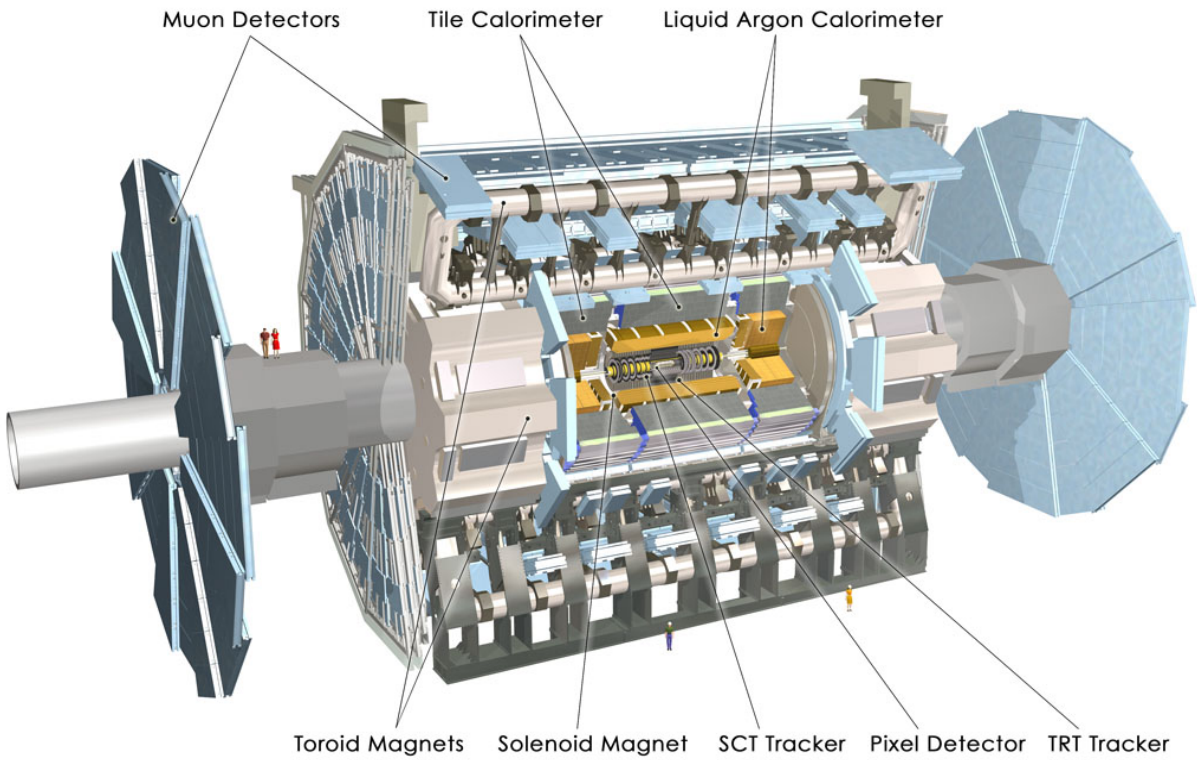
- Peak Luminosity: $3.4 \cdot 10^{33} \mathrm{cm}^{-2} \mathrm{s}^{-1}$
- 1331 bunches - 50ns spacing
- ~ 12 collisions per bunch crossing

already $\sim 5 \mathrm{fb}^{-1}$ collected





The ATLAS detector



Magnets:

- Solenoid (inner): 2T
- Toroid (outer):0.5 T

EM Calorimeter:

- $\sigma/E \sim 10\%/\sqrt{E} \oplus 0.7\%$
- excellent e/γ identification
- good energy resolution

Precision Muon spectrometer:

- $\sigma/P_T \sim 10\%$@ 1TeV
- fast trigger response
- good momentum resolution
- standalone muon identification

Hadron Calorimeter:

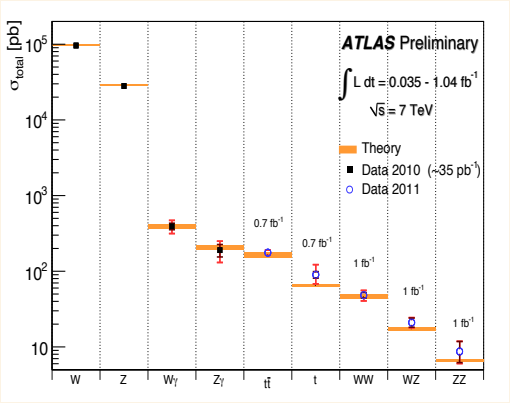
- $\sigma/E \sim 50\%/\sqrt{E} \oplus 3\%$
- good jet resolution
- good $\not{E}_T$ resolution

Inner Detector:

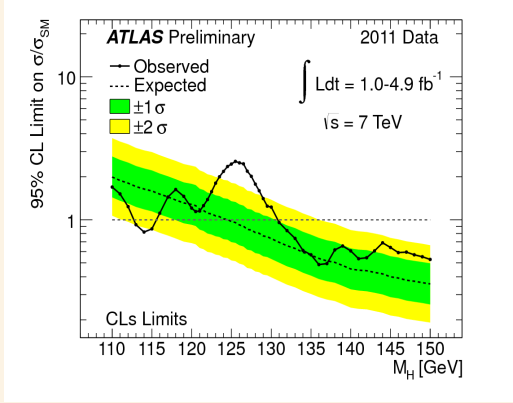
- Si pixel and strips; TRT
- $\sigma/P_T \sim 5 \cdot 10^{-4} P_T \oplus 0.001$
- Excellent vertex reconstruction

ATLAS physics program

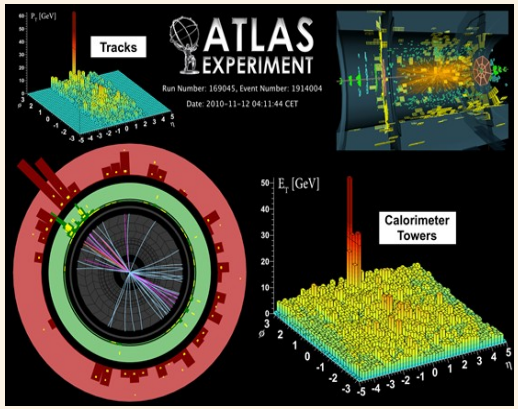
Standard Model (re)discovery



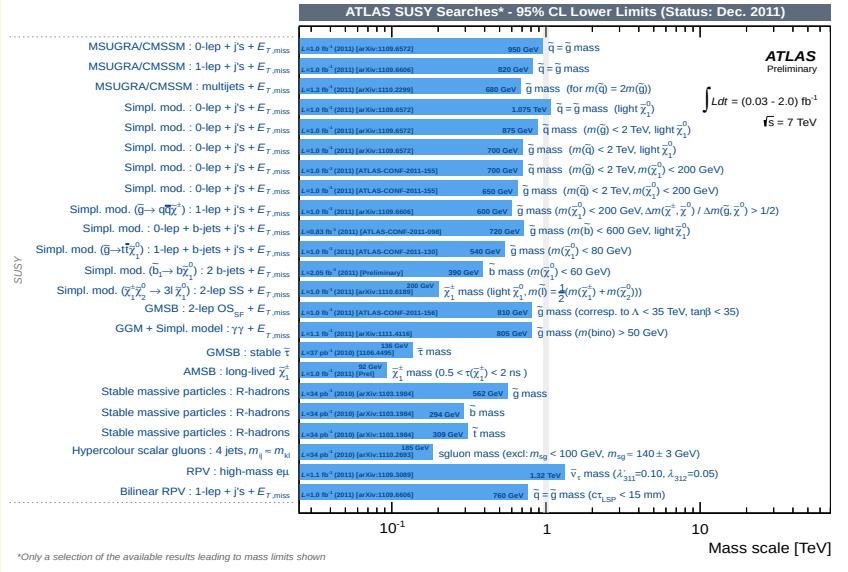
The EWSB mechanism

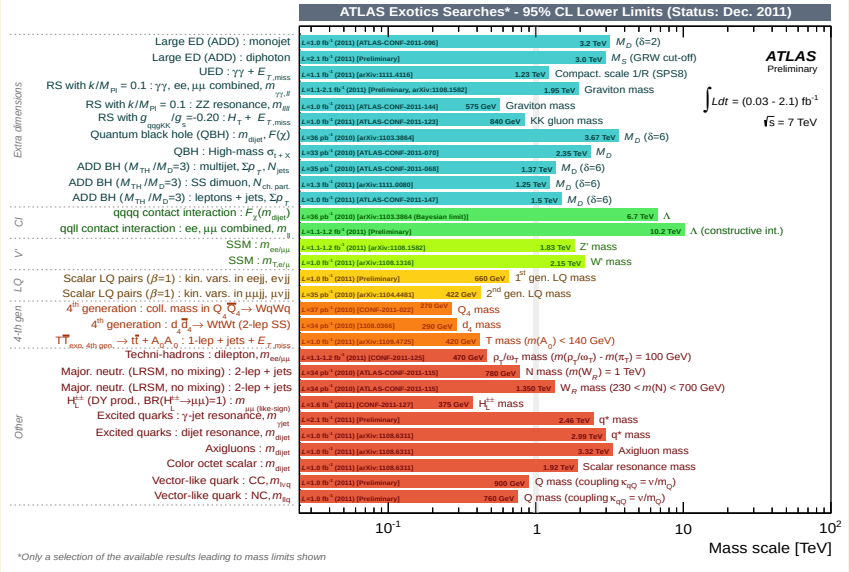


Heavy Ion physics



Beyond the Standard Model Searches





Why a General Search?

- Many possible models for new physics
- Too many analyses to setup
- And NP needs not be in one of those
- **What if we are missing it?**

Why a General Search?

- Many possible models for new physics
- Too many analyses to setup
- And NP needs not be in one of those
- **What if we are missing it?**

We need a model-independent pointing tool

Why a General Search?

- Many possible models for new physics
- Too many analyses to setup
- And NP needs not be in one of those
- **What if we are missing it?**

We need a model-independent pointing tool

Going Model Independent:

- Be open for everything and expect the unexpected!
- Trade sensitivity for a larger coverage

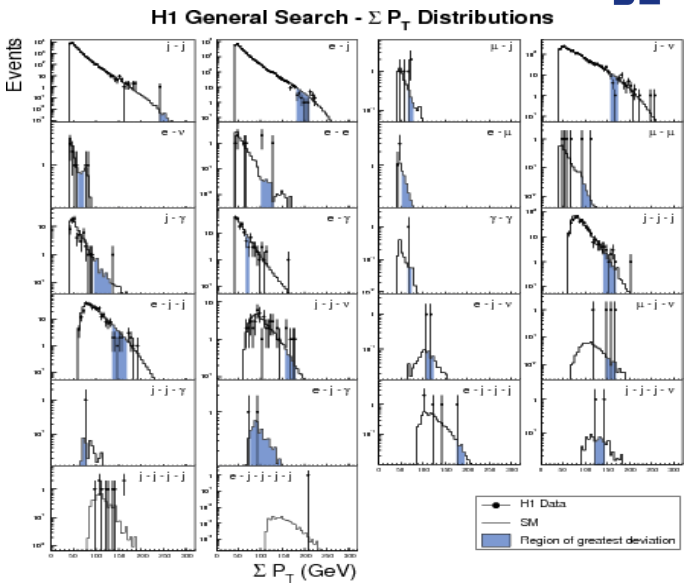
General Searches in a Nutshell - I

- A General Search is not meant to substitute many dedicated analyses;
  $\rightarrow$ ***It's a pointing tool!***
- The aim is to reveal discrepancies between data and Montecarlo looking at all possible events topologies
- Only assumption is that NP would appear at high mass/high P_T
- Divide events into exclusive classes
- Perform an automated scan of distributions
- Still needs a dedicated analysis to interpret excesses

General Searches in a Nutshell - II

A History of General Searches

- SLEUTH at D0 Phys.Rev.Lett.86(2001)
- VISTA+SLEUTH at CDF Phys. Rev. D 78 (2008)
- H1 Phys. Lett. B 602 (2004)
- Music at CMS CMS-PAS-EXO-10-021
- D0 Submitted to PRD



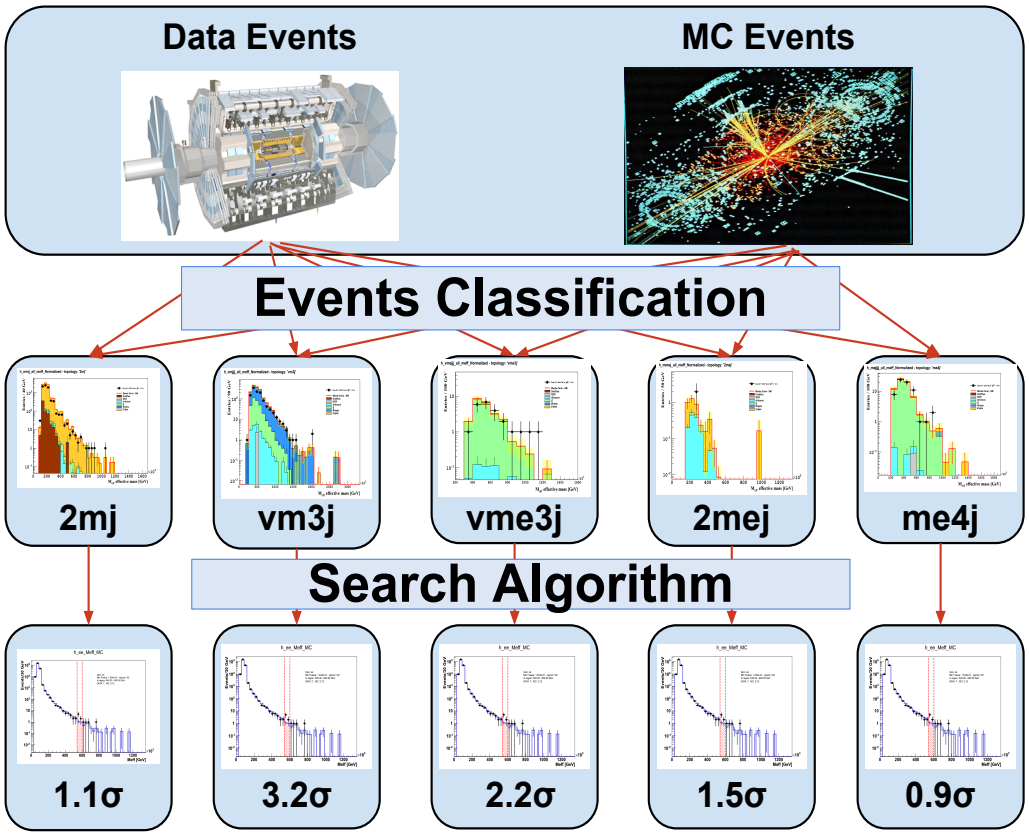
What we added:

- Use of a merging algorithm not to be flood by channels we are not sensitive to:
  - Reducing a potential infinite number of channels to a smaller, manageable one
  - Increasing our sensitivity by reducing the Look Elsewhere Effect
- Look at the interesting channels with an independent dataset; without the big penalty given by the trial-factors.

General Searches in a Nutshell - III

Events Classification

- Channels are defined according to the particle content of the event.
  e.g. an event with three jets and an electron goes into the "3je" channel



Events Selections

- a "ν" if $\not{E}_T > 125$ GeV
- a "m" for each muon with $P_T > 20$ GeV
- a "e" for each electron with $P_T > 25$ GeV
- a "j" for each jet satisfying:
  - $P_T > 300, 40, 40, \ldots$ GeV if topology has only jets
  - $P_T > 60, 25, 25, \ldots$ GeV if leptons are present

Data and MC

- We run over all recorded events containing a Muon, a Jet or an Electron
- The total integrated luminosity used amounts to $1080.116\ \mathbf{pb}^{-1}$
- To keep the analysis as simple as possible the lowest unprescaled single object triggers have been used.

Triggers

Muon:

- EF_mu18 && L1_MU11 period $\leq$I
- EF_mu18_medium period J

EGamma:

- EF_e22_medium

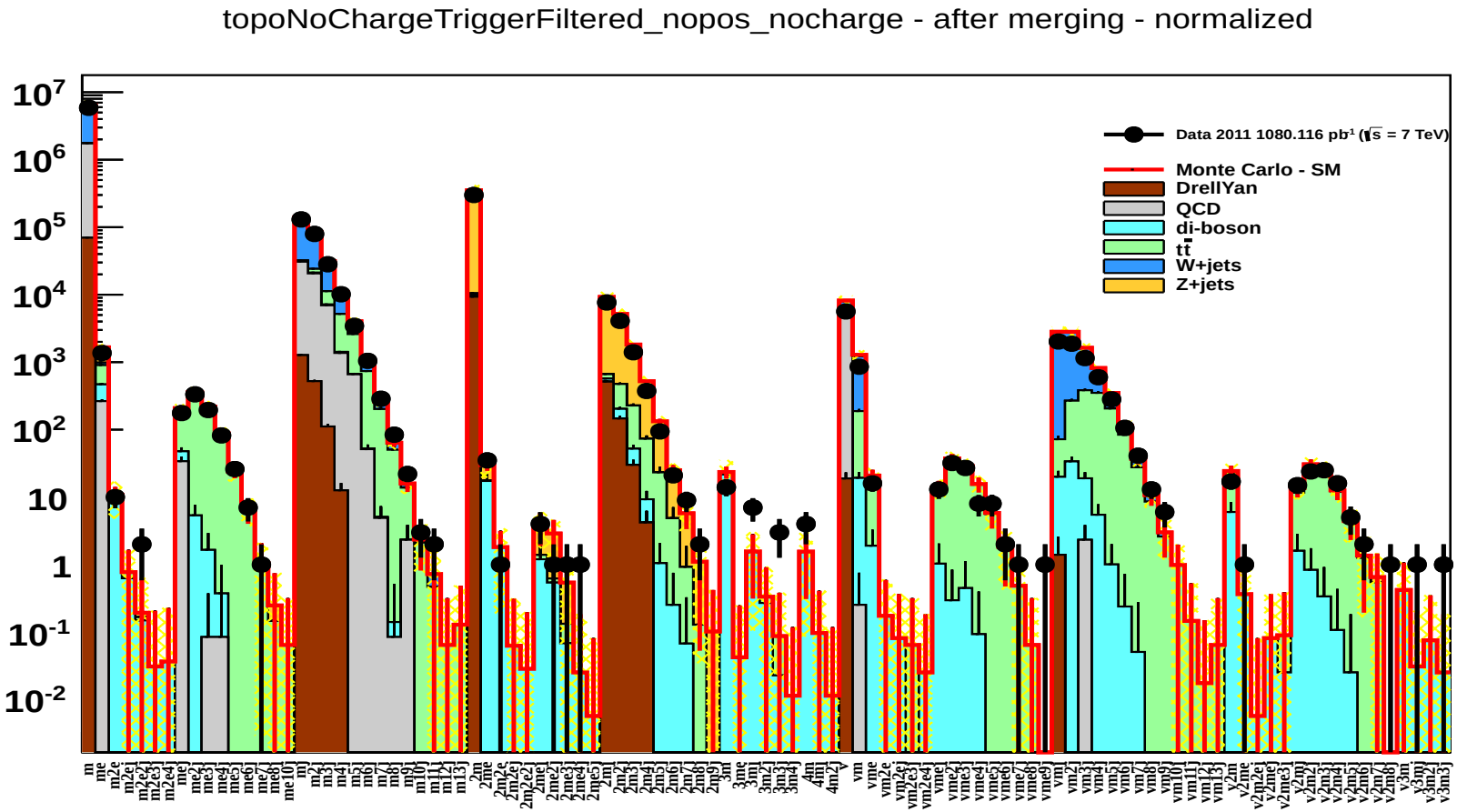
JetTauEtmiss:

- EF_j180 period B
- EF_j180 period $\geq$ D

MC samples

Sample	Generator
QCD	Pythia
Top	MC@NLO
Zll	Alpgen+Jimmy
Znunu	Alpgen+Jimmy
Wbb	Alpgen+Jimmy
Wlnu	Alpgen+Jimmy
DiBoson	Herwig
Drell-Yan	Pythia

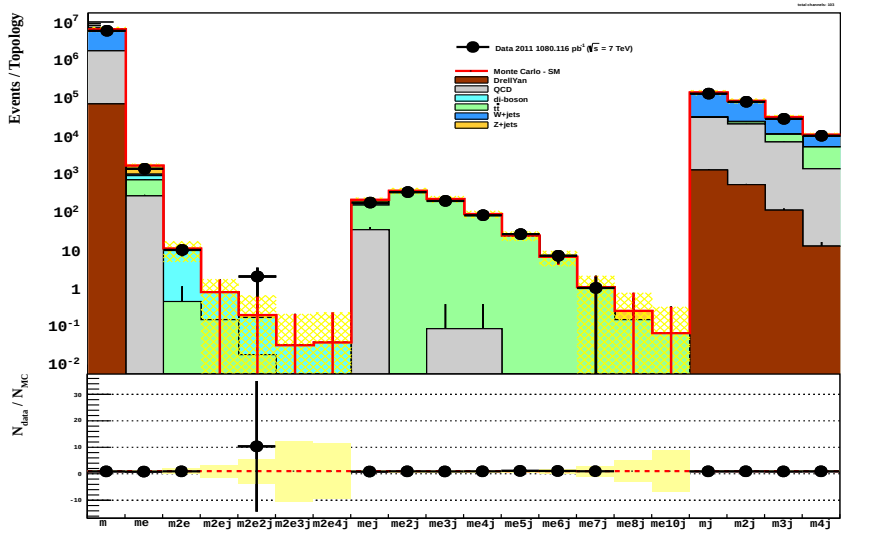
Global Plot - Muons



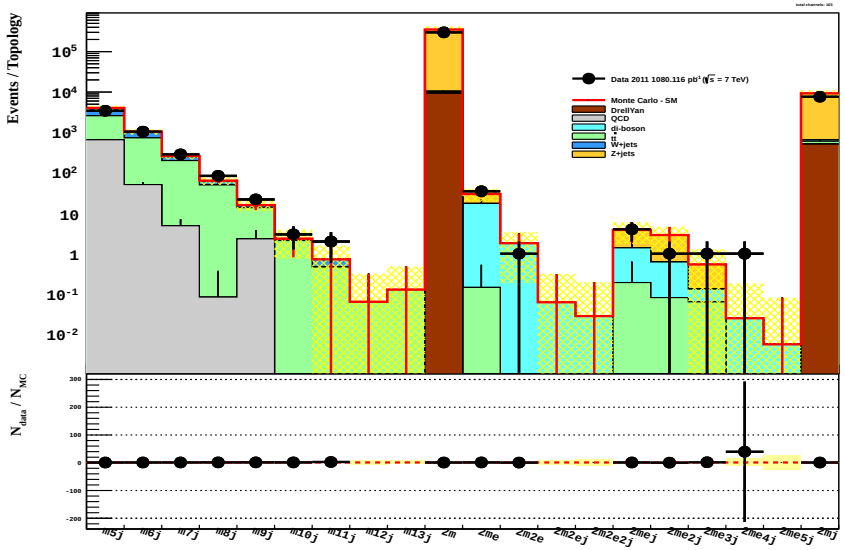
- Outstanding DATA/MC agreement over many order of magnitudes
- Even for high-multiplicity channels (up to 11 jets)

Global Plot - Muons (Zoom)

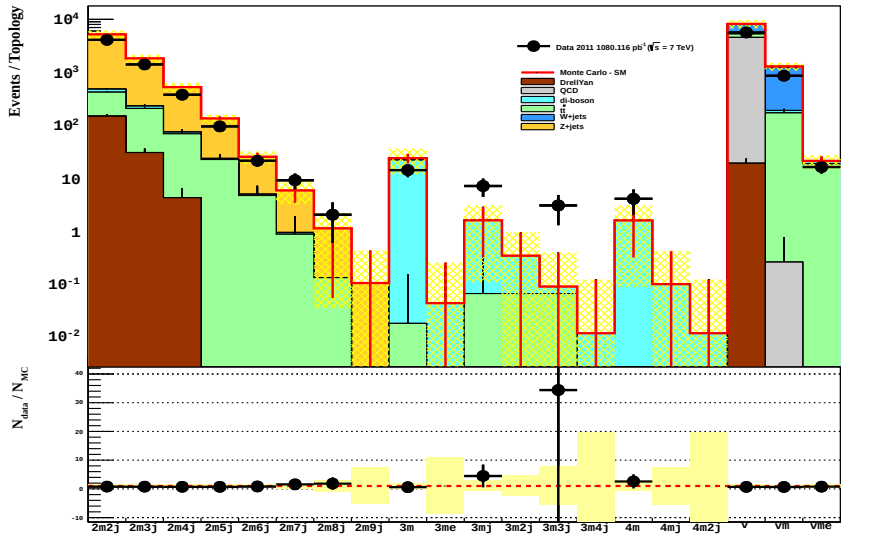
Topologies - view 1/6



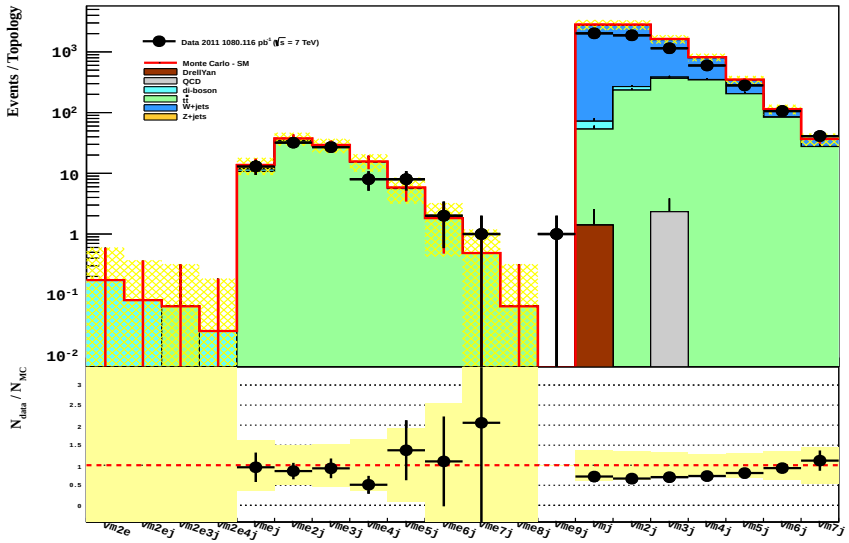
Topologies - view 2/6



Topologies - view 3/6

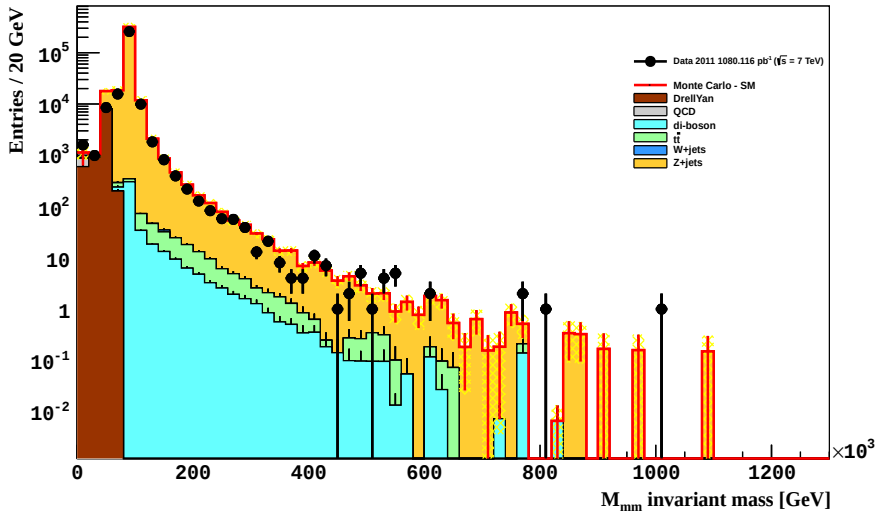


Topologies - view 4/6

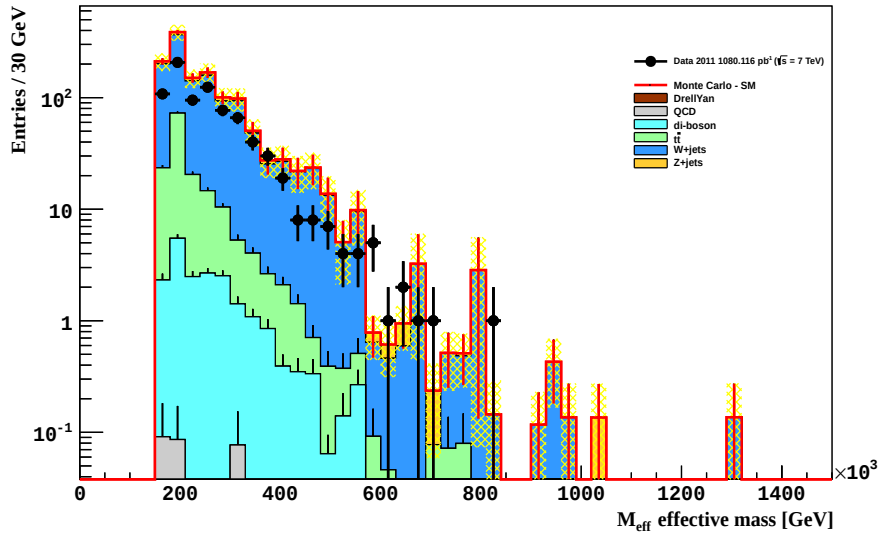


Control Plots - Muons

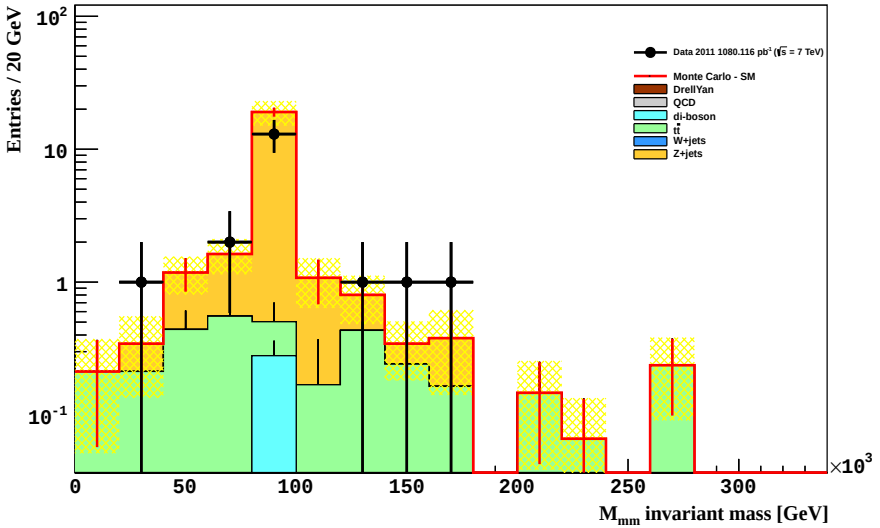
h_mm_Mmumu_Normalized - topology: '2m'



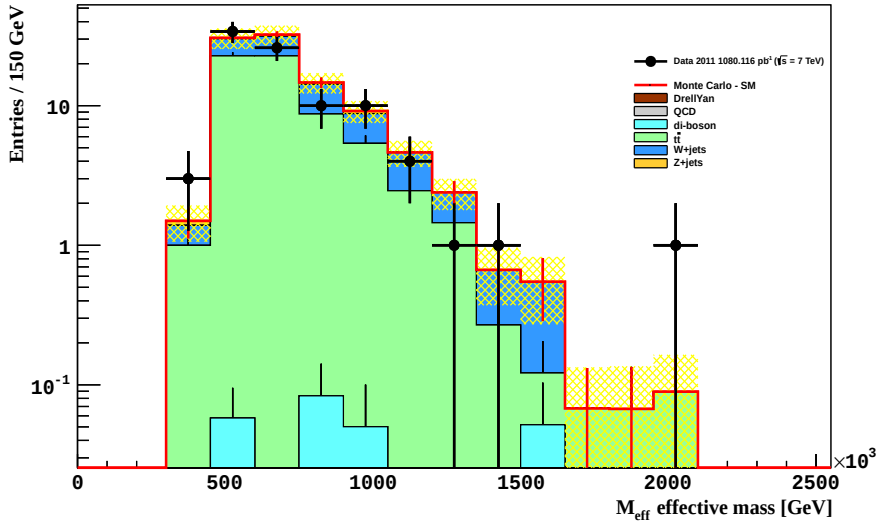
h_vm_all_meff_Normalized - topology: 'vm'



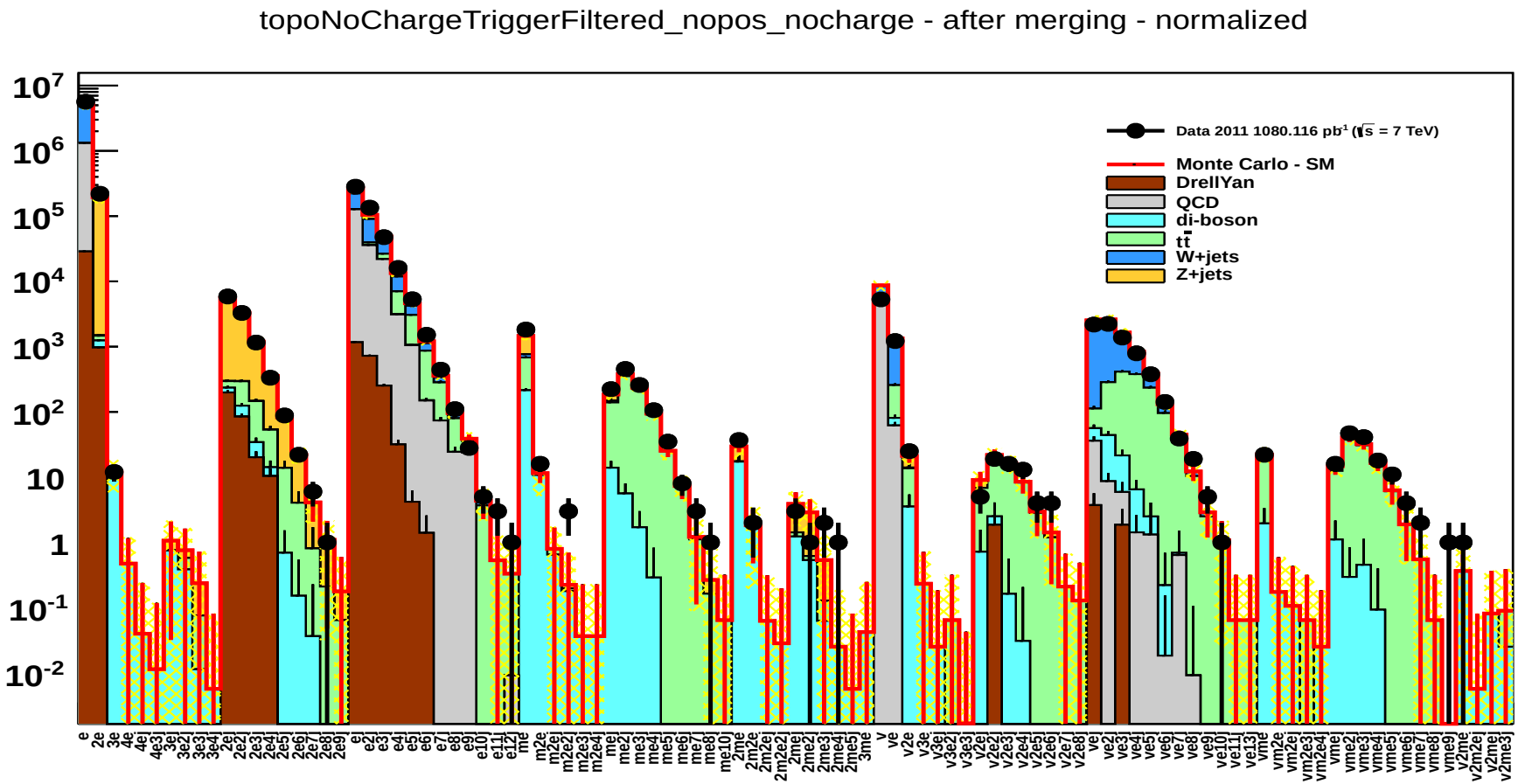
h_mmjjjjjj_Mmumu_Normalized - topology: '2m6j'



h_vmjjjjjj_all_meff_Normalized - topology: 'vm6j'

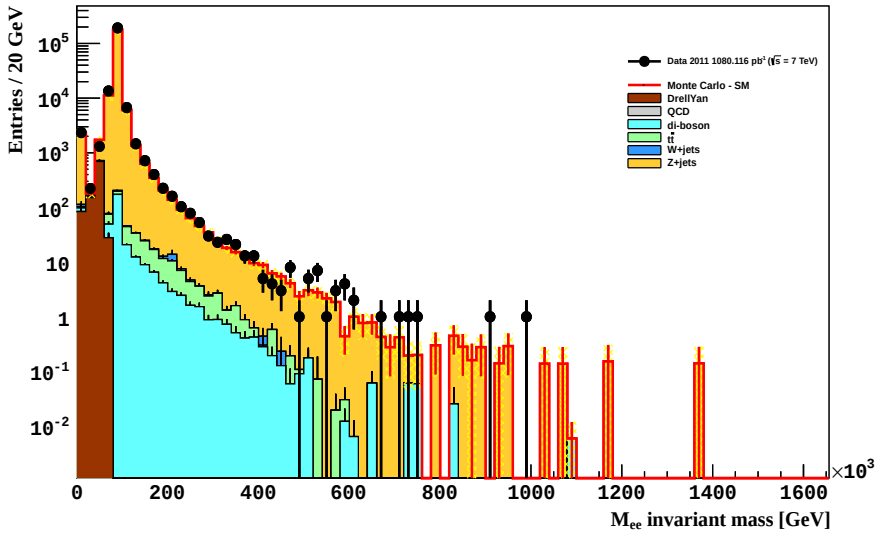


Global Plot - Electrons

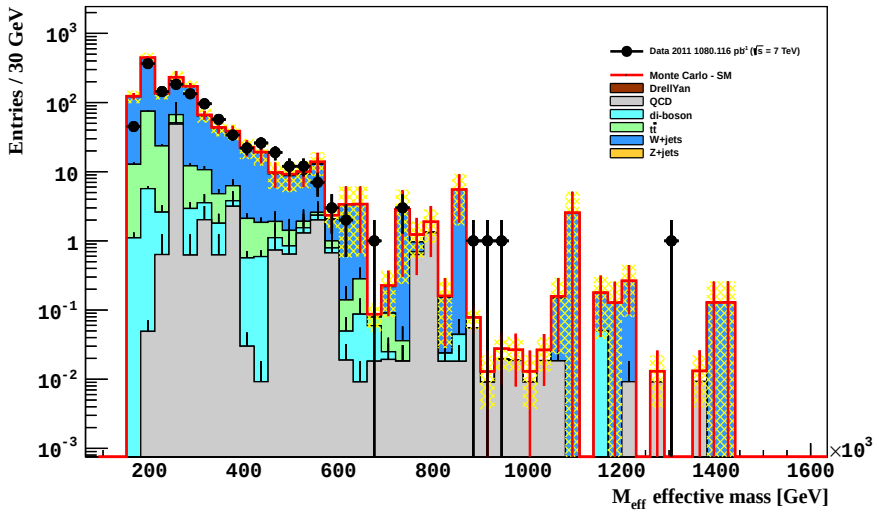


Control Plots - Electrons

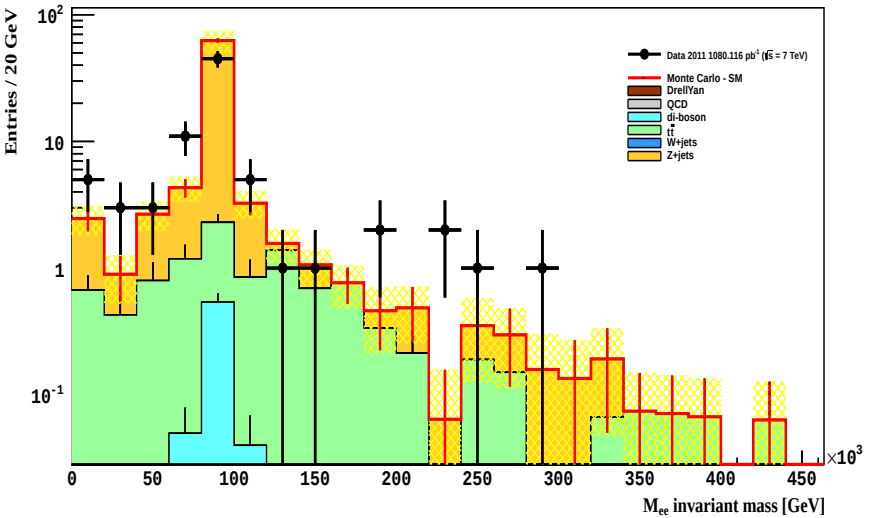
h_ee_Mee - topology: '2e'



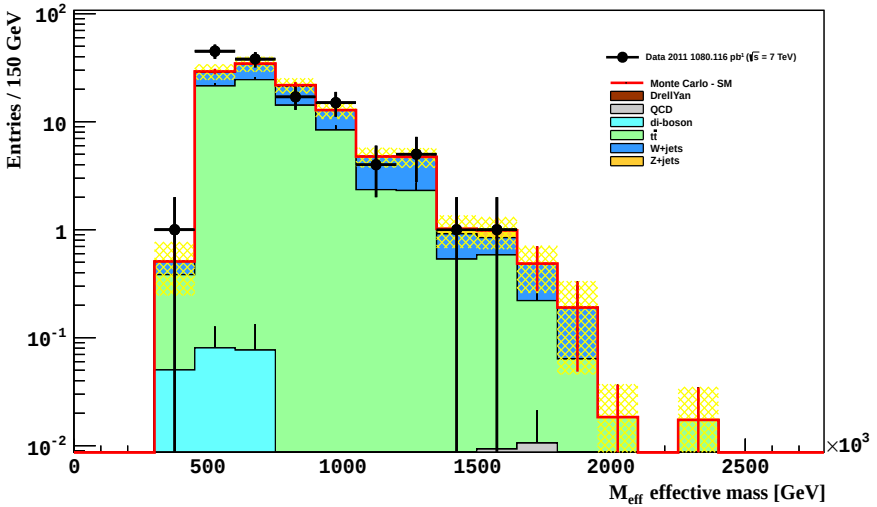
h_ve_all_meff - topology: 've'



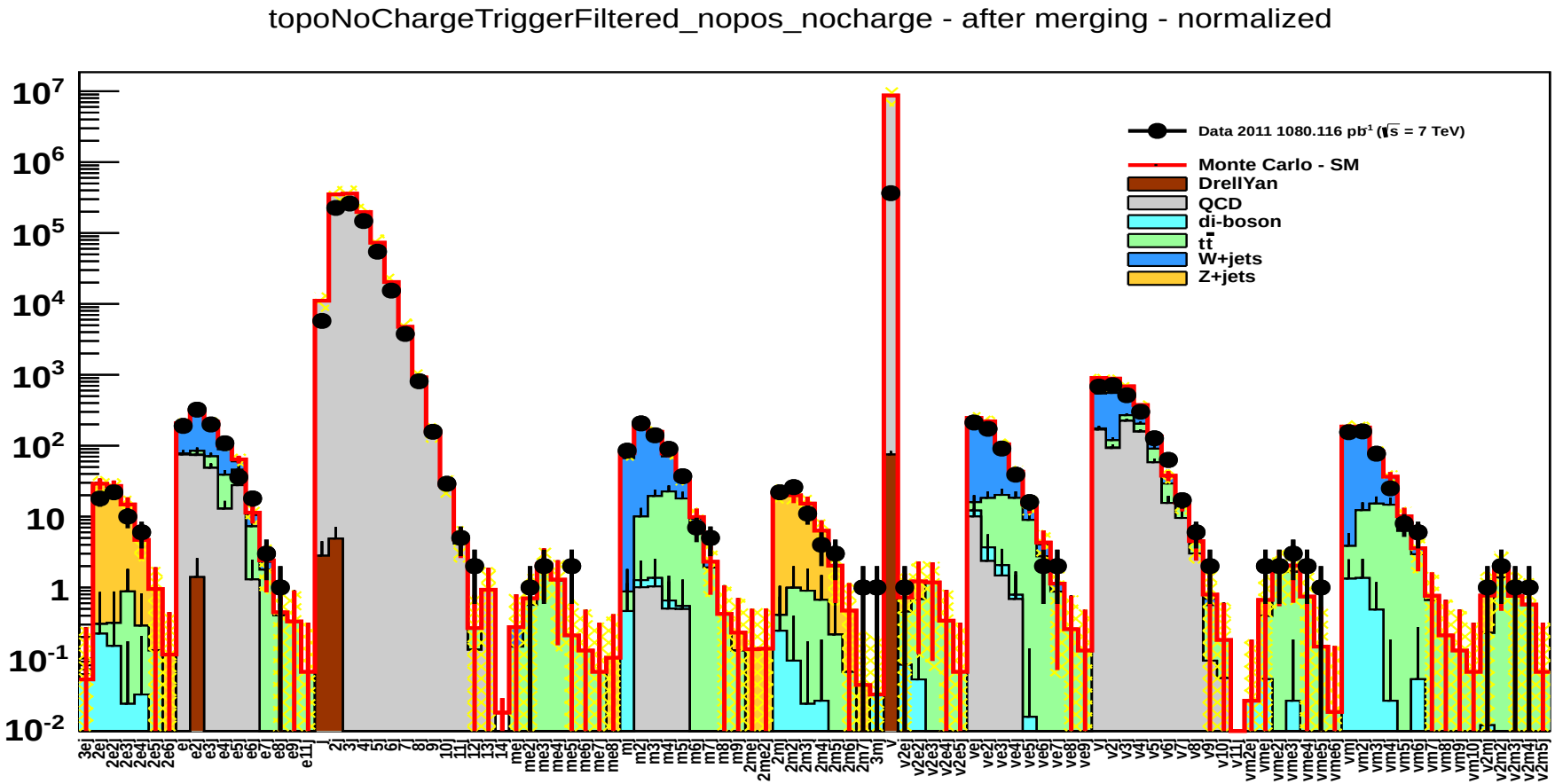
h_eejjjjj_Mee - topology: '2e5j'



h_vejjjjjj_all_meff - topology: 've6j'

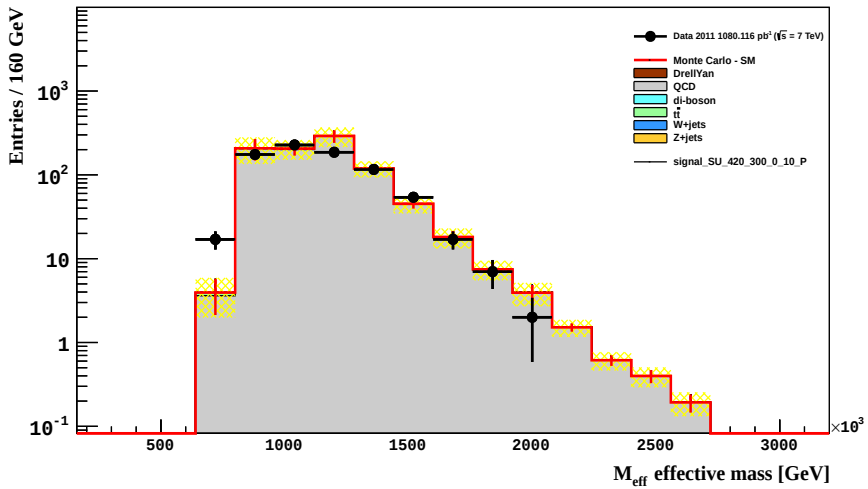


Global Plot - Jets

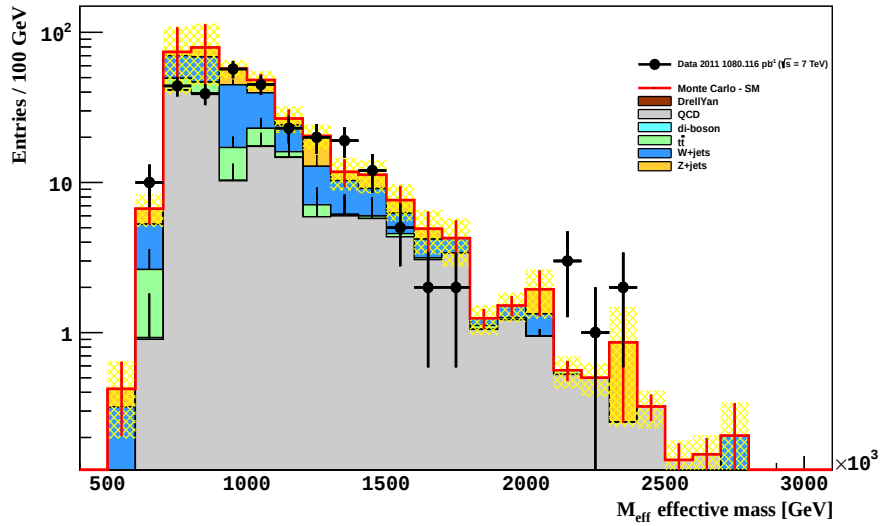


Control Plots - Jets

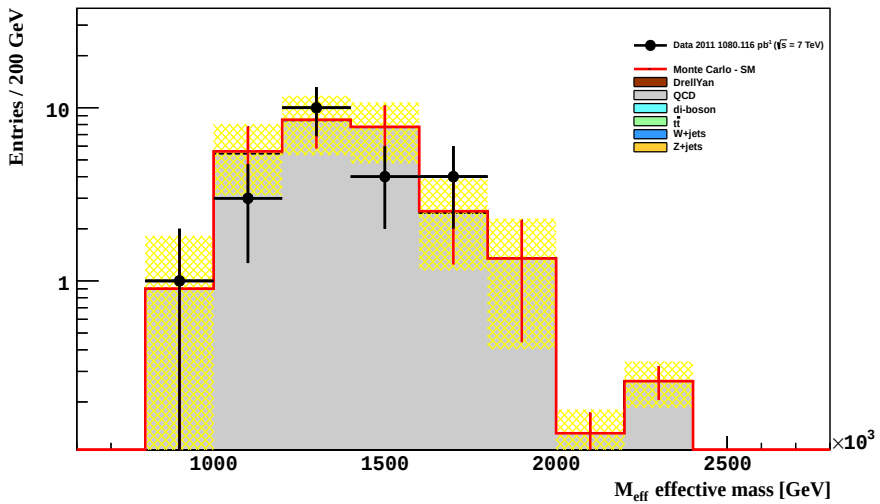
h_jjjjjjj_all_meff - topology: '8j'



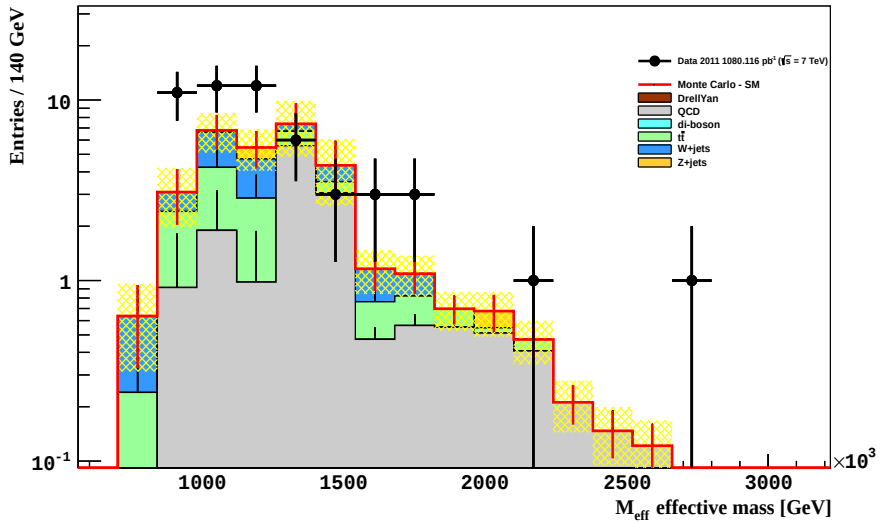
h_vjjjj_all_meff - topology: 'v4j'



h_jjjjjjjjj_all_meff - topology: '10j'



h_vjjjjj_all_meff - topology: 'v6j'

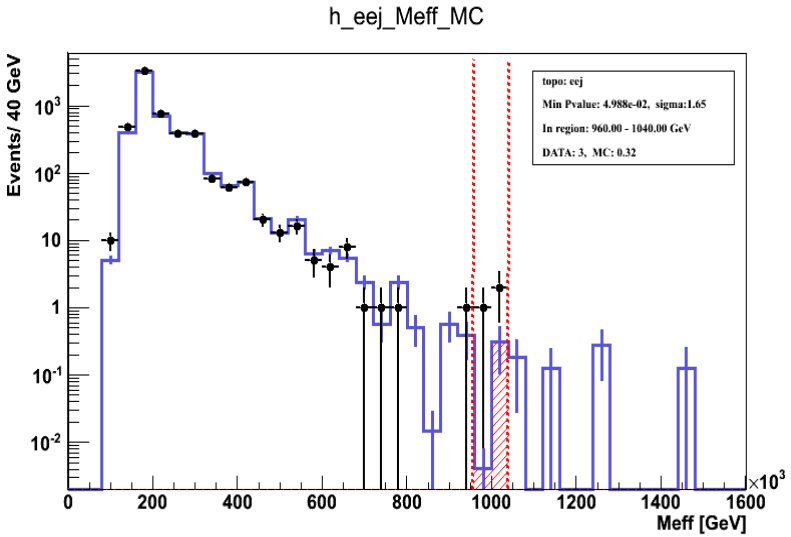


H1 Search Algorithm

We perform a scan for deviations over the M_{eff} distribution using the H1 search algorithm

$$M_{eff} = \sum P_T + \not{E}_T$$

1. Loop over all connected regions
2. Calculate p-value, with a Poisson pdf (statistics) convoluted with a Gaussian (systematics)
3. Keep the region with the smallest p-value



Correcting for the Look Elsewhere Effect

The Scan algorithm needs to be corrected for the trial factors due to the many tests being performed

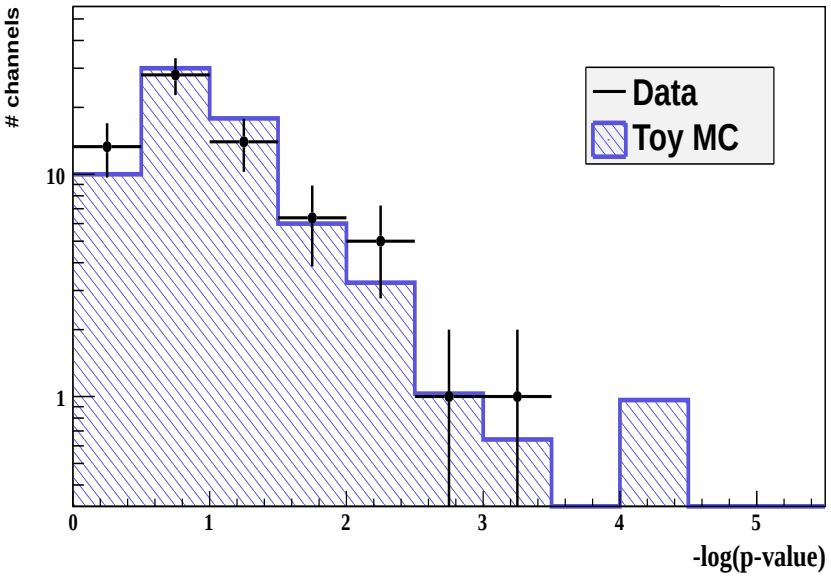
1. Generate toy experiments according to the BKG-only hypothesis
2. Compute a corrected p-value, $\tilde{p}$ as:

$$\tilde{p} = \frac{\text{toys with } p \leq p_{DATA}}{\text{total generated toys}}$$

Scan Results

Global Trial-Factors

1. Still need to account for the multiple channels we are looking at!
2. Comparing the $\tilde{p}$ distribution observed in data with the one expected from generating toys for all channels
3. Identify the fraction of experiments that would see a deviation at least as significant as the one observed



Analysis Strategy

- Select the channels with the biggest deviations
- Setup a dedicated analysis for each of those (using more refined bgk. estimation techniques)
- Confirm (or disprove) the excesses with the new unbiased dataset

Conclusions and Perspectives

- The Analysis is now solid enough and showing very promising results
- Unfortunately no sign of new physics yet

Ongoing Improvements

- Sensitivities studies for some specific NP models (W'/Z', SUSY reference points)
- Detailed accounting of systematics in the scan (JES, fake leptons)
- b-jet and photon topologies to be added soon

– Identify the list of interesting channels –

Conclusions and Perspectives

- The Analysis is now solid enough and showing very promising results
- Unfortunately no sign of new physics yet

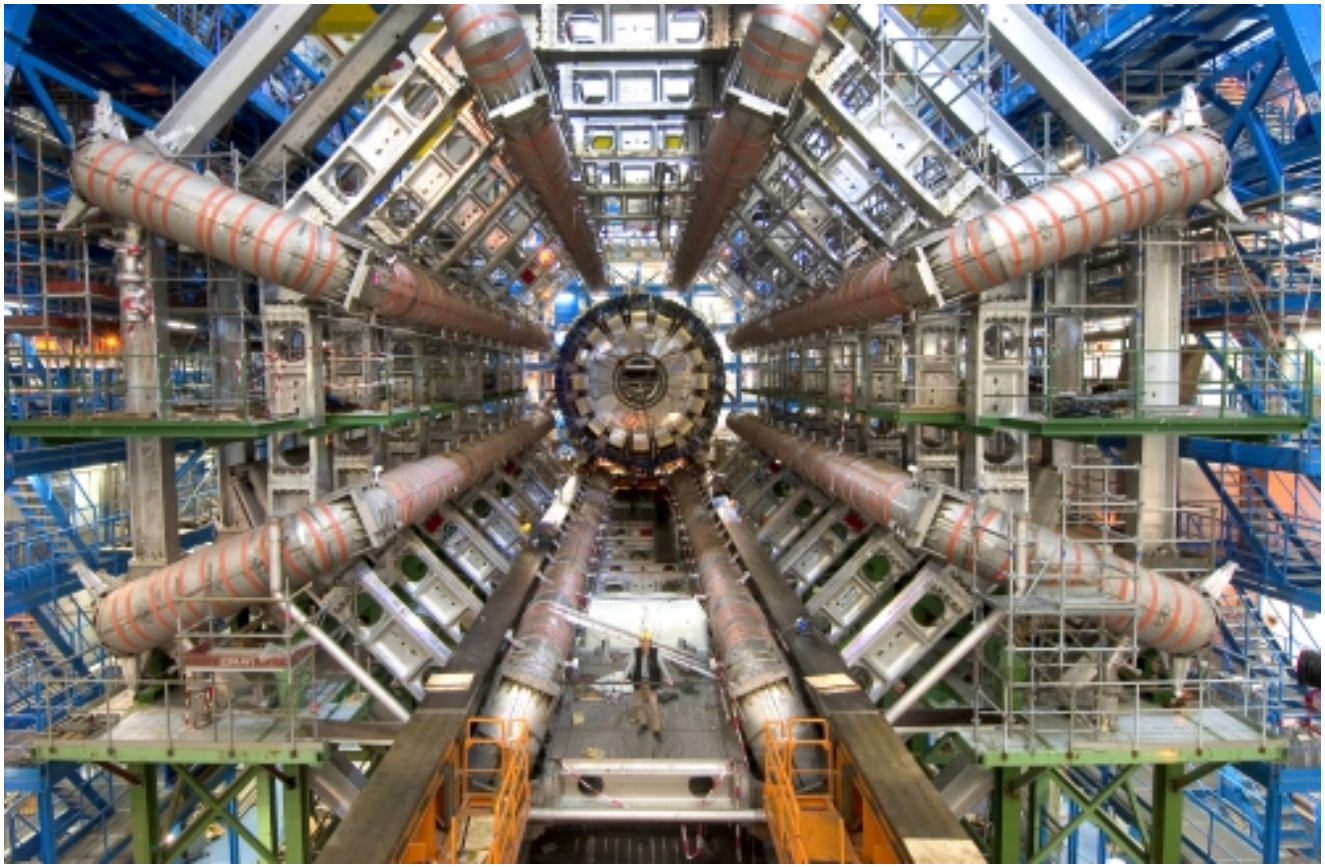
Ongoing Improvements

- Sensitivities studies for some specific NP models (W'/Z', SUSY reference points)
- Detailed accounting of systematics in the scan (JES, fake leptons)
- b-jet and photon topologies to be added soon

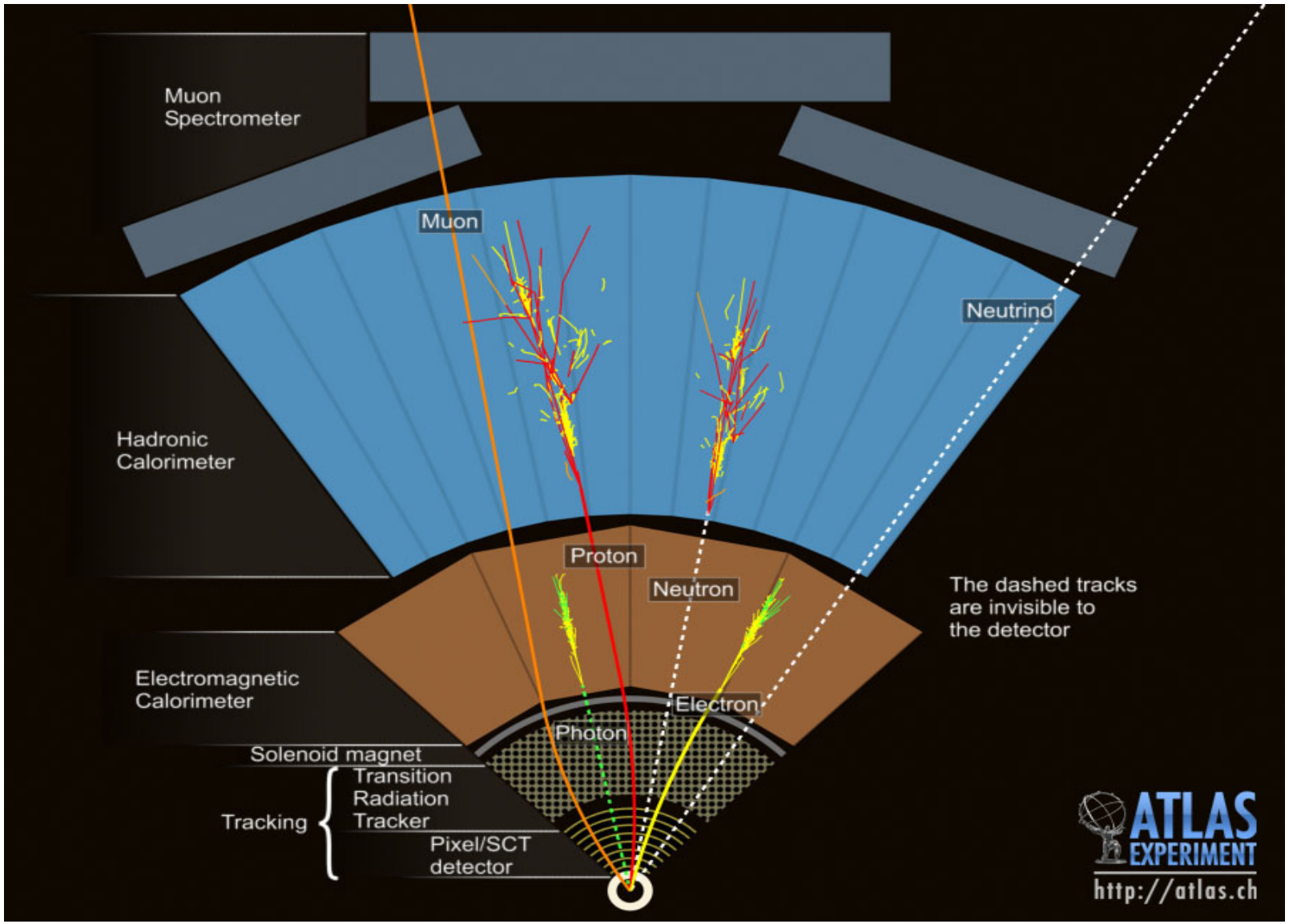
– Identify the list of interesting channels –

STAY TUNED!!

Backup Slides

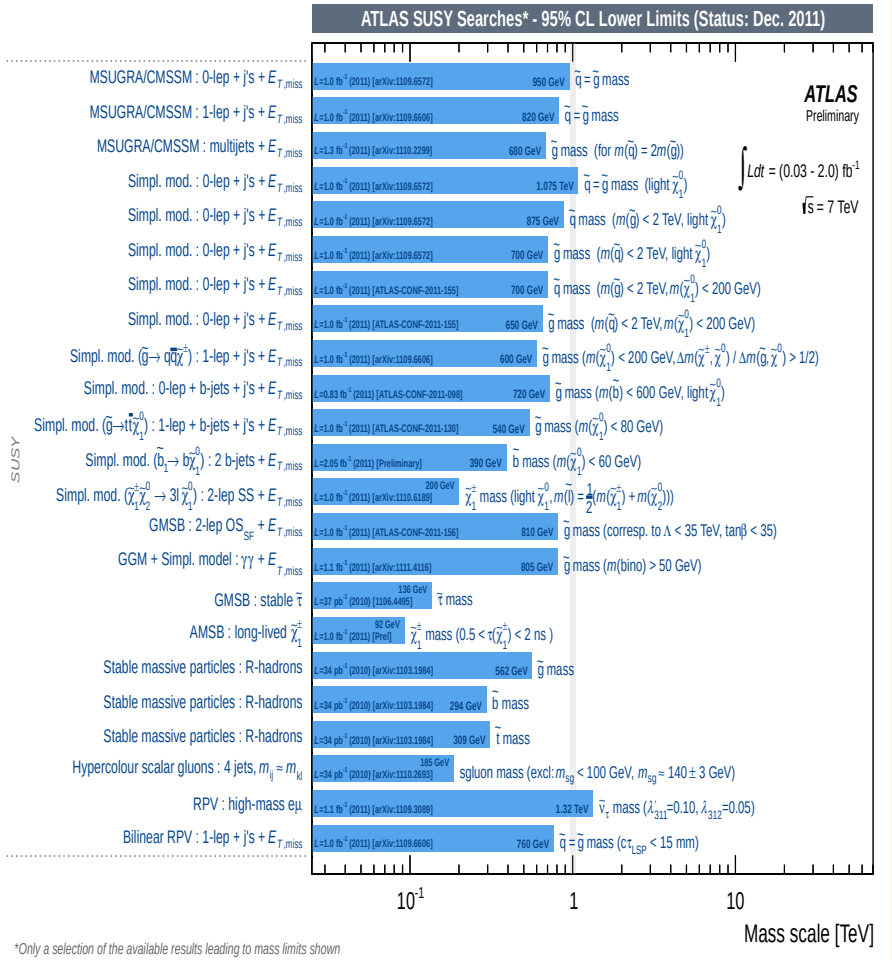


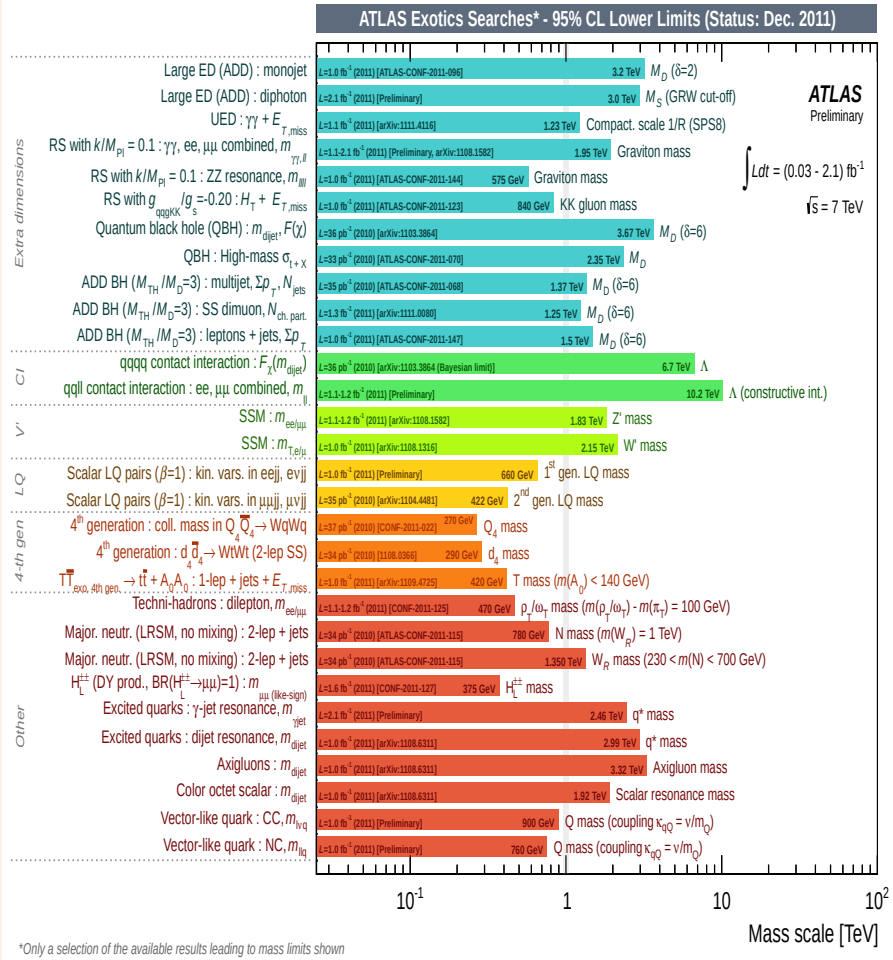
Objects Reconstruction



BSM Searches Summary

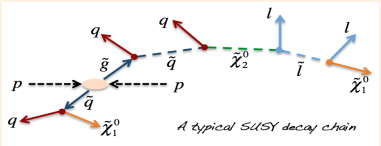
Beyond the Standard Model Searches

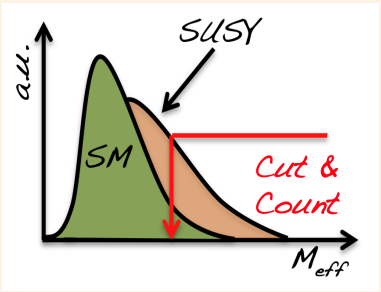




A typical analysis

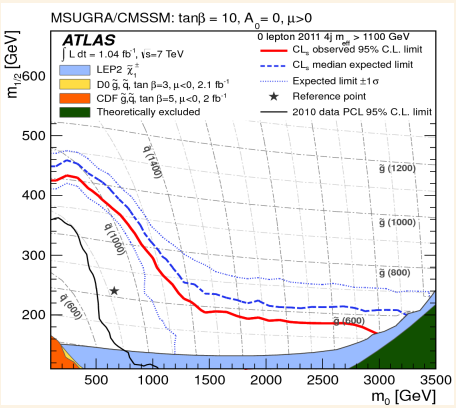
- Take your favourite model for NP (or ask your best theory friend!)
- Generate simulated events
- Choose your final states according to the specific model signatures, and optimize selections and acceptance for maximum sensitivity
- Test the observed data against the null (SM) hypothesis





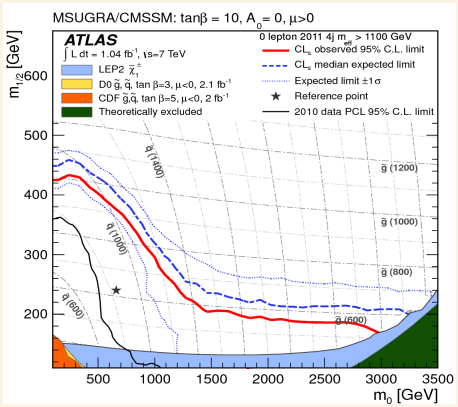
$P(data|SM) < 10^{-7}$

You found New Physics!!!



$P(data|SM) >> 10^{-7}$

Set 95% CL limits



Object Definitions

Object Definitions

- **electron:**
  - isTight = True
  - etaMax = 2.47
  - ptMin = 25 GeV
  - ZVertexVeto = True
- **muon:**
  - isTight = True
  - etaMax = 2.4
  - ptMin = 20 GeV
  - ZVertexVeto = True
- **jet:**
  - DeltaPhi($\not{E}_T$, first 3 jets) > 0.4 Leptons Streams
  - DeltaPhi($\not{E}_T$, first 4 jets) > 0.2 Jet Stream
  - etaMax = 2.8
- **missing energy:**
  - $\not{E}_T > 125$ GeV
  - Corrected for leptons and jets contribution

Cleaning Cuts

- Jet Cleaning Cuts
- veto events with a cosmic muon
- clean vertex (more than 4 tracks)
- Smart LAr Hole event veto
- PileUp and lepton reconstruction reweighting is applied to MC

The Merging algorithm

- Two treshold are set given the Background expectations
  - Threshold_Low **0.001**
  - Threshold_Up **0.05**
- We loop over all channels populated either by DATA or BKG events
- For each channel, we look if BKG expectation > Threshold_Low
  - If yes we take the channel as it is
  - If not we start mergin it
    - Remove an object from the topology string, in this order: "j", "e", "m", "v"
    - Check again if BKG expectation > Threshold_Low
    - If yes check if the new channel has BKG exp. > Threshold_Up
    - If not: merge the channel into the new one
    - If yes: take the old channels as it is. This avoids burying interesting channels under very populated channels, losing sensitivity about them
    - If not remove another object from the Dynamic topology iteratively, until one of the following conditions is fullfilled:
    - We find a channel with BKG exp. > Threshold_Up
    - We take the channel as it is
    - After removing all objects the Dynamic topology goes empty

p-value definition

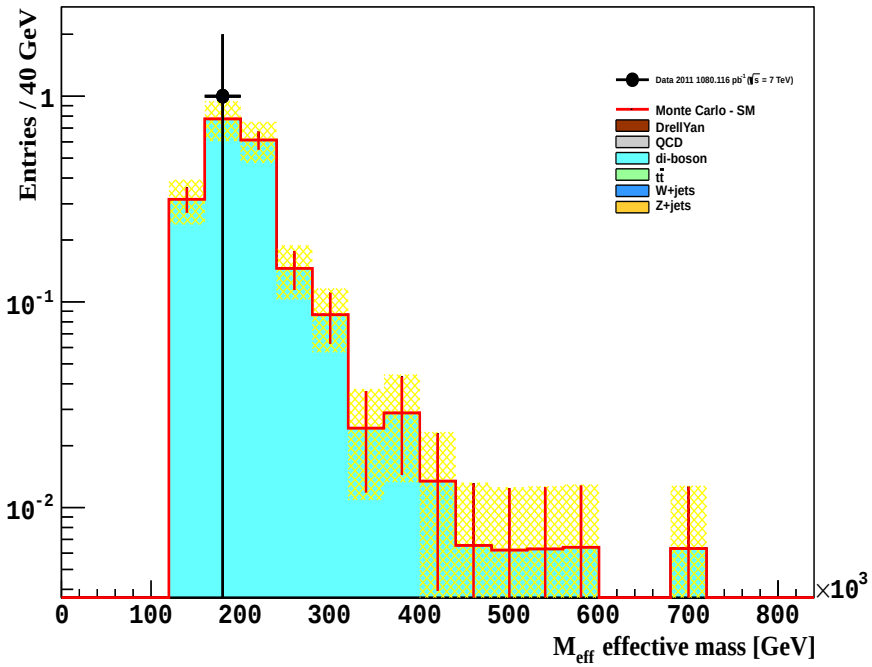
Following the H1 statistical analysis the following p-value has been defined

$$p = \begin{cases} \sum_{i=N_{data}}^{\infty} A \cdot \int_0^{\infty} db \; exp\left(\frac{-(b-N_{SM})^2}{2(\delta N_{SM})^2}\right) \cdot \frac{e^{-b} b^i}{i!} & N_{data} > N_{SM} \\ \sum_{i=0}^{N_{data}} A \cdot \int_0^{\infty} db \; exp\left(\frac{-(b-N_{SM})^2}{2(\delta N_{SM})^2}\right) \cdot \frac{e^{-b} b^i}{i!} & N_{data} < N_{SM} \end{cases}$$

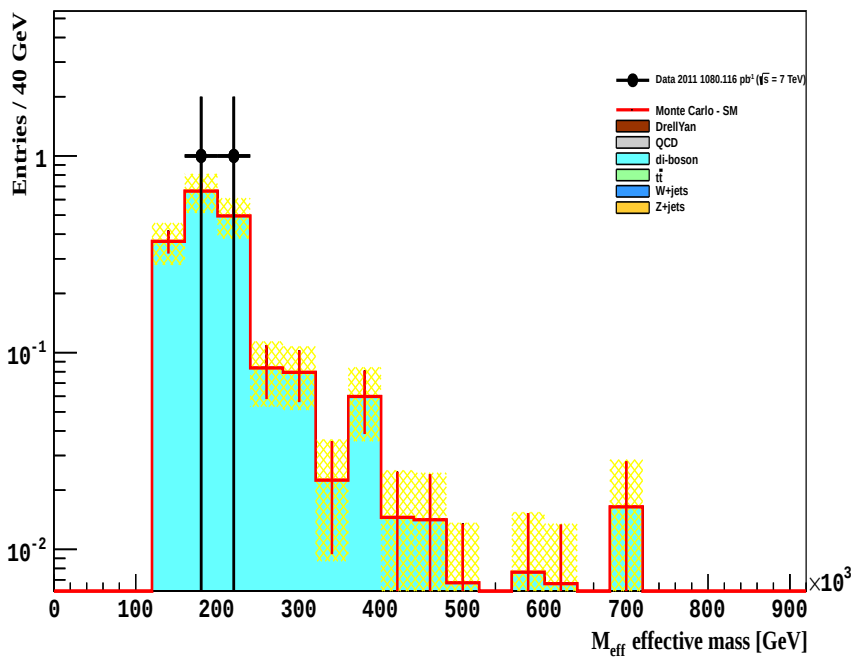
- Convolution of a Poisson term for statistical fluctuations, plus a Gaussian to account for systematics uncertainties.
- It is a hybrid frequentist/bayesian approach (Cousins - Highlands)

Diboson Events

h_mmee_all_meff_Normalized - topology: '2m2e'



h_mmmm_all_meff_Normalized - topology: '4m'



- Our first diboson candidates events