MUSiC

Model Unspecific Search in CMS

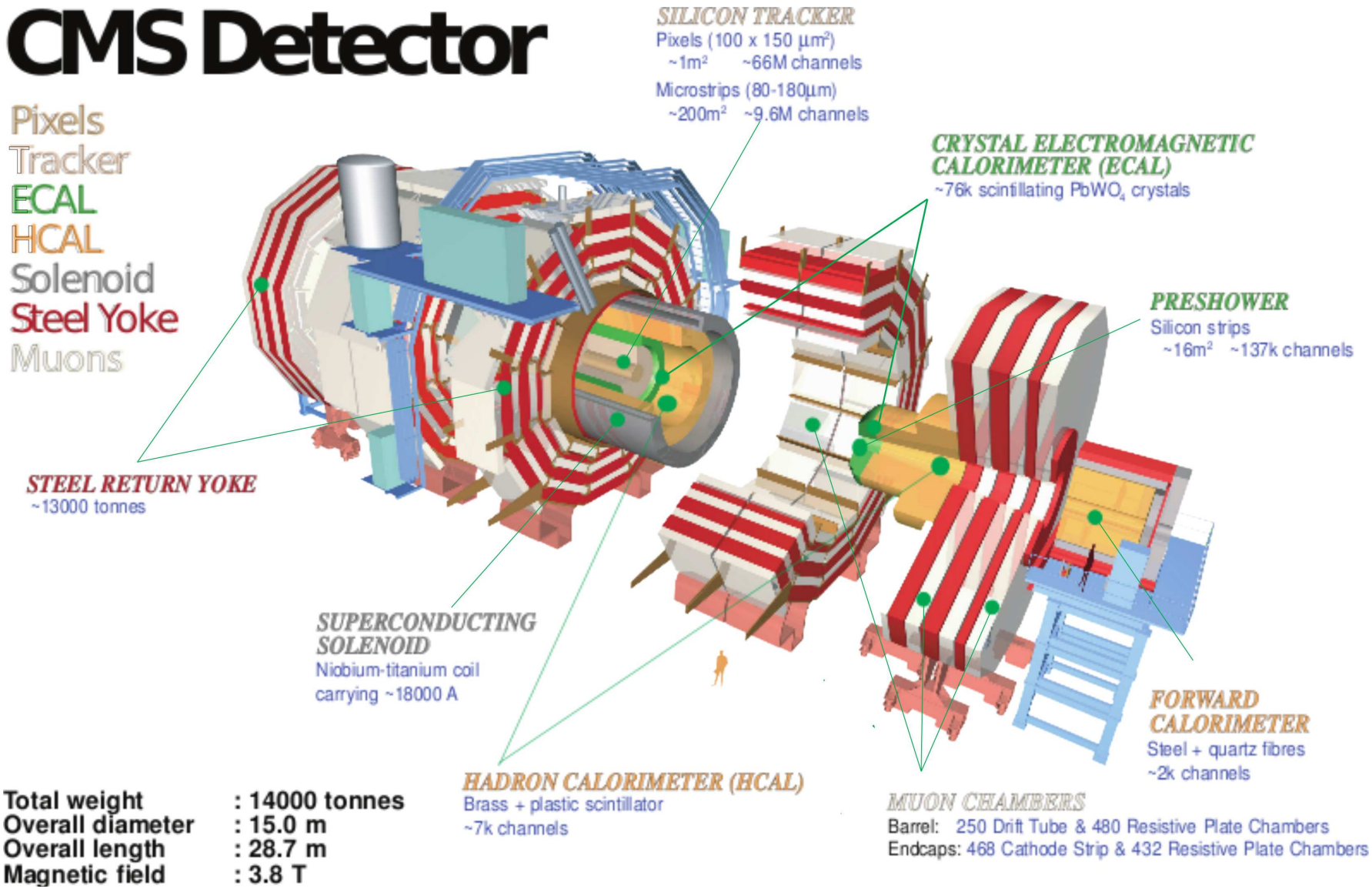
Holger Pieta on behalf of the CMS collaboration

Physikalisches Institut III A, RWTH Aachen University

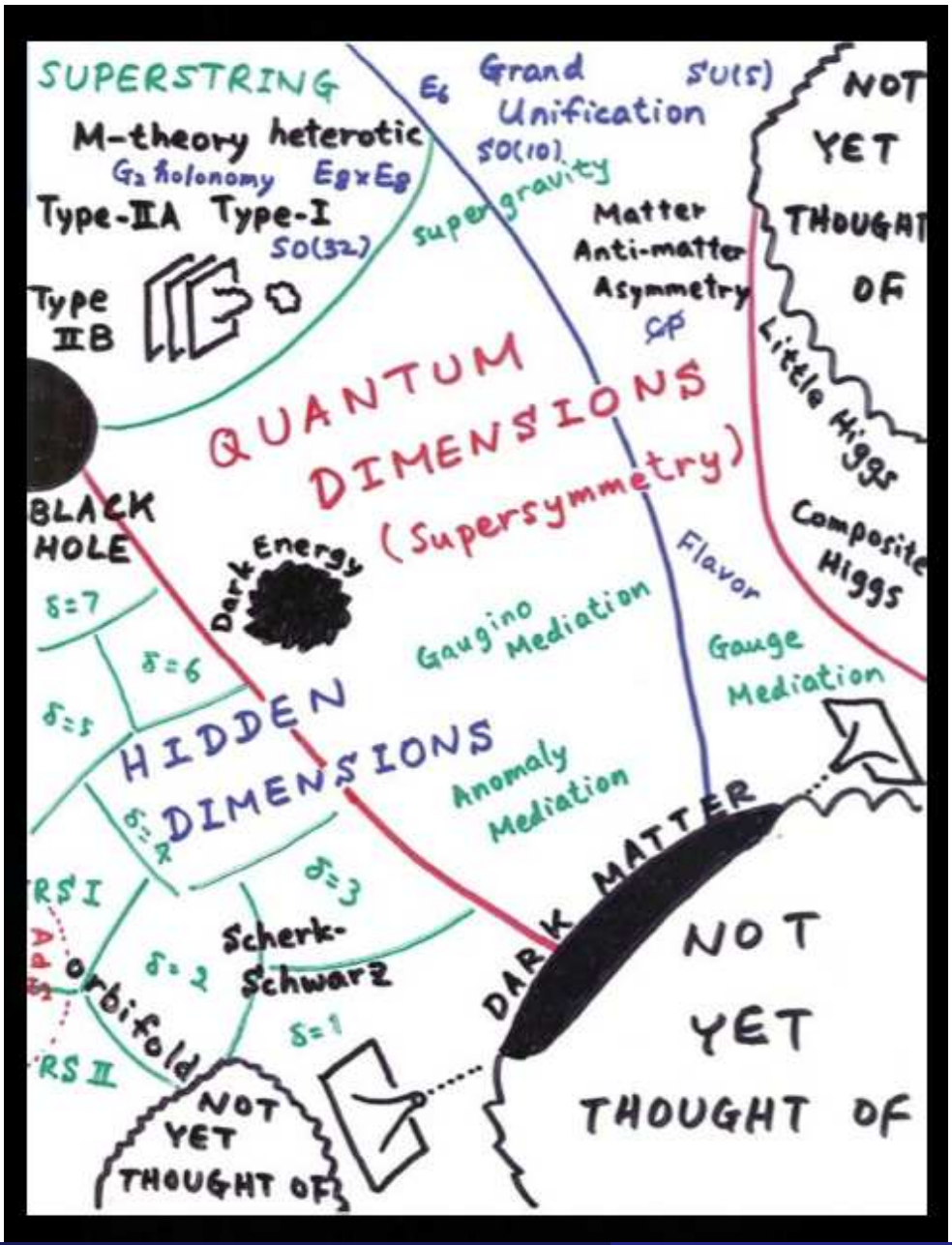
August 27, 2010

Introduction

CMS - The Compact Muon Solenoid



Theory landscape



Challenge

Can we have analyses for all of that?
Even the not-yet-thought-off?

Example: SUSY

Not one model, but multitude of parametrizations and free parameters.

Idea

Model independent analysis

The approach

Minimize theoretical bias

- Assume just one model: The Standard Model
- Look at (almost) all events
- Look for deviations from the Standard Model expectation

Not new

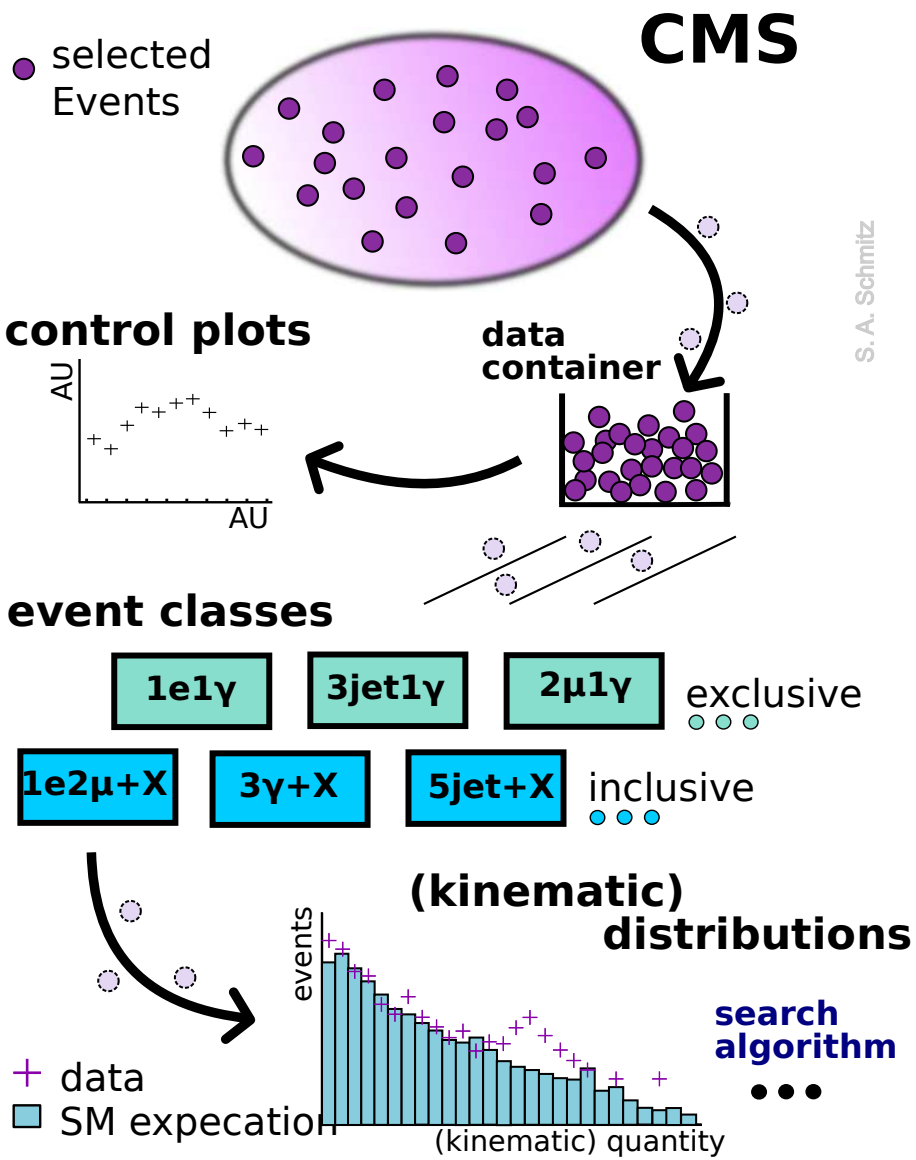
Successfully performed at: L3, D0, H1, CDF, ...

Details in note: CMS PAS EXO-08-005

`http://cms-physics.web.cern.ch/cms-physics/public/EXO-08-005-pas.pdf`

Implementation details

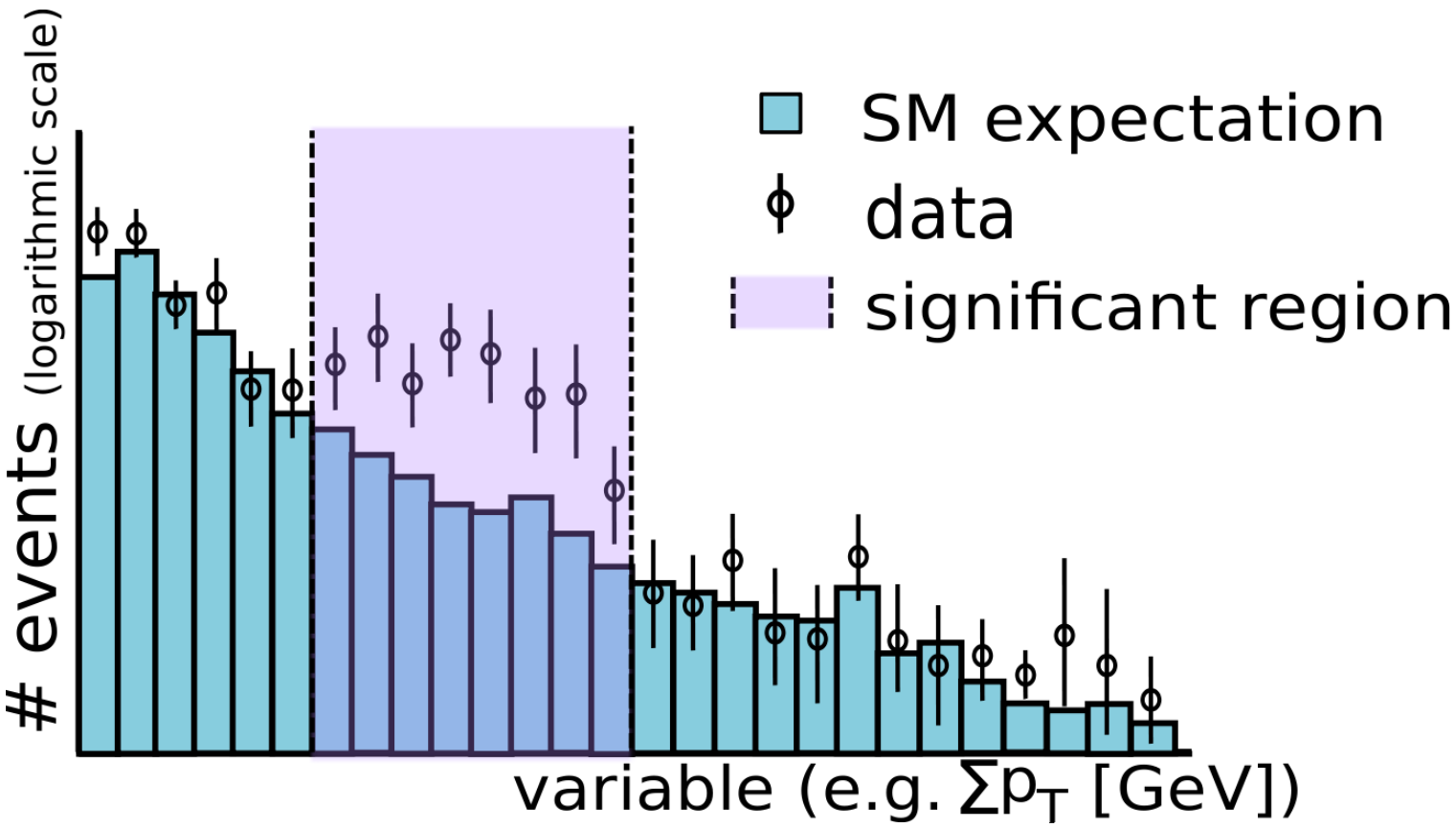
Workflow



Workflow

- Select events and objects with accepted cuts
- Generate control plots
- Sort events into classes depending on their object content: μ, e, γ, jets, MET
- Generate kinematic distributions of these classes: $\sum p_T$, M_{inv} and MET
- Scan distributions for the most significant region of connected bins

Region of Interest: Connected bins/Sliding window



Most significant region

Select region with lowest probability of MC to yield even more events, commonly called p_{data} or simply "p-value".

p-value explained

Default: Gauss-Poisson convolution

$$p_N = \sum_{i=n_d}^{\infty} \int_0^{\infty} Gauss(\mu = n_{MC}, \sigma = \sigma_b, b) \cdot Poisson(\lambda = b, i) db$$

In case of $n_d < n_{MC}$: $\sum_{i=0}^{n_d}$

Alternative: LogNormal-Poisson convolution

$$p_N = \sum_{i=n_d}^{\infty} \int_0^{\infty} Gauss(\mu = \log(n_{MC}), \sigma = \log(\sigma_b), \log(b)) \cdot Poisson(\lambda = b, i) db$$

Better treatment of some uncertainties, worse for others

Good cross check

Systematic uncertainties

Monte-Carlo

- Standard model cross sections (often guesswork, e.g. 10 %)
- Parton density functions (as recommended by e.g. CTEQ)
- MC statistics

Detector

- Luminosity (11 %)
- Reconstruction efficiencies (few % usually)
- Fake probabilities (safe side: 100 %)
- Jet energy scale (10 %)

Statistical challenge

p-value

Best effort: Include all knowledge about statistical and systematic errors

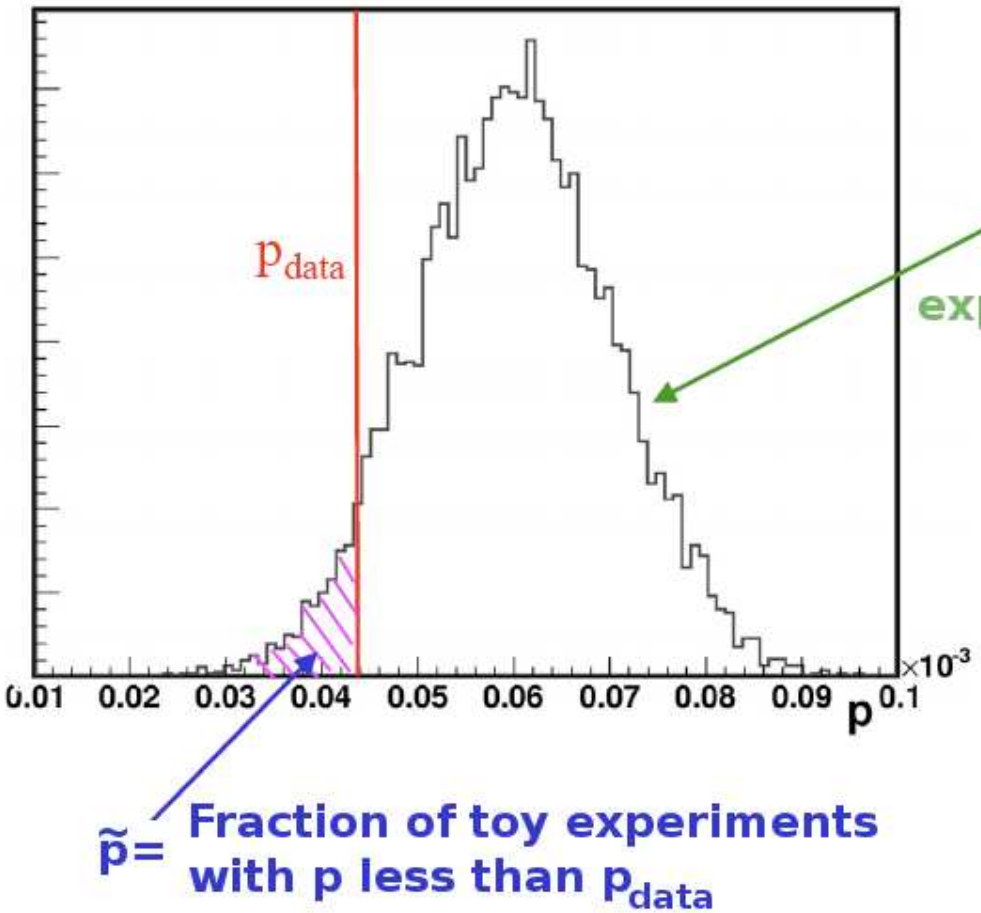
Selection bias: Look-elsewhere-effect

p-value distribution expected to be flat in a random selection
Selecting the most significant region: Bias towards smaller p-values
Consequence: Most-significant-region p-value not suited as significance estimator

A way out

Take the selection bias into account: $\tilde{p}$
Hard to impossible to calculate analytically
Solution: Do toy experiments

$\tilde{p}$ explained



- Randomize MC expectation, taking all known uncertainties into account
- Scan for most significant region
- Count toy experiments with higher significance than data
- Fraction of toy experiments with $p_{toy} > p_{data}$: $\tilde{p}$

Scope of MUSiC

What is MUSiC

- Alarm system
- Physics and detector monitor
- Complementary to dedicated analysis
- No automated discovery and PhD thesis tool

MUSiC found something?

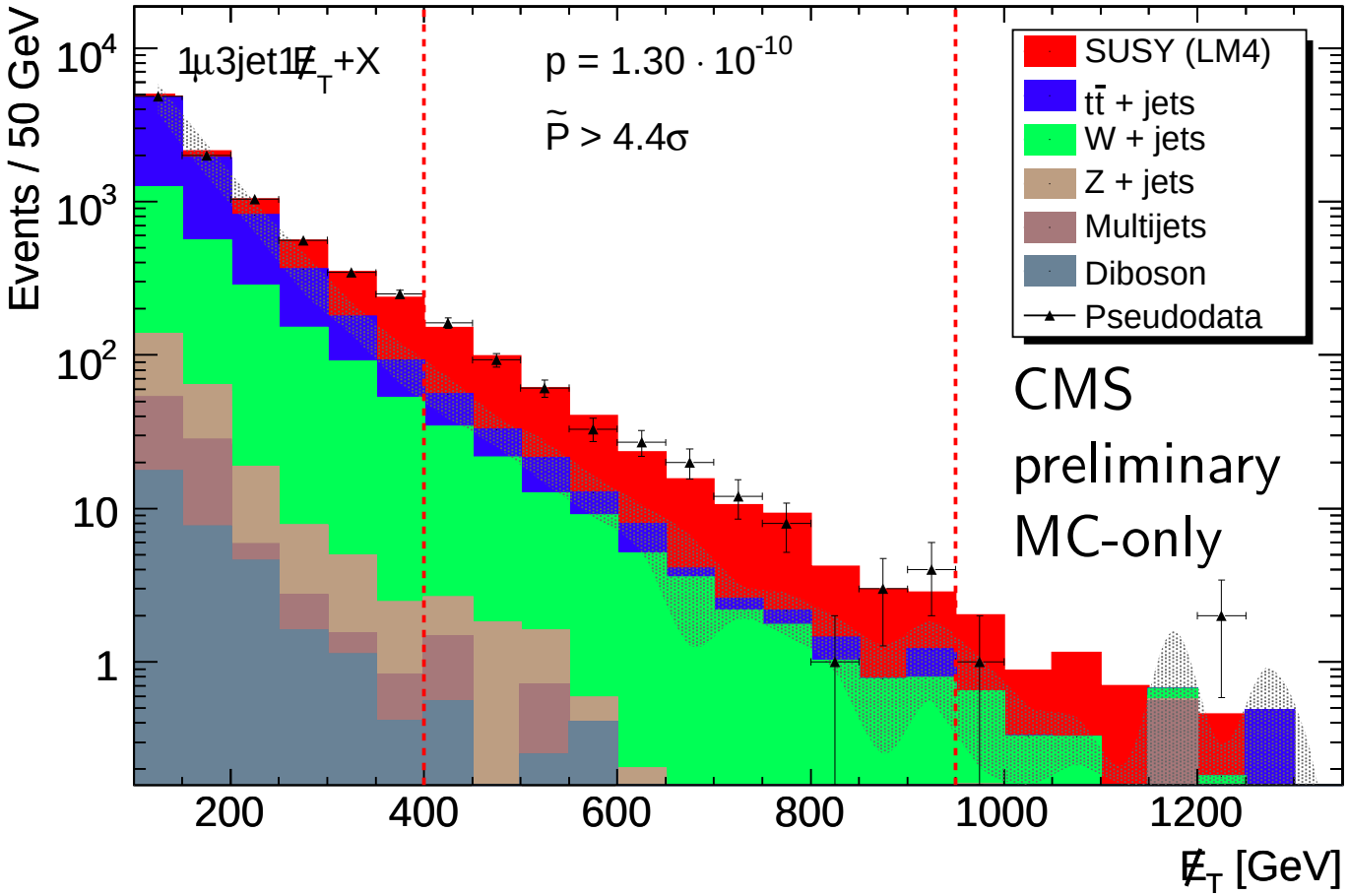
- Detector effects?
- Monte-Carlo problems (i.e. tuning)?
- New physics?
- $\rightarrow$ Careful investigations needed

Benchmarks (MC-only)

mSUGRA Supersymmetry - 1

CMS Low Mass benchmark point 4

$m_o = 210$ GeV, $m_{1/2} = 285$ GeV, $\tan\beta = 10$, $\mathrm{sgn}\mu = +$, $A_0 = 0$



Result

Many significant deviations, typically with high multiplicities

mSUGRA Supersymmetry - 2

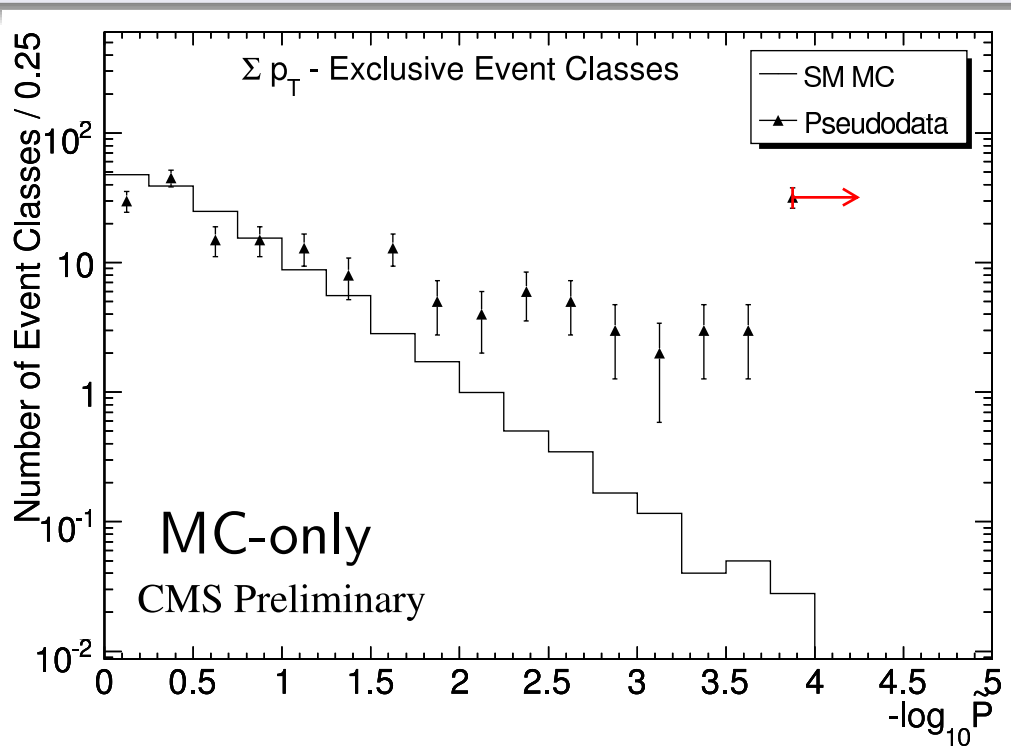
$\tilde{p}$ ditribution

$\tilde{p}$ is a probability

$\rightarrow$ Flat distribution between 0 and 1, if just statistical fluctuations

$\rightarrow$ Linear in a log-log plot

$\rightarrow$ Deviation from this expectation: Systematic effect (e.g. discovery)



Discovery

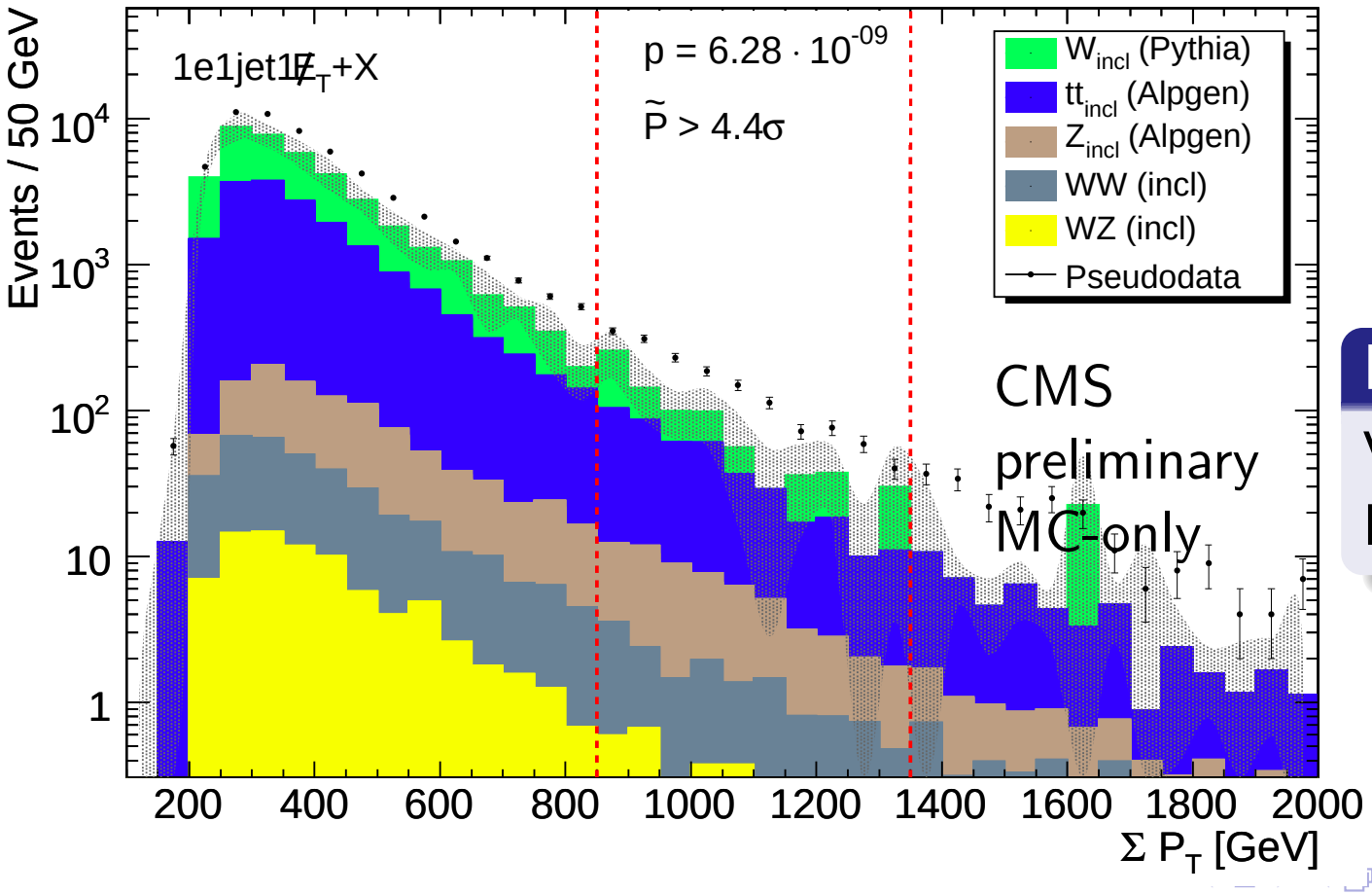
Distribution shiftet to smaller values

$\rightarrow$ discovery

Monte-Carlo tuning

ALPGEN vs. Pythia

ALPGEN standard model W samples as signal, compared to Pythia standard model W and other backgrounds, so no new physics

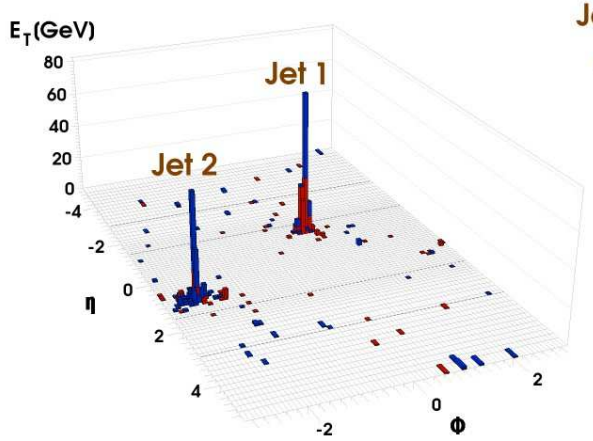


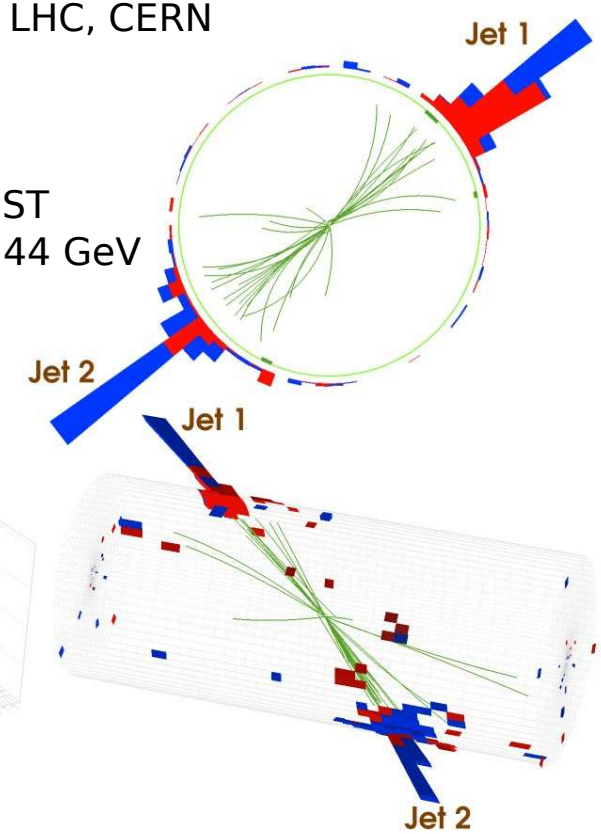
Result

Visible in Lepton+Jets classes

Short look at first data

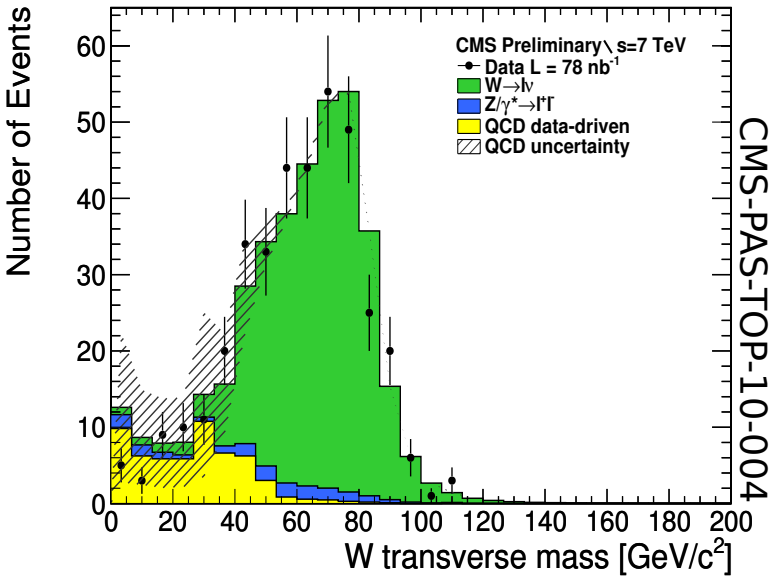
CMS Experiment at LHC, CERN
Run 133450
Event 16358963
Lumi section 285
Sat Apr 17 2010, 12:25:05 CEST
Jet1: pT=253 GeV; Jet2: pT=244 GeV
Dijet mass: 764 GeV

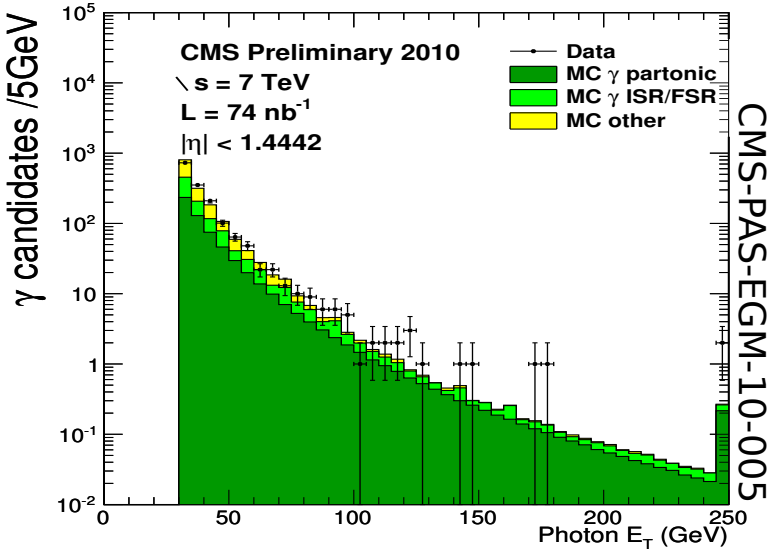


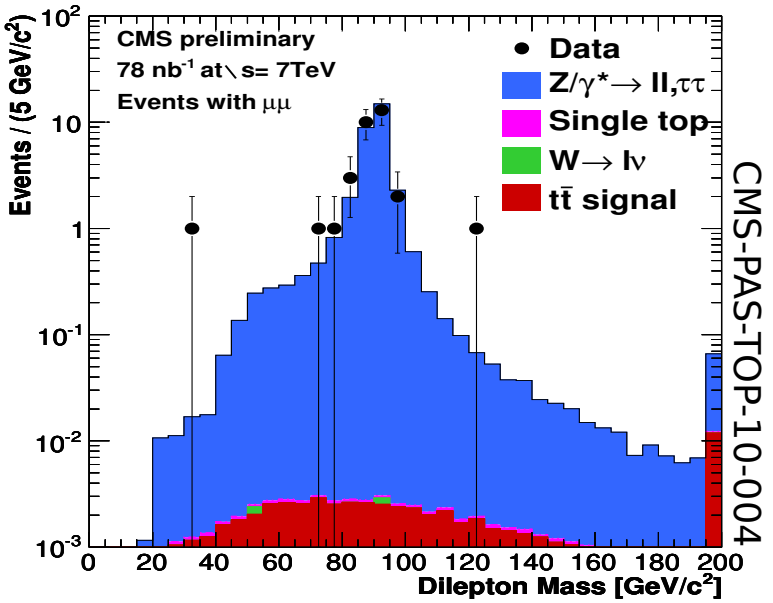


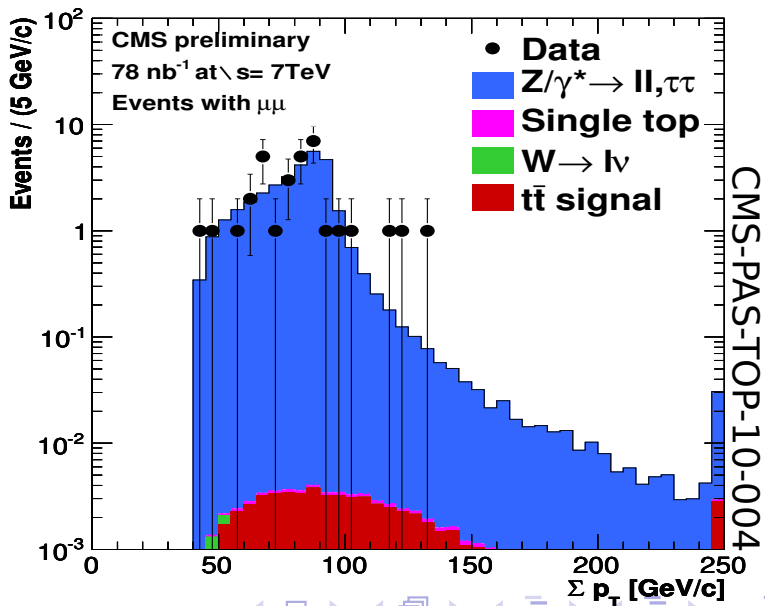
All following plots originate from dedicated analyses.
However, they are very similar to MUSiC plots.

Leptons and photons: Looking good



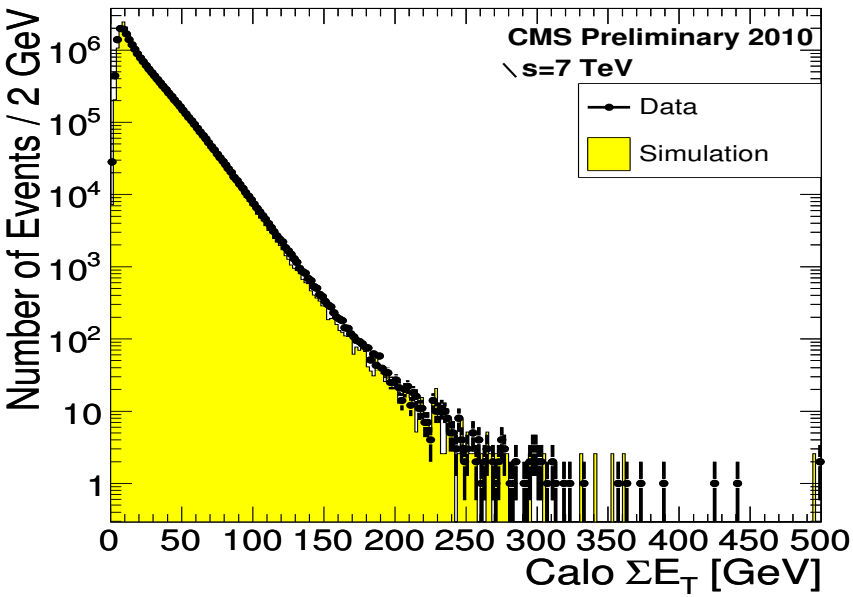


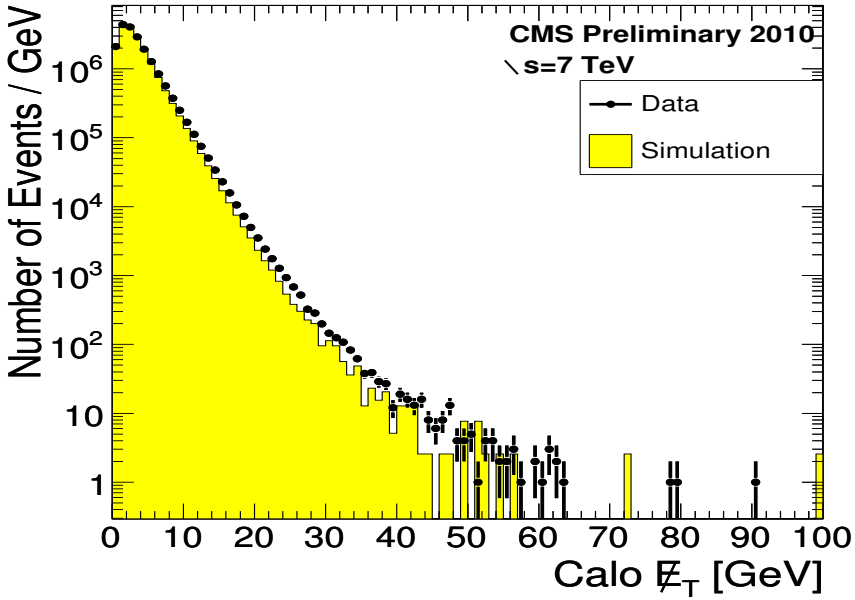


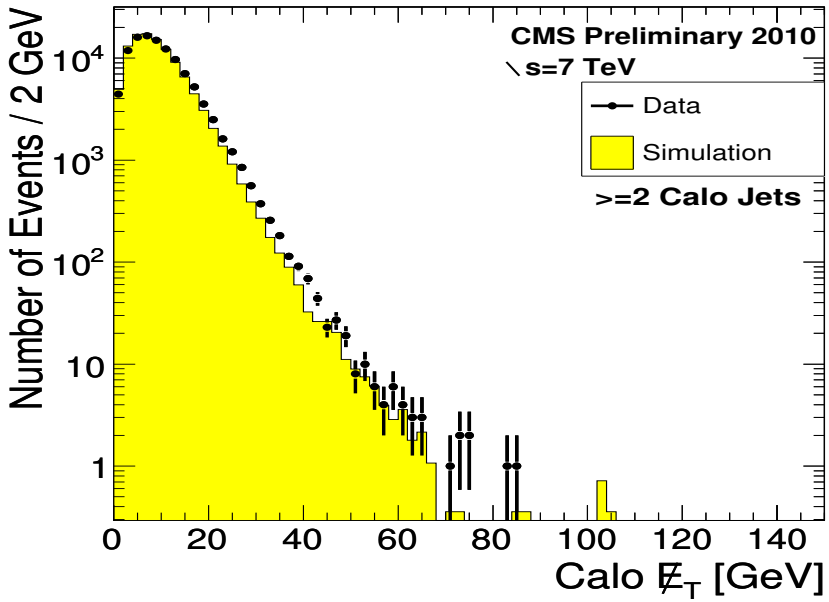


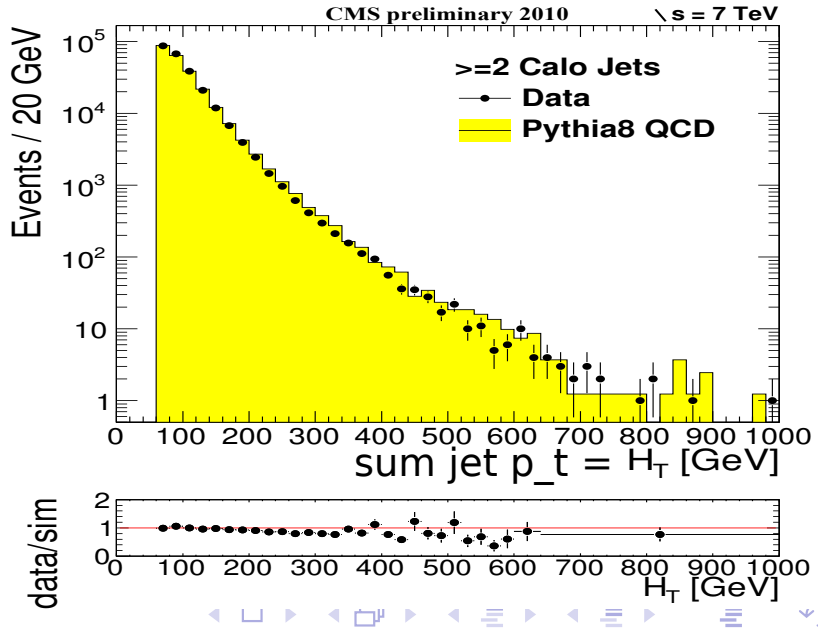
Jets and MET: Looking good

All Plots: CMS-PAS-JME-10-004









Conclusions

- Model independent approach that looks for deviations from the standard model
- Complementary to dedicated analyses
- Possible deviations still need detailed investigation
- Method works in MC examples
- First view at recent data looks promising
- Let's see where nature leads us