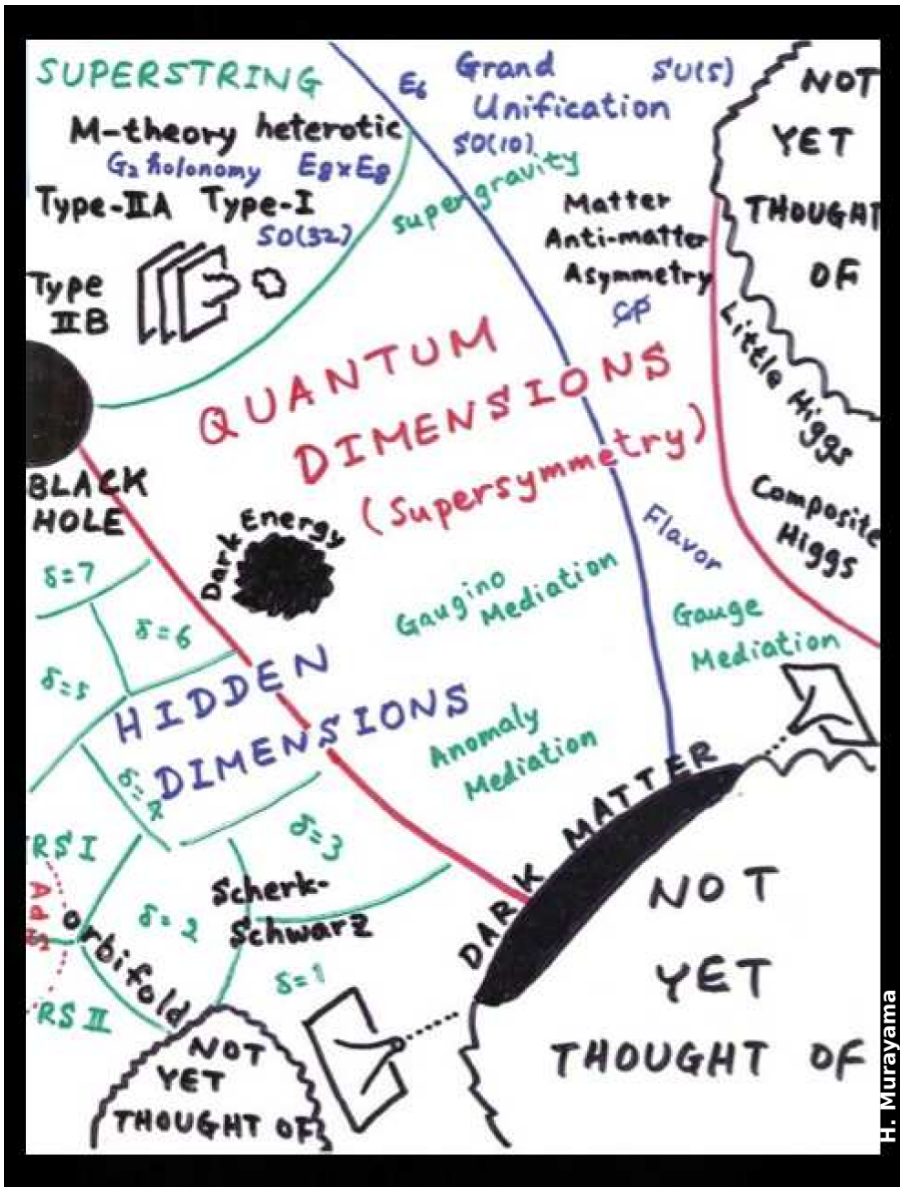
Model independent search for new physics (MUSiC) in CMS

Mark Olschewski, Michael Brodski, Thomas Hebbeker, Arnd Meyer, Paul Papacz, Holger Pieta

Terascale Alliance Workshop, Bonn 07.-09.12.2011

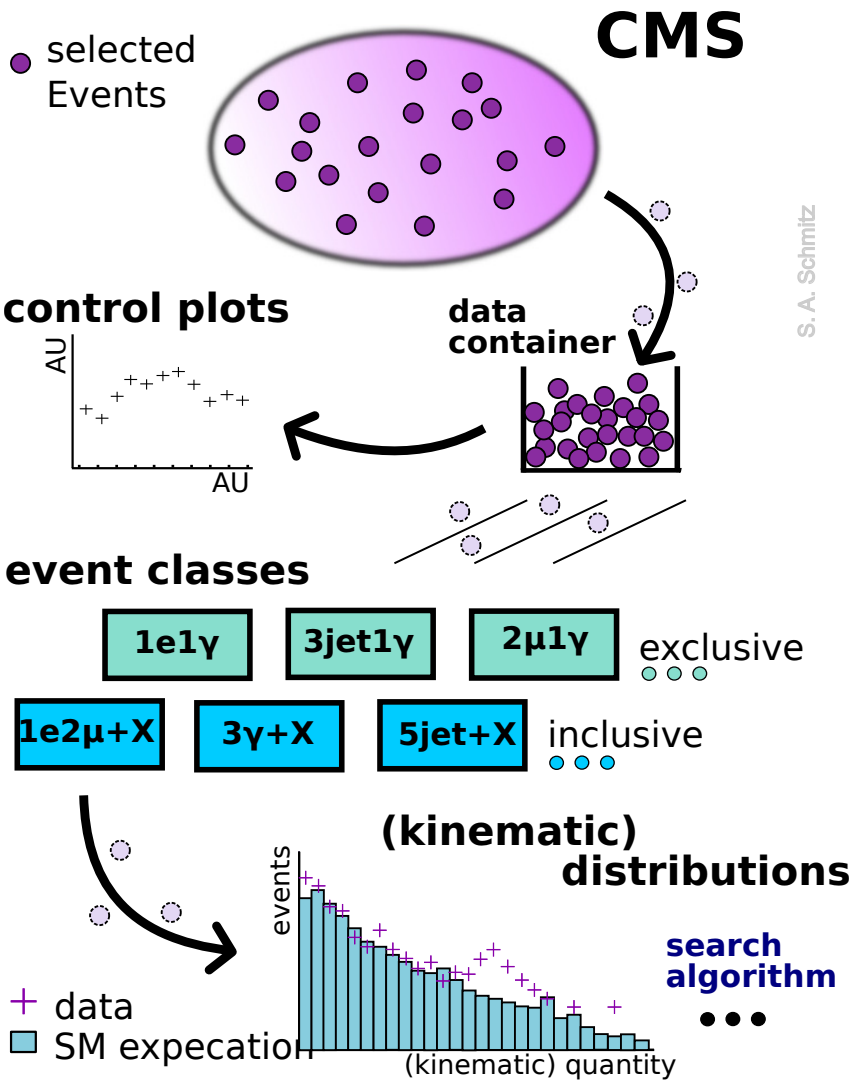
Bundesministerium für Bildung und Forschung

Motivation for a model unspecific search



- Huge variety of different, yet untested models.
- Each has its own motivation.
- A complementary approach to testing specific models is to scan the data for deviations from the SM prediction.

MUSiC workflow

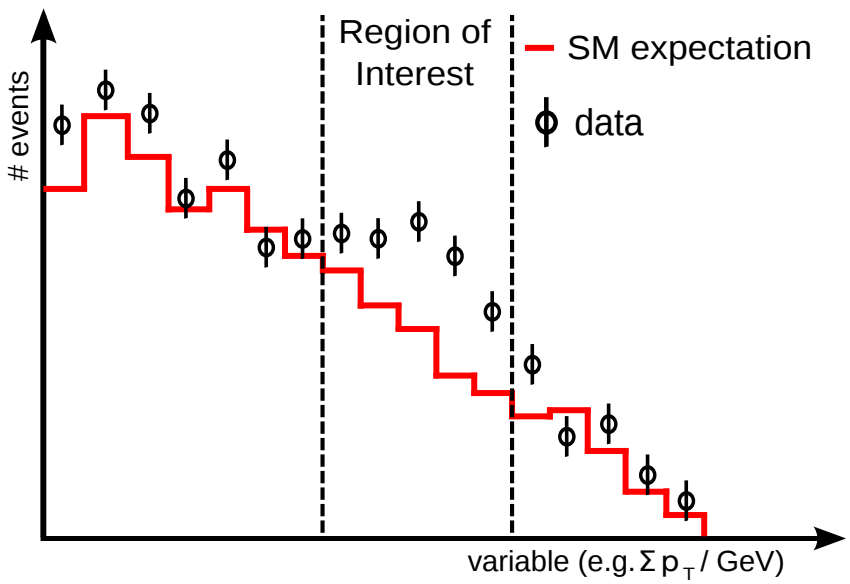


- Sort events into classes depending on their physics object content: μ, e, γ, jet, MET.
- Create kinematic distributions for each class: M_{inv}, $\sum p_T$, MET.
- Do the same for **Standard Model** Monte Carlo.
- Apply generic search algorithm to find most significant deviation.

P-value

$$p = \sum_{i=n_{\text{data}}}^{\infty} A \cdot \int_0^{\infty} dx \exp\left(\frac{-(b-x)^2}{2\sigma_b^2}\right) \cdot \frac{e^{-x}x^i}{i!}$$

in case of $n_{\text{data}} \geq b$.



- Number of data events n, Monte Carlo expectation b, uncertainty σ_b.
- p = probability to see n events or more, when Monte Carlo expectation is b.
- Region of Interest: Connected bin region with the smallest p-value.
- Also sensitive to deficits.

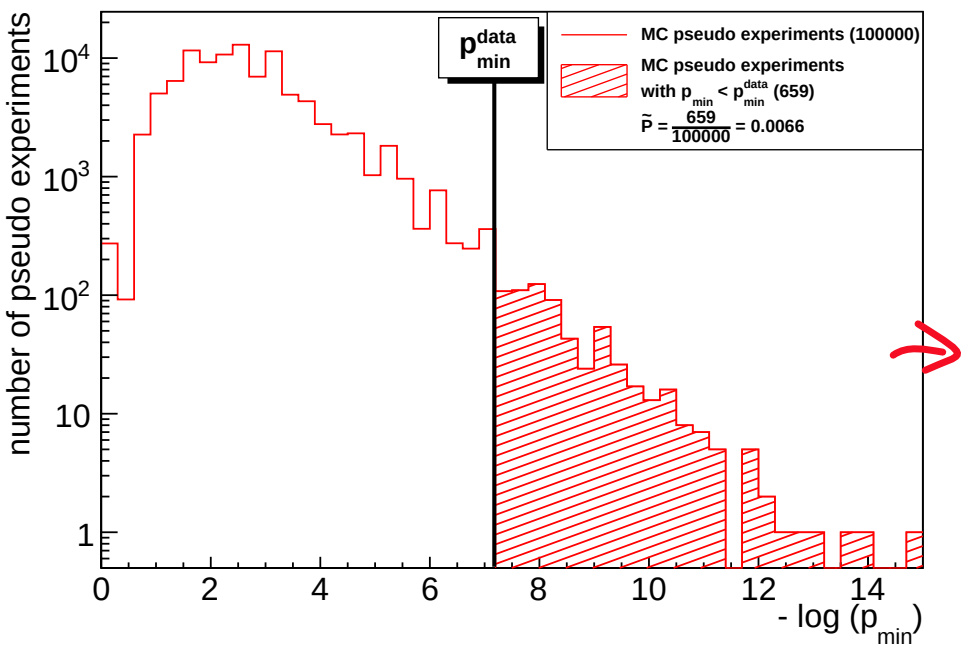
Look Elsewhere Effect

When calculating the significance of a deviation, the number of different bin regions as well as the correlation between them have to be taken into account.

CMS data

pseudo data from Monte Carlo

Calculate p-value for each region and look for the region with the minimal p-value p_{min}.



$$\tilde{p} = \frac{\#\text{ pseudo experiments with } p^{MC}_{min} \leq p^{data}_{min}}{\text{total } \#\text{ pseudo experiments}}$$

Selection criteria for 2011 analysis

Event selection

Require at least one **muon** with $p_T > 32\,\text{GeV}$ or one **electron** with $p_T > 82\,\text{GeV}$ (above single lepton trigger thresholds).

Object selection

Object	$p_T^{\text{min}}/\text{GeV}$	$\lvert\eta\rvert^{\text{max}}$
Muon	18	2.1
Electron	25	2.5
Photon	25	1.442
Jet	50	2.5
MET	30	

Additional object dependent quality criteria are applied.

Luminosity

$4.6\ \text{fb}^{-1}$

Monte Carlo samples

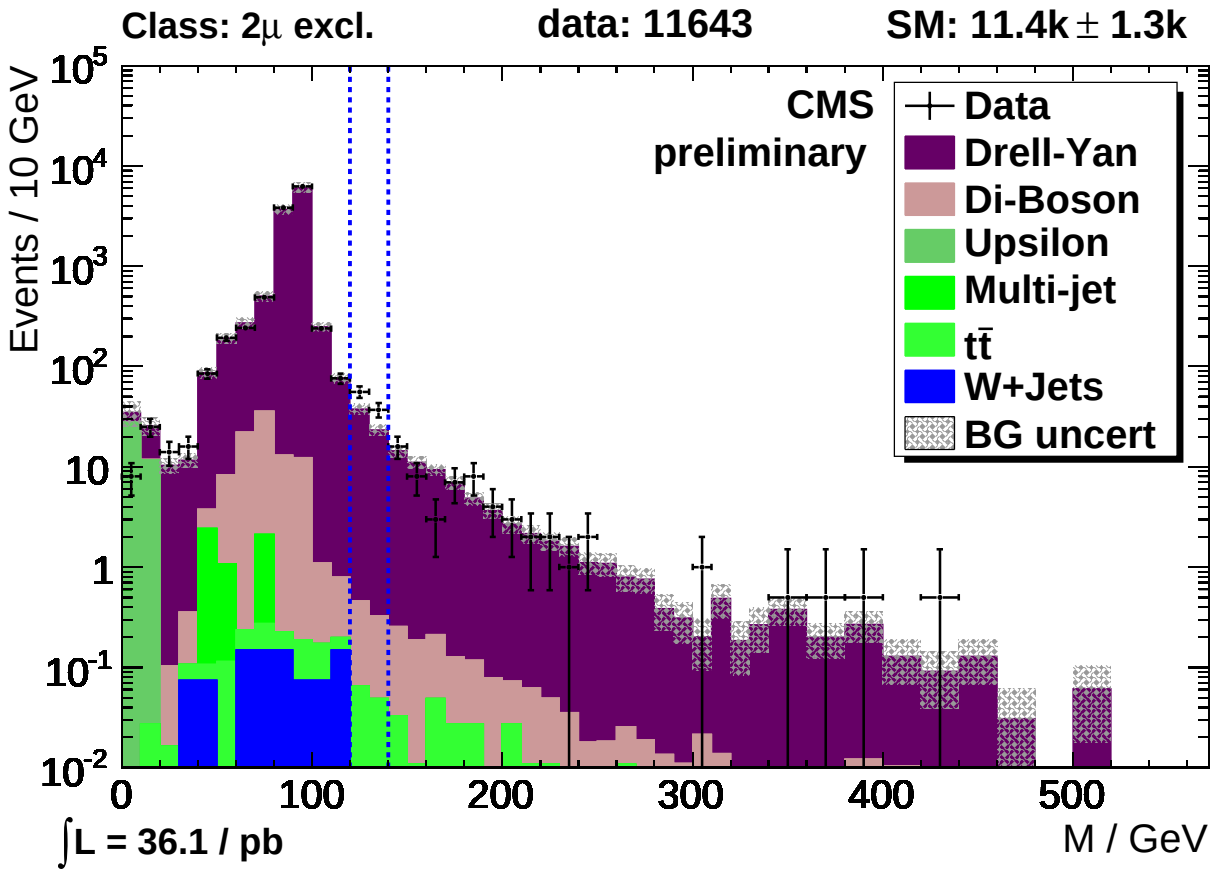
80 different samples from various generators.

- W boson
- Drell-Yan
- tt
- Multi-boson
- Υ
- Multi-jet
- Photon+jets

Uncertainties – Work in progress

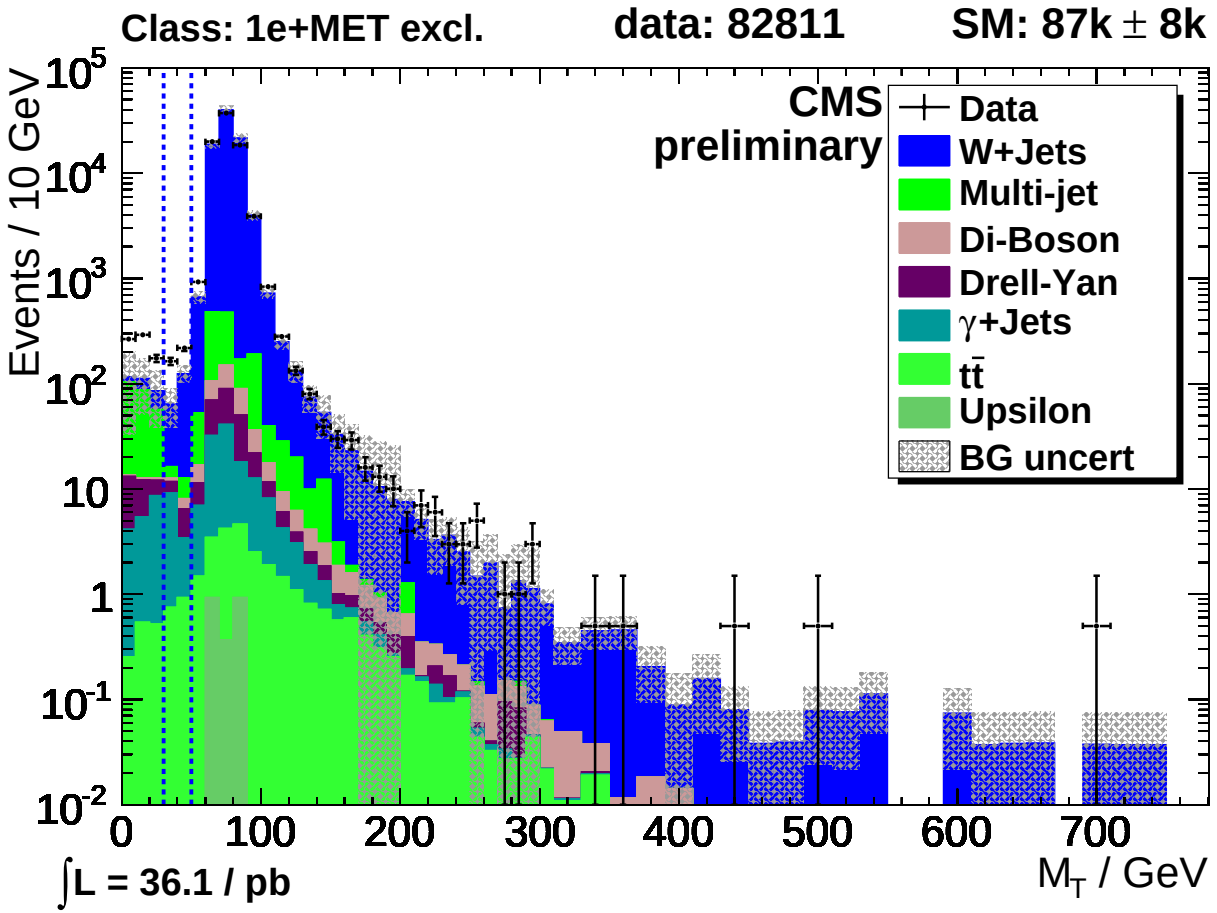
Contribution	**Value**	**Remarks**
MC statistics	various	sample dependent
Luminosity	4.5%	
PDF	various	PDF reweighting method
jet energy correction	3% to 5%	p_T dependent
reconstruction efficiencies	1% to 4%	object dependent
misreconstruction probabilities	30% to 100%	object dependent
W-boson cross sec.	5%	NNLO
DY cross sec.	5%	NNLO
tt cross sec.	10%	NNLL
Multi-boson cross sec.	10%	LO/NLO
Υ cross sec.	30%	measured
QCD-multijet cross sec.	50%	LO
photon+jets cross sec.	50%	LO

Drell-Yan dominated distribution



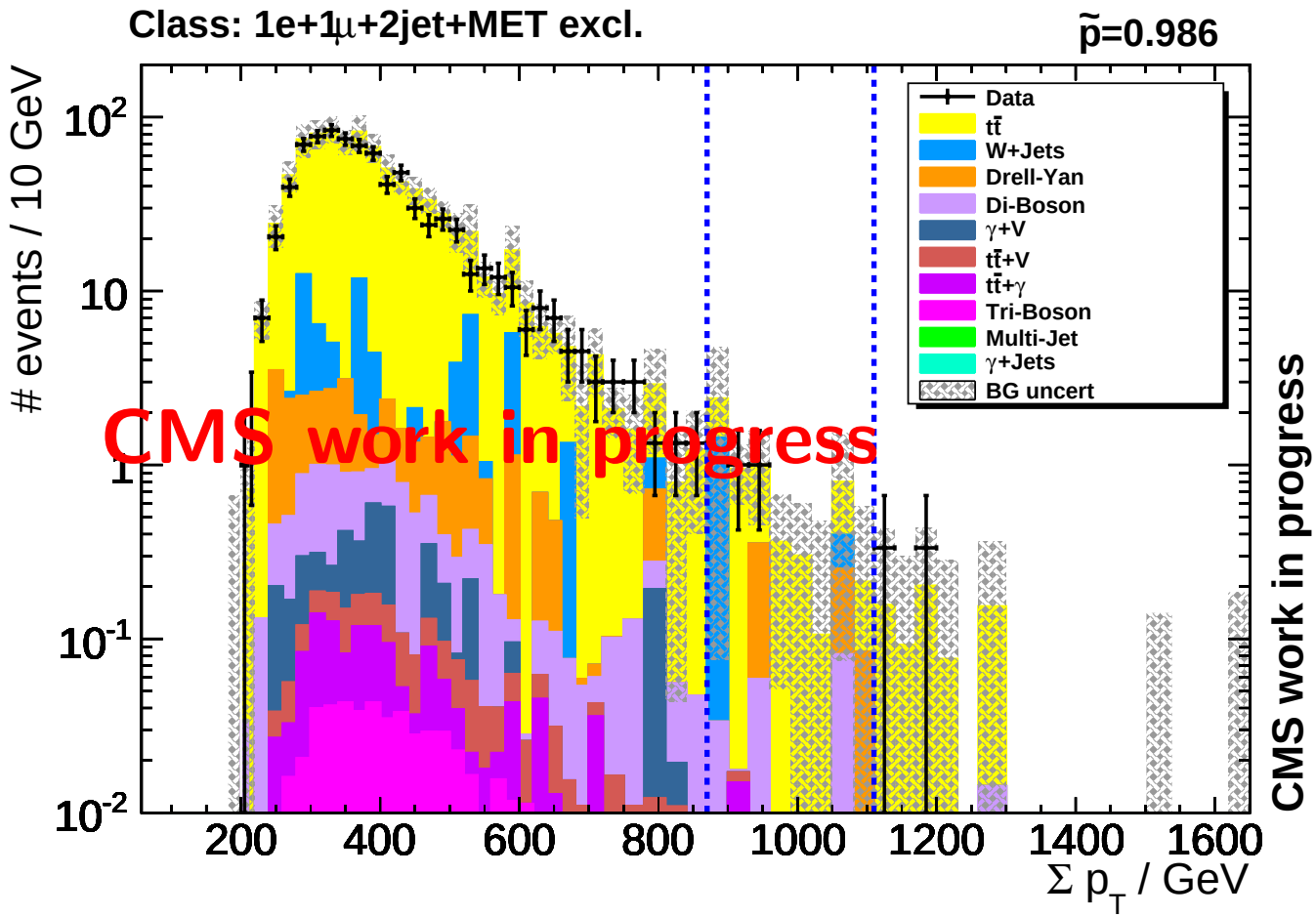
Overall good agreement in 2010 data.

W dominated distribution



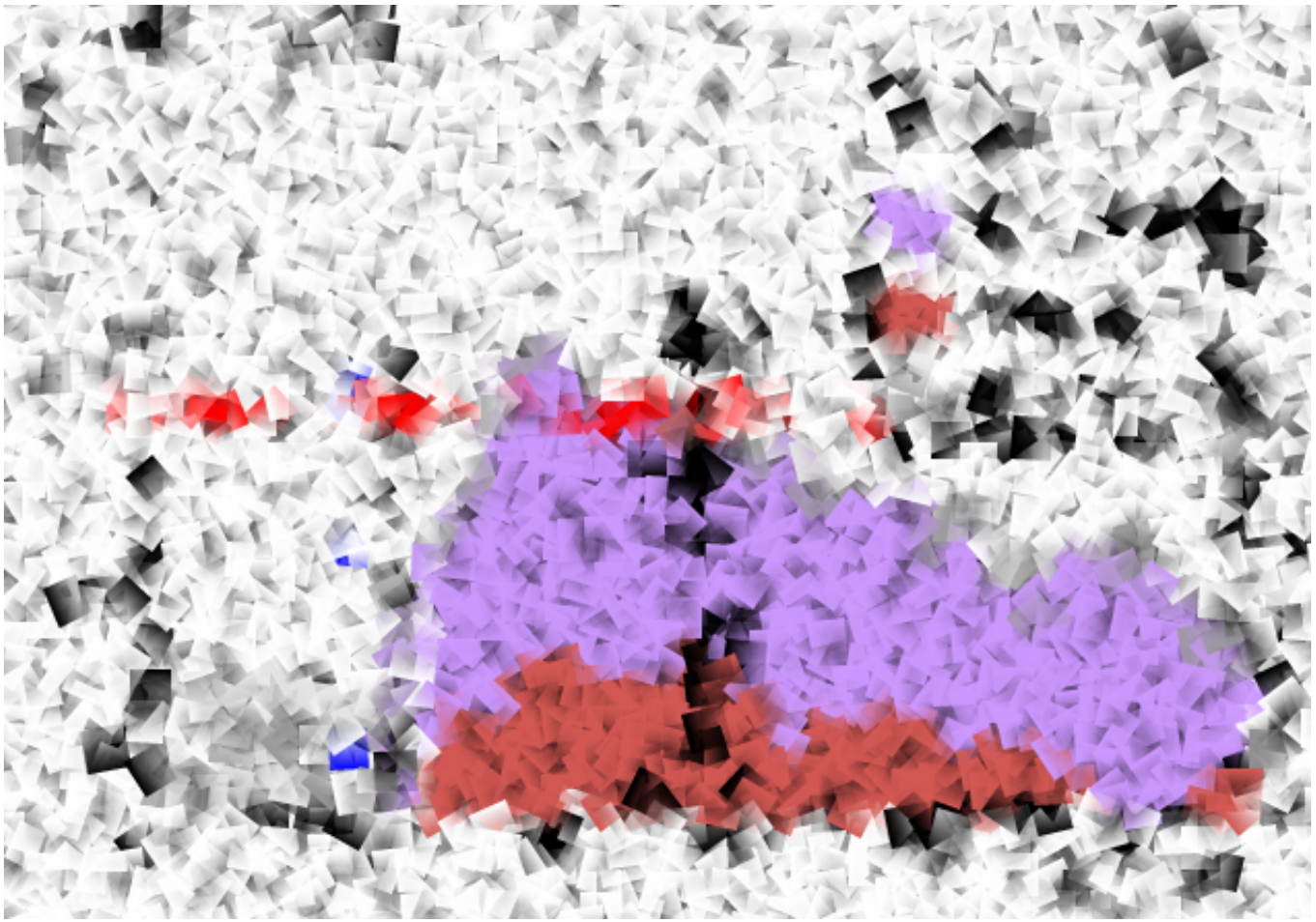
Overall good agreement in 2010 data. Some MC samples have small statistics.

$t\bar{t}$ dominated distribution

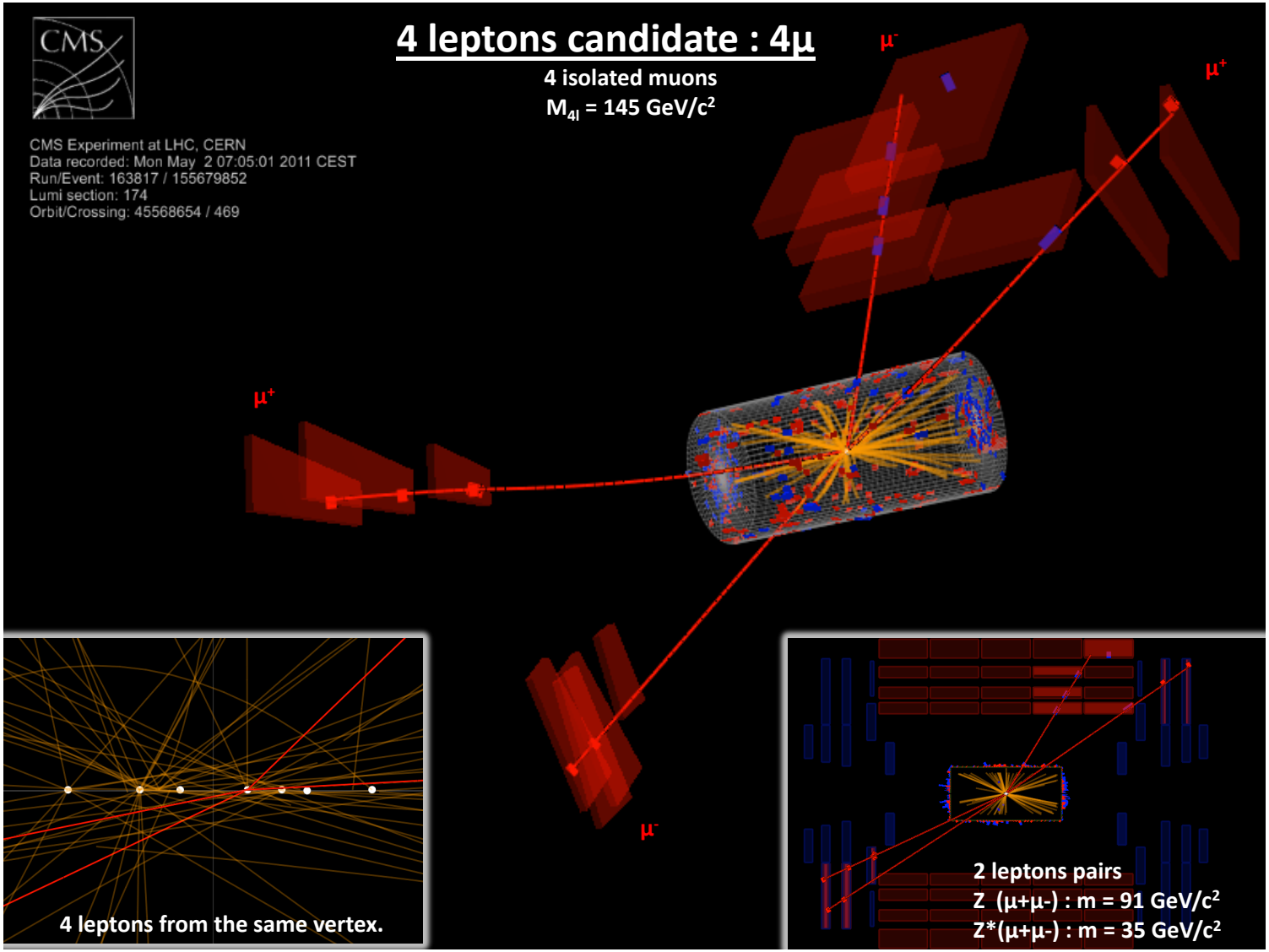


Deficits also detectable by MUSiC.

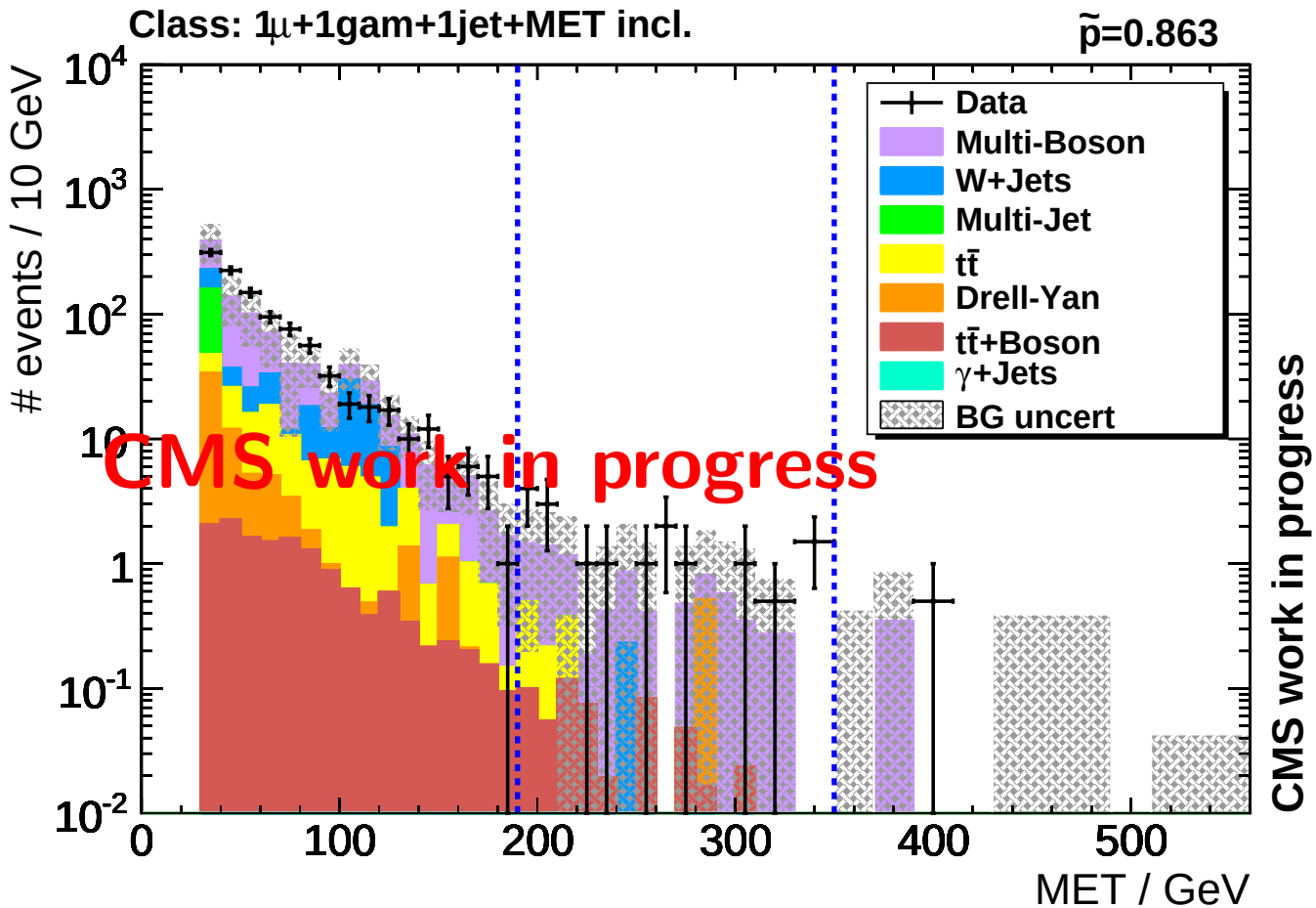
Quadruple-lepton distribution



4μ event display

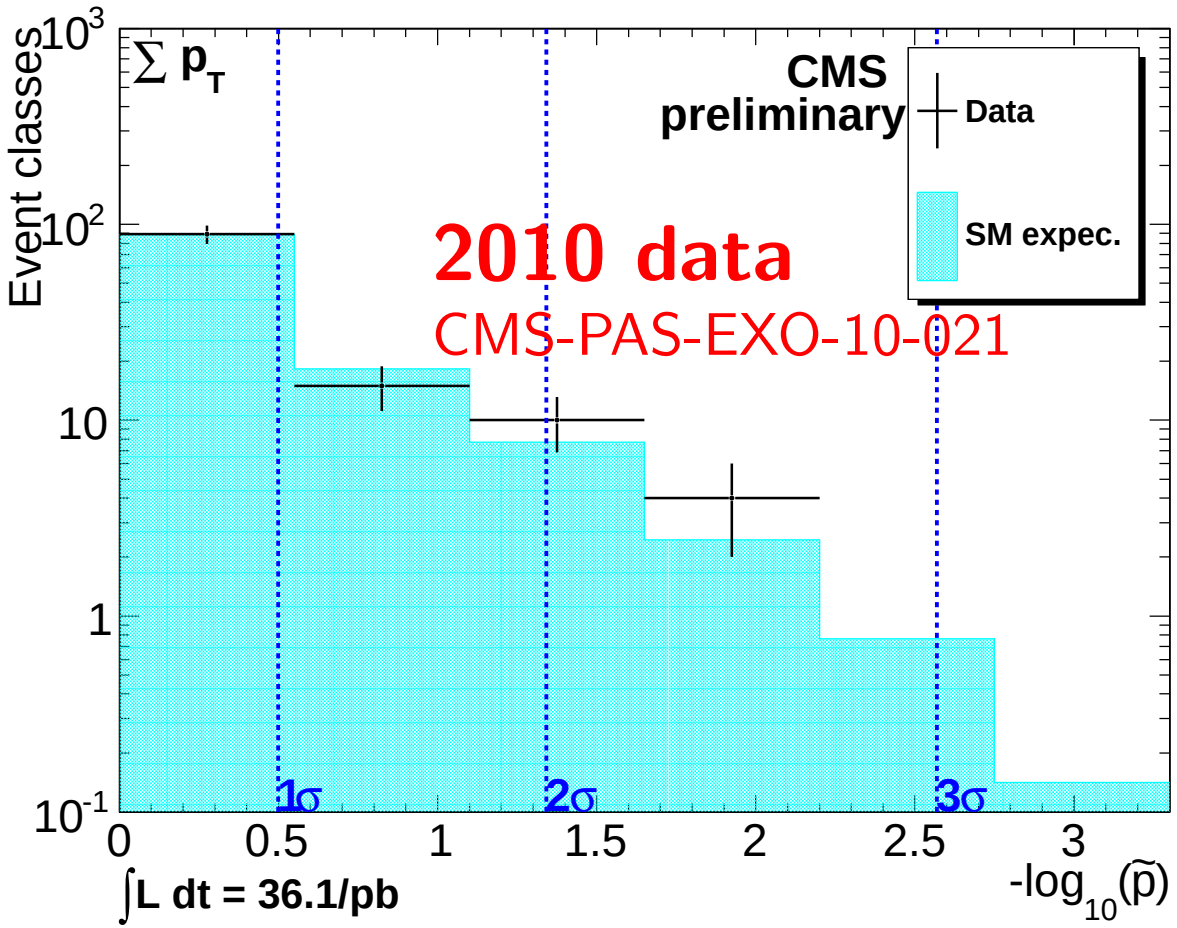


Multiple object distribution



Samples with small cross sections become increasingly important in 2011.

2010 $\tilde{p}$ distribution



Last year's data has shown an overall good agreement with expectation.

Summary

MUSiC – Model Unspecific Search in CMS

- Looks for deviations in a large number of classes.
- Complementary to dedicated analyses.
- Good agreement between data and Monte Carlo in 2010.

2011 data

- Analysis is underway.
- Challenging conditions due to higher (instantaneous) luminosity.
  - Handling for higher pile-up needed.
  - MC prediction for (very) rare SM processes becomes increasingly important in certain final states.