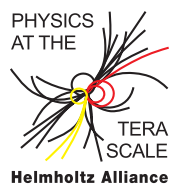
Model Unspecific Search in CMS (MUSIC) at 13 TeV

Recent Efforts with 2016 data

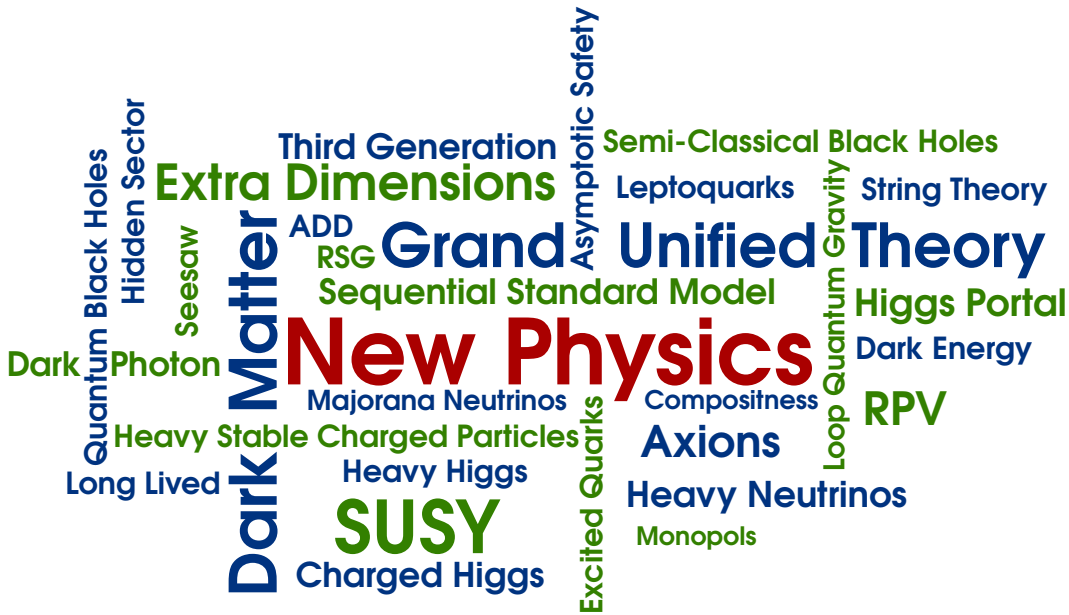
Deborah Duchardt, Simon Knutzen, Jonas Lieb, **Tobias Pook**, Jonas Roemer

Terascale Workshop - November 28th 2017



sponsored by
Bundesministerium
für Bildung
und Forschung

Outline



MUSiC Workflow

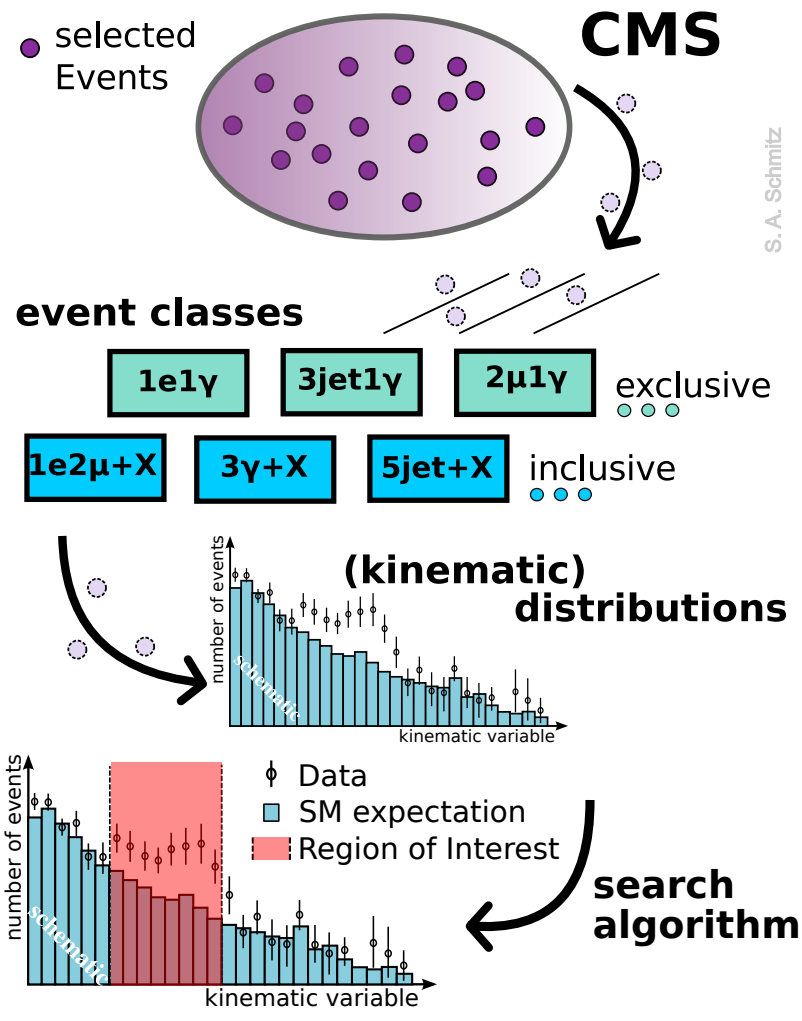
Goal: Scan hundreds of final states for deviations and find most significant regions

Classification:

- Select well reconstructed + isolated physics objects in event (e, μ, γ, b, jet, MET)
- Sort each event based on physics object content into event class

Scanning:

- Compare data to MC in each connected bin region $\rightarrow$ p-value
- Correct for look-elsewhere-effect using pseudo experiments $\rightarrow$ $\tilde{p} - value$



MUSiC Workflow: Scan

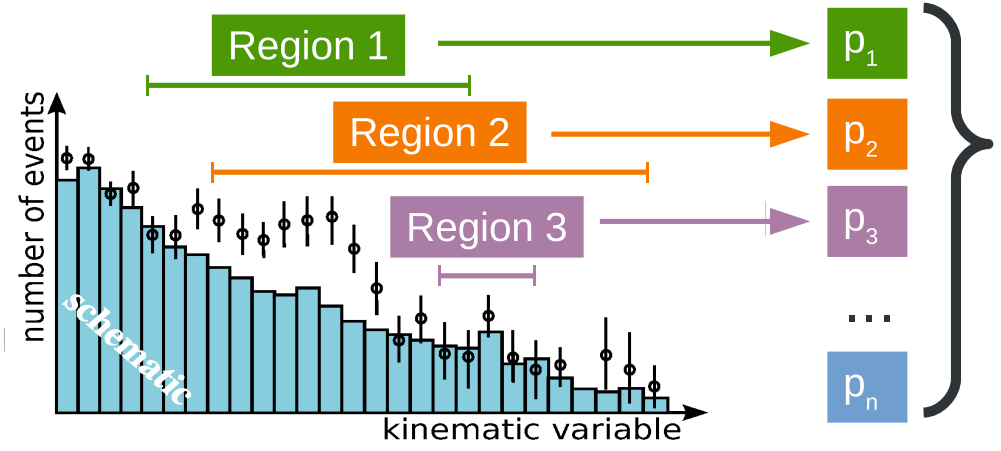
Definition of the MUSiC p-value

$$p_{\text{data}} = \begin{cases} \sum_{i=N_{\text{data}}}^{\infty} \underbrace{C \cdot \int_0^{\infty} \mathrm{d}\lambda}_{\text{normalisation}} \underbrace{\exp\left(-\frac{(\lambda - N_{\text{SM}})^2}{2\,\sigma_{\text{SM}}^2}\right)}_{\text{systematics}} \cdot \underbrace{\frac{\mathrm{e}^{-\lambda}\,\lambda^i}{i!}}_{\text{statistics}} & \text{if } N_{\text{data}} \geq N_{\text{SM}} \\ \sum_{i=0}^{N_{\text{data}}} C \cdot \int_0^{\infty} \mathrm{d}\lambda \exp\left(-\frac{(\lambda - N_{\text{SM}})^2}{2\,\sigma_{\text{SM}}^2}\right) \cdot \frac{\mathrm{e}^{-\lambda}\,\lambda^i}{i!} & \text{if } N_{\text{data}} < N_{\text{SM}} \end{cases}$$

N_{SM} : Sum of weighted MC events in region

N_{data} : Sum of data events in region

$$\sigma_{SM} = \sqrt{\sigma^2_{MC,\,stat} + \sigma^2_{MC,\,sys}}$$



Region with lowest p-value
=
Region of Interest

Toy experiments

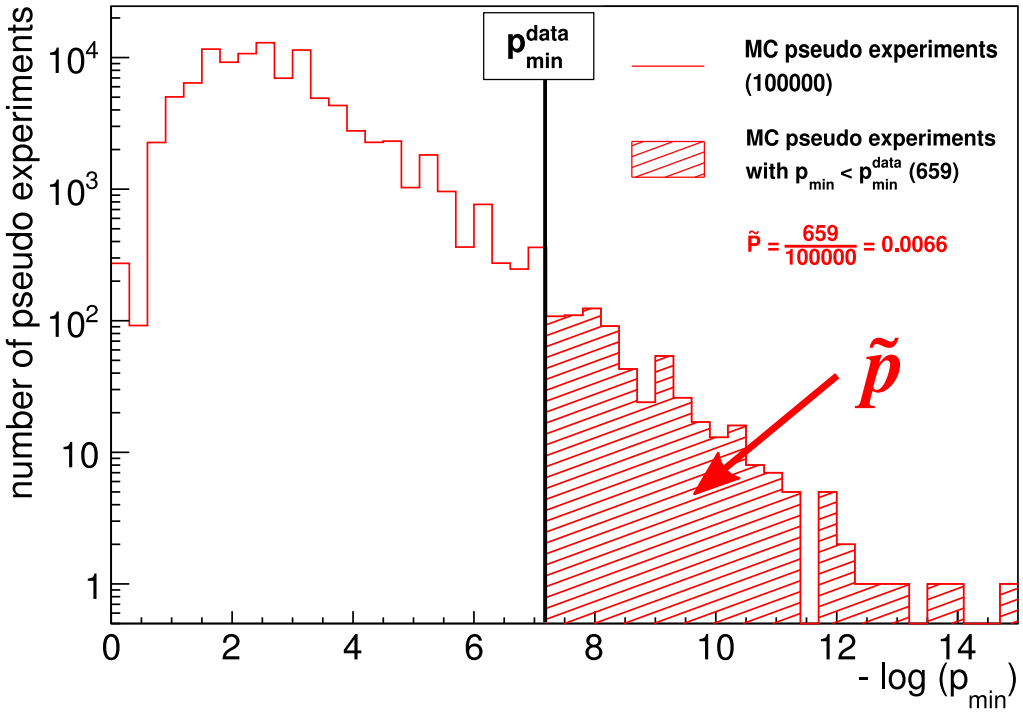
Look-Elsewhere effect correction

Look-Else-Where Effect Correction

Considering many regions in **many distributions** it becomes more probable to see a large deviation somewhere in the distribution only by chance due to statistical fluctuations.

Approximation with Toy Experiments

- Randomize MC expectation bin by bin, taking all known uncertainties into account (up to 10^5 times)
- Scan for most significant region
- Count pseudo exp. with smaller p-value than data



$\tilde{p}$ Definition

$$\tilde{p} = \frac{\text{number of pseudo experiments with } p_{pseudo} < p_{data}}{\text{number of pseudo experiment}}$$

Dataset & MC 2016

Dataset: Full 2016 Dataset with $35.9\mathrm{fb}^{-1}$

Triggerstreams:

		$p_{T,1}$ (GeV)	$p_{T,2}$ (GeV)
Single	μ	50	
	e	115	
	γ	175	
Double	μ	17	8
	e	33	33

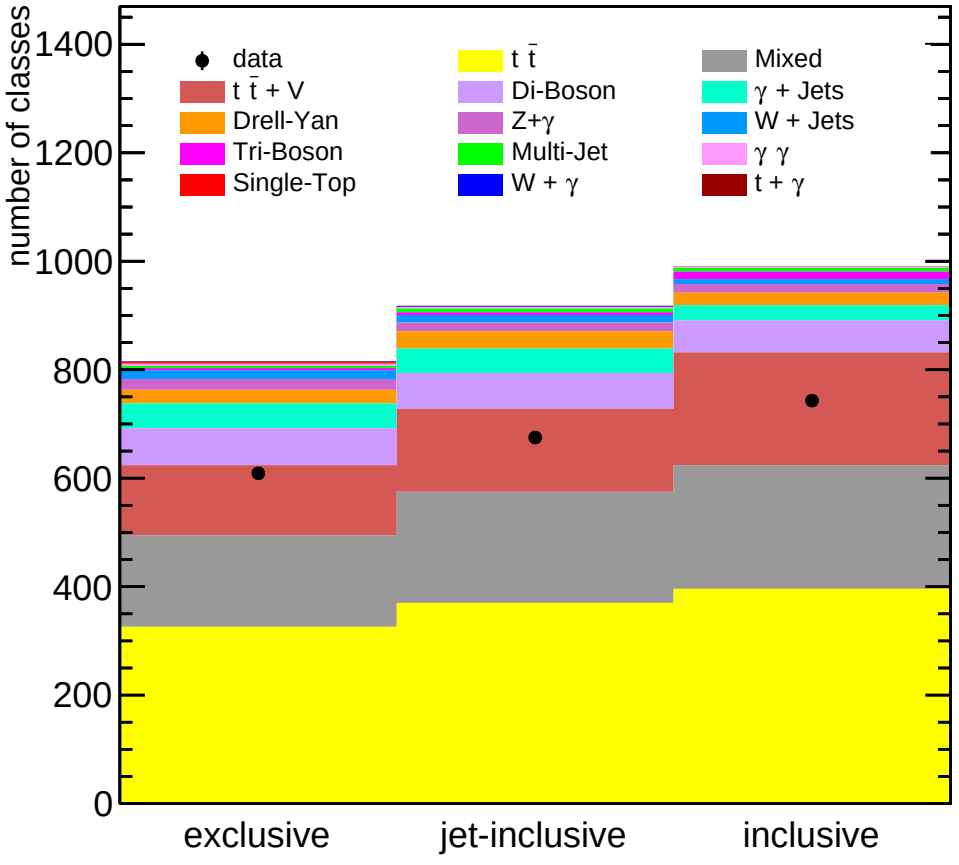
Monte Carlo: Full list on next slide

Monte Carlo Sets

Pythia8			
$W \to l\nu$	high mass tails		
Powheg			
$ZZ \to 2l2\nu$		$ZZ \to 4l$	
$WZ \to 3l\nu$		$WZ \to l\nu qq$	
$WW \to 4q$		$WW \to 2l2\nu$	bulk & mass binned
$Z \to 2l$	high mass binned	$t\bar{t}$	bulk & mass binned
single-top	tW-channel t & $\bar{t}$	single-top	t-channel t & $\bar{t}$
Sherpa			
$\gamma\gamma$+jets	mass binned		
Madgraph			
$W\gamma \to l\nu\gamma$		γ+Jets	HT binned
$ttZ + \text{jets}$		QCD	HT binned
AMC@NLO			
$W + \text{Jets}$	bulk & Pt binned	single-top	s-channel
$t\gamma$+ jets		$t\bar{t}\gamma$+ jets	
$t\bar{t}\gamma\gamma$		$t\bar{t}t\bar{t}$	
$t\bar{t}W\text{jets} \to qq$		$ZZ \to 4q$	
$ZZ \to 2l2q$		$ZZ \to 2q2\nu$	
$WZ \to 2l2q$		$WZ \to 1l3\nu$	
$WZ \to 1l1\nu2q$		WWW	
WWZ		ZZZ	
$WW\gamma$		WZγ	
$W\gamma\gamma$		WZγ	

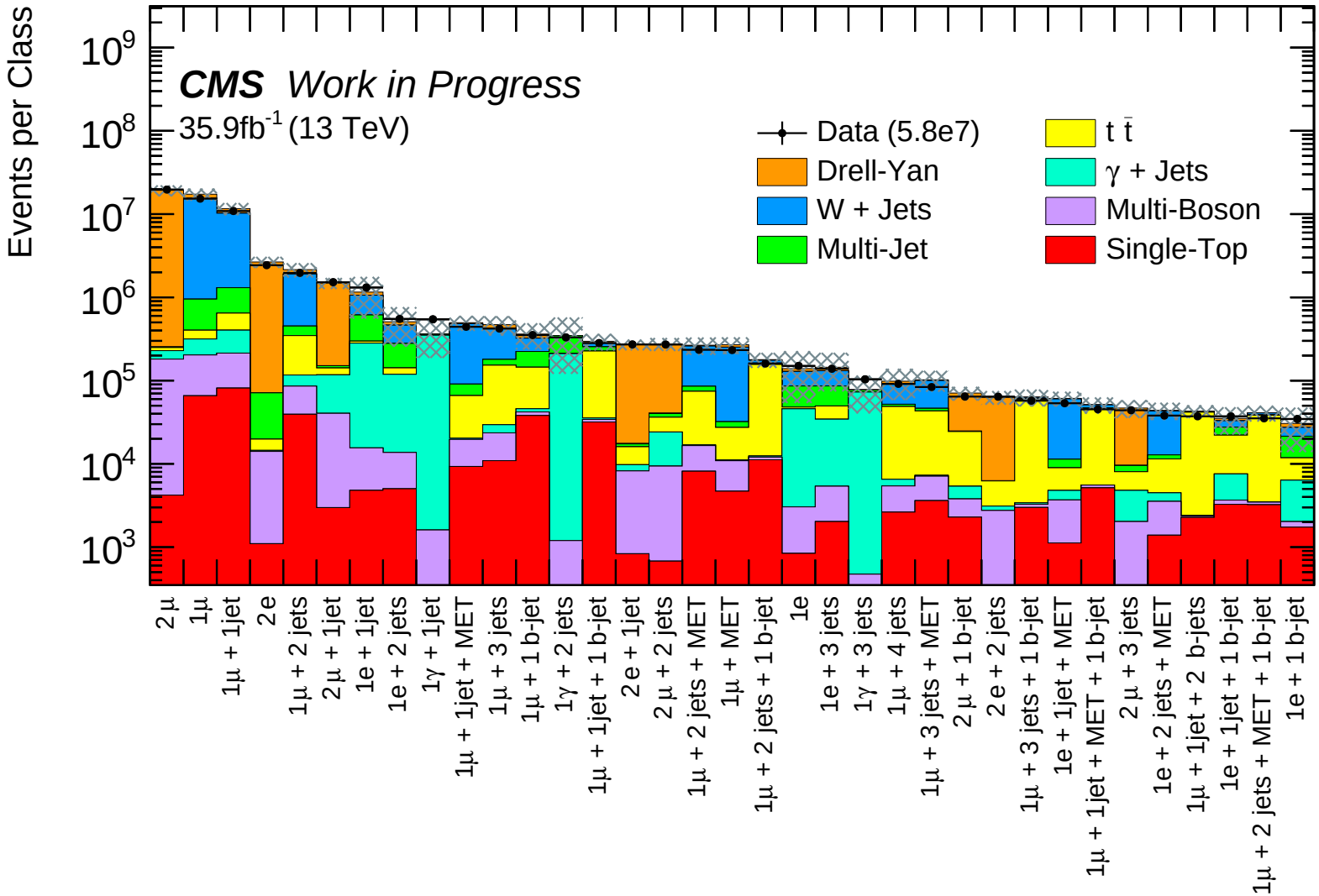
Number of classes

Do we really need to consider so many processes ?

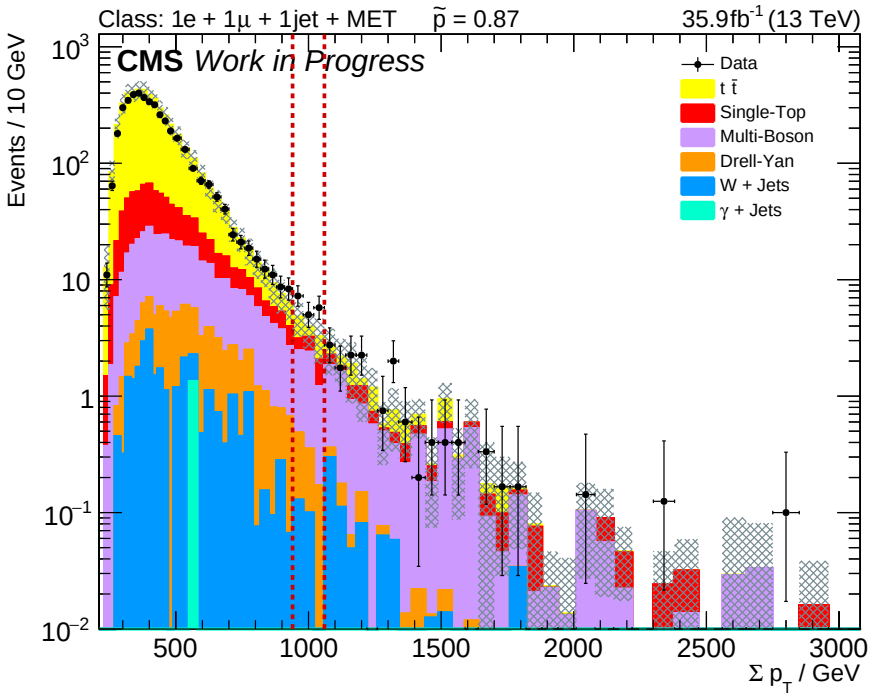


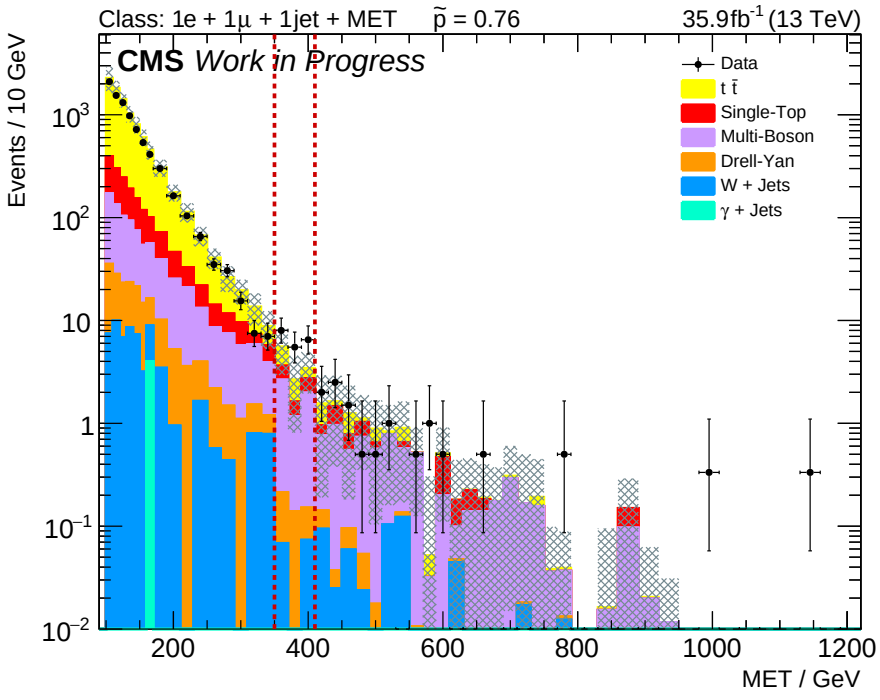
- Only MC classes with $N_{MC} > 0.01$
- Color indicates process with $N_{MC} > 50\%$ of total yield
- Number of classes with data smaller than, due to $N_{MC} < 1$ cases

Integral of distributions



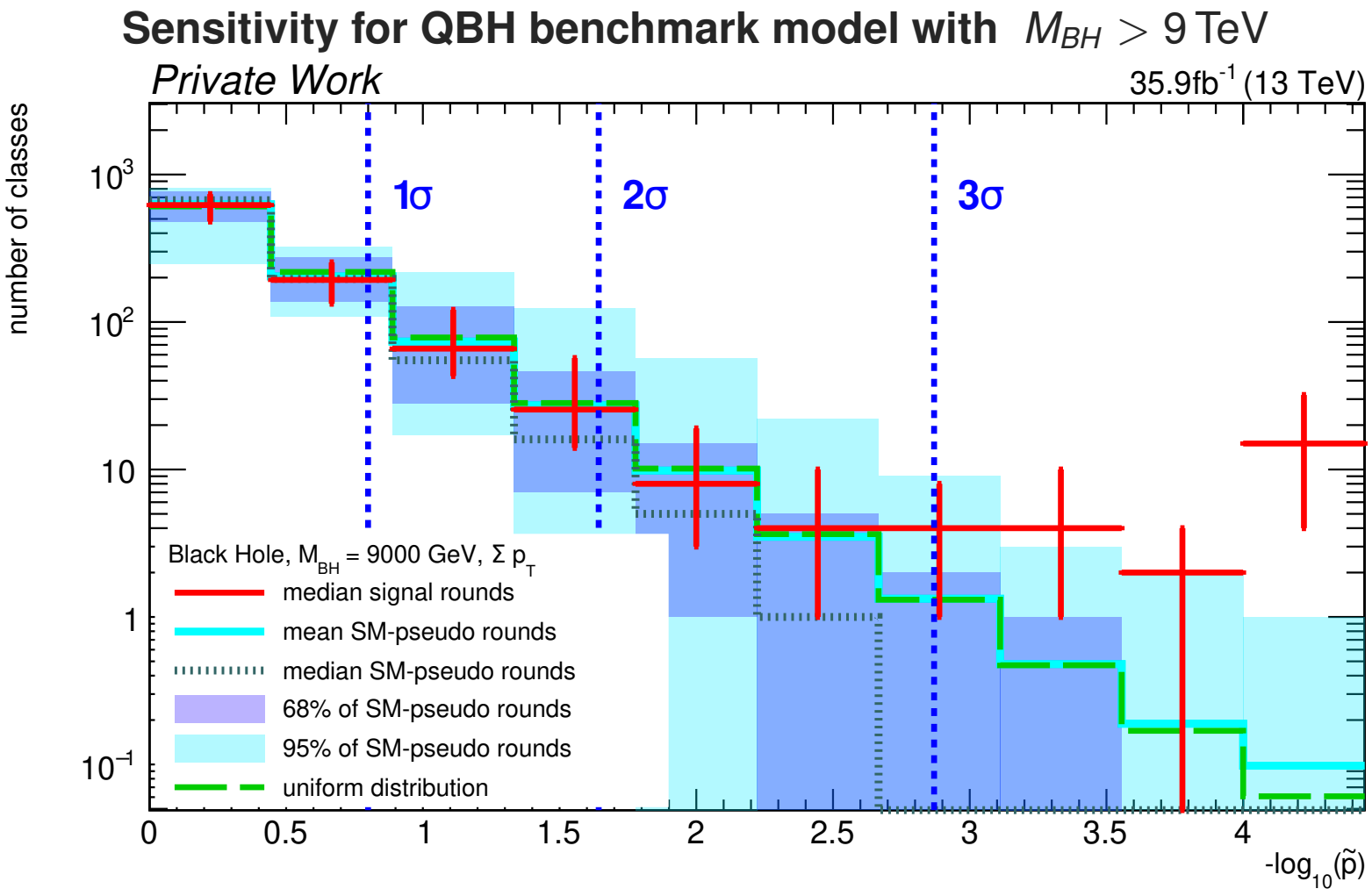
Examples: $1e\,1\mu\,1\text{jet} + \text{MET}$ incl.





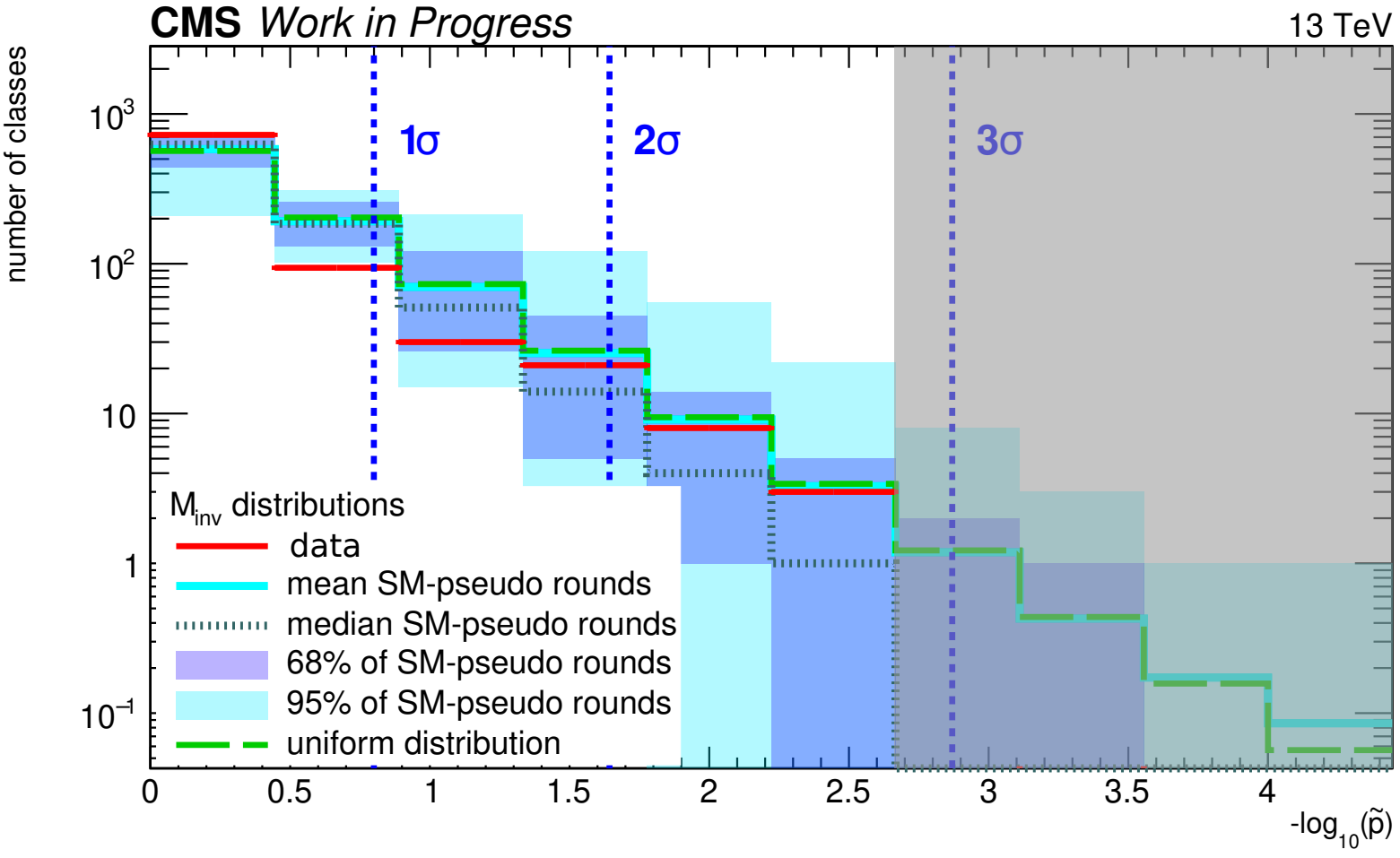
Example for two non-significant excesses found by scan algorithm

Scan Sensitivity

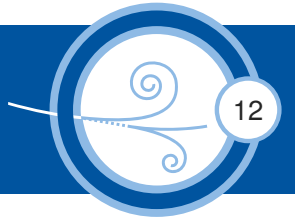


Sensitivity for several models studied in master thesis by J. Lieb

$\tilde{p}$-Distribution M_{inv}



Conclusion



- MUSiC efforts are ongoing to analyze 2016 dataset
- Set of considered objects extended with b-jets
- Increase in number of considered events using single photon data stream
- MUSiC analysis for 2017 dataset will start soon with additional manpower

Backup

CMS Germany BMBF FSP-CMS
DFG
PHYSICS AT THE TERA SCALE Helmholtz Alliance
sponsored by Bundesministerium für Bildung und Forschung