MUSiC: Model Unspecific Search in CMS

Lorenzo Vigilante on behalf of the CMS Collaboration

ICHEP 2020 Conference

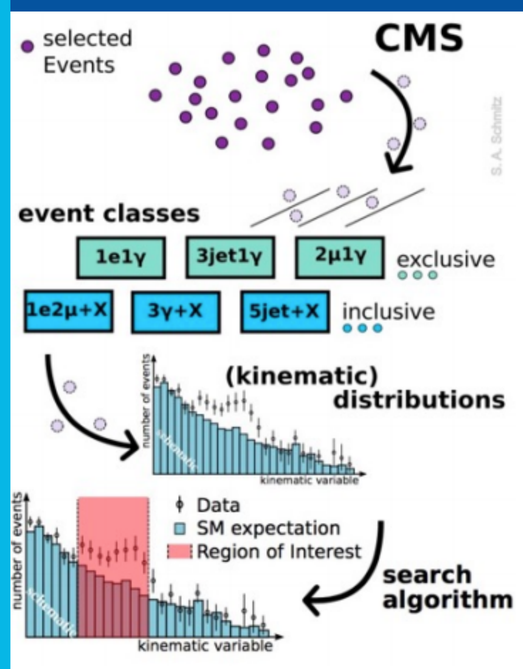
RWTH AACHEN UNIVERSITY
Physics Institute III A

SPONSORED BY THE
Federal Ministry of Education and Research

MOTIVATION

- Search for new physics phenomena.
- Search in multiple diverse final states.
- No specific input of any particular new physics model.
- Search for deviations from the Standard Model (SM) only hypothesis.
- Complementary approach to dedicated analyses.
- Prevent possible detectable signs of new physics from being overlooked: models with no dedicated analyses, unconsidered new physics phenomena.

MUSiC IN A NUTSHELL



Event selection

- Use CMS data and simulation of SM background from O(100) processes.
- Event and object selection based only on reconstruction quality.

Classification

- Sort simulation and data into event classes (final states).
- Consider three distributions per event class: $S_T = \sum |p_T|$, invariant mass (M) and p_T^{miss}.

Scan

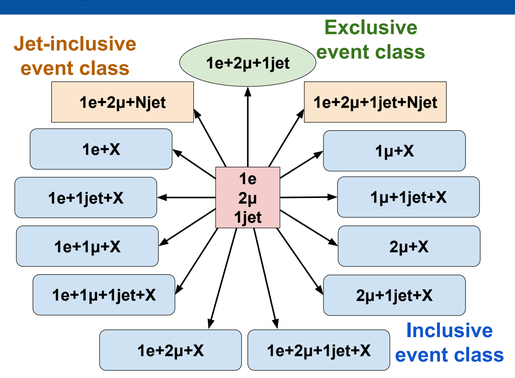
- Find region with most significant deviation.
- Determine p-value with look-elsewhere correction ($\tilde{p}$).

EVENT SELECTION

- Use proton collision data collected by CMS in 2016 (35.9 fb^{-1}) with $\sqrt{s} = 13\ TeV$.
- Physics objects considered : electrons (e), muons (μ), photons (γ), jets, b-jets and missing transverse momentum (p_T^{miss}).
- Scan final states containing at least one lepton (electron or muon).

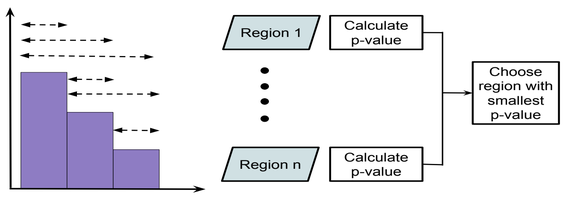
CLASSIFICATION

- Determine event class based on well reconstructed objects in an event.
- Sort event into exactly one “exclusive” event class.
- Sort event also into several “inclusive” and “jet-inclusive” event classes.
- Determine systematic uncertainties in an automated way.



SCAN

- Scan for deviation between data and SM simulation in different event classes.
- Scan for deviations in kinematic distributions (S_T, M, p_T^{miss}) : Region of Interest (RoI) scan along with total event yield.



- Consider all connected bin regions.
- Region of Interest (RoI) is defined as the region with the smallest p-value.

Significance calculation and correction of Look-Elsewhere effect

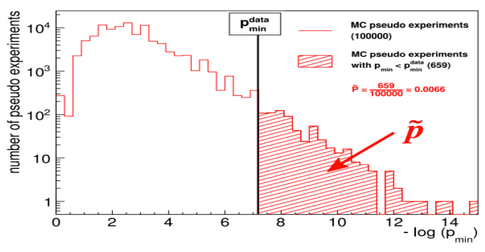
- Use Gaussian prior to model systematic uncertainties on the mean of the Poisson counting experiment.

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

N_{SM}: Sum of weighted MC event in region.

N_{data}: sum of data events in region.

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



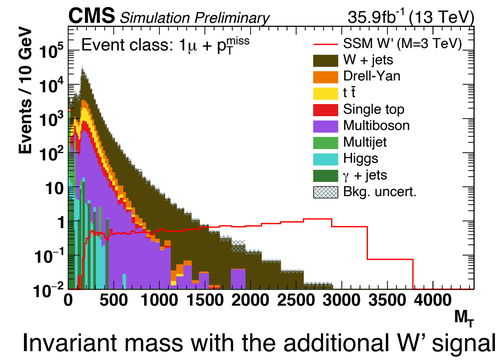
- Significance for RoI is overestimated due the large number of considered regions.
- Use pseudo-experiments to correct for look-elsewhere effect.

SENSITIVITY STUDIES

Two different approaches used to demonstrate capability of the analysis to identify deviations in the comparison between data and simulation.

Inject a BSM signal

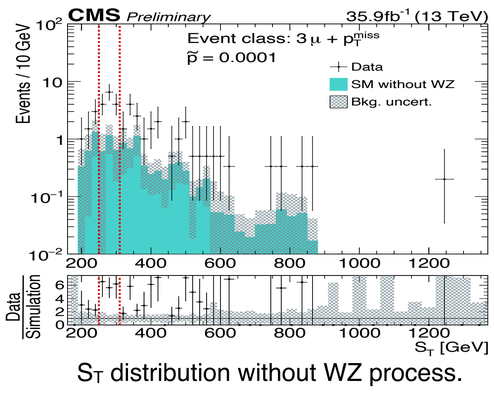
- W’ simulated samples are added on top of SM processes.
- Different final states (e.g. $1\mu + p_T^{miss}$) with significant deviations beyond the expectation are found.



Invariant mass with the additional W’ signal.

Remove a SM signal

- WZ process is removed from SM background.
- Scan against recorded CMS data.
- Several final states (e.g. $3\mu + p_T^{miss}$) show large and significant deviations.



S_T distribution without WZ process.

RESULTS

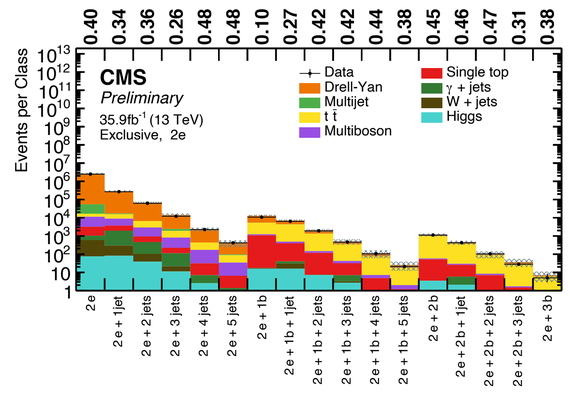
Total yield scans and object group representation

- Evaluate p-value for each event class based on total yield.
- Event classes grouped by their object content e.g. 2e object group consists of all classes with exactly two electrons and any number of (b-)jets.

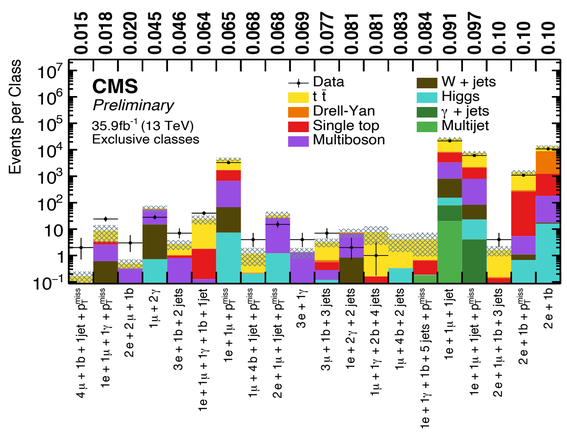
Number of event classes	
Exclusive	498
Inclusive	571
Jet-inclusive	530

- No particular event class being found to have a very significant deviation.

2e object group representation.

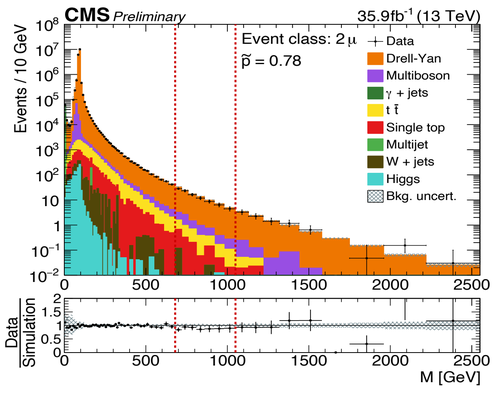


Event classes with most significant deviation based on the total yield scan.

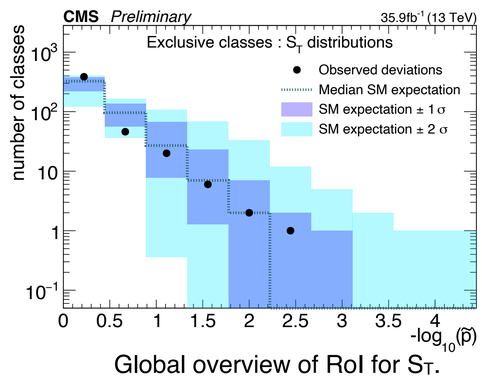


Results of RoI scan

- Distributions of S_T , M, and missing p_T are scanned to search for deviations.
- Due the large number of different event classes a global overview of the scan is required.
- $\tilde{p}$-values calculated for each kinematic distribution are summarized in a single histogram and compared with SM only expectation obtained from pseudo experiments.
- No event classes with an outstanding deviation have been found.
- Largest deviations are along expectation within uncertainties from SM only hypothesis.



Invariant mass distribution for 2μ class.



Global overview of RoI for S_T.

REFERENCES

- CMS-PAS-EXO-19-008

vigilante@physik.rwth-aachen.de