Studies with Matrix Element Method

Suzan Basegmez, Giacomo Bruno

Université catholique de Louvain - CP3

May 28, 2013

Outline

MEM and Normalization

Likelihood function and normalization of Matrix Element Method outputs

Validation

Basic validation and application studies to a search for heavy di-muon resonances.

MEM and Normalization of Outputs

Analysis Level Selections

In an experimental set-up one usually measures the physics objects with **detection efficiencies** after some online/offline **selections**.

X : full phase space of an experimental event $\rightarrow$ X' : reduced phase space of the detected object.
Y: full phase space at parton level $\rightarrow$ Y': sub-phase space

It can be proven that likelihood function of underlying model parameters M, for one experimentally observed events **x** in the reduced phase space X' of selectable events (analysis-driven) is given by;

Likelihood Function :

$$L(\alpha) = pdf(x|\alpha\, AND\, x\, in\, X^{'}) = \frac{1}{\epsilon^{'}\sigma^{'}} \int_{Y^{'}} \frac{d\sigma_{\alpha}}{dy} w(x|y) dy$$

Likelihood Function with MEM

$$L(\alpha) = pdf(x|\alpha\, AND\, x\, in\, X^{'}) = \frac{1}{\epsilon^{'}\sigma^{'}} \int_{Y'} \frac{d\sigma_\alpha}{dy} w(x|y)dy$$

Where;

- $\frac{d\sigma_\alpha}{dy}$: is parton level cross-section for a given model
- $w(x|y)$: "transfer functions" (Note to remark on next slide)
- $Y^{'}$: reduced parton level phase-space in which $w(x|y) = 0$ for any x in X' and y outside Y'
- $\sigma^{'}$: reduced cross-section at the parton level of Y'
- $\epsilon^{'}$: efficiency of events in Y' to end up in reduced phase space X'
- $\sigma^{'}$: the reduced cross-section for parton level events in Y'

Remarks on Normalisation of Outputs

Programs like MadWeight returns only the so-called "weight"

$$W = \int_{Y'} \frac{d\sigma_\alpha}{dy} w(x|y) dy$$

- Where $w(x|y)$ is what the user has defined, i.e. often unit normalized function of x and y.
- Factorization of single object is assumed, so y and x will indicate in the following, for simplicity, a single object (parton and reco-level)

$w(x|y)$ must satisfy $\int_X w(x|y)dx = 1$ for any y in Y

- Where X is full experimental event space, which includes non-reconstructed objects.
- In other words if $u(x|y)$ is unit-normalized function of x for given y, then one should use $w(x|y) = \varepsilon(y)u(x|y)$ where $\varepsilon(y)$ (usually < 1 for leptons) is the efficiency of the reconstructed object

Likelihood Function with MEM

$$L(M) = \frac{1}{\epsilon' \sigma'} \int_{Y'} \frac{d\sigma_\alpha}{dy} \varepsilon(y) u(x|y) dy$$

Overall Normalization :

- Fundamental for parameter estimation:

$$L(\{x_i\}|\mu, M) = \prod_i \frac{\mu}{\mu\sigma_S\varepsilon_S + \sigma_B\varepsilon_B} W_S(x_i, M) + \frac{1}{\mu\sigma_S\varepsilon_S + \sigma_B\varepsilon_B} W_B(x_i|B)$$

- For simplified hypothesis testing there is no problem because one usually ends up with a monotonous function of the likelihood ratio. Still it has an impact for significance computation using PLR test statistics.

$$K_D = \frac{W_S}{W_S + W_B} = \frac{1}{1 + \frac{\varepsilon_B \sigma_B L_B}{\varepsilon_S \sigma_S L_S}}$$

Likelihood Function with MEM

$$L(M) = \frac{1}{\epsilon'\sigma'} \int_{Y'} \frac{d\sigma_\alpha}{dy} \varepsilon(y) u(x|y) dy$$

Normalization of transfer function :

- Only matters if $\varepsilon(y)$ varies over an interval centered around the experimentally measured value x and as large as few times the resolution. If this is not the case, it goes out of integration and:
  - In Hypothesis testing it cancels out event-by-event in the likelihood ratio.
  - In parameter estimation ends up a overall multiplicative factor of a likelihood function.
- Special relevant case is $H \to ZZ \to 4l$, where it would require attention : Higgs properties with the same channel, if one pulls down the lepton P_T very low

Hypothesis Testing and Significance

Likelihood function can be used to compute the significance of a discovery or excluded cross-sections.

We use the asymptotic formulae with likelihood ratio test statistics to reduce computation time (G.Cowan, K.Cranmer, E. Gross,O.Vitells, arxiv:1007.1727)

Profile Likelihood Ratio for μ signal strength : $\lambda(\mu) = \frac{L(\mu,\hat{\hat{\theta}})}{L(\hat{\mu},\hat{\theta})}$

Test statistics for discovery of a signal :

$$q_0 = \begin{cases} -2ln\lambda(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases}$$

$\lambda(0)$ is the profile likelihood ratio for $\mu = 0$

Significance of the test is $Z = \sqrt{q_0}$

Validation

Validation of the method with

A search for $Z^{'}$ high mass resonance

Overview of the Validation

- MadGraph generator used for exclusive signal and background event generation.
  - Signal Events : $Z^{'} \rightarrow \mu^{-}\mu^{+}$ (Seq. Standard Model)
  - Background Events : $DY \rightarrow \mu^{-}\mu^{+}$

  Generator-level inv-mass range $(Y^{'})$: $600\,GeV \leq M \leq 1400\,GeV$
  Selection-level inv-mass range $(X^{'})$: $800\,GeV \leq M \leq 1200\,GeV$

- Estimation of parameters: resonance mass, signal fraction.
- Significance of signal hypothesis
- Comparison with CMS $Z^{'}$ shape analysis was performed (with RooFit/RooStats)

Generation, Selection, Normalization

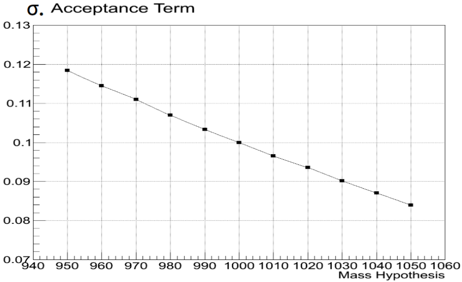
Signal mass $= 1\,TeV/c^2$
Signal width $= 3\%$

5% detector resolution effect via Gaussian smearing on muon energy

Analysis Selections :

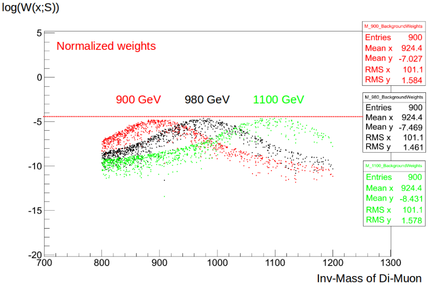
- $800\,GeV \leq M \leq 1200\,GeV$
- $P_t^{\mu} > 45\,GeV/c$
- $|\eta| < 2.4$

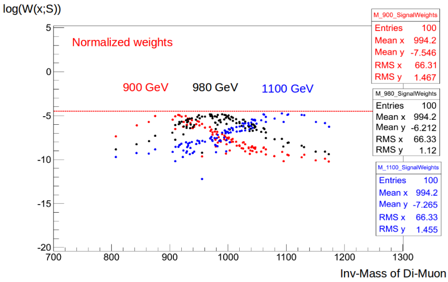
$\sigma_s.\varepsilon_s$
for different mass hypotheses.



Likelihood Function Validation

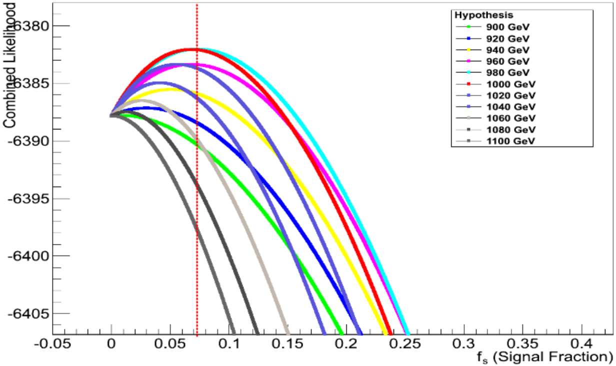
log-L(M) for DY and Signal events under signal hypothesis of 900, 980, 1100 GeV





Parameter Estimation

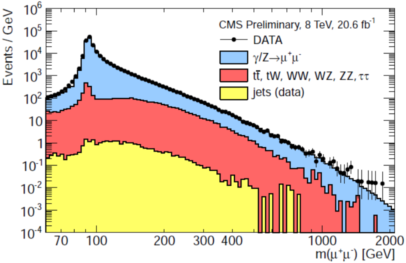
log-likelihood for Signal + Background events for one toy experiment

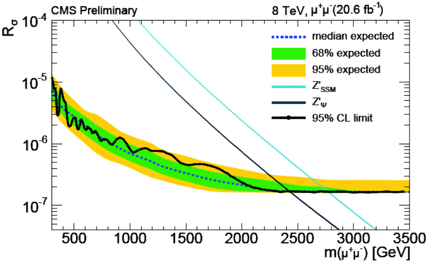


Signal fraction (f_s) and the resonance mass $M_{z'}$ are estimated from maximization of the combined likelihood function.

Comparison with CMS classical shape analysis

Classical analysis: search in the di-muon invariant mass distribution a peak compatible with a "narrow" resonance on top of a smooth background.





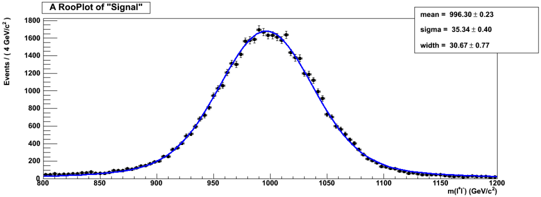
Comparison with shape analysis

- Fit for signal fraction for a fixed mass hypothesis
- Simultaneous fit for signal fraction and mass.

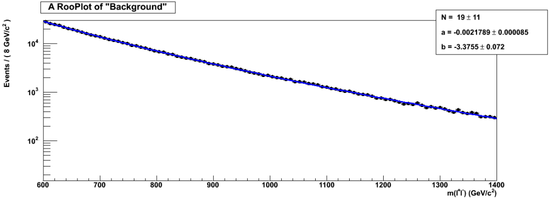
Classical shape based analysis

MEM with MAD-Weight

Shapes for Classical Analysis



Signal Model :
Breit-Wigner $\otimes$Gaussian

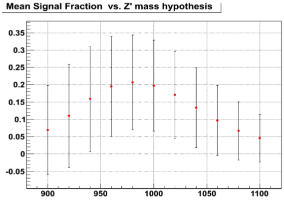


Background Model :
$Nexp^{am}mb$

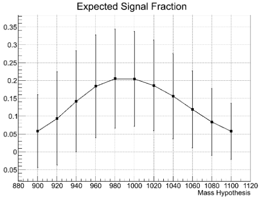
Comparison : Signal fraction and significance

Pseudo experiments with 20B+5S events for fixed mass points

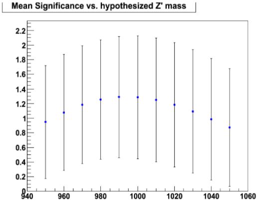
Classical Shape Analysis



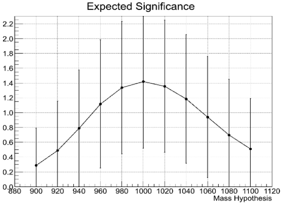
Matrix Element Method



Classical Shape Analysis



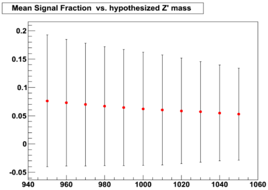
Matrix Element Method



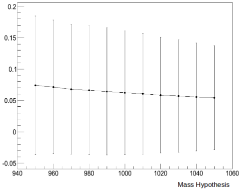
Comparison : Background-only experiments

Signal fraction and significance for 20 background-only experiments

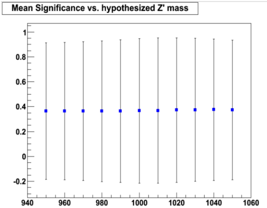
Classical Shape Analysis



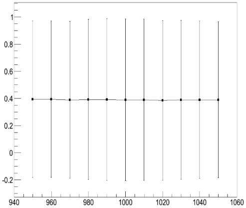
Matrix Element Method



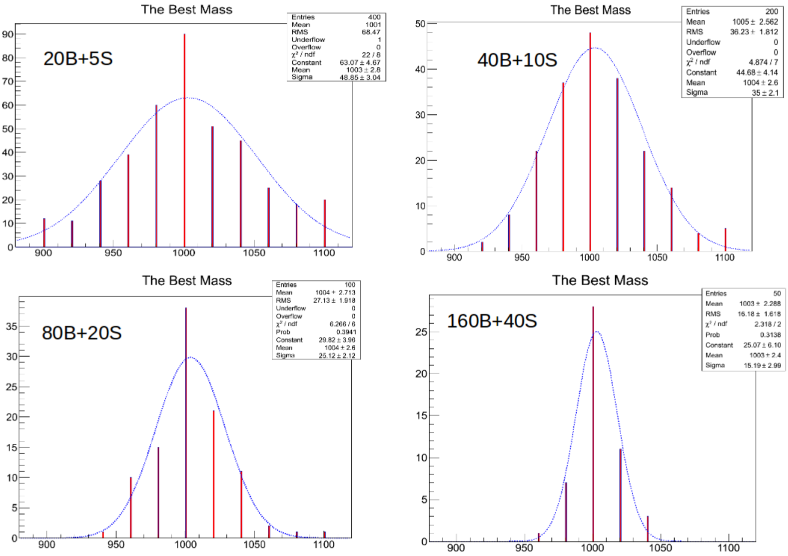
Classical Shape Analysis



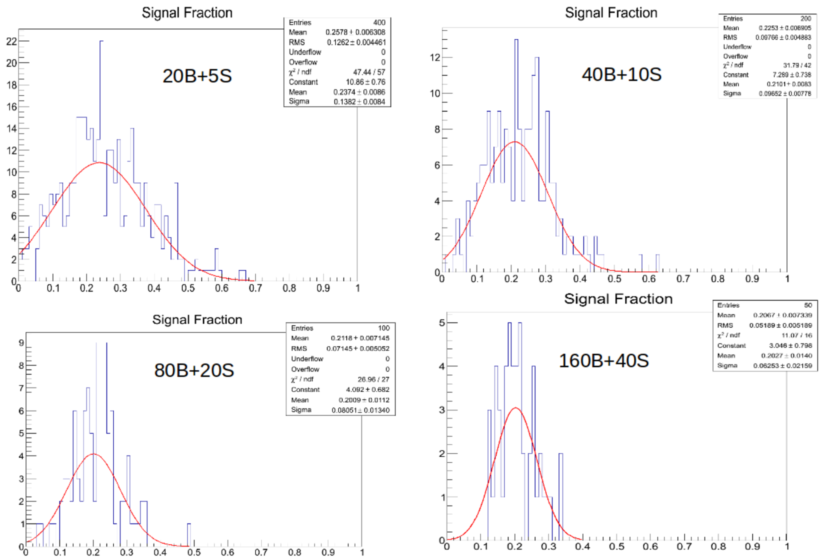
Matrix Element Method



MEM - simultaneous fit - mass

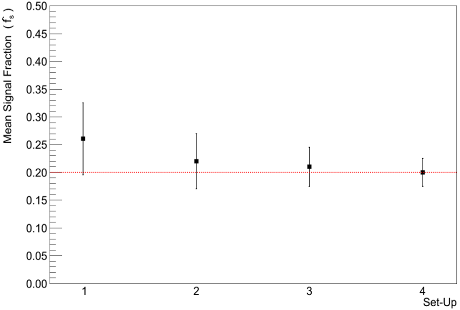


MEM - simultaneous fit - signal fraction

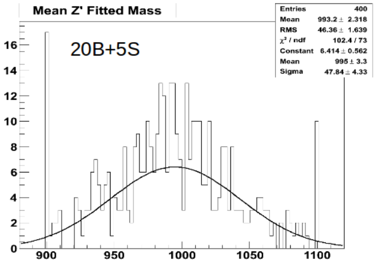


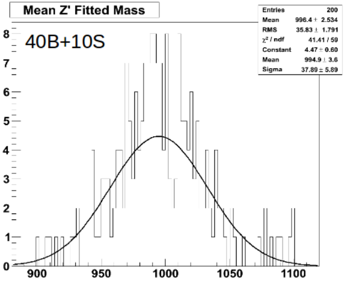
MEM - simultaneous fit - signal fraction

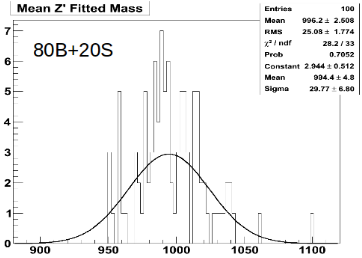
Estimated signal fraction with different pseudo experiments tests with different statistics (S+B ensemble) for a fixed signal fraction



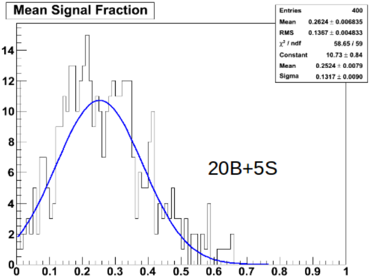
Classical analysis - simultaneous fit - mass

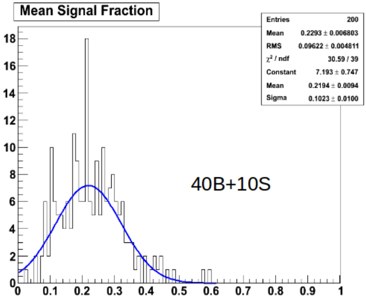


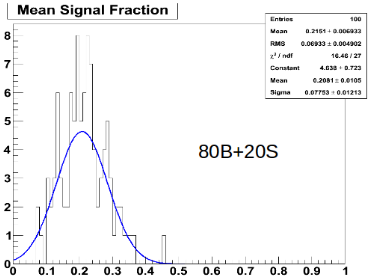




Classical analysis - simultaneous fit - signal fraction







The same bias observed !

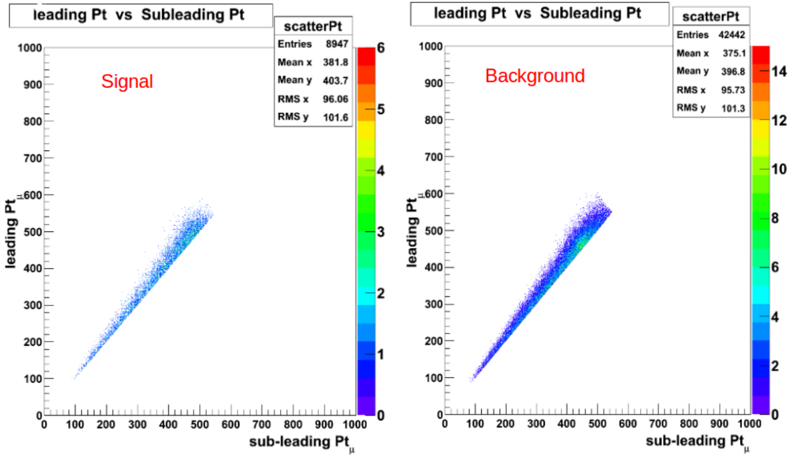
Conclusion

- A generic likelihood-based analysis approach with Matrix element method has been developed.
- The method can be used for parameter estimation (mass, cross-section) or hypotheses testing (significance, exclusion limits, spin-parity etc.).
- It is adapted and validated on MC for a search analysis $Z^{'} \rightarrow \mu^{-}\mu^{+}$
  - Mass estimator is ok! estimator is biased for signal fraction in low statistics cases.
  - The same bias was observed with shape analysis.
  - The bias disappears when we increase the size of S+B numbers in the samples.
  - Sensitivity gain $\approx 10 - 20\%$ comparison to shape analysis
- Application to another process like $H \rightarrow ZZ \rightarrow 4l$ can follow a very similar way.

Back-Up

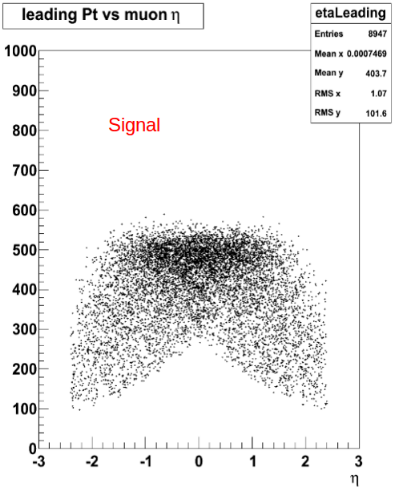
Kinematical Distributions

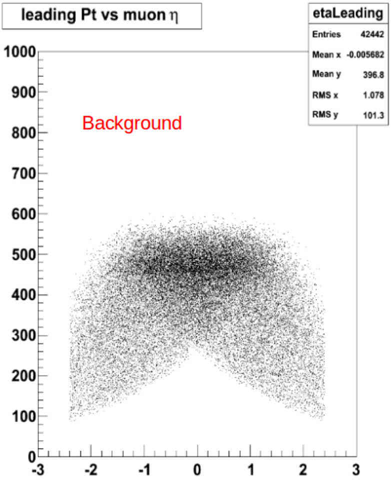
Leading P_t^μ vs. subleading P_t^μ for selected events.



Kinematical Distributions

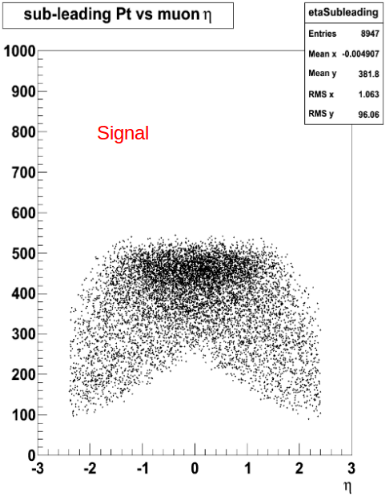
Leading P_t^{μ} vs. η^{μ} for selected events.

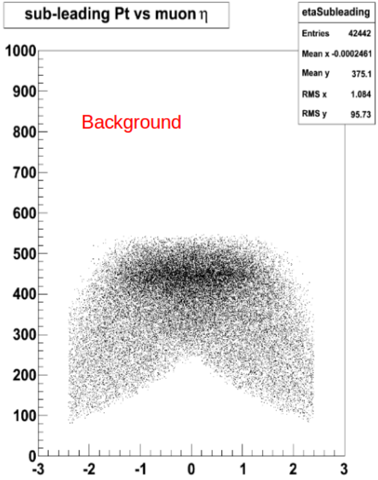




Kinematical Distributions

Sub-leading P_t^{μ} vs. η^{μ} for selected events.





MEM : Fitted Fraction vs. Fitted Mass