MATRIX ELEMENT METHOD

Florencia Canelli
Universität Zürich

May 28, 2013

Definition

- Commonly referred as the method that includes the matrix elements calculations for signal and background events to evaluate probability densities in an event-by-event basis

- Sometimes referred as any analysis that uses matrix element as variables, weights or other way
  - ideogram, neutrino weighting, etc. could fit in this category

Measurements and searches

- Originally developed to measure the top quark mass and W helicity in top events ($\alpha = m_t, f_{0,+,-}$)

$$P(\alpha)$$

- Later adapted to searches
  - Main of the effort was made for observation of single top and Higgs

$$P$$

Building P in an ideal case

- Probability density function for ONE event is characterized by a set of measurements x, for a set of parameters α

$$P(x \mid \alpha)$$

for a perfect detector

$$P(x \mid \alpha)\,dx = \frac{d\sigma(x \mid \alpha)}{\sigma(\alpha)}$$

measured variables

parameters

$$d\sigma \approx \frac{|M(\alpha)|^2}{\text{flux factor}} \times \text{ phase space}$$

Building P with detector resolutions

partonic integral

partonic variables

mapping between partonic and measured variables

$$P_W(x \mid \alpha) = \frac{1}{\sigma(\alpha)} \int d\sigma(y \mid \alpha) W(y, x)$$

Note that W(y,x) can be a function of parameters like JES

Building P - Acceptance

Acceptance limited by the physical properties of the detector and the event selection

$$P(x|\alpha) = \frac{A(x)\int d\sigma(y|\alpha)W(y,x)}{\sigma(\alpha) < A(\alpha) >}$$

Mean acceptance

$$< A(\alpha) >= \int A(x)P_W(x|\alpha)dx$$

Background

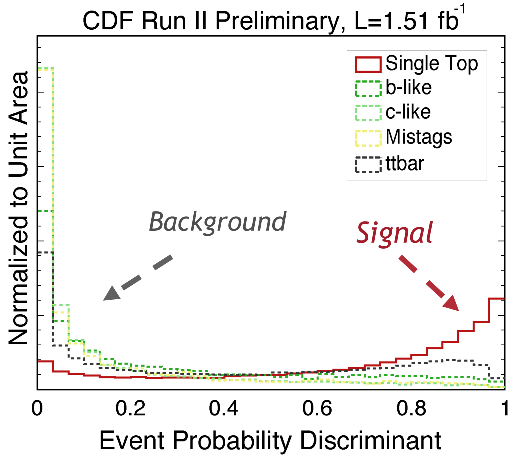
- Measurement: sum over all states that can lead to the set of measurements *x*

$$P(x \mid \alpha) = \sum_{i=states} c_i \, P_i(x \mid \alpha)$$

- Search: method is turned into a single-variable template analysis

$$EPD = \frac{P_{signal}}{P_{signal} + P_{background}}$$

In principle the extraction of signal is not different than any other template analysis



Likelihood for searches

Example single top at CDF

$$\mathcal{L}(\beta_1,\ldots,\beta_5;\delta_1,\ldots,\delta_{10}) = \underbrace{\prod_{k=1}^{B} \frac{e^{-\mu_k}\cdot \mu_k^{n_k}}{n_k!}}_{Poisson\ term} \cdot \underbrace{\prod_{j=2}^{5} G(\beta_j|1,\Delta_j)}_{Gauss\ constraints} \cdot \underbrace{\prod_{i=1}^{10} G(\delta_i,0,1)}_{Systematics}$$

Expected mean in bin k:

$$\mu_k = \sum_{j=1}^{5} \beta_j \cdot \underbrace{\left\{\prod_{i=1}^{10} [1+|\delta_i|\cdot(\epsilon_{ji+}H(\delta_i)+\epsilon_{ji-}H(-\delta_i))]\right\}}_{Normalization\ Uncertainty}$$

$$\underbrace{\cdot\ \alpha_{jk}}_{Shape\ P.} \cdot \underbrace{\left\{\prod_{i=1}^{10} (1+|\delta_i|\cdot(\kappa_{jik+}H(\delta_i)+\kappa_{jik-}H(-\delta_i)))\right\}}_{Shape\ Uncertainty}$$

$\beta_j = \sigma_j/\sigma_{SM}$ parameter
single top (j=1)
W+bottom (j=2)
W+charm (j=3)
Mistags (j=4)
ttbar (j=5)

k = Bin index

i = Systematic effect

δ_i = Strength of effect

$\varepsilon_{ji\pm}$ = ±1σ norm. shifts

$\kappa_{jik\pm}$ = ±1σ shift in bin k

®Correlation between Shape/Normalization uncertainty included (δ_i)

®Profile Likelihood with respect to all nuisance parameters

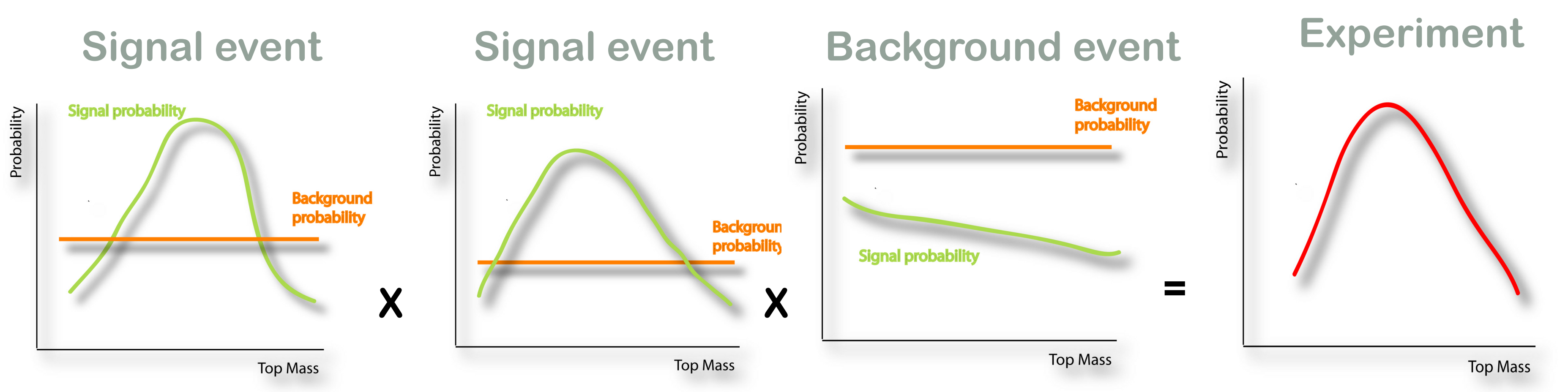
Likelihood for measurements

- Extracting a set of parameters α given N events is obtained by maximizing

$$L(\alpha, c_s) \propto \prod_{i=1}^{N} P(x_i \mid \alpha) \qquad \left(\text{with } \int P(x \mid \alpha)\, dx = 1\right)$$

- Fraction of signal events is obtained simultaneously

$$P(x \mid \alpha) = c_s\, P_s(x \mid \alpha) + (1 - c_s) P_b(x)$$



History

- Introduced in 2000 to measure m_t
  - Very limited statistics, 2 to 1 background contamination
  - Until then m_t was measured using a template of the reconstructed top mass
  - D0 experiment had no b-tagging. CDF much better precision on m_t
  - The goal of the MEM analysis was to use more information with less dependence on the MC

- Before 2000 there were similar ideas
  - K. Kondo: J. Phys. Soc. Japan, 57, 4126 (1988), J. Phys. Soc. Japan, 60, 836 (1991), J. Phys. Soc. Japan, 62, 1177 (1993)
  - R.H. Dalitz and G.R. Goldstein: Phys. Rev. D45, 1541 (1992), Phys. Lett. B287, 225 (1992), Phys. Rev. D47, 967 (1993) (w/ K. Sliwa), J. Mod. Phys. A9, 635 (1994), Proc. Roy. Soc. Lond. A455, 2803 (1999)

History

- In 2004, D0 experiment Nature 429,638 (2004)
    - First complete measurement, including: 1) all detector effects (e.g. reconstruction efficiencies, cuts, trigger, …), 2) correct normalization, 3) background probabilities, 4) MC tests of linearity, 5) pull calculations and 6) estimation of systematic effects.

- After 2004, it has been applied to all ttbar channels, single top, WH, H to WW at the Tevatron and some LHC use (H to ZZ (4l and 2l2j), VH, H to WW)

Technical details – Matrix element

- Matrix Element
  - In ttbar LO (Mahlon-Parke)
  - For searches Madgraph is usually used for single top and Higgs (HELAS)

$$P(x \mid \alpha) = \frac{1}{\sigma(\alpha)} \int d\sigma(y \mid \alpha) W(y, x \mid JES) f(x^1_{Bj}) f(x^2_{Bj}) dx^1_{Bj} dx^2_{Bj}$$

$$d\sigma \approx \frac{|M(\alpha)|^2}{\text{flux factor}} \times \text{ phase space}$$

The ME @ LO limits the size of the sample that can be described well by P
For example: ttbar in l+jets has 4 jets at LO, we could calculate P for the 4 leading jets in the >4 sample, but not optimally

Technical details - Parton Distribution Functions

- Use CTEQ routines (LO vs NLO)

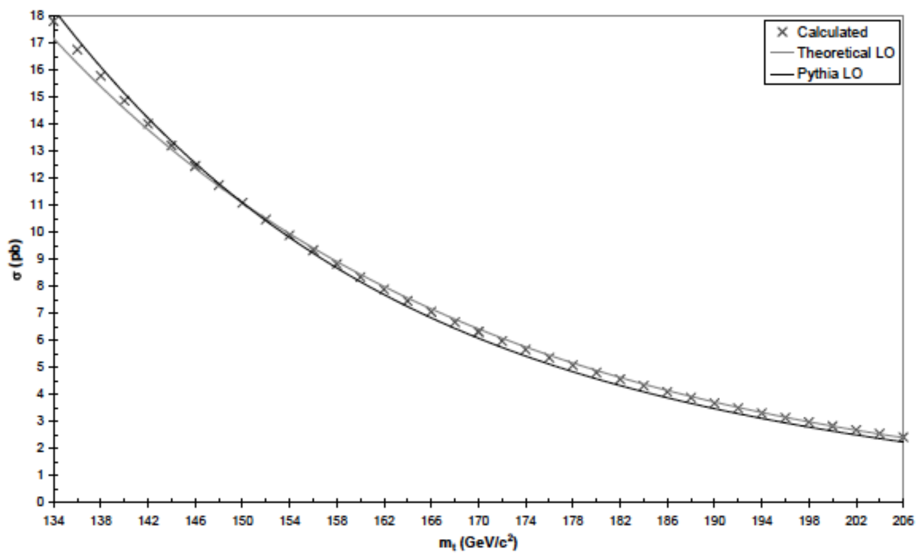
$$P(x|\alpha) = \frac{1}{\sigma(\alpha)} \int d\sigma(y|\alpha) W(y,x|JES) f(x_{Bj}^{1}) f(x_{Bj}^{2}) dx_{Bj}^{1} dx_{Bj}^{2}$$

PDF

Technical details - Normalization I

- Cumbersome and CPU intensive
  - Allows to test assumptions and to debug method
  - Sometimes used the σ(α) or σ obtained from MC

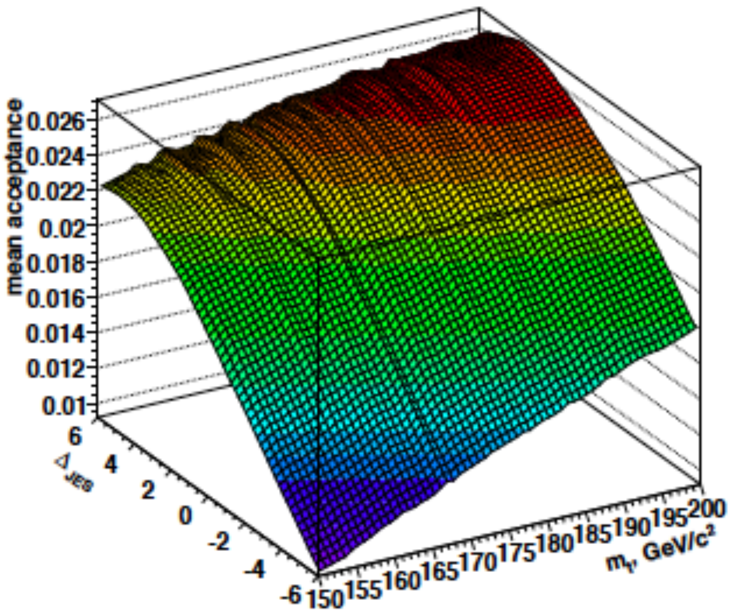
$$P(x|\alpha) = \frac{1}{\sigma(\alpha)} \int d\sigma(y|\alpha) W(y,x|JES) f(x^1_{Bj}) f(x^2_{Bj}) dx^1_{Bj} dx^2_{Bj}$$

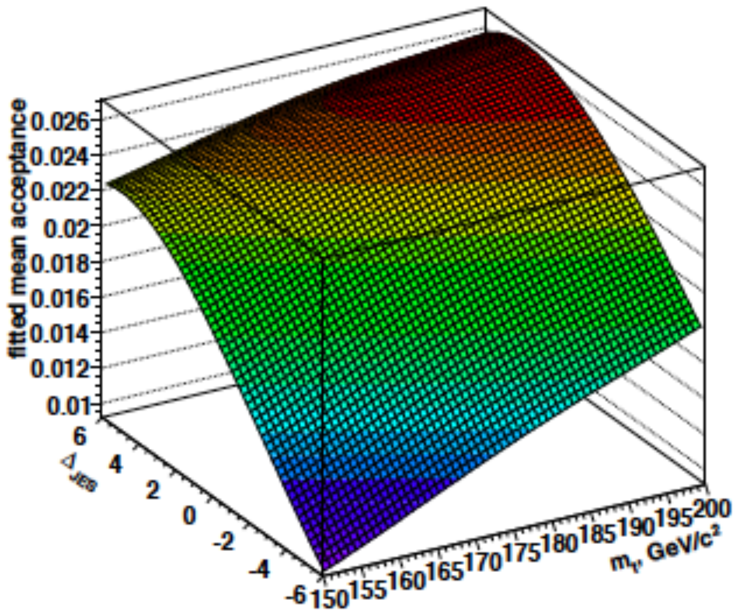


Technical details - Normalization II

- Mean acceptance calculated as function of the parameters to measure

$$< A(\alpha) > = \int A(x) P_W(x \mid \alpha) dx \qquad < A(\alpha) > = \frac{n_{accepted}(\alpha)}{n_{produced}(\alpha)}$$





Technical details - Background

- Use MC matrix element generators

$$P_b(x) = \frac{1}{\sigma_b} \int d\sigma_b(y) W(y,x)$$

- Note that not all the background processes can be calculated in P_b (fakes, multijets)
    - Creative ways have been used to develop pdfs for these backgrounds

Technical details - Integration

- Careful choice of integration variables
- Careful choice of assumptions in transfer functions
- Use narrow width approximation to smooth out integrand
- Convergence tests

- Methods
  - VEGAS most used for background (larger number of integrations)
  - Radmul (adaptive quadrature) (used for low number of integrations)
  - DIVONNE/CUBA

- CPU (2011) (depends on precision required) in a 2.0 GHz
  - Top mass with 5 integrations: 4 sec per event per point per jet parton-assignment (12 comb x 31 mass points x 17 JES points = 7 hours/event)
  - Single top (3 integrations): 1 s to 10 s per event all processes but ttbar in 3 jets (6 integrations) 5 mins
  - Both analyses were close to a million CPU-hours

Transfer functions

- Describe hadronization, detector resolution and reconstruction effects, including the deposition of energy from a parton outside the corresponding jet algorithm and extra energy from underlying event

$$W_{jet}(E_{jet},E_{parton}) = \frac{1}{\sqrt{2\pi}(p_1 + p_2 p_5)}[\exp\frac{-(\delta_E - p_1)^2}{2p_2^2} + p_3 \exp\frac{-(\delta_E - p_4)^2}{2p_5^2}]$$

where: $p_i = a_i + b_i E_{parton}$ $\delta E = (E_{parton} - E_{jet})$

  - Assume lepton well measured, angles and energy (delta functions)
  - Jes, parameterized function for energies (and angles)
    - Use Pythia MC
    - Done for different jet flavors and different eta regions using the same event selection
  - Use more information by including tracks, fraction of had and em energy, p_T of jet, etc. in NN to obtain a new "E_{jet}"

Jet-parton permutations and neutrino solutions

- One of the more appealing advantages of the MEM method is that each jet-parton permutation (or lepton) enters with a different weight

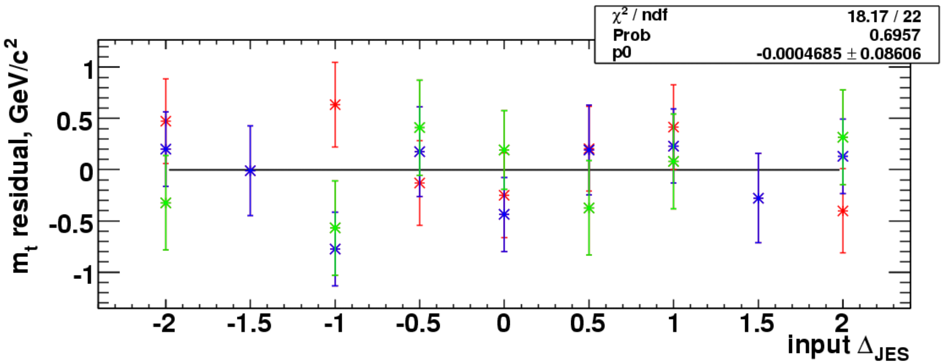
$$P(x\,|\,\alpha) = \frac{1}{\sigma(\alpha)<\mathrm{A}(\alpha)>}\sum_{jet-parton}^{2,6,12}\int |M(\alpha)|^2\, W(y,x\,|\,JES)\, f(x_{Bj}^1)\, f(x_{Bj}^2)\, dx_{Bj}^1\, dx_{Bj}^2\, d\Phi$$

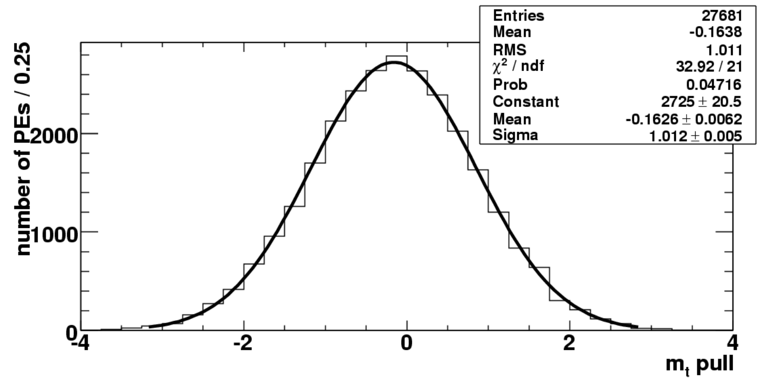
- For example in ttbar:
  - Most of the permutations do not contribute to P
  - The right permutation always contribute (among the first 3)

- The neutrino solutions are part of the integration

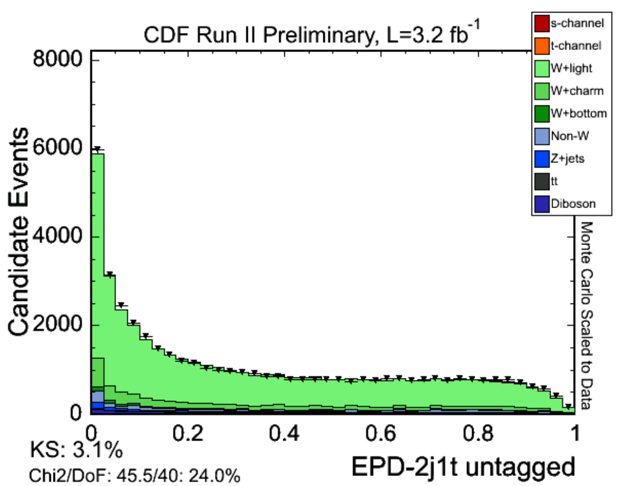
Tests and checks

- Linearity and pull tested with MC





- EPD data/MC in different control regions



Is it a better method ?

- In measurements with $\delta_{stat} > \delta_{sys}$ benefit from MEM approach
  - In original mt analysis had a factor of 2 improvement using the same data !
  - If statistics is large, MEM or a simple Mt reco reach the similar sensitivity That being said, more statistical power can help decrease systematics

- The current case of the top mass
  - All uncertainties between 0.1-0.6 GeV
  - Total uncertainty in one analysis is about 1 GeV (0.5 stat + 0.8 syst)
  - Going below will require more understanding of color reconnection, initial/ final state radiation, etc.
  - The statistical power of the MEM could help decrease systematics

Is it a better method ?

- In searches the ME has performed between 0% and 10% gain over the other multivariate analysis (NN and BDT)
  - Found to be 50-70% correlated (later combined)
  - MEM as template ?
  - NLO information not totally used ?
  - More detector information used by NN/BDT ?
  - Analyzer dependence ?

- In general systematic uncertainties have been found to be smaller in MEM results than with other methods

Summary and conclusions

- MEM has provided experiments a different way to analyze data
  - Beyond cut and count and one-variable templates (larger stat power)
  - With complete analyzers control (vs NN)

- Many improvements and optimizations could be envisioned

- Experimentally is a very expensive analysis
  - It is delicate and slow
  - The pace at the LHC might not be right for MEM

- A more collaborative effort among experimentalist and theorists on creating MEM tools would be very beneficial