Applications of matrix element methods

Adrien Caudron (UCLouvain – CP3)
on behalf of ATLAS and CMS experiments

LCWS 2014
6-10 October 2014
Belgrade, Serbia

Outline:

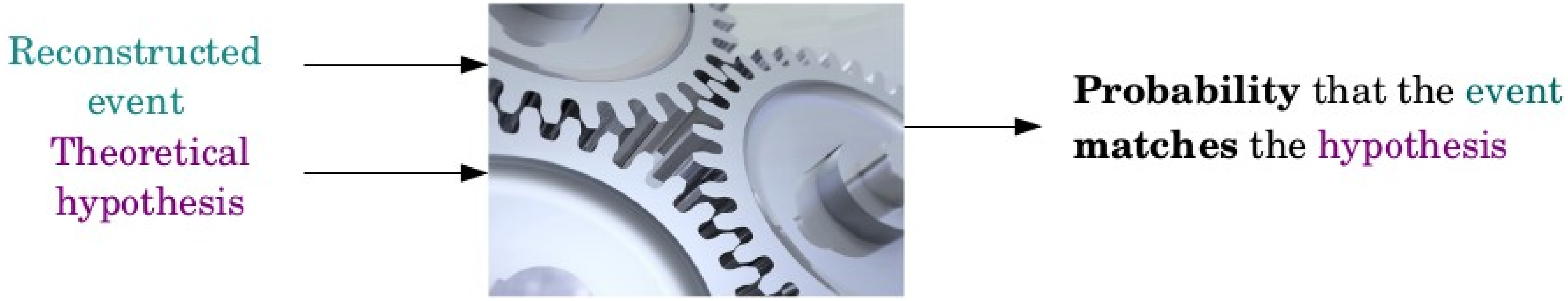
- Introduction to Matrix Element Method (MEM)
- Application to ATLAS and CMS:
  * H → WW search
  * ttH search with H → bb
  * H → ZZ search and characterisation

The Matrix Element Method

- Event by event discriminator based on matrix elements

- Usage of a maximal amount of theoretical information available from the hard process

- Combined with reconstruction level information

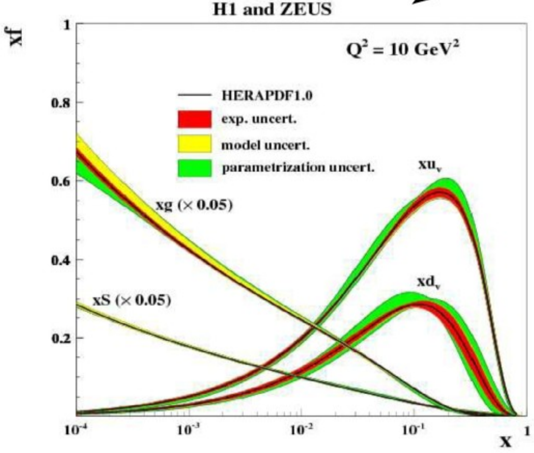


How does it work?

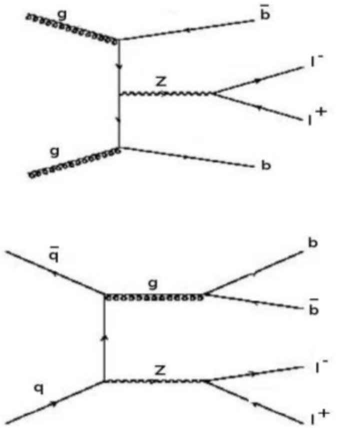
Building the discriminant

Probability that the event with reconstructed kinematics x matches the hypothesis α:

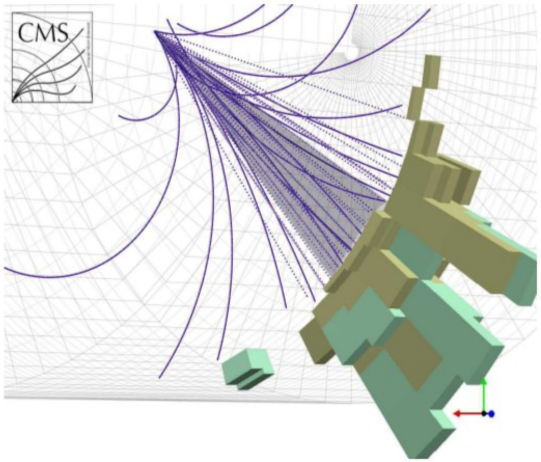
$$P(x^{vis}|\alpha) = \frac{1}{\sigma_\alpha} \int dx_1 dx_2 f(x_1) f(x_2) \int d\phi |M(p)_\alpha|^2 W(p^{vis}, p)$$



Proton density functions



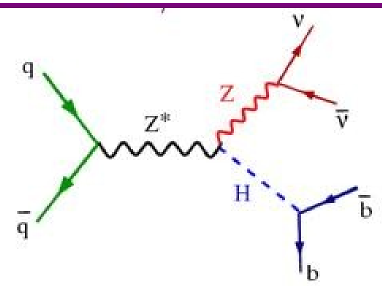
Matrix Element at LO of the hypothesis



Transfer functions extracted from simulation

Why is it interesting?

What you have: 4-vector of particles in final state



E_{miss}^{T}

p_T, η, φ, E (x n. particle)

↓

Which variables to use? How combining them?

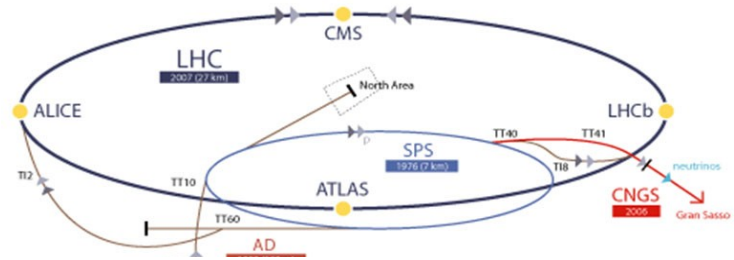
↓ **Matrix element method**

- **All kinematics** → 1 **discriminating variable** by hypothesis
  - Use almost **all you know** from the event
  - Based on **physics** knowledge
  - you can still **combine different hypotheses** and eventually other information (b-tagging, extra-objects...)

↓

Final discriminant

ATLAS and CMS

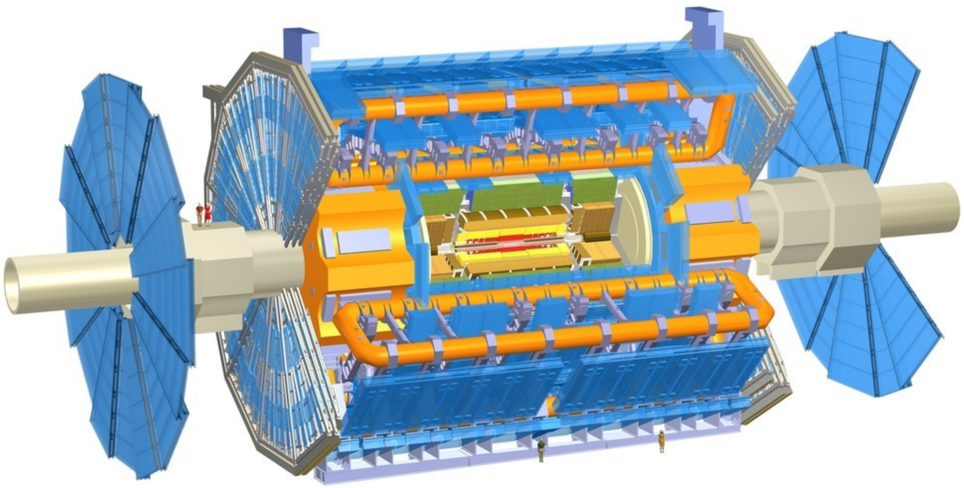


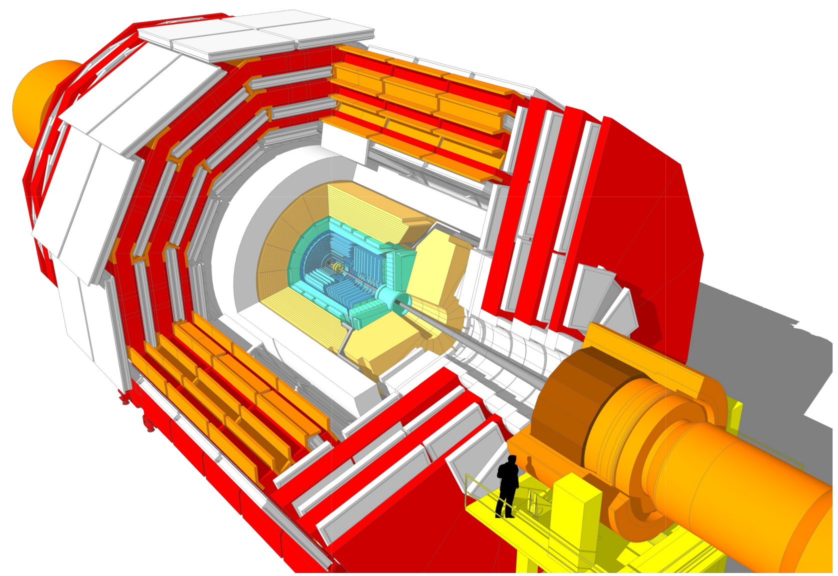
- Experiments at **LHC**

- **Discovery** of the **Higgs** boson in 2012

- ATLAS during LHC Run1:
  - 2010-2012: 4.7/fb + 20.3/fb at 7 TeV and 8 TeV

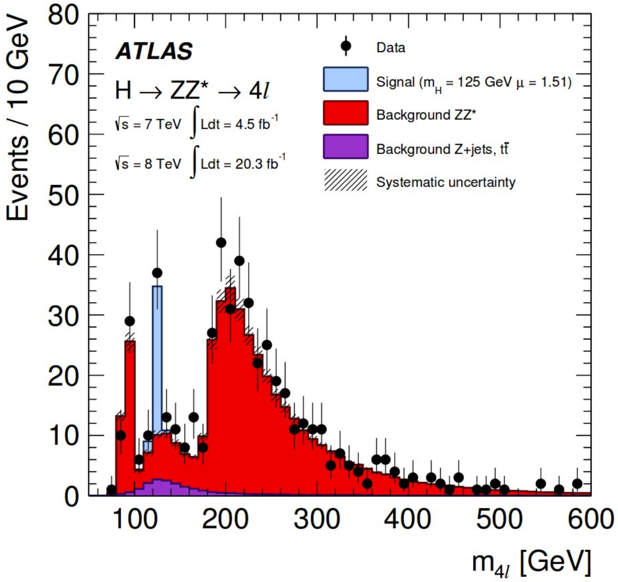
- CMS during LHC Run1:
  - 2010-2012: 5.1/fb + 19.7/fb at 7 TeV and 8 TeV

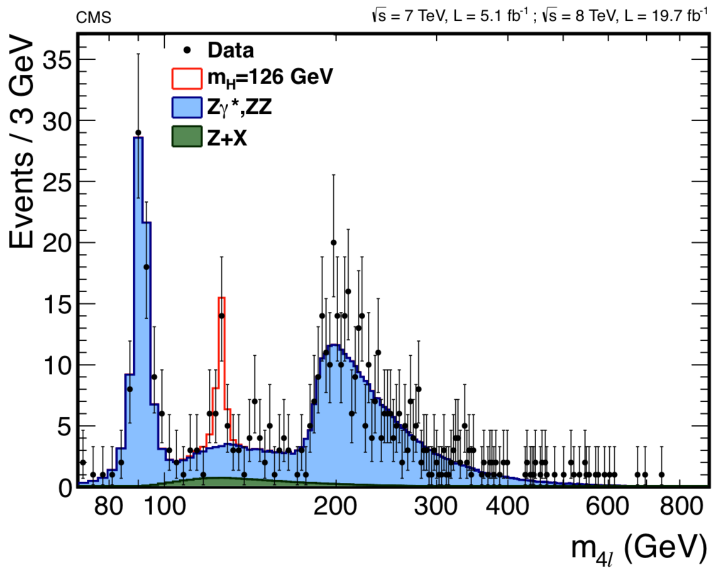




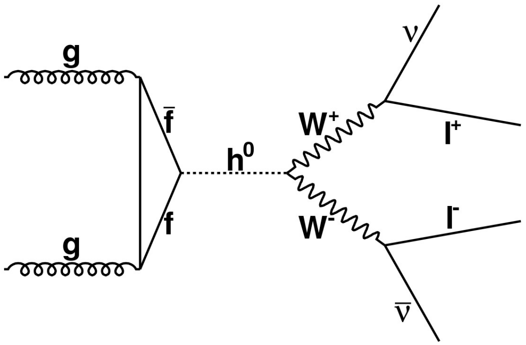
Application of Matrix Element Method

- H → WW search
- ttH(bb) search
- H → ZZ
  - Higgs search
  - Spin/parity determination
  - HZZ vertex tensor structure





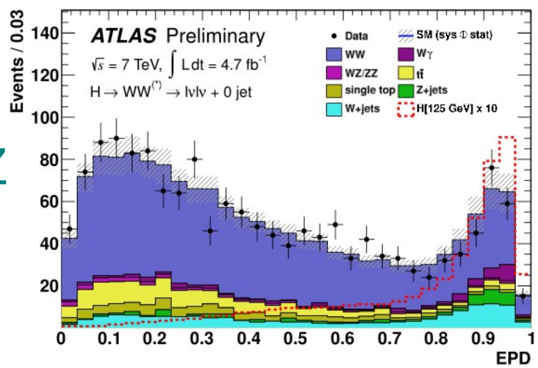
Search for H → WW

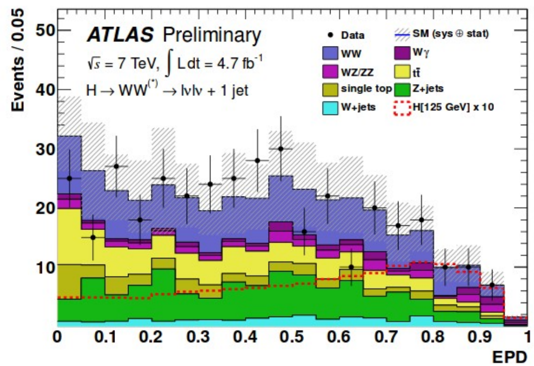


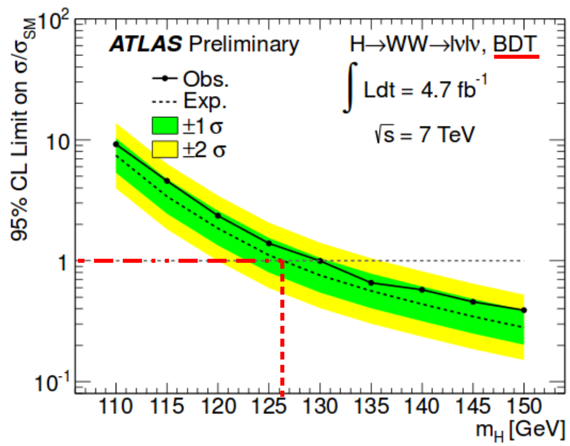
- H → WW → lνlν (l = e, μ):
  - 2 well reconstructed leptons
  - Missing energy from neutrinos
    - Poor mass resolution
- MEM used as a cross-check to the BDT analysis
  - Background hypotheses:
    - WW for categories 0/1 jet
    - ttbar for category with 1 jet
  - Difficulty: missing transverse energy
    - Integrate over the allowed phase space of the undetected neutrinos
  - Transfer function on jet energy

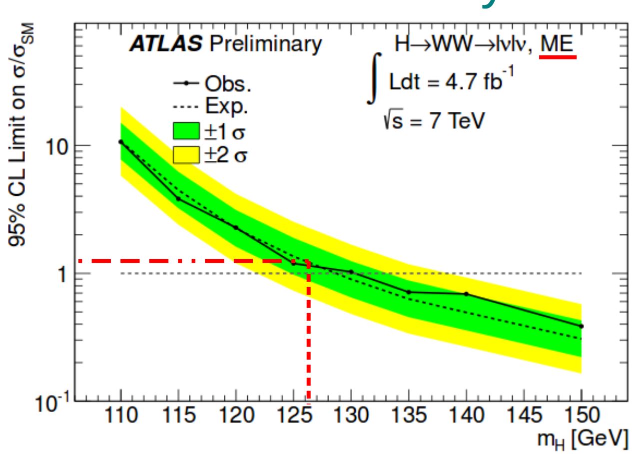
Search for H → WW
Results

- Define P_s and P_b as event probability densities for the signal and the background
  - Build a discriminant: $EPD = P_s/(P_s + P_b)$
  - Limited discrimination power for W/Z+jets and ZZ/WZ
  - The discriminant is used to subdivide the events in bins of different sensitivity
- This simple use of ME is nevertheless competitive
  - Only 15-20% less sensitive than the BDT analysis





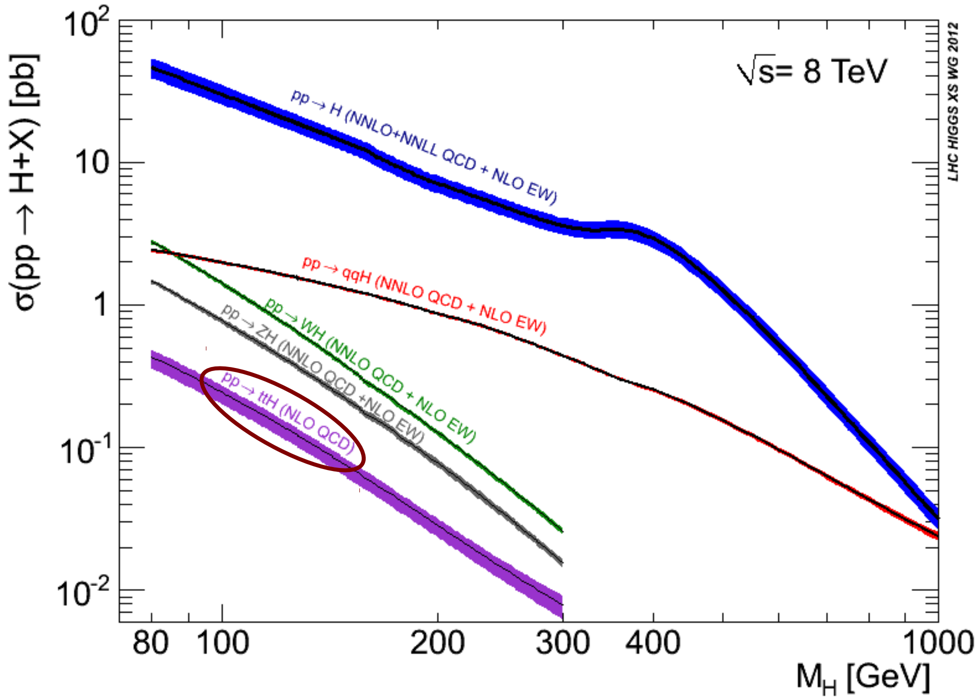


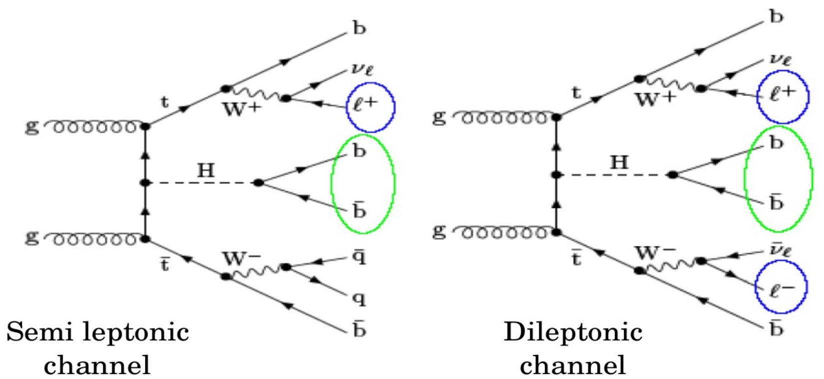


ATLAS-CONF-2012-060

ttH, H → bb search based on MEM

- One of the most challenging channel:
  - low production rate
  - H → bb highest BR but low M(bb) resolution
- Final states with lot of objects:
  - very distinctive signature
  - MEM:
    - Handles final state combinatorics
    - Discriminates against irreducible ttbb background





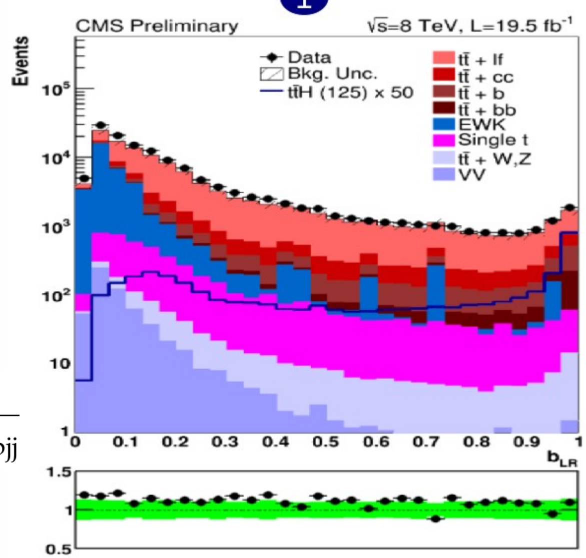
ttH, H → bb search based on MEM

1. Build likelihood ratio discriminant to distinguish signal and tt+bb/cc (irreducible bkg) from tt+light flavour
2. Discrimination Signal and tt+bb/cc improved by the MEM
3. Build the final discriminant combining the MEM weights with b-tagging information

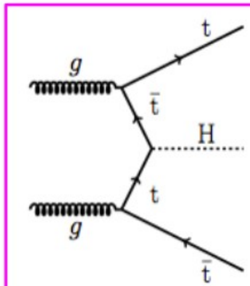
1

Test 2 versus
4 b-jets
hypotheses

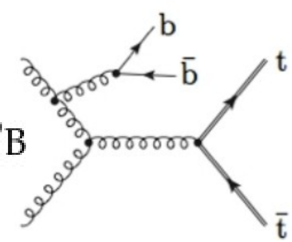
$$b_{\mathrm{LR}} \equiv \frac{\mathcal{L}_{\mathrm{bbbb}}}{\mathcal{L}_{\mathrm{bbbb}} + \mathcal{L}_{\mathrm{bbjj}}}$$



2



w_{S} **vs** w_{B}

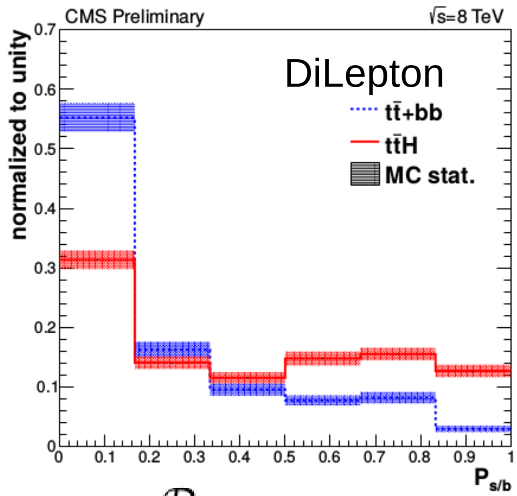


3

Probabilities: $\mathcal{P}_{\mathrm{S}}(\mathbf{y},\xi) \equiv w_{\mathrm{S}}(\mathbf{y})\,\mathcal{L}_{\mathrm{bbbb}}(\xi)$

$$\mathcal{P}_{\mathrm{B}_1}(\mathbf{y},\xi) \equiv w_{\mathrm{B}}(\mathbf{y})\,\mathcal{L}_{\mathrm{bbbb}}(\xi)$$

$$\mathcal{P}_{\mathrm{B}_2}(\mathbf{y},\xi) \equiv w_{\mathrm{B}}(\mathbf{y})\,\mathcal{L}_{\mathrm{bbjj}}(\xi)$$

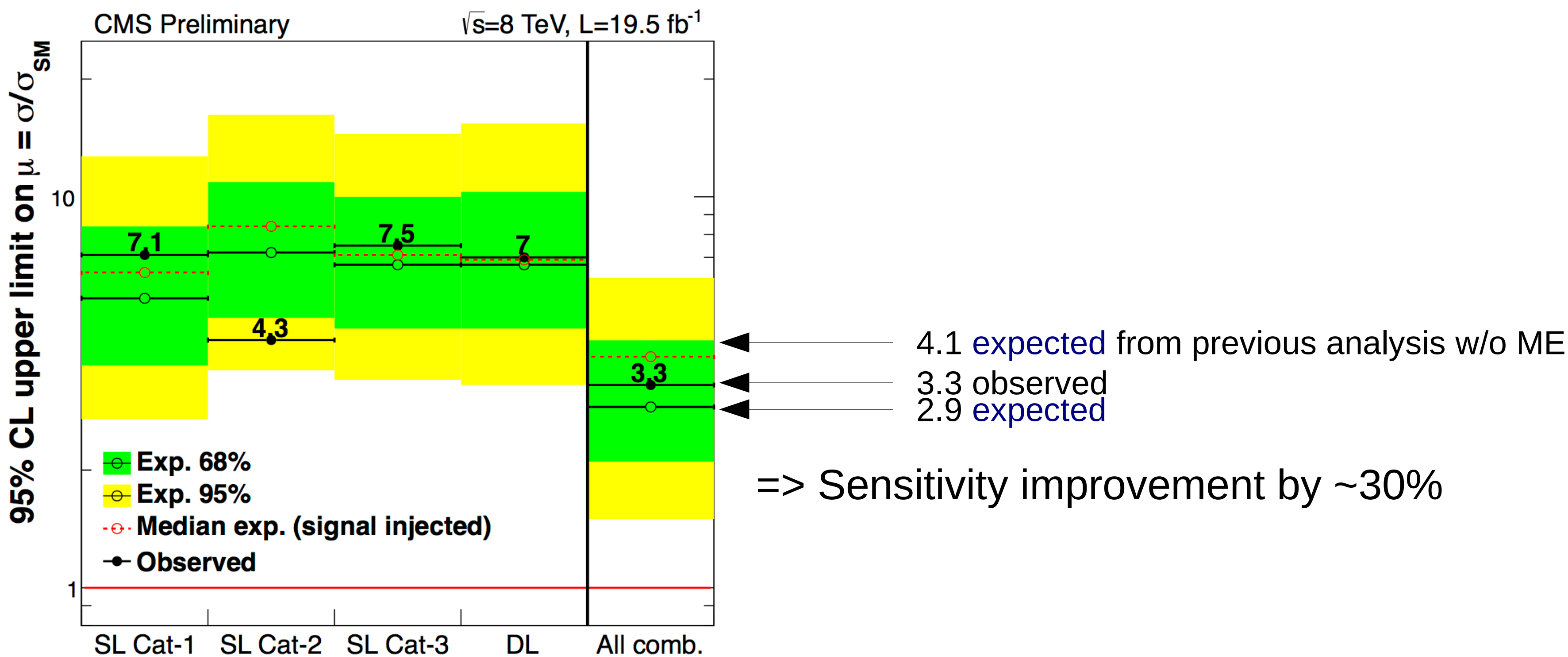


$$P_{s/b} = \frac{\mathcal{P}_{\mathrm{S}}}{\mathcal{P}_{\mathrm{S}} + \lambda_{b/j}\mathcal{P}_{\mathrm{B}_1} + (1-\lambda_{b/j})\mathcal{P}_{\mathrm{B}_2}}$$

$\lambda_{b/j}$ = fraction ttbb / ttjj

ttH, H → bb search based on MEM

- **Results**:
  - Categorisation depending on the number of leptons and jets
  - Exclusion limits on x times the σ_{SM} :

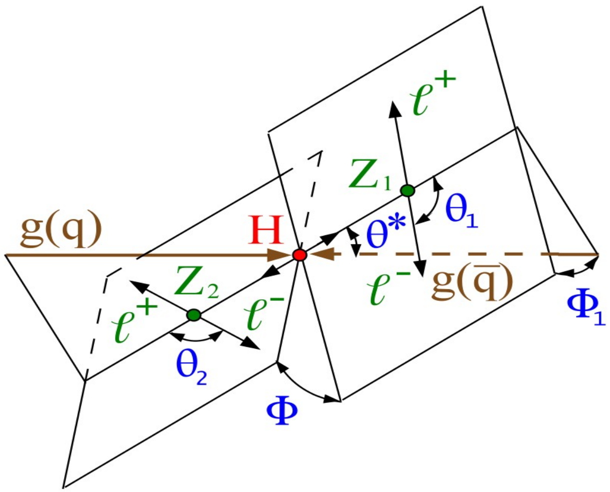


=> Sensitivity improvement by ~30%

CMS-PAS-HIG-14-010

H → ZZ analysis

- H → ZZ → 4 leptons (e, μ)
  - Clean channel → leptons very well measured
    - Main background: ZZ production
  - One of the main channel leading to the Higgs discovery
  - Well suited channel for studying Higgs properties
- Kinematic configuration of the 4-leptons in the centre of mass:
  - Helicity angles: $\vec{\Omega} \equiv (\theta^*, \Phi_1, \theta_1, \theta_2, \Phi)$
    - Sensitivity to the spin and parity
  - Z_1 and Z_2 masses

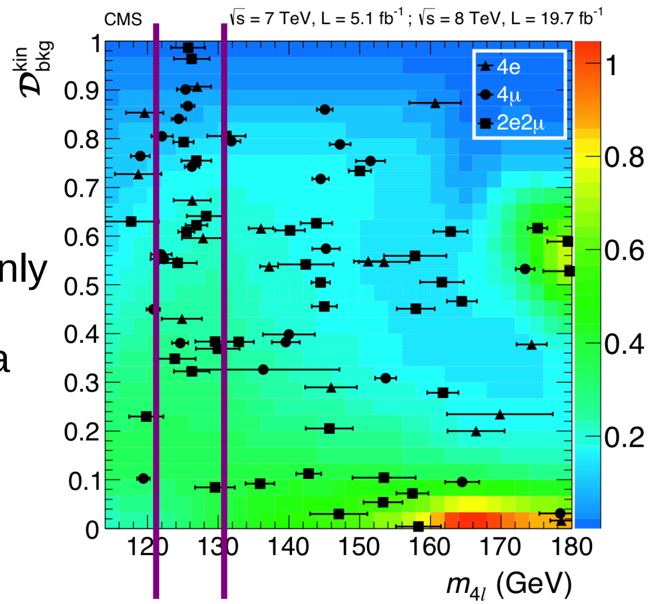


- Matrix Element Likelihood Approach (MELA):
  - Make use of the **whole kinematic configuration** of the system $\mathcal{P}^{\text{kin}}(m_{Z_1}, m_{Z_2}, \vec{\Omega}|m_{4\ell})$
  - Test of **several hypotheses**
    - Signal vs background (H → ZZ vs ZZ)
    - Signal vs signal (test of spin and parity)
  - **No integration** → simplified approach
    - Possible thanks to good lepton reconstruction

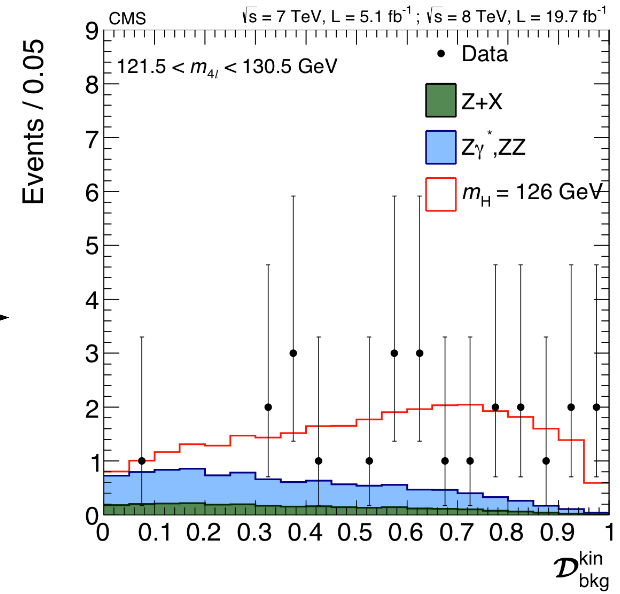
$H \to ZZ \to 4l$
MEM for signal extraction

- Build discriminant H → ZZ (0^+) vs ZZ (bkg): $$\mathcal{D}^{\text{kin}}_{\text{bkg}} = \frac{\mathcal{P}^{\text{kin}}_{0^+}}{\mathcal{P}^{\text{kin}}_{0^+} + \mathcal{P}^{\text{kin}}_{\text{bkg}}}$$

Bkg only vs data



$121.5 < m_{4l} < 130.5$



- Combined in a likelihood function
  - with m(4l) and p_T(4l) in 0/1 jet categories
  - and with m(4l) and a discriminant for VBF production in the 2 jet category

Phys. Rev. D 89, 092007

$H \rightarrow ZZ \rightarrow 4l$

MEM for signal hypothesis test

- Same principle but test the spin and parity properties:

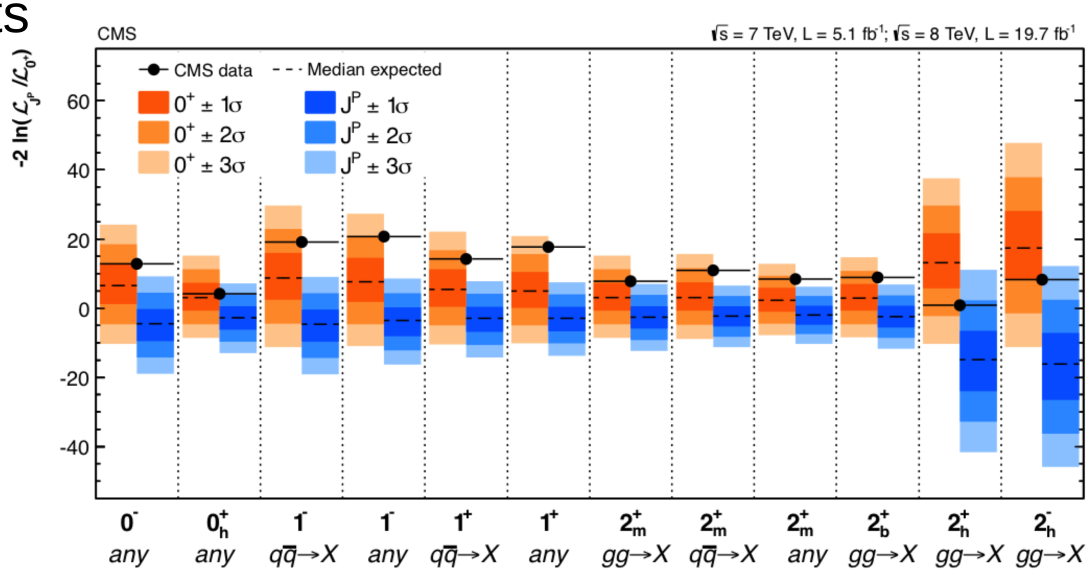
$$\mathcal{D}_{J^P} = \left[1 + \frac{\mathcal{P}^{\rm kin}_{J^P}(m_{Z_1}, m_{Z_2}, \vec{\Omega}|m_{4\ell})}{\mathcal{P}^{\rm kin}_{0^+}(m_{Z_1}, m_{Z_2}, \vec{\Omega}|m_{4\ell})}\right]^{-1}$$

J^P refer to the alternative hypothesis

- Spin 1 and spin 2 → reduce assumption on production mechanism: integrate the ME squared over the production angles cosθ* and Φ_1
- Background is identified with: $\mathcal{D}_{\rm bkg} = \left[1 + \frac{\mathcal{P}^{\rm kin}_{\rm bkg}(m_{Z_1}, m_{Z_2}, \vec{\Omega}|m_{4\ell}) \times \mathcal{P}^{\rm mass}_{\rm bkg}(m_{4\ell})}{\mathcal{P}^{\rm kin}_{0^+}(m_{Z_1}, m_{Z_2}, \vec{\Omega}|m_{4\ell}) \times \mathcal{P}^{\rm mass}_{\rm sig}(m_{4\ell}|m_{0^+})}\right]^{-1}$
- Combined the 2 discriminants in a likelihood function:
  - A test statistics is performed:

$$q = -2\ln(\mathcal{L}_{J^P}/\mathcal{L}_{0^+})$$

  - Hypotheses tested:



Phys. Rev. D 89, 092007

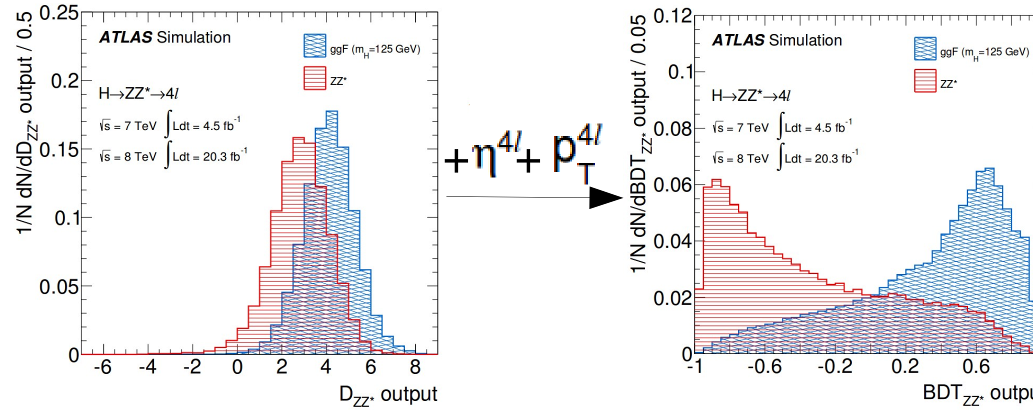
$H \rightarrow ZZ \rightarrow 4l$

ATLAS analyses

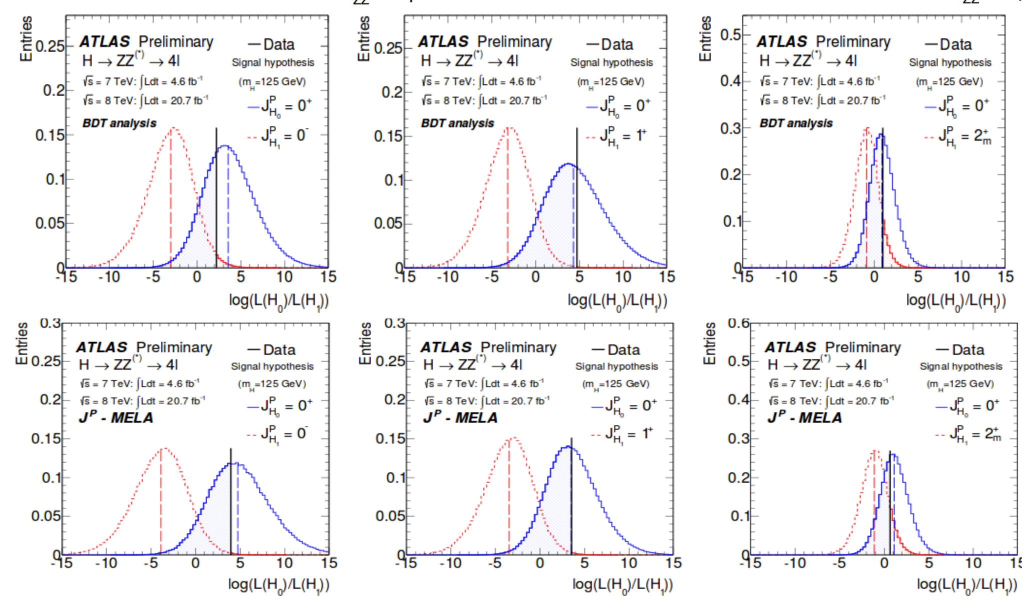
- Signal extraction:
  - Use MELA to build a discriminant: D_{ZZ}
  - BDT with $D_{ZZ} + \eta^{4l} + p_T^{4l}$
    - used in a fit to extract the signal strength and mass

arXiv:1408.5191

$$D_{ZZ^*} = \ln\left(\frac{|\mathcal{M}_{sig}|^2}{|\mathcal{M}_{ZZ}|^2}\right)$$



- Spin/parity determination:
  - MELA discriminant used as an alternative method to a BDT
  - Test statistics 0^+ vs J^P
  - MEM and BDT have similar sensitivity

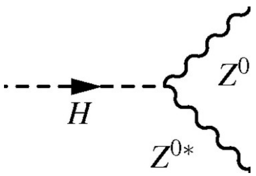


ATLAS-CONF-2013-013

$H \rightarrow ZZ \rightarrow 4l$

Prospects: test HZZ vertex tensor structure

- Prospective studies at 14 TeV for HL-LHC
  - With 300 fb^{-1} and 3000 fb^{-1}
    - Highly challenging environment: high PU...
- Test sensitivity to non-SM contribution to HZZ vertex:



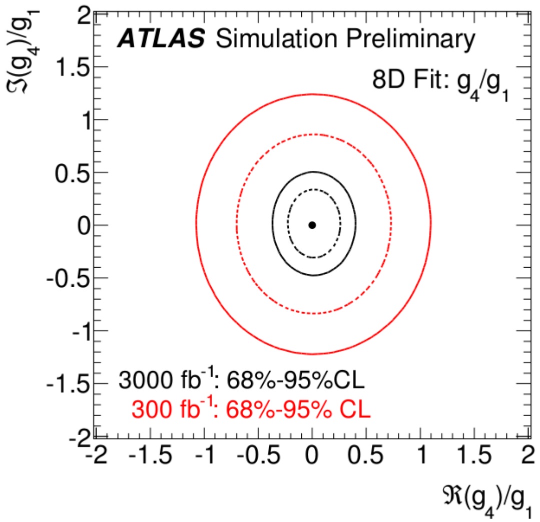
  - Amplitude depends on 4 coupling constants:
    - *g1:* SM expectation
    - *g2*: beyond SM contribution in loop
    - *g3* ~ 0
    - *g4*: CP-odd components
- Two strategies based on MEM:
  - Define set of sensitive variable built from the ME
  - ME analytical form in a likelihood (see next slide)
    - Both methods are competitive

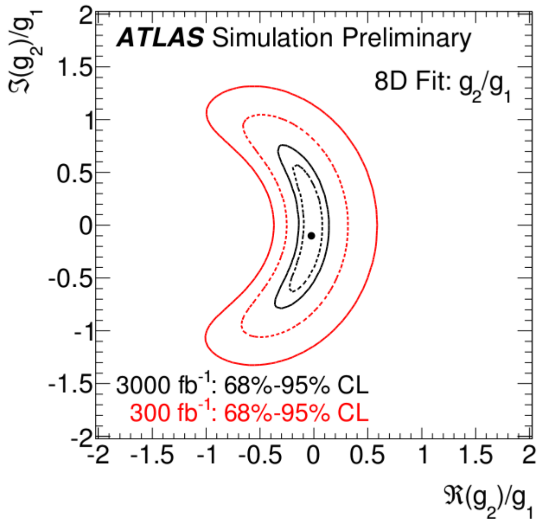
Observable	Sensitivity
$\ln \frac{\|ME(g_1=1,g_2=0,g_4=0)\|^2}{\|ME(g_1=0,g_2=0,g_4=1)\|^2}$	$\|g_4\|/g_1$
$\ln \frac{\|ME(g_1=1,g_2=0,g_4=-2+2i)\|^2}{\|ME(g_1=1,g_2=0,g_4=2+2i)\|^2}$	$\Re(g_4)/g_1$
$\ln \frac{\|ME(g_1=1,g_2=0,g_4=2-2i)\|^2}{\|ME(g_1=1,g_2=0,g_4=2+2i)\|^2}$	$\Im(g_4)/g_1$
$\ln \frac{\|ME(g_1=1,g_2=0,g_4=0)\|^2}{\|ME(g_1=1,g_2=1,g_4=0)\|^2}$	$\|g_2\|/g_1$
$\ln \frac{\|ME(g_1=1,g_2=-1+i,g_4=0)\|^2}{\|ME(g_1=1,g_2=1+i,g_4=0)\|^2}$	$\Re(g_2)/g_1$
$\ln \frac{\|ME(g_1=1,g_2=1-i,g_4=0)\|^2}{\|ME(g_1=1,g_2=1+i,g_4=0i)\|^2}$	$\Im(g_2)/g_1$

Test HZZ vertex tensor structure

ME output as direct input to the fitted likelihood

- Define a pdf function from the ME:
  - Use this pdf in the likelihood function
    - Sensitive to coupling constants g_i (real and imaginary parts)
- Perform a fit in the plan: $(\Re(g_i)/g_1, \Im(g_i)/g_1)$
- 2D contour plots of the exclusion limits are derived
- Results expressed as limit on $f_{g_i} = \frac{|g_i|^2\sigma_i}{|g_1|^2\sigma_1 + |g_2|^2\sigma_2 + |g_4|^2\sigma_4}$





→

Luminosity	f_{g_4}	f_{g_2}
300 fb^{-1}	0.20	0.29
3000 fb^{-1}	0.06	0.12

Expected 95% CL upper limits

ATL-PHYS-PUB-2013-013

Conclusion

- Matrix Element Method powerful method
  - Help to improved signal/background discrimination
  - Lot of nice features but not easy tool (integration, automation, cpu...)
  - Lively field: improvements in the future

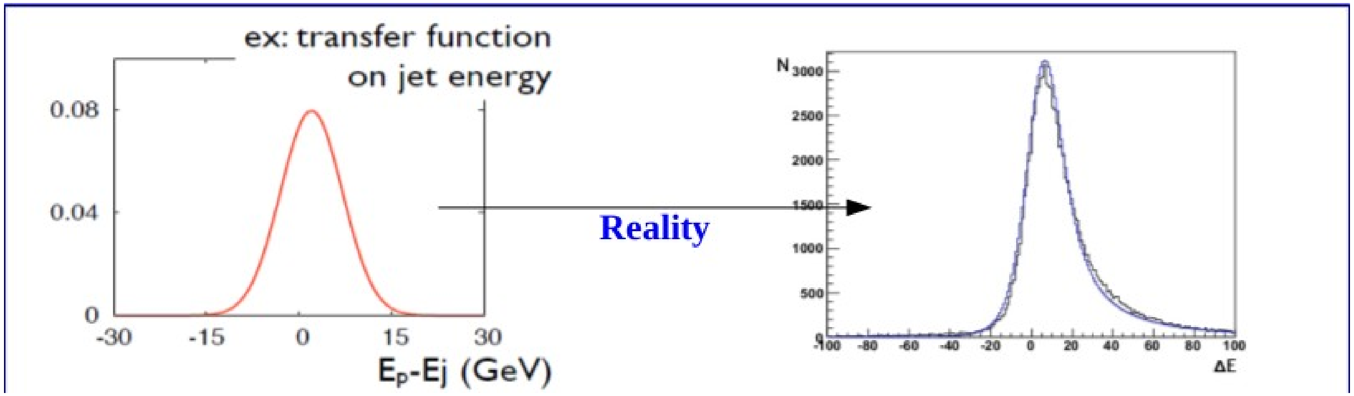
- ATLAS and CMS experiments used this method in different analyses:
  - Search and discovery of the Higgs boson:
    - $H \rightarrow WW$: no gain observed
    - ttH(bb): clear gain obtained compared to previous analysis
    - $H \rightarrow ZZ$: successfully combined with other discriminating variables
  - Higgs spin/parity determination in $H \rightarrow ZZ$
  - Measure of the HZZ vertex tensor structure

- Becoming a standard tool in HEP:
  - To be taken into account for future analysis perspective
    - Useful for precision measurements and for search with few free parameters

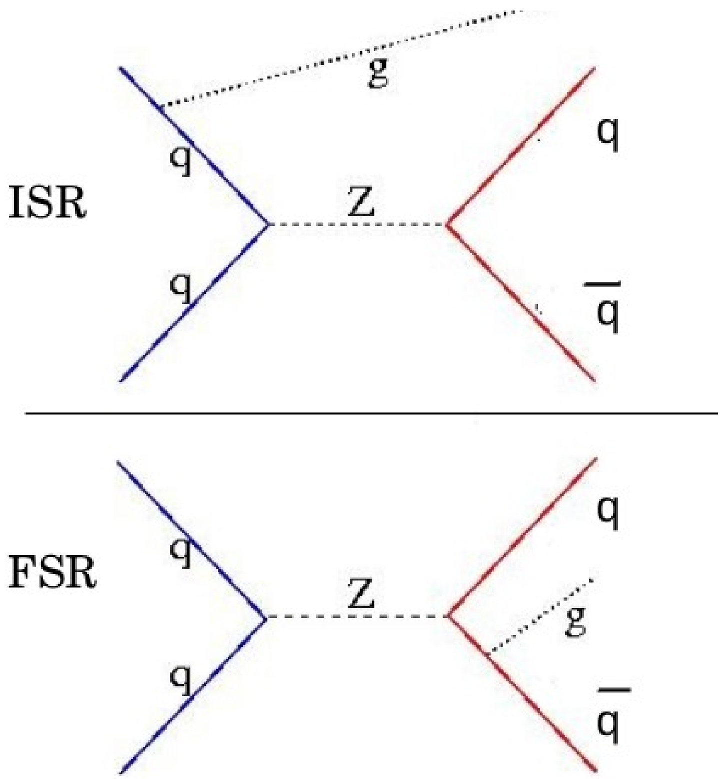
Backup

Transfer function

- Transfer function:
  - showering/hadronisation effects
  - + experimental resolution/reconstruction
    - P(x, α) convoluted with a TF W(p vis ,p)
- Example: Likelihood fit on Δ(E parton – E jet)
  - Can use another variable (ex: muon → dependence in $1/p_T$)
- Imply particles not correlated
- No dedicated TF for neutrinos
- Example:



NLO effects

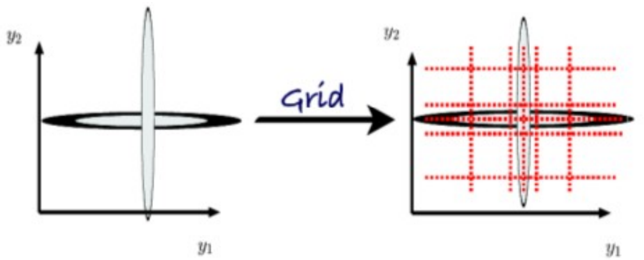


- The Z produced is not at rest
  - induce a transverse boost of the system
  - Matrix Element doesn't match anymore
  - ➢ Correct for the boost

- Instead of having 2 particles in the final state → 3
  - Matrix Element doesn't match anymore
  - ➢ Apply another Matrix Element: signal+extra object
  - ➢ Or recombine the extra jet+particle before applying MEM

Challenges

- Integral computation → **not trivial** ! Sharp functions:
  - Breit-Wigner
  - Transfer functions
- Integral **convergence:** need adaptive Monte Carlo Technique to pickspoint in interesting areas



- This is model dependent → ideal case: **automatic**, model independent, **fast**
- Real experiment: different configuration possible
  - need to handle combinatorics → average between several weights (each parton-jet matching combination possible)
- Missing energy → P(x, α) must be summed over the unobserved degrees of freedom.

Pros and cons of the method

✔ Maximize the amount of **theoretical information** for your discrimination

✔ **No "training"** as for most MVA methods

✔ **Many** potential **applications** (Tevatron: top mass measurement, single top discovery, CMS: Higgs search, spin correlation measurement, ...)

✗ Depending on your model, the computation can be **CPU demanding**

Time to compute the weight of one event, using MadWeight 5 [1]

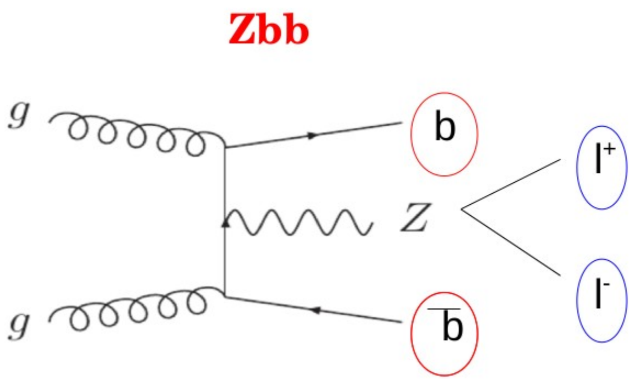
ZH	< 5 s
tt fully leptonic	10 s
Zbb	18 s
tt semi-leptonic	41 s
ttH fully leptonic	1 min

x thousand of events !

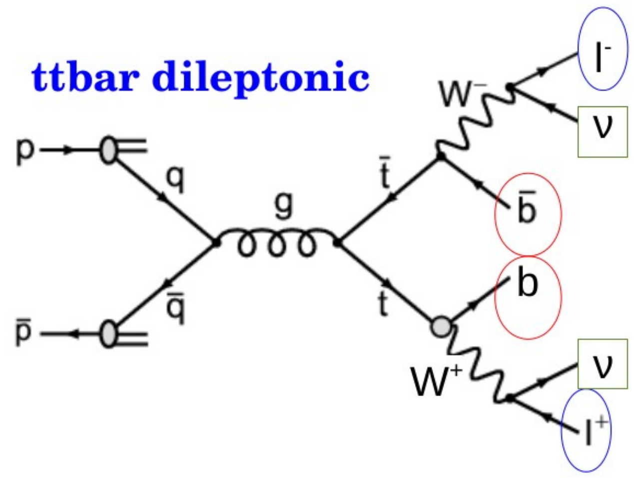
✗ **ME at LO** only (assignment between reconstructed jets and partons can be ambiguous beyond LO)

[1] P. Artoisenet, V. Lemaître, F. Maltoni, OM: JHEP 1012:068

Z+bb analysis: automatised MEM



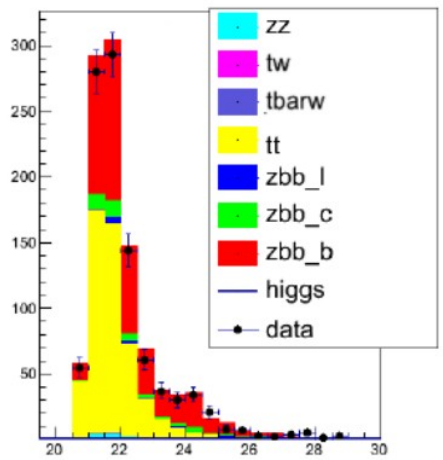
VS



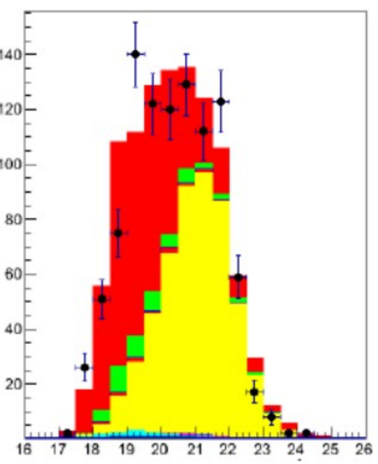
- MEM used as cross check for standard analysis to discriminate Z+bb and ttbar dileptonic processes
  - very similar processes, same objects in final state
    - After cut on M(ll) and E_{miss}^{T}
  - Use discrimination power to estimate the ttbar fraction in the data
- Use MadWeight to compute probability that an event α matches Z+bb/ttbar hypotheses
  - fully automatized procedure
    - All processes can be calculated in principle
  - transfer function for electrons, muons, bjets
  - correction for ISR jets

Z+bb analysis: automatised MEM

MEM output

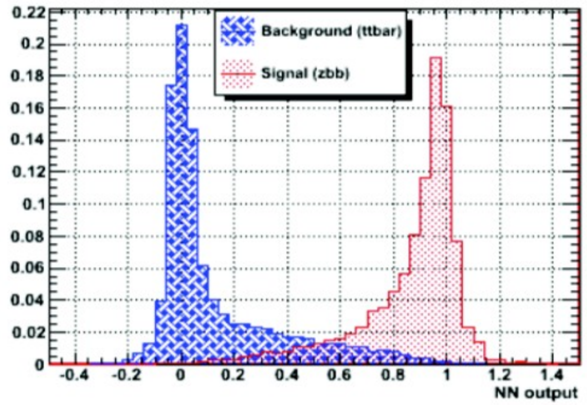


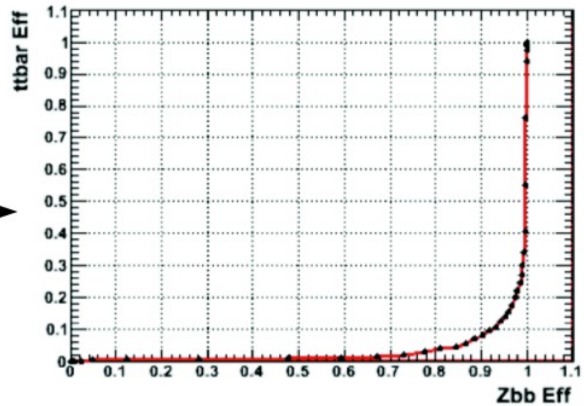
ttbar hypothesis



Zbb hypothesis

MVA output





→ Can achieve excellent separation