

# Matrix Element Techniques in the CMS Higgs->4l studies

#### Guenakh Mitselmakher University of Florida

Kruger-2014

# **Matrix Element Method (MEM)**

Definition (from the talk by F. Canelli, Zurich, May 2013)

 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

## History (very incomplete)

• MEM developed around 2000 to measure the top quark mass and W helicity in top events. Later adapted to searches

- Observation of single top, Higgs searches and studies

 Most extensively used for the H→4l golden channel" searches and properties measurements

This talk will discuss some applications of the Matrix Element Method (MEM) to the H→4I "golden" channel studies by CMS

## July 4, 2012. A new boson discovery announced



# July 4, 2012. CMS results



Analysis of 5 main expected Higgs decays channels presented. Experimental observation of a new boson with the mass ~ 125 Gev, consistent with the SM Higgs

 $H \rightarrow ZZ \rightarrow 4l$  is expected to be the most sensitive channel in broad range of masses

Significance by decay channel:

Decay mode	Expected	Observed
ZZ	3.8σ	$3.2\sigma$
$\gamma\gamma$	$2.8\sigma$	$4.1\sigma$
WW	$2.5\sigma$	$1.6\sigma$
bb	$1.9\sigma$	$0.7\sigma$
au au	$1.4\sigma$	—

An over 3 sigma evidence at the mass of ~ 125 Gev in this channel alone, despite very small statistics

Matrix Element Technique used to enhance sensitivity in this channel, provides better signal/background separation

## 2010-2012 (Run 1): LHC integrated luminosity

#### CMS Integrated Luminosity, pp





LHC status & prospects Frédérick Bordry Large Hadron Collider Physics (LHCP) conference – New York - 2nd June 2014

# LHC run 1 (2010-2012)

- Results discussed in this talk are based on the statistics accumulated by CMS in Run 1
- Only the 2011 +2012 data (most of the Run 1 statistics) used
   2011: 7Tev c.m. LHC energy, ~ 5 fb<sup>-1</sup>
   2012: 8Tev LHC c.m. energy, ~20 fb<sup>-1</sup>

# **CMS:** The Compact Muon Solenoid



## Higgs production and decay modes

Production at ~125 Gev

**Decay modes and branching ratios** 



- Intermediate/high mass:  $H \rightarrow WW$ ,  $H \rightarrow ZZ$ 

## **Higgs particles produced in ATLAS + CMS (estimate)**

- Number of SM Higgs particles produced in ATLAS and CMS in 2011-2012 = 1,100,000 (total Cross Section 22 pb) x (25 fb-1) x (2experiments)
- Contribution of different production mechanisms (wrt the total Cross Section)
- ggF = 87%
- VBF = 7%
- VH = 5%
- ttH = 0.6%
- Decay modes (I = e or mu)
- BR(bb) = 57%
- BR(tautau) = 6%
- BR(WW->2l2v) = 22% x (0.22)^2 = 1.1%
- BR(gamgam) = 0.23%
- BR(ZZ->4I) = 2.8% x (0.06)^2 = 0.013%
- BR(mumu) = 0.022%

#### • Discussion in this talk is limited to the rare, but most convenient $ZZ \rightarrow 4I$ mode

# $H \rightarrow ZZ \rightarrow 4I$ "golden" decay channel

#### used for: discovery, mass, width, spin/parity, coupling to bosons...



#### Analysis strategy:

- Four leptons (muons or electrons) final state
- Reach and fully measured decay kinematics
- four-lepton mass is the key observable
- split events into 4e, 4μ, 2e2μ channels
- Backgrounds:
  - ZZ (dominant): from Monte Carlo (MC)
  - reducible (with non-isolated or "fake" leptons): from control region

Exploit differences in differential distributions (Matrix Element Method, or MEM) for signal/ bkgd separation, and use to improve significance, and for the spin-parity and other properties measurements

#### Analysis features:

- high S/B-ratio (~2:1)
- but very small event yield
- excellent mass resolution = 1-2%"
- Z→4l decay peak conveniently nearby,
   natural validation of the discovered peak

# Higgs $\rightarrow$ ZZ $\rightarrow$ 4I: signal and backgrounds



Kruger2014, December 1, 2014

# $H \rightarrow ZZ \rightarrow 4I$ , backgrounds

Irreducible backgrounds:

- $q\bar{q} \rightarrow ZZ \rightarrow 4\ell$  is the dominant background:
  - Electroweak process known to NLO.
  - Estimated from MC (5% uncertainty).

•  $gg \rightarrow ZZ \rightarrow 4\ell$ :

- Appears at NNLO only.
- ~10% contribution for  $m_{4\ell} > 180$  GeV.
- Negligible at  $m_{4\ell} \approx 125$  GeV.
- Estimated from MC (40% uncertainty).

The reducible background: Much smaller contribution than reducible bkgd

- "Z"+X:
  - Consists of  $Zb\bar{b}$ ,  $t\bar{t}$ , Z + jets, etc. Has nonprompt leptons.
  - MC cannot be trusted.
  - Estimated from data (large uncertainty, small contribution).

# Expected and observed statistics (Run 1) around Higgs peak



Source         Count           ZZ $6.8$ Z + X $2.5$ $m_{\rm H}$ = 125.6         18.4           Observed         25	$121.5 < m_{4\ell} < 130.5 \text{ GeV}$ :		
ZZ $6.8$ Z + X $2.5$ $m_{\rm H} = 125.6$ $18.4$ Observed $25$	Source	Count	
$Z + X$ 2.5 $m_{\rm H} = 125.6$ 18.4         Observed       25	ZZ	6.8	
$m_{\rm H} = 125.6$ 18.4 Observed 25	Z + X	2.5	
Observed 25	$m_{\rm H} = 125.6$	18.4	
	Observed	25	

Very small statistics. Can do better than just counting events? The answer is yes: by using kinematical information and Matrix Element

# Signal and background ME calculations

- The Matrix Element based KinematicalDiscriminant allows to built an additional (to the trigger and selection cuts) filter, suppressing further a calculable dominant background
- At the discovery time the analytically calculated MELA (\*) Kinematical Discriminant used: no interference between leptons in the 4e and 4mu final state has been included yet
- CMS now uses generator-based signal/bckgd ME calculations, with interference fully included in the calculations:
  - generator based MELA: JHUGen + MCFM based package.
  - MEKD(\*\*): MadGraph+FeynRules based package, used for cross-checks and validation.

(\*) in CMS, MELA refers to Matrix Element Likelihood Approach (\*\*) MEKD refers to Matrix Element Kinematical Discriminant software package

# **Matrix Element Based Discriminant**

The signal  $gg \to H \to ZZ \to 4\ell$  and the background  $q\bar{q} \to ZZ \to 4\ell$  are

- Well known processes from the theory.
- Have multi-body (four leptons) final state.
- All final state objects are well reconstructed.

By Neyman–Pearson lemma, the optimal discriminant between two hypotheses (H  $\rightarrow$  4 $\ell$  and ZZ  $\rightarrow$  4 $\ell$  here) is

$$d = \frac{pdf(\text{event}|\text{H})}{pdf(\text{event}|\text{ZZ})} \sim \frac{|\mathcal{M}(p_1, p_2, p_3, p_4|\text{H})|^2}{|\mathcal{M}(p_1, p_2, p_3, p_4|\text{ZZ})|^2},$$

#### which is the idea of the Matrix Element Method (MEM).

CMS uses Kinematical Discriminant, taking into account "calculable" Irreducible background. It is close to optimal signal/background filtering, exploiting kinematics differences between signal and background

# H--> 4I observables and MEM

Kinematics of the  $4\ell$  system at LO can be fully described by  $m_{4\ell}$ ,  $m_1$ ,  $m_2$ , and  $\vec{\Omega} = (\theta_1, \theta_2, \Phi, \Phi_1, \theta^*)$ : a way to interpretate a set of 4-momenta ( $p = p_1 + p_2 + p_3 + p_4$ ).

- In production-independent case: Φ<sub>1</sub> and θ\* dependecy is removed.
- Matrix-element-method-based (MEM) approach compresses the information into unnormalized probability density functions ( $\mathcal{P}^{kin} = |\mathcal{M}|^2$ ):

g(q)

Φ

$$\mathcal{P}_{SM} = \mathcal{P}_{SM}^{kin}(\vec{\Omega}, m_1, m_2 | m_{4\ell}) \times \mathcal{P}_{sig}^{mass}(m_{4\ell} | m_H)$$

$$\mathcal{P}_{\mathbf{J}^{\mathbf{P}}} = \mathcal{P}_{\mathbf{J}^{\mathbf{P}}}^{\mathrm{kin}}(\vec{\Omega}, m_1, m_2 | m_{4\ell}) \times \mathcal{P}_{\mathrm{sig}}^{\mathrm{mass}}(m_{4\ell} | m_H)$$

$$\mathcal{P}_{q\bar{q}ZZ} = \mathcal{P}_{q\bar{q}ZZ}^{kin}(\vec{\Omega}, m_1, m_2 | m_{4\ell}) \times \mathcal{P}_{q\bar{q}ZZ}^{mass}(m_{4\ell})$$

# It was shown by the MEKD team that the same-flavor interference is crucial for distringuishing some exotic states.

## Kimematical Discriminants, definitions in this study

- The simplest discriminant would be  $\mathcal{D}_{simple} = \frac{\mathcal{P}_A}{\mathcal{P}_B}$  to separate A from B.
- Discriminants here are defined as follows:

$$\mathcal{D}_{bkg} = \frac{\mathcal{P}_{SM}}{\mathcal{P}_{SM} + c \times \mathcal{P}_{bkg}} = \left[ 1 + c(m_{4\ell}) \times \frac{\mathcal{P}_{bkg}^{kin}(m_1, m_2, \vec{\Omega} | m_{4\ell}) \times \mathcal{P}_{bkg}^{mass}(m_{4\ell})}{\mathcal{P}_{SM}^{kin}(m_1, m_2, \vec{\Omega} | m_{4\ell}) \times \mathcal{P}_{sig}^{mass}(m_{4\ell} | m_H)} \right]^{-1},$$
  
$$\mathcal{D}_{JP}^{kin} = \frac{\mathcal{P}_{SM}^{kin}}{\mathcal{P}_{SM}^{kin} + c_{JP} \times \mathcal{P}_{JP}^{kin}} = \left[ 1 + c_{JP} \times \frac{\mathcal{P}_{JP}^{kin}(m_1, m_2, \vec{\Omega} | m_{4\ell})}{\mathcal{P}_{SM}^{kin}(m_1, m_2, \vec{\Omega} | m_{4\ell})} \right]^{-1}.$$

- For convenience we compress discriminants between 0 and 1 and use c to avoid overcompression.
- Production-independent discriminants D<sup>dec</sup> are constructed identically out of P<sup>dec</sup>, which have production information removed.

#### Discriminants dramatically improve the sensitivity!

#### How does KD work for signal/bkgd separation for SM Higgs



Main reason why kinematics information helps in  $H\rightarrow 4l$ :

For low-mass Higgs, most of the information in  $m_{Z2}$ : mass of the "virtual" Z

- $H \rightarrow ZZ^* \rightarrow 4\ell$
- $q\bar{q} \rightarrow Z\gamma^* \rightarrow 4\ell$

Use of the full event kinematics in MEM allows for even better signal/bkgd separation



# **Templates and Likelihoods**

Each observed event has a different likelihood according to the templates:

$$\mathcal{L} = \exp\left(-n_{\rm sig}(\vec{\zeta}) - n_{\rm bkg}\right) \prod_{i}^{N} \left(n_{\rm sig} \times \mathcal{P}_{\rm sig}(\vec{x}_{i}; \vec{\zeta}) + n_{\rm bkg} \times \mathcal{P}_{\rm bkg}(\vec{x}_{i})\right).$$

where  $\vec{\zeta}$  can be thought as a model and  $\vec{x}_i$  as an observed event.



- The important part of statistical analysis in our tests are templates.
- Templates serve as probabilistic pmfs:
  - In MEM studies, KDs serve as basis.
  - Templates are filled using MC events:

## Combining information in 4l state from mass and KD for Higgs signal extraction

Invariant mass  $(m_{4\ell})$  + discriminant  $(\mathcal{D}_{bkg}^{kin}) \Rightarrow 2D$ .



Data events follow Higgs+background expected distribution.

## H→4l signal significance using full available statistics and ME based Kinematical Discriminant



- The excess near 126 GeV is of  $7\sigma$ .
- The SM Higgs is excluded everywhere else up to 800 GeV.
- ME discriminant brings a 20% improvement in expected significance.
- Signal strength  $\mu = \frac{\sigma}{\sigma_{SM}} = 0.93^{+0.29}_{-0.25}$ , compatible with the SM Higgs.

#### Higgs mass and (direct) width measurements using KD



# Spin-parity studies of new resonance using Kinematical Discriminants

What are we testing?

- Testing the SM Higgs vs. alternative spin-parity (spin-0, spin-1, and spin-2) states.
- Testing a spin-0 particle with anomalous couplings.
- Testing for two nearly mass-degenerate states for all spin-0, spin-1, and spin-2 models with SM Higgs.

# Spin-parity studies (J<sup>p</sup>) using Kinematics in 4l decays



• Done in the past:  $\pi^0 \to \gamma^* \gamma^* \to 4e$ 

 Two different hypotheses can be separated using Neyman–Pearson lemma:

$$d = \frac{|\mathcal{M}(p_1, p_2, p_3, p_4 | J^P)|^2}{|\mathcal{M}(p_1, p_2, p_3, p_4 | \mathbf{H})|^2}.$$

⇐ SM Higgs versus pseudoscalar.

# **Amplitude structure**

• Spin 0:  

$$A_{J=0} = v^{-1} \left( \left[ a_{1} - e^{i\phi_{\Lambda_{1}}} \frac{q_{1}^{2} + q_{2}^{2}}{(\Lambda_{1})^{2}} \right] m_{v}^{2} \epsilon_{1}^{*} \epsilon_{2}^{*} + a_{2} f_{\mu\nu}^{*(Z)} f^{*(Z),\mu\nu} + a_{3} f_{\mu\nu}^{*(Z)} \tilde{f}^{*(Z),\mu\nu} \right)$$

$$(a_{1} = 2: \text{SM Higgs}) \quad 0_{m}^{+} \quad 0_{h}^{+} \quad 0^{-}$$
• Spin 1:  

$$A_{J=1} = b_{1} \left[ (\epsilon_{1}^{*}q) (\epsilon_{2}^{*} \epsilon_{X}) + (\epsilon_{2}^{*}q) (\epsilon_{1}^{*} \epsilon_{X}) \right] + b_{2} \epsilon_{\alpha\mu\nu\beta} \epsilon_{X}^{\alpha} \epsilon_{1}^{*\mu} \epsilon_{2}^{*\nu} \tilde{q}^{\beta}.$$

$$1^{-} \qquad 1^{+}$$

• Spin 2 has 10 terms and the following models:

$$2_{\rm m}^+$$
,  $2_{\rm h2}^+$ ,  $2_{\rm h3}^+$ ,  $2_{\rm h}^+$ ,  $2_{\rm b}^+$ ,  $2_{\rm h6}^+$ ,  $2_{\rm h7}^+$ ,  $2_{\rm h}^-$ ,  $2_{\rm h9}^-$ , and  $2_{\rm h10}^-$ .

Guenakh Mitselmakher Kr

# Scalar vs Pseudoscalar



• To perform a hypothesis testing we use the following test statistic (TS):

$$q = -2 \times \ln \frac{\mathcal{L}(\text{data}|J^P)}{\mathcal{L}(\text{data}|0_m^+)}$$

one q for all the events in a dataset.

 We perform pseudoexperiments to fill expected TS distributions for the null (SM Higgs) and the alternative (J<sup>P</sup>) hypotheses.

Data is consistent with orange (SM Higgs) and very unlikely for blue  $(0^{-})$ .

Pseudoscalar is excluded at 99.9% CL.

## **Alternative spin-parity states testing summary**



All the models are excluded at 95% CL or better.

Data is consistent with the SM Higgs.

# Validation of the $J^{P}$ studies with $Z \rightarrow 4I$ (1)



- $Z \rightarrow 4\ell_{2}$
- The null hypothesis:  $Z \rightarrow 4\ell$ .
- Alternatives hypotheses:

• 
$$gg \rightarrow H \rightarrow ZZ \rightarrow 4\ell$$
.

• 
$$gg \to 0^+_Z \to 4\ell$$
.

$$\circ gg \to 1_Z^+ \to 4\ell.$$

• 
$$q\bar{q} \rightarrow 2_Z^+ \rightarrow 4\ell$$
.

• 
$$m_{\rm X} = m_{\rm Z}, \, \Gamma_{\rm X} = \Gamma_{\rm Z}.$$

The SM Z is a mixture of an axial and a vector parts.

Guenakh Mitselmakher

Kruger2014, December 1, 2014

# Validation of the $J^{P}$ studies with $Z \rightarrow 4I$

The "null" hypothesis: SM Z $\rightarrow$ 4l, a "mixture" of an axial and a vector The alternative hypothesis:  $gg \rightarrow H \rightarrow ZZ \rightarrow 4\ell$ 



# Search for mass-degenerate states with different J<sup>P</sup>

Search for **two** nearly mass-degenerate noninterfering states under the observed 125-GeV peak:  $H + J^P$ .

Experimental examples:

- Positronium (0<sup>-</sup>, 1<sup>-</sup>):  $\frac{\Delta m}{m} \sim O(10^{-10})$ .
- $\chi_b (0^{++}, 1^{++}, 2^{++})$ :  $\frac{\Delta m}{m} \sim O(10^{-3})$ .

We probe fractions by fitting likelihoods and evaluating the

$$f(J^{P}) = \frac{\sigma_{J^{P}}}{\sigma_{H} + \sigma_{J^{P}}},$$
  
assuming  $\Gamma_{J^{P}}$  and  $\Gamma_{H} << \Delta m_{4l} << \text{resolution} \sim 1 \text{ GeV}.$ 

# **Degenerate states testing (summary)**



- We cannot exclude smaller than 50% fraction (limited statistics).
- Most of the tests prefer pure SM Higgs hypothesis.

## Indirect (off-shell) Higgs width measurement (1)

The Breit–Wigner part in a differential cross section:

$$\frac{d\sigma_{gg \rightarrow H \rightarrow ZZ}}{dm_{ZZ}^{2}} \sim g_{ggH}^{2} g_{HZZ}^{2} \frac{F(m_{ZZ})}{(m_{ZZ}^{2} - m_{H}^{2})^{2} + m_{H}^{2} \Gamma_{H}^{2}} :$$

$$\sigma_{gg \rightarrow H \rightarrow ZZ}^{\text{on-shell}} \sim \frac{g_{ggH}^{2} g_{HZZ}^{2}}{\Gamma_{H}} \\ \sigma_{gg \rightarrow H \rightarrow ZZ}^{\text{off-shell}} \sim g_{ggH}^{2} g_{HZZ}^{2} \end{pmatrix} \frac{\sigma_{gg \rightarrow H \rightarrow ZZ}^{\text{off-shell}}}{\sigma_{gg \rightarrow H \rightarrow ZZ}^{\text{on-shell}}} \sim \Gamma_{H}.$$
For the SM H, ratio is 10%. Great for experimental  $\Gamma_{H}$  determination.

In the off-shell region,  $gg \rightarrow ZZ$  is comparable to  $gg \rightarrow H$ , thus needs to be considered (interference is strong and negative!).

## Indirect (off-shell) Higgs width measurement (2)



#### Analysis strategy:

- Large  $\Gamma_{\rm H} \Rightarrow$  more events at large  $m_{\rm ZZ}$ .
- Using ME discriminant (gg → H/box → ZZ vs. qq̄ → ZZ) to improve sensitivity.
- Add  $H \rightarrow ZZ \rightarrow 2\ell 2\nu$  to improve sensitivity.

Result:  $\Gamma_{\rm H} < 5.4 \ \Gamma_{\rm H}^{\rm SM}$ , i.e.,  $\Gamma_{\rm H} < 22 \ {\rm MeV}.^{ab}$ 

 ${}^{a}\Gamma_{\mathrm{H}}^{\mathrm{SM}} = 4.15 \text{ MeV.}$  ${}^{b}\Gamma_{\mathrm{H}}^{\mathrm{direct}} < 3.4 \text{ GeV:} > 100 \times \text{ worse.}$ 

# Summary

- Matrix element techniques has been extensively used in Higgs discovery and properties studies, particularly for studies of the H→ZZ→4l decays, where the event kinematics is reach, fully measured with good precision and both the sygnal and the background is largely calculable
- Matrix element techniques in many cases allowed for significant improvements in precision of the studies
- Still plenty of room for deviations from SM Higgs (within errors) exist, providing many opportunities for continuing use of the Matrix Element techniques for studies of the Higgs decay channels. We are at the beginning of the program of precision measurements of the H125 GeV particle with a hope to find BSM physics via deviations from the SM predictions.

# **Backup slides**

# Spin-0 Amplitude measurements (CMS)

The spin-0 amplitude is restricted to

$$A_{J=0} = \nu^{-1} \left( \left[ a_1 - e^{i\phi_{\Lambda_1}} \frac{q_1^2 + q_2^2}{(\Lambda_1)^2} \right] m_{\nu}^2 \epsilon_1^* \epsilon_2^* + a_2 f_{\mu\nu}^{*(Z)} f^{*(Z),\mu\nu} + a_3 f_{\mu\nu}^{*(Z)} \tilde{f}^{*(Z),\mu\nu} \right) \right)$$

Allowed $(\Lambda_1 \sqrt{|a_1|}) \cos(\phi_{\Lambda_1})$  $a_2/a_1$  $a_3/a_1$ Observed $[-\infty, -119 \text{ GeV}] \cup [104 \text{ GeV}, \infty]$  $[-2.28, -1.88] \cup [-0.69, \infty]$ [-2.05, 2.19]Expected $[-\infty, -50 \text{ GeV}] \cup [116 \text{ GeV}, \infty]$  $[-0.77, \infty]$ [-3.85, 3.85]

- No significant inconsistencies at 95% CL.
- Results are statistically limited.
- Non-*a*<sub>1</sub> terms appear in the SM via loops—minute.

# Spin-0 amplitude expected SM coefficients



## Spin-parity tests adding H->WW and $H\rightarrow\gamma\gamma$

#### H→ZZ→4I

- 4l system is fully reconstructed
- use leptons momenta to construct discriminants







ATLAS: MVA-based discriminant

#### H→WW→lvlv

dilepton angle is sensitive to spin of the original H-boson



#### Η→γγ

- J=1 forbidden (Landau-Yang theorem)
- $-\cos\theta^*$  is the only variable sensitive to  $J^P$  information at leading order



- Left plot before acceptance and reconstruction.
- after acceptance x reconst., discrim. power lessens
- poor S:B makes the measurement very difficult

# Spin-parity results (ATLAS)

- Use MVA-based discriminants, find sensitive observables
- Test several alternative spin-parity hypotheses J<sup>P</sup> (0<sup>-</sup>,1<sup>+</sup>,1<sup>-</sup>,2<sup>+</sup>) compared to SM hypothesis: 0<sup>+</sup>
- Production modes
  - spin-2 : test production mechanism via combination of ggF & qqbar annihilation
  - spin-1 :signal produced via qqbar annihilation (ggF forbidden)
  - spin-0 : ggF (qqbar annihilation negligible)

0<sup>+</sup> vs 0<sup>-</sup> (only ZZ): 97.8% CL 0<sup>+</sup> vs 1<sup>+</sup> (ZZ +WW): 99.97% CL 0<sup>+</sup> vs 1<sup>-</sup> (ZZ+WW): 99.7% CL 0<sup>+</sup> vs 2<sup>+</sup> (γ γ +ZZ+WW)>99.9% CL

All tested alternative spin hypotheses excluded at > 97.8% CL (June 2014)





# Higgs couplings scale ~ mass

• Scale SM couplings by measured scale factors, and plot modified couplings vs particle masses:



The magnitude of couplings range by several orders of magnitude!  $H \rightarrow \mu\mu$  is not on the left plot, but is much lower if we use existing upper limits (CMS and ATLAS), an evidence for the non-universality of H interactions with  $\tau$  and  $\mu$  (3<sup>rd</sup> and 2<sup>nd</sup> generations). Will be much improved as HL LHC (right plot)