

Searches of the Higgs boson in the diphoton decay mode at CMS

Tommaso Tabarelli de Fatis Università and INFN - Milano Bicocca

2014, Dec 5th, LAL Orsay

BICOCCA

Talk's overview

[EPJ C (2014) 74:3076]

- 2012, Jul Higgs boson discovery by ATLAS and CMS
- 2013, Feb Preliminary CMS results on the full dataset (Moriond)
 - 5.1 fb⁻¹ at √s=7 TeV and 19.7 fb⁻¹ at √s=8 TeV
- 2014, Jul improved analysis of the full dataset:
 - Optimized performance of energy reconstruction
 - New method for modelling the background
 - Sensitivity +30%
 - Accurate modelling and study of the photon response
 - Systematic uncertainty on mass reduced by 1/3
 - Selection optimization including tags for exclusive production modes
 - Improved sensitivity to specific couplings
 - With details from photon reconstruction paper being submitted



Higgs production and $H \rightarrow \gamma \gamma$ decay



3

The CMS ECAL

- ▶ Homogeneous, compact, hermetic, fine grain PbWO₄ crystal calorimeter
 - → Emphasis on energy resolution
 - → No longitudinal segmentation (except preshower)



Measurement strategy

- Photon energy:
 - Clustering independent of particle type
 - Calibration and corrections
 - Estimate per-photon resolution

2. Photon identification

- Shower shapes and isolation variables
 - Estimate per-photon quality

Signal model from simulation:

- Same corrections on e[±] and γ
- Z → ee events to model detector response (i.e. tune MC simulation): photon energy scale, resolution, and efficiencies

3. Vertex identification:

- Select vertex among ~20 concurrent collisions (*pileup*)
- Per-event probability of correct vertex assignment

Event selection and analysis categories:

- All event information in one BDT discriminating variable
 - **Event kinematics, resolution, photon quality**
- Dijet BDT + Exclusive mode tags (VBF, VH, ttH)

 Global fit: signal model + parametric background

+ systematic uncertainties

BDT = Boosted Decision Tree

Energy reconstruction at CMS

$$E_{e,\gamma} = \mathcal{G} \mathcal{F}_{e,\gamma} \sum_{i} c_i s_i(t) \mathcal{A}_i$$

Response uniformity

- Crystal light yield (LY) spread ~ 10%
- Endcap VPT response spread ~ 25%
 - Intercalibration

Response stability

- LY variation with temperature: -2.2%/°C
- Gain variation (Barrel APDs):
 - $\Delta G/\Delta T = -2.4\%/°C; \quad \Delta G/\Delta V = 3.1\%/V$
- Transparency variation with radiation
 - Stabilization and response corrections

Geometry, tracker material and B field

- Photon conversions
- Energy spread along ϕ at \approx constant η
 - Clustering and energy corrections (+ global scale G)





(2)

(3)

Single crystal calibrations

- Intercalibration and transparency corrections
 - One transparency measurement (laser light) each crystal every 40 min
 - Monthly intercalibration: $\pi^0/\eta \rightarrow \gamma\gamma$ mass, ϕ -invariance of energy flow
 - Once per year: electron E/p and $Z \rightarrow$ ee mass
- By construction $\langle c_i \rangle = 1$, $s_i(t=0) = 1$ Intercalibration and corrections affect resolution, not the energy scale! $\times 10^{3}$ **Relative E/p scale CMS Preliminary 2012** 1.02 Ge/ ECAL barre **CMS Preliminary 2012** no corrections √s = 8 TeV L = 19.6 fb 180 1.01 Events / 0.5 √s = 8TeV, L = 19.6 fb⁻¹ 160 Intercalibrations (IC) 140 IC + LM corrections 0.99 RMS < 0.1% ECAL barrel 120 0.98 100 0.97 **80** 0.96 60 0.95 σ_{ci}~0.5% **40**⊢ 0.94 without laser monitoring correction 20 0.93 with laser monitoring correction 20/04 20/05 19/06 18/08 17/09 19/07 17/1016/1160 80 100 120 date (day/month) M_{ee} (GeV/c²)



Clustering and energy corrections



- Dynamic clustering to recover energy radiated upstream of ECAL
 - Super-clusters of clusters along \u03c6 (bending direction)
 - Ratio of energy in fixed array to super-cluster energy convenient to classify e/γ with radiation/conversions (e.g. R9=E_{3x3}/E_{sc})

Energy corrections (F_{e/γ}) based on MC simulation

 Global (material distribution) and local (voids between crystals) coordinates, shower shape variables (radiation effects), pileup variables, ...

Photon energy corrections

Improved input:

- Time-dependent simulation (8 TeV only) → improved shower shapes
 - *Pileup* in extended time window, equivalent noise and *pileup* evolution

Improved BDT algorithm:

- Predict parametric E_{true}/E_{raw} p.d.f. (Gauss+ Crystal-Ball wings)
- Corrected energy from the mode
- Per-event resolution "from the p.d.f. spread" >> input to diphoton BDT



Test resolution with $Z \rightarrow e^+e^-$



Photostatistics and constant term according to test beam results

Example of imperfect simulation



- Data/MC Tracker material thickness in x₀ from:
 - Bremsstrahlung of electrons $f = (p_{vtx}-p_{out})/p_{vtx}$
 - Multiple scattering of low momentum pions
 - Differences in TIB/TOB support structure (InI~0.5) and services in the TK barrel/endcap transition



More examples...

Conversions in the tracker volume:



Energy flow in the preshower in $Z \rightarrow ee$ events:







Cables and connectors at the back of the Tracker (not-'mapped' by-tracks)--

Tuning of calibration and MC resolution

Multi-step procedure:

- Correct scale for residual time dependence of response in Run × IŋI bins
 - Small: <0.1 (0.2) % in EB (EE)
- Simultaneous fit of scale in data and additional Gauss smearing in MC in (4 Inl × 2 R9) bins
 - In barrel at 8 TeV: $\sigma_{Gauss} = a/\sqrt{E_T} (+) c$
- Residual correction scale correction in E_T × InI × R9 bins
 - Only at 8 TeV in the barrel

There is a systematic uncertainty associated

NOTE: Consistent non-linearity seen in 7 TeV data, but dataset insufficient to derive corrections





Photon energy scale check: $Z \rightarrow \mu \mu \gamma$



- Inclusive Data/MC calibration agree to about 1 o
 - $\Delta E = (0.25 \pm 0.11_{stat} \pm 0.17_{syst}) \% = (0.25 \pm 0.20) \%$
 - Systematic uncertainty include fit reproducibility, and selections
- No strong constraint on the Higgs boson mass uncertainty
 - Precision in individual R9, $|\eta|$ bins ~0.3%, $<p_T> = 28$ GeV

Effect of smearing added to MC



ĮNFŃ

INFN ALISSIANNU DI MILANO Photon resolution after MC tuning



R

Photon identification

Loose (trigger) pre-selection + identification BDT



Uncertainty on the photon ID



BDT score of electrons in Z→ee events

- Reconstructed as photons
- Systematic uncertainty applied to cover any data/MC discrepancies:
 - BDT score shift by ±0.01

BDT score of photons in $Z \rightarrow \mu \mu \gamma$ events

Systematic uncertainty band as from Z→ee study

Vertex identification

- Resolution unaffected if vertex within 10 mm of true position
- 1. BDT to identify vertex
 - Hardness of interactions p_T
 - Balance of diphoton system and charged tracks
 - Conversion information



2. BDT to assign per-event probability of correct vertex id

- BDT score and distance of three most likely vertices
- Number of vertices in the event
- Number of conversion tracks



>> Input to diphoton BDT



Event selections and analysis categories



Transformed combined BDT classifier score

INFN MILSHEI STUDI Exclusive tag mode summary

Mutually exclusive classes

Acceptance × efficiency

	No. of	classes		48.6%			
Label	7 TeV	8 TeV	Main requirements \rightarrow 8 TeV at m _H = 125 GeV: 4	49.3%			
tīH lepton tag	*	1	$p_{\rm T}^{\gamma 1} > m_{\gamma \gamma}/2$ 1 b-tagged jet + 1 electron or muon				
VH tight ℓ tag	1	1	$p_{\rm T}^{\gamma 1} > 3m_{\gamma\gamma}/8$ [e or μ , $p_{\rm T} > 20$ GeV, and $E_{\rm T}^{\rm miss} > 45$ GeV] or [2e or 2μ , $p_{\rm T}^{\ell} > 10$ GeV; $70 < m_{\ell\ell} < 110$ GeV]				
VH loose ℓ tag	1	1	$p_{\rm T}^{\gamma 1} > 3m_{\gamma \gamma}/8$ e or μ , $p_{\rm T} > 20$ GeV				
VBF dijet tag 0-2	2	3	$p_{\rm T}^{\gamma 1} > m_{\gamma \gamma}/2$ 2 jets; classified using combined diphoton-dijet BDT				
VH $E_{\rm T}^{\rm miss}$ tag	1	1	$p_{\rm T}^{\gamma 1} > 3m_{\gamma \gamma}/8$ $E_{\rm T}^{\rm miss} > 70 {\rm GeV}$				
tīH multijet tag	*	1	$p_{\rm T}^{\gamma 1} > m_{\gamma \gamma}/2$ 1 b-tagged jet + 4 more jets				
VH dijet tag	1	1	$p_{\rm T}^{\gamma 1} > m_{\gamma \gamma}/2$ jet pair, $p_{\rm T}^{\rm j} > 40 \text{GeV}$ and $60 < m_{\rm ii} < 120 \text{GeV}$				
Untagged 0-4	4	5	The remaining events, classified using diphoton BDT				
	LabeltīH lepton tagVH tight ℓ tagVH loose ℓ tagVH loose ℓ tagVBF dijet tag 0-2VH $E_{\rm T}^{\rm miss}$ tagtīH multijet tagVH dijet tagUntagged 0-4	LabelNo. of 7 TeVtīH lepton tag \star VH tight ℓ tag1VH tight ℓ tag1VH loose ℓ tag1VBF dijet tag 0-22VH E_T^{miss} tag1tīH multijet tag \star VH dijet tag1Untagged 0-44	LabelNo. of classes 7 TeV8 TeVtĒH lepton tag \star 1VH tight ℓ tag11VH loose ℓ tag11VBF dijet tag 0-223VH E_T^{miss} tag11tĒH multijet tag \star 1VH dijet tag11Untagged 0-445	$\begin{array}{c c c c c c c c c c c c c c c c c c c $			

* For the 7 lev dataset, events in the ttH lepton tag and multijet tag classes are selected first, and combined to form a single event class.

against tagged

Events tested



- Parametric shape (sum of Gaussians)
 - Fit diphoton invariant mass in MC simulation
 - Nine m_H values 110-150 GeV; linear interpolation in between
 - Separate fits for each each category



About 20% improvement in resolution compared to Moriond 2013

Expected signal (SM) breakdown

		Expected SM Higgs boson signal yield ($m_{\rm H}$ =125 GeV)							Bkg.		
E	vent classes	Total	ggH	VBF	WH	ZH	tīH	$\sigma_{\rm eff}$ (GeV)	(GeV)	(GeV ⁻¹)	
	Untagged 0	5.8	79.8%	9.9%	6.0%	3.5%	0.8%	1.11	0.98	11.0	
	Untagged 1	22.7	91.9%	4.2%	2.4%	1.3%	0.2%	1.27	1.09	69.5	
	Untagged 2	27.1	91.9%	4.1%	2.4%	1.4%	0.2%	1.78	1.40	135.	
1	Untagged 3	34.1	92.1%	4.0%	2.4%	1.3%	0.2%	2.36	2.01	312.	
14	VBF dijet 0	1.6	19.3%	80.1%	0.3%	0.2%	0.1%	1.41	1.17	0.5	
2	VBF dijet 1	3.0	38.1%	59.5%	1.2%	0.7%	0.4%	1.65	1.32	3.5	
lev	VH tight ℓ	0.3	_		77.2%	20.6%	2.2%	1.61	1.31	0.1	
2	VH loose ℓ	0.2	3.6%	1.1%	79.1%	15.2%	1.0%	1.63	1.32	0.2	
	VH E _T ^{miss}	0.3	4.5%	1.1%	41.5%	44.6%	8.2%	1.60	1.14	0.2	Co
	VH dijet	0.4	27.1%	2.8%	43.7%	24.3%	2.1%	1.54	1.24	0.5	unt
	tīH tags	0.2	3.1%	1.1%	2.2%	1.3%	92.3%	1.40	1.13	0.2	
	Untagged 0	6.0	75.7%	11.9%	6.9%	3.6%	1.9%	1.05	0.79	4.7	- Bo
	Untagged 1	50.8	85.2%	7.9%	4.0%	2.4%	0.6%	1.19	1.00	120.	- Bo
	Untagged 2	117.	91.1%	4.7%	2.5%	1.4%	0.3%	1.46	1.15	418.	- Bo
	Untagged 3	153.	91.6%	4.4%	2.4%	1.4%	0.3%	2.04	1.56	870.	-
7	Untagged 4	121.	93.1%	3.6%	2.0%	1.1%	0.2%	2.62	2.14	1400.	
ਿੰਦ	VBF dijet 0	4.5	17.8%	81.8%	0.2%	0.1%	0.1%	1.30	0.94	0.8	
9.7	VBF dijet 1	5.6	28.5%	70.5%	0.6%	0.2%	0.2%	1.43	1.07	2.7	
41	VBF dijet 2	13.7	43.8%	53.2%	1.4%	0.8%	0.8%	1.59	1.24	22.1	
Te	VH tight ℓ	1.4	0.2%	0.2%	76.9%	19.0%	3.7%	1.63	1.24	0.4	
∞	VH loose ℓ	0.9	2.6%	1.1%	77 .9 %	16.8%	1.5%	1.60	1.16	1.2	
	VH E _T ^{miss}	1.8	16.3%	2.7%	34.4%	35.4%	11.1%	1.68	1.17	1.3	
	VH dijet	1.6	30.3%	3.1%	40.6%	23.4%	2.6%	1.31	1.06	1.0	
	ttH lepton	0.5	_	_	1.6%	1.6%	96.8%	1.34	1.03	0.2	
	tīH multijet	0.6	4.1%	0.9%	0.8%	0.9%	93.3%	1.34	1.03	0.6	<u>[E</u>

Composition of untagged categories:

- Boosted diphoton pair
- Both in barrel & R9>0.94
- Both in barrel

[EPJC (2014) 74:3076]

23

Expected signal breakdown at 8 TeV



For the parallel processors in human brains

24

ĮNFŇ

Statistical methodology

- Simultaneous binned Max-likelihood fit to diphoton invariant mass in all the categories:
 - $\Box \quad \mathcal{L} = \mathcal{L} \ (\text{data } \mid s(p, m_{\gamma\gamma}) + b(m_{\gamma\gamma}))$
 - s = signal parametric model from MC simulation
 - Systematic effects included as nuisance parameters on the signal model
 - b = background function: shape unknown
 - ightarrow ightarrow Wants a general description with negligible bias in the fit



25

Background description

- Previous analysis (each category):
 - Bernstein polynomials with polynomial order set by:
 - Fit bias < 1/5 σ_{fit} (**) [i.e. add *D.O.F.* until $\sigma_{syst}^2 + \sigma_{stat}^2 \approx \sigma_{stat}^2$]

New 'envelope method' (each category):

- Max-likelihood to select the function and the order which fit the best
 - Bernstein polynomials, Laurent polynomials, power law, and exponential families

Add penalty for different number parameters

good coverage and negligible bias with

 $-2log\lambda' = min\{-2log\lambda_i + N_{par}\}_i$

[Fits satisfying bias condition (**)]

'Profile' the likelihood over the function choices

- Account for arbitrariness and uncertainty of the choice
- Overall sensitivity of the analysis improves by +7%
 - Best fit functions typically have fewer parameters





Inclusive sum of all events selected

Sum weighted by sensitivity



M DEGLI STUDI MILANO BICOCCA Results: signal yield

Dataset	Significance (obs)	σ/σ _{SM}	m _H (GeV)
7 TeV	4.7 σ	2.22 +0.62 -0.55	124.2
8 TeV	4.0 σ	0.90 +0.26 -0.23	124.9
7+8 TeV	5.7 σ	1.14 +0.26 -0.23	124.7





Systematic uncertainties on the signal yield

Source	Effect
Theory	±0.11
Diphoton BDT mismodelling	±0.06
Energy and resolution corrections	±0.02
Other experimental	±0.04



Largest uncertainty of instrumental origin

- Studied by modifying inputs that distort the diphoton BDT score and affect categorization
 - I. Photon BDT score
 - II. Energy resolution estimate

Test of the uncertainty on the BDT score with Z→ee events

- → Slightly overestimated
- [Shift of the two inputs chosen to cover discrepancies in the tails of the respective distributions]

29

Results: mass measurement



Energy corrections (at m_Z) and linearity

- Photon energy corrections : $\delta m_H = 0.05 \text{ GeV}$
 - Method stability against R9 reweighting, selections, fit range
- Residual non linearity : $\delta m_H = 0.10 \text{ GeV}$
 - Dielectron invariant mass vs $H_T = \frac{1}{2} (E_{T,1} + E_{T,2})$ in boosted $Z \rightarrow ee$
 - E/p vs E_T with electrons from Z and W decays



Photon/ electron response difference

- Imperfect simulation of e/γ difference : $\delta m_H = 0.10 \text{ GeV}$
 - Per-photon effect from *double ratio of e/y response difference* in *modified* and *default* simulation
 - Longitudinal non-uniformity of light collection :
 - Imperfect EM shower simulation :
 - G4 modified with Seltzer-Berger model

Imperfect description of material :



0.02 GeV (next) 0.05 GeV

0.07 GeV

Photon/ electron response difference (2)

- Imperfect simulation of e/γ difference : $\delta m_H = 0.10 \text{ GeV}$ (cont'd)
 - ► Longitudinal non-uniformity of light collection : 0.02 GeV
 - Residual non-uniformity from lab tests (all crystals!): 0.14%/X₀
 - e/γ response difference from difference in shower depth
 - → Per-photon scale change (including radiation effects): ~0.05% anti-correlated between R9>0.94 and R9<0.94 photons</p>



Result cross checks: compatibility

- Three alternative analyses
 - OK at 1σ level
- Compatibility with preliminary result
 - OK at <2o using jackknife technique to account for correlations
- Compatibility among channels
- Compatibility of 7 and 8 TeV $_{VH}^{VH}$
 - At 2σ (most from VBF)



More results on Higgs properties

Production mechanism and coupling modifiers



 $\mu_{ggH+ttH} = 1.13^{+0.37}_{-0.32}$

 Higgs boson decay width, search for additional states, high mass search, spin hypotheses



Reported CMS analysis in search for SM $H \rightarrow \gamma \gamma$ Dataset: 5.1 fb⁻¹ at $\sqrt{s}=7$ TeV and 19.7 fb⁻¹ at $\sqrt{s}=8$ TeV

- Discussed the steps to optimize the performance of photon energy reconstruction, the accurate modelling of the photon response for the H $\rightarrow\gamma\gamma$ analysis, and the associated systematic uncertainties
 - Additional details in CMS photon performance paper being submitted
- Described the new analysis optimization with tags for exclusive production modes



Trigger selections and efficiencies

Trigger selections and efficiencies



High Level Trigger

- Asymmetric E_T thresholds on two photons
- Photons pass EITHER loose calo ID (shower shape + iso) OR high R₉ (i.e. unconverted)
- Trigger efficiency for preselection is 99.4 %

Time dependent simulation





events

ECAL endcap [MeV] CMS Preliminary 2011-2012 2.8 600 2.6 Noise 500 er shutdow 2.4 400 2.2 300 2 200 1.8 100 1.6 04/11 07/11 10/11 01/12 04/12 07/12 10/12 12/12 Date [month/year] 4000 Data RunD 35000 Static MC

Out-of-time pileup collisions simulated in an extended window

- Collisions in previous beam crossings impact on the amplitude reconstruction (dynamic pedestal subtraction)
- II. Dynamic evolution of the per-channel noise level
 - ▶ Three run periods in 2012
- Improved Data/MC agreement





Imperfections in simulations

Effect of inter-crystal voids:







Monitoring and correction of imperfect monitoring corrections

Phi-simmetry of energy flow to monitor the evolution with time of the spread in intercalibration constants



'Measurement of energy corrections'

- Energy correction in data: E_{corr}/E_{raw}
 - Electrons from $W \rightarrow ev$ events



43

Efficiency and acceptance



Signal composition: 7 TeV data







Diphoton BDT and photon quality

Mapping of diphoton BDT score to R9 / η regions



7 TeV data

INFN BICOCCA Background: bias studies



Signal strength



Most (1.5σ) from one VBF category



Cross-check – different analyses



Systematic	Value	Error
Pagaling (aba)	1 1 /	+0.26
baseline (obs)	1.14	-0.23
Cut based (abs)	1 20	+0.29
Cut-based (obs)	1.29	-0.26
Sideband (obs)	1.06	+0.26
Sideband (ODS)	1.00	-0.23
2D Dijet (aha)	1.04	+0.25
2D Dijet (obs)	1.04	-0.25





Cut based analysis

- less model dependence in selections
- No exclusive tag categories
- Fit signal strength in InI,R9 categories and five bins of spin sensitive angle
 - Collins-Soper

50





[CMS-PAS-HIG-14-009]

Higgs properties all decay modes





52

Modifiers of fermion and vector-boson couplings



53