Inclusive Search for the $H \to \gamma\gamma$ with Unbinned Maximum-Likelihood (ML) Technique at CMS

Serguei GANJOUR
CEA-Saclay/IRFU, Gif-sur-Yvette, France

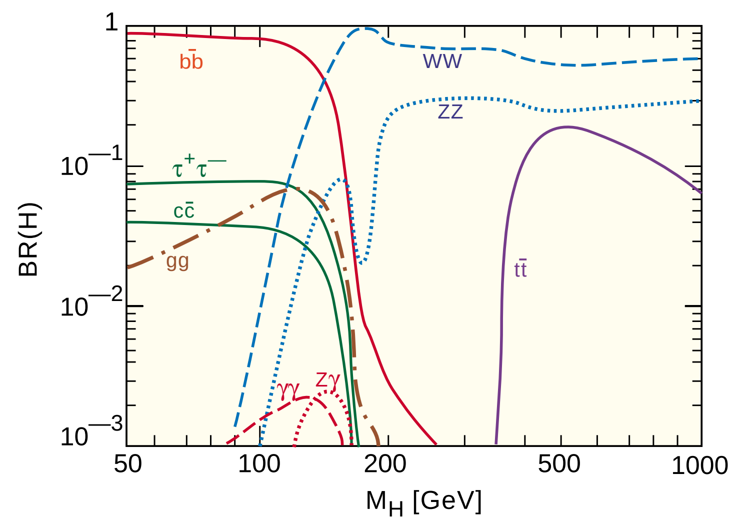
Grégory SCHOTT
Karlsruhe Institute of Technology, Karlsruhe, Germany

Anna FALKIEWICZ
IPNL, Lyon, France

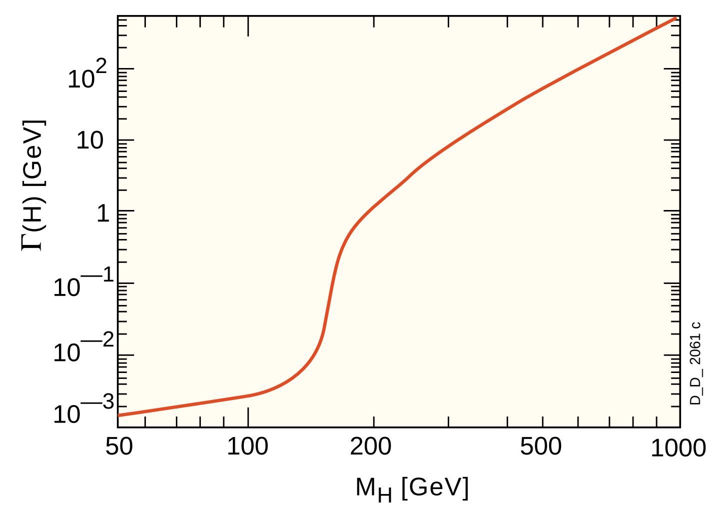
CMS Physics France
CEA/Saclay
April 2, 2010

Benchmark Channel: $H \to \gamma\gamma$

- Discovery channel at masses above LEP limit ($114.4\,\mathrm{GeV}/c^2$) and below about $150\,\mathrm{GeV}/c^2$
  - low signal rate $\mathcal{B} \sim 10^{-3}$
    - decay involves q,W loops;
  - clean signature (contrarily to $H \to b\bar{b}$);
  - identified as a narrow peak on the top of continuous background



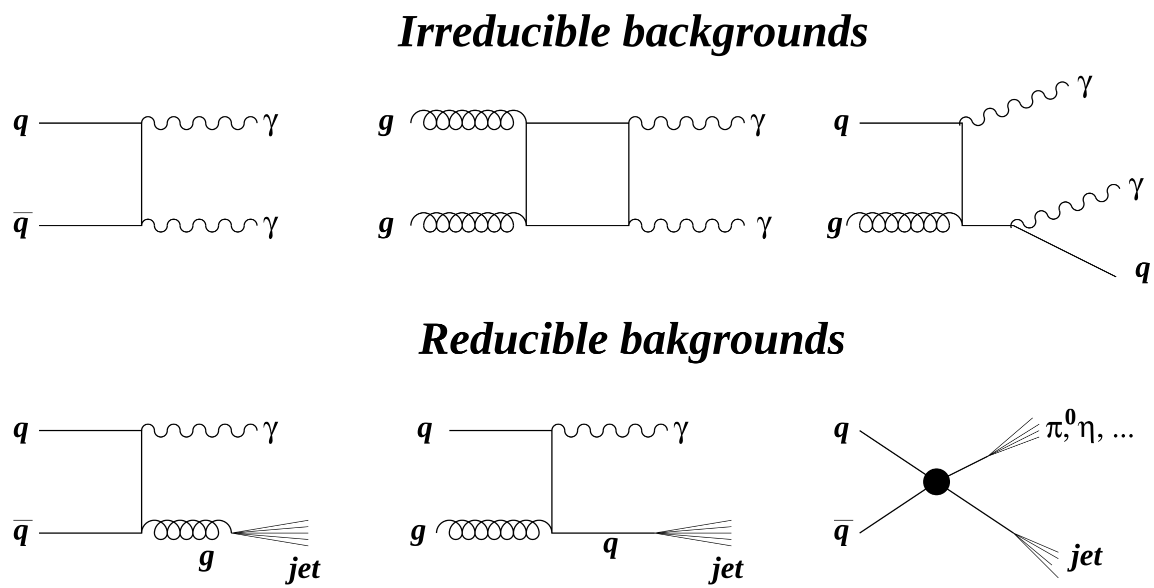
- Fully detected signature
  - $\Gamma \sim \mathcal{O}(10^{-3}\,\mathrm{GeV})$ for $m_H \leq 2m_W$
    - detector resolution is crucial
  - $\Gamma \sim \mathcal{O}(1-100)\,\mathrm{GeV}$ for $m_H \geq 2m_Z$
    - requires efficient background rejection



- Inclusive search allows any Higgs production mechanism to pass event selections

- Irreducible backgrounds
  - born, box and isolated bremsstrahlung;
  - diff. rates at 120 GeV $d\sigma/dm_{\gamma\gamma} \sim 100\,\text{fb}/\,\text{GeV}/c^2$
  - required mass resolution: $\Delta M_{\gamma\gamma} \leq 1\,\text{GeV}/c^2$
- Reducible backgrounds
  - QCD high p_T jet processes;
  - neutral hadrons (π^0, η) may lead to the fake photons;
  - required jet suppression

$\epsilon_{jets} \sim 10^{-3}$



Process	p_T (GeV/c)	σ_{LO} (pb)
$H \to \gamma\gamma$ (120 GeV/c^2)	—	0.057 (NLO)
pp$\to \gamma\gamma$ (born)	> 25	32
pp$\to \gamma\gamma$ (box)	> 25	22
pp$\to \gamma$+jet	> 40	4×10^4
pp$\to$jets	> 60	6.6×10^6

Work Considered

- Unbinned maximum-likelihood (ML) approach is a new technique in CMS
  - incorporates $m(\gamma\gamma)$ and other discriminating variables
    - similar to recent ATLAS publication
  - study assumes $\sqrt{s} = 10$ TeV LHC energy regime
- Derive basic principles that should drive the ML analysis
  - model has to be designed to facilitate analyses at low statistics
    - provide a simple flexible model that can be expended step by step
  - expose best kinematical discriminators
    - how to deal with possible correlations?
    - design neural network discriminator for correlated variables
    - test justifiability of NN usage
  - feed the likelihood model to `RooStatsCms` package
    - estimation of expected limits and significance;
    - profile likelihood, frequentist (or modified frequentist CLs), Bayesian methods

Signal Samples at $\sqrt{s} = 10$ TeV

Process	Generator	#Events	σ^{gen}	σ^{sim}	$\int Ldt$, (fb^{-1})
H120 (all)	PYTHIA (LO)	20k	38 fb	38 fb	526
H120 (gluonfusion)	MCatNLO	20k	40.8±0.7 fb	40.8 fb	490

- M(H) values: 120, 130, 150 GeV/c^2
- Combine different generators and production mechanisms
  - Pythia samples include a mixture of all productions (XS weights)
  - compute k-factor with NLO event generator for the gluon fusion
  - assume same k-factor (1.51) for all production types

PYTHIA Signal Cross Section for $M(H) = 120\,\mathrm{GeV}/c^2$

$\sqrt{s}$	gg fusion	WW fusion	ZZ fusion	WH	ZH	ttH
14 TeV	49 fb	9 fb	3.4 fb	3.7 fb	2.1 fb	1.5 fb
10 TeV	27 fb	4.9 fb	1.8 fb	2.5 fb	1.2 fb	0.6 fb

NLO Cross Sections

$M(H)$	σ (fb)
120 GeV	57.4
130 GeV	49.9
150 GeV	23.2

Background Samples at $\sqrt{s} = 10$ TeV

Process	Generator	$\sqrt{s}$	$\widehat{p}_T(H_t)$	#Events	σ^{gen}	σ^{sim}	Lum (fb^{-1})
$gg \rightarrow \gamma\gamma$ (box)	PYTHIA	10 TeV	$25 \div \mathrm{Inf}$	500k	22 pb	22 pb	22.7
$gg \rightarrow \gamma\gamma$ (box)	PYTHIA	10 TeV	$10 \div 25$	500k	580 pb	580 pb	0.86
$qq \rightarrow \gamma\gamma$ (born)	MADGRAPH	10 TeV	$10 \div \mathrm{Inf}$	1M	210 pb	210 pb	4.8
$pp \rightarrow \gamma + jets$	MADGRAPH	10 TeV	$40 \div 100$	1.96M	40.6 nb	40.6 nb	0.0483
$pp \rightarrow \gamma + jets$	MADGRAPH	10 TeV	$100 \div 200$	730k	8.3 nb	8.3 nb	0.0877
$pp \rightarrow \gamma + jets$	MADGRAPH	10 TeV	$200 \div \mathrm{Inf}$	2.18M	0.99 nb	0.99 nb	2.2
QCD Jets	PYTHIA	10 TeV	60	245k	6.6 ub	6.6 nb	0.037

☞ `MADGRAPH` includes the effect of high-p_T jets;

☞ Require further re-weighting with NLO generators JETPHOX, DIPHOX, GAMMA2MC

☞ Private production of the QCD jets background with pre-filter

☞ Combine two highest p_T photons in the event to form the Higgs boson candidate:

➠ $p_T{}^{\gamma 1,2} = (20, 20)$ GeV (DoublePhoton HLT path)

☞ Correct for the primary vertex:

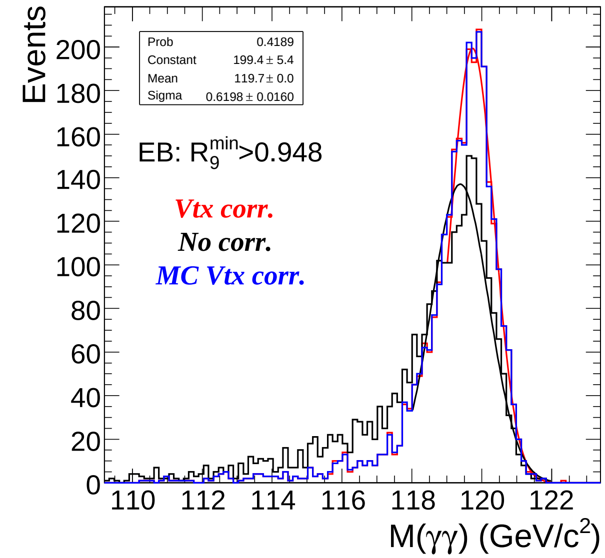
➠ no pile-up samples (efficiency?)

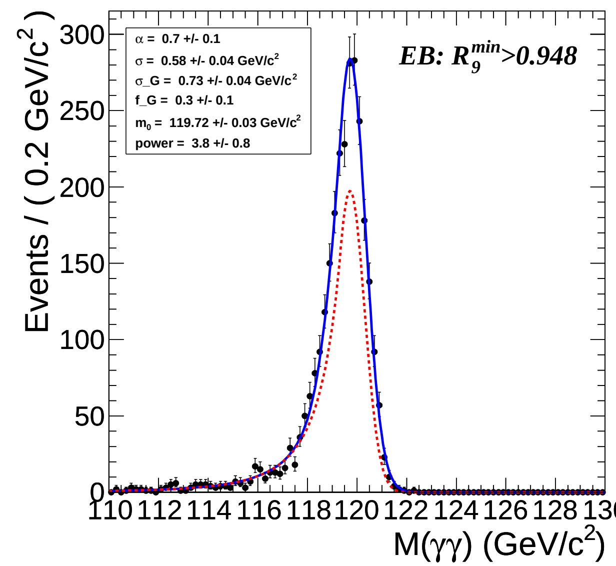
➠ negligible effect of vertex resolution

➠ simplify study with control sample $Z^0 \rightarrow e^+e^-$

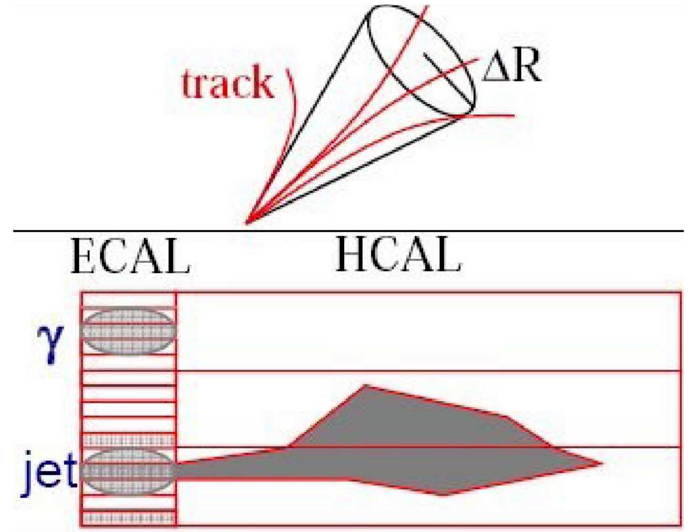
☞ Resolution model: sum of Crystal Ball and Gaussian

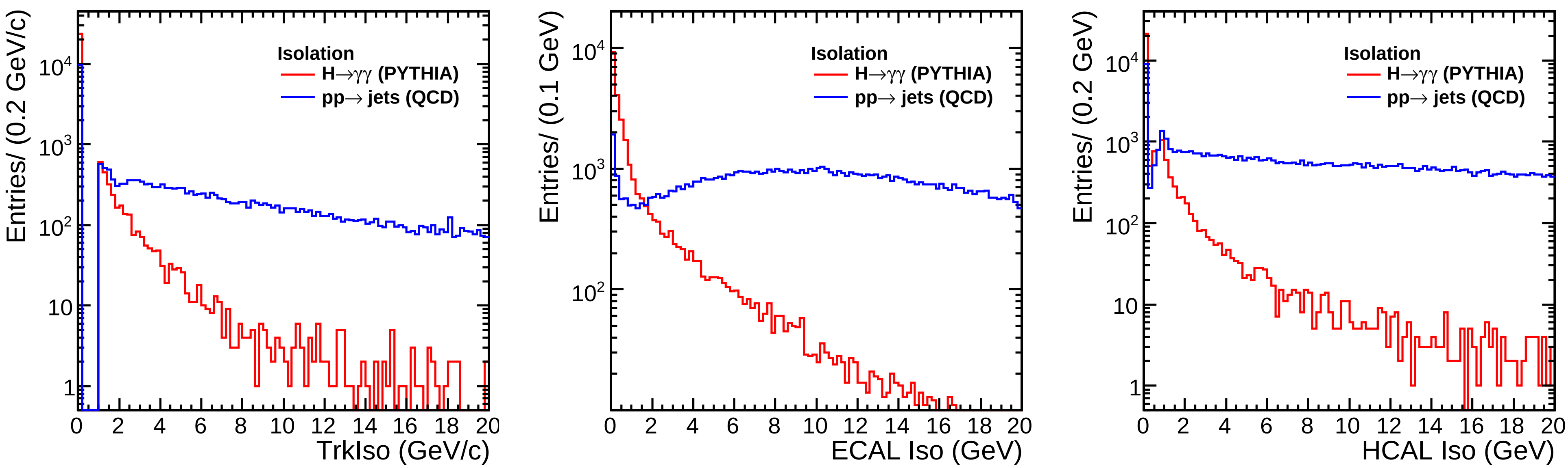
➠ best obtained resolution: $\Delta M_{\gamma\gamma} \simeq 0.6\,\mathrm{GeV}/c^2$, for EB non-converted di-photon events





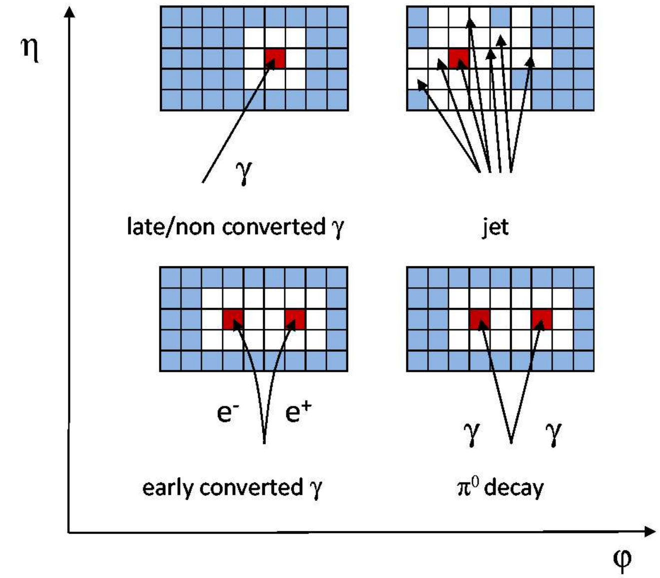
- ☞ π^0 from fragmentation processes fake photons
  - ➠ accompanied by other particles
  - ➠ isolation exploit 3 subdetectors variables computed inside a cone: $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$
    - ➜ $\Sigma\, {p_T}^{trk} < 2\,\mathrm{GeV}/c$ in tracker
    - ➜ $\Sigma\, E_T - E_T^{\gamma} < 2\,\mathrm{GeV}$ in ECAL
    - ➜ $\Sigma\, H_T < 4\,\mathrm{GeV}$ in HCAL





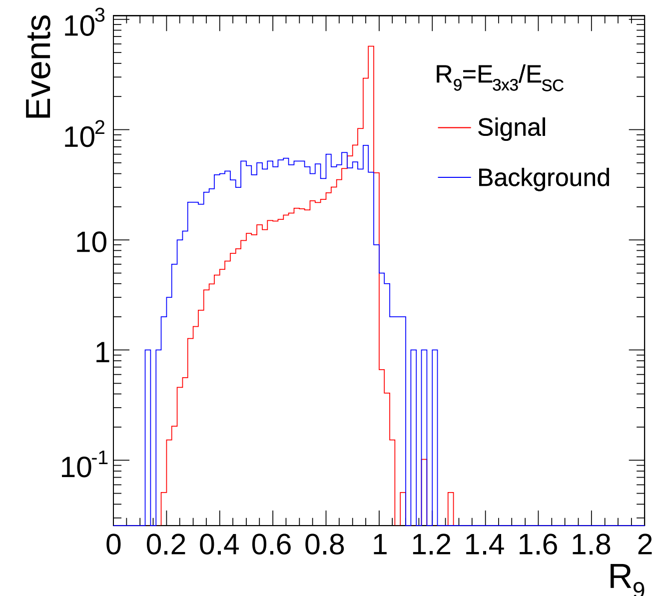
☞ High value of $R_9 = E_{3x3}/E_{SC}$ readily identifies non-converted photons

- automatically selects against π^0;
- converted category remains background enriched;



☞ Divide events according to signal purity: R_9 and η

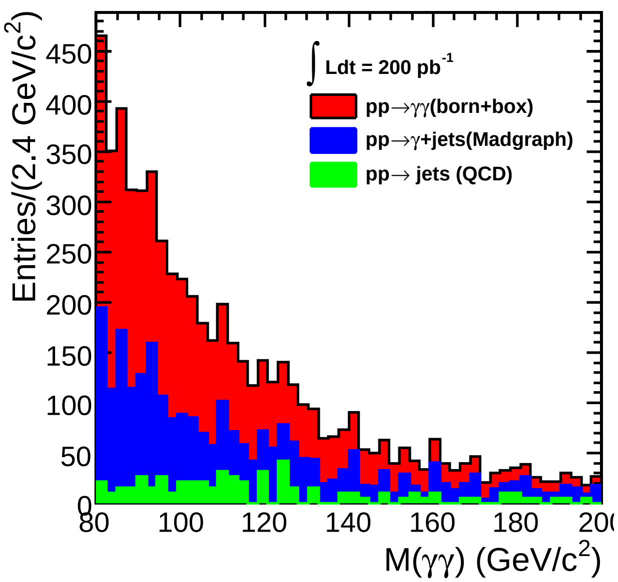
- S/B varies with η;
- different resolution for EB and EE;
- conversion degrades resolution and raises bkg.

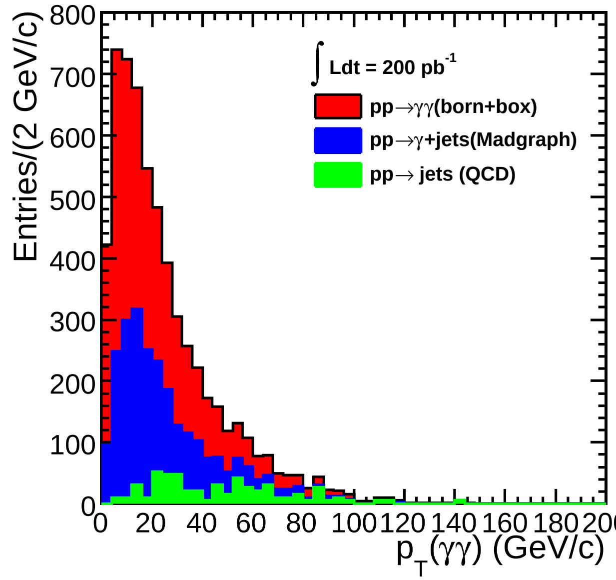


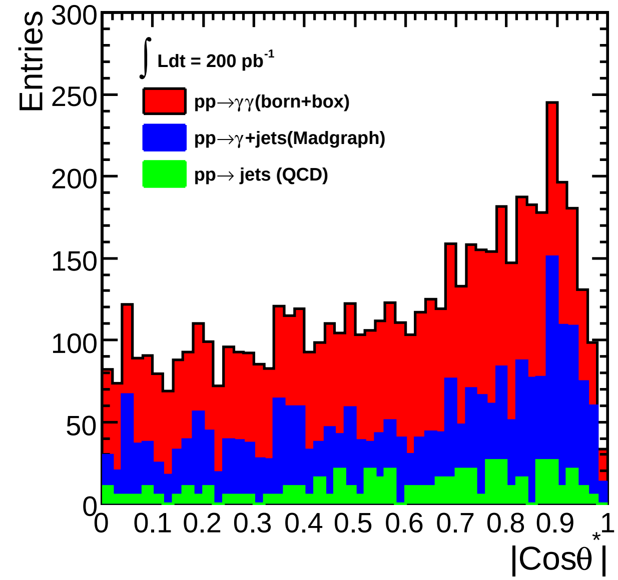
4 categories

	$\lvert\eta^{max}\rvert < 1.479$	$\lvert\eta^{max}\rvert > 1.479$
$R_9^{min} > 0.93$	25.4%	25.5%
$R_9^{min} < 0.93$	26.6%	22.5%

☞ Consider 4, 6 and 12 category splitting







- Signal assumption
  - apply k-factor 1.2 for box (`Pythia`);
  - no k-factor for the born (`Madgraph`);
- Limited statistics to estimate background yields per categories
- Isolation cuts are included

Event yields for $\int Ldt$ =200 pb^{-1} and $\sqrt{s} = 10$ TeV

Process	$\lvert m_{\gamma\gamma} - M_{\rm H}\rvert < 5$	$100 < m_{\gamma\gamma} < 150$
H→ $\gamma\gamma$	4.6	5.0
pp→ $\gamma\gamma$(born)	266	1291
pp→ $\gamma\gamma$(box)	74	353
pp→ γ+jets	165	825
pp→ QCD	70	281

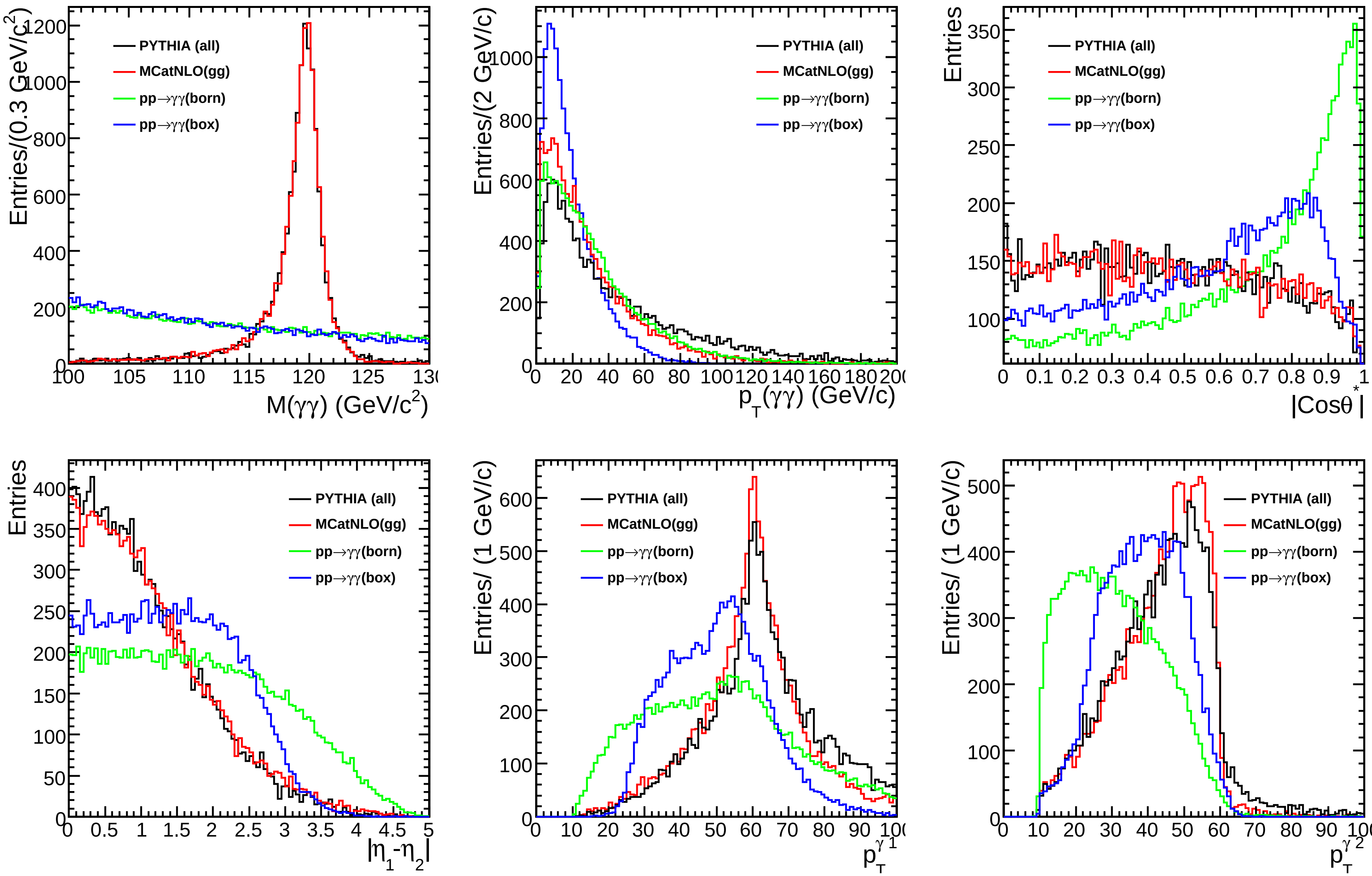
Observables and Correlations

- **Consider 6 kinematical parameters:**
  - Inv.mass $\boldsymbol{m_{\gamma\gamma}}$
  - Transv. momentum $\boldsymbol{p_{T\gamma\gamma}}$
  - Pseudo-helicity $|\cos\boldsymbol{\theta^*}|$
  - Pseudo-rapidity $|\boldsymbol{\eta^{\gamma1} - \eta^{\gamma2}}|$
  - $\gamma 1$ transv. momenta: $\boldsymbol{p_T}^{\gamma1}$
  - $\gamma 2$ transv. momenta: $\boldsymbol{p_T}^{\gamma2}$
- Correlation of $m_{\gamma\gamma}$ with others is $\leq 5\%$ ($\leq 15\%$) for signal (background)
- Construct unbinned likelihoods based on
  - 5 vars with TMVA: all except $m_{\gamma\gamma}$
    - ➔ test different classifiers
  - uncorrelated: $m_{\gamma\gamma}$, p_T, $|\cos\theta^*|$
- Model optimization: 3 event types $\{\texttt{signal}, \texttt{born}, \texttt{box}\}$;

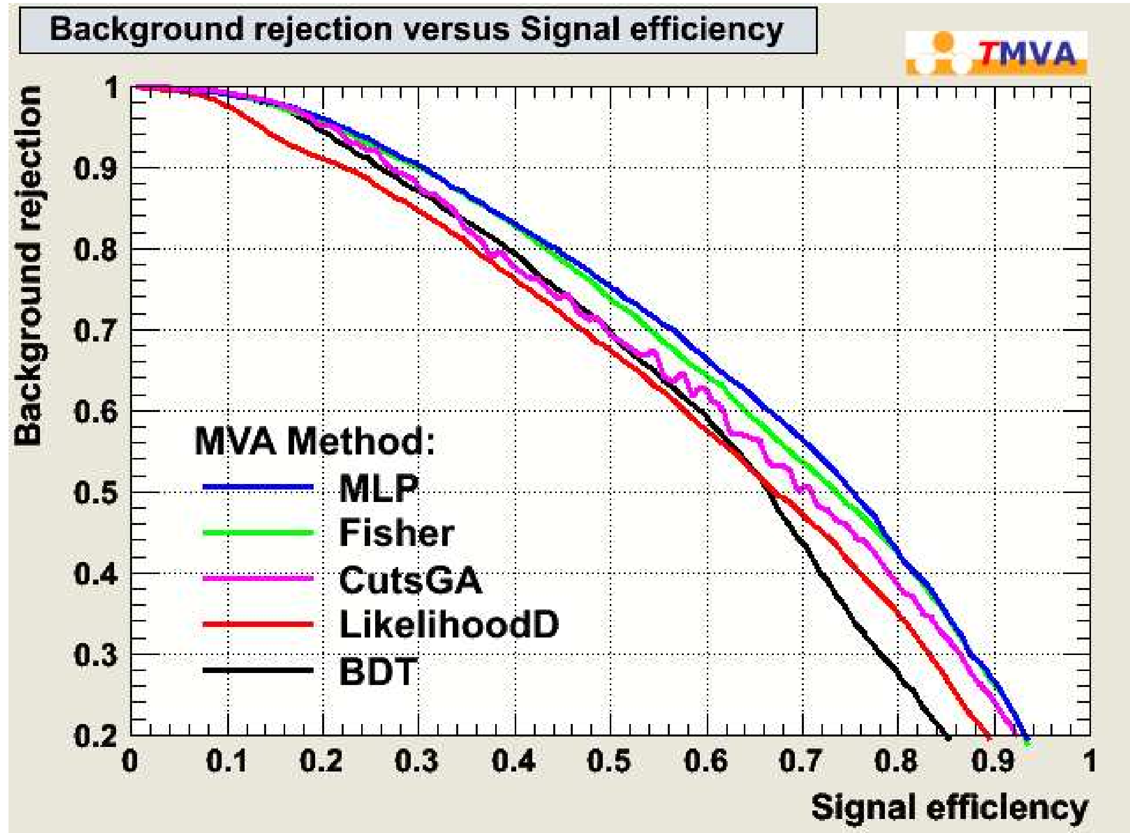
lin. corr. coef. (%)		| cos θ* |	PT (H)	| Δη (γ) |	PT (γ1)	PT (γ2)
Signal	m (H)	3	0	3	2	2
Box		3	4	5	12	13
Born		2	5	5	10	10
γ+jets		4	5	9	7	7
QCD γγ		6	2	9	5	8
Signal	PT (γ2)	-53	35	-57	29	
Box		-72	7	-90	66	
Born		-59	-11	-70	15	
γ+jets		-46	-6	-64	15	
QCD γγ		-80	-3	-83	45	
Signal	PT (γ1)	-6	87	-43		
Box		-59	63	-87		
Born		-21	81	-71		
γ+jets		-15	87	-78		
QCD γγ		-57	63	-80		
Signal	| Δη (γ) |	53	-14			
Box		74	-30			
Born		54	-30			
γ+jets		39	-48			
QCD γγ		84	-21			
Signal	PT (H)	-1				
Box		-19				
Born		-2				
γ+jets		-3				
QCD γγ		-7				

lin. corr. coef. (%)		m (H)	trkIso (γ1)	ecalIso (γ1)	hcalIso (γ1)	trkIso (γ2)	ecalIso (γ2)	hcalIso (γ2)
Signal	nnet5	0	3	16	3	4	4	4
Box		6	1	10	2	2	7	2
Born		6	5	4	9	1	1	0
γ+jets		1	11	7	13	3	2	2
QCD γγ		3	9	8	26	3	5	29

Discriminating Variables



- Build the following samples to probe TMVA classifiers
  - signal: `Pythia`;
  - background: born and box;
  - each sample is splitted in two parts: training and test;
- Test algorithms:
  - multi-layer-perceptron (MLP) NNet;
  - Fisher discriminant;
  - Boosted Decision Trees;
  - likelihood



- All classifiers are similar
  - MLP NNet looks slightly better

- Likelihood provides a general approach for parameter estimation and statistical inference
  - likelihood $\mathcal{L}$ is a product of p.d.f. functions $f^i \equiv f(x^i)$: $\mathcal{L}(\vec{\theta}) = \Pi_{i=1}^{N} f^i(\vec{\theta})$
  - p.d.f. function is a sum over different types of events: $f^i = \Sigma_{j=\{s,b,...\}} \epsilon_j f_j^i$

$$\mathcal{L}(\vec{\theta}) = \frac{e^{-(n_s+n_b)}(n_s+n_b)^N}{N!} \prod_{i=1}^{N} n_s f_s^i(\vec{\theta}) + n_b f_b^i(\vec{\theta})$$

- Several observables $x \rightarrow \vec{x} = \{m_{\gamma\gamma}, p_{T\gamma\gamma}, |\cos\theta^*|, ...\}$ leads to the product p.d.f.

$$f^i = \mathcal{M}^i(m_{\gamma\gamma}) \cdot \mathcal{P}^i(p_{T\gamma\gamma}) \cdot \mathcal{T}^i(|\cos\theta^*|) \cdot ...$$

- One possibility to account correlations between observables:
  - Multi-Variate Analysis (MVA) technique: NN, BDT, Fisher,...

$$f^i = \mathcal{M}^i(m_{\gamma\gamma}) \cdot \mathcal{N}^i(NN_{out})$$

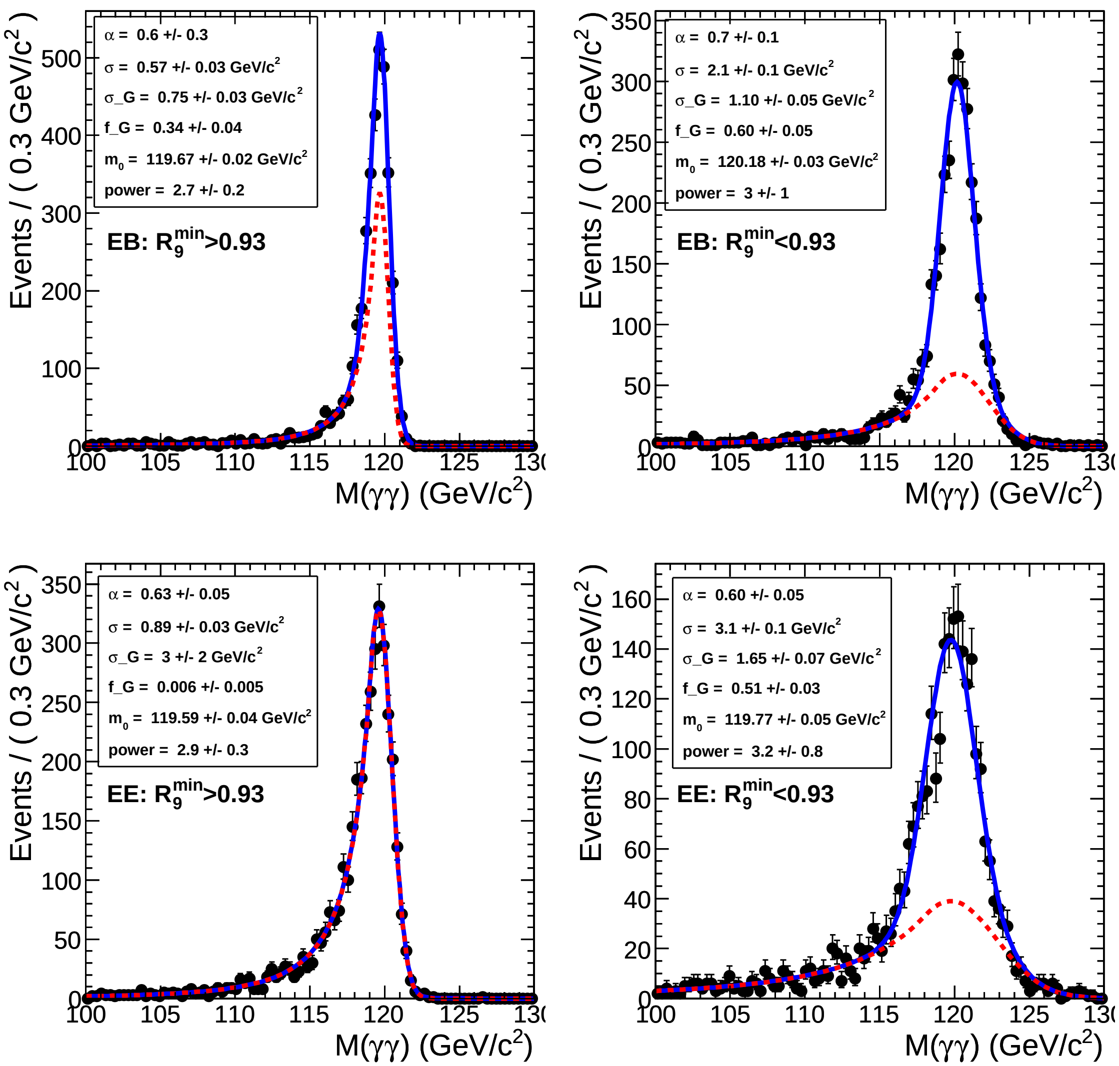
- Discrete variable or category split the p.d.f. as

$$\mathcal{M}^i(m_{\gamma\gamma}) = \prod_{k=cat} \mathcal{M}_k^i(m_{\gamma\gamma}, cat(R_9^{min}, |\eta|^{max}))$$

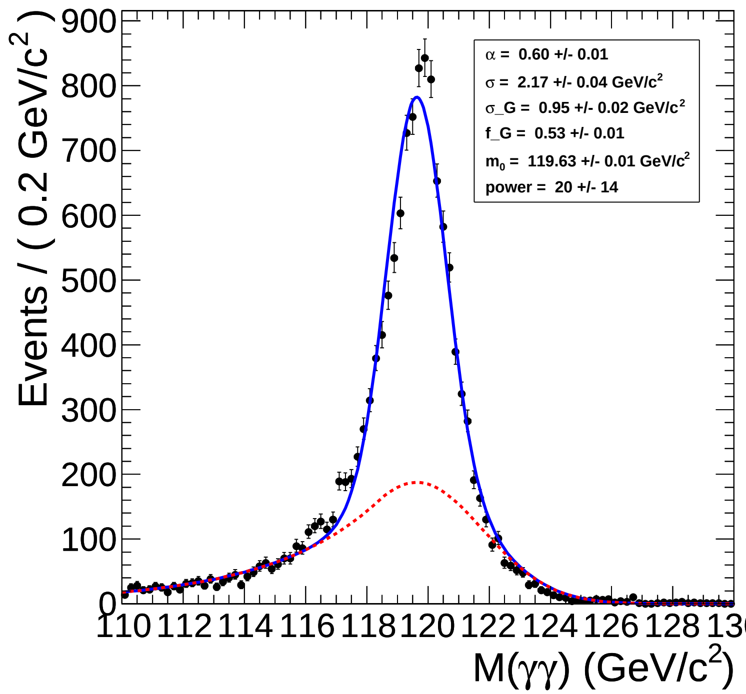
- f^i are different for each event type {`signal`, `born`, `box`, `gamjet`, `qcd`}

Signal Resolution Model

Model: Crystal Ball plus Gaussian



Fit of all categories



	Barrel	Endcap
$R_9^{min} > 0.93$	25.4%	25.5%
$R_9^{min} < 0.93$	26.6%	22.5%

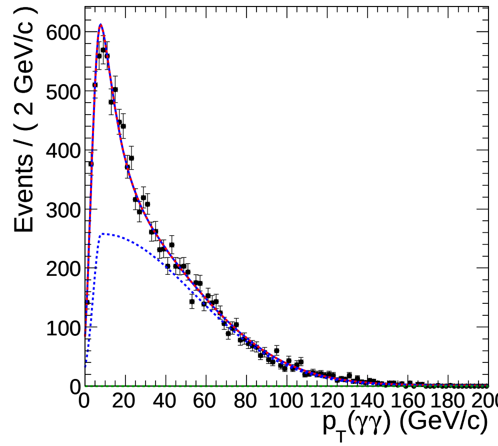
$\Delta M_{\gamma\gamma} \sim 0.6 \div 2.2\,\mathrm{GeV}/c^2$

P.D.F.s for Fit Variables

- Higgs boson p_T exhibits a long tail beyond the maximum
- $|\cos\theta^*|$ is uniform for scalar Higgs boson
- Suppress $|\cos\theta^*|$ towards one due to acceptance effect
- Enhancement of bkg. phase space towards $|\cos\theta^*| \sim 1$

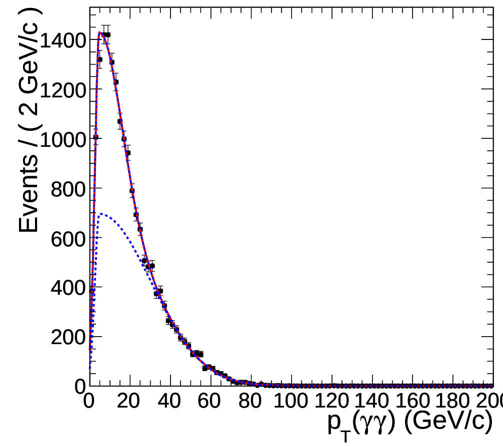
Variable	Signal	Background
$m_{\gamma\gamma}$	CB+Gauss	Exp
$p_{T\gamma\gamma}$	Landau+Bif.Gauss	2 +Bif.Gauss
$\lvert\cos\theta^*\rvert$	2 Gauss+Pol1	2 Gauss+Pol1
MLP	Sum of Gaussian	Sum of Gaussian
$Fisher$	Sum of Gaussian	Sum of Gaussian

Signal

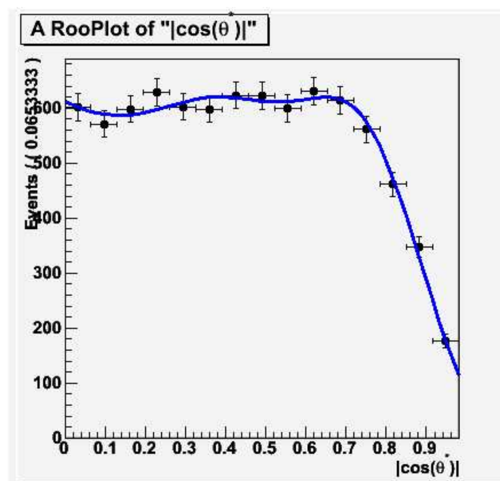


Landau plus Bifurc. Gaussian

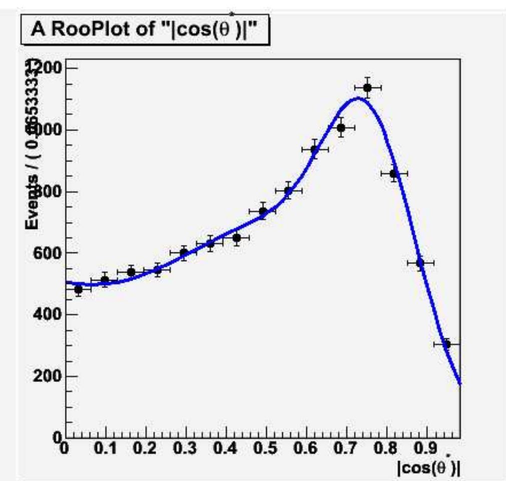
Background



2 Bifurc. Gaussians



2 Gaussians plus Pol1



2 Gaussians plus Pol1

Model Optimization

☞ Run number of toy experiments

- test different models;
- extended unbinned ML fit;
- estimate significance;

$$S_L = \sqrt{-2(\ln \mathcal{L}_{S+B} - \ln \mathcal{L}_B)}$$

Likelihood model provides maximal performances exploiting the event topology and 3 discriminating variables

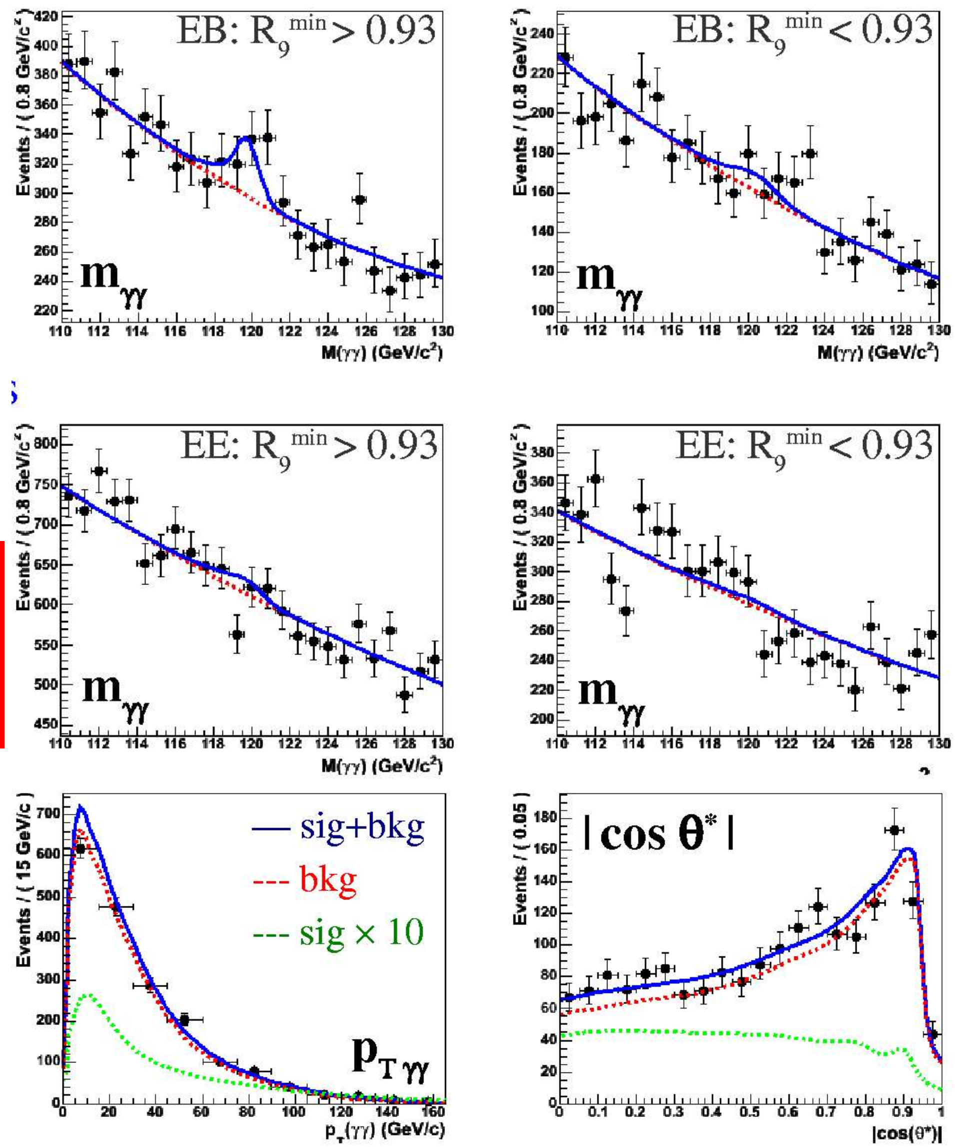
$10\ \mathrm{fb}^{-1}$ and $\sqrt{s} = 10$ TeV

Model	$\langle S_L \rangle$
$\mathcal{L}(m_{\gamma\gamma})$	2.14±0.03
$\mathcal{L}(m_{\gamma\gamma}, p_T)$	2.66±0.03
$\mathcal{L}(m_{\gamma\gamma}, \cos\theta^*)$	2.47±0.03
$\mathcal{L}(m_{\gamma\gamma}, {p_T}^{\gamma\gamma}, \cos\theta^*)$	3.00±0.03
$\mathcal{L}(m_{\gamma\gamma}, MLP)$	2.87±0.03
$\mathcal{L}(m_{\gamma\gamma}, Fisher)$	2.76±0.03
$\mathcal{L}(m_{\gamma\gamma}, cat4)$	2.35±0.06
$\mathcal{L}(m_{\gamma\gamma}, {p_T}^{\gamma\gamma}, \cos\theta^*, cat4)$	3.60±0.06

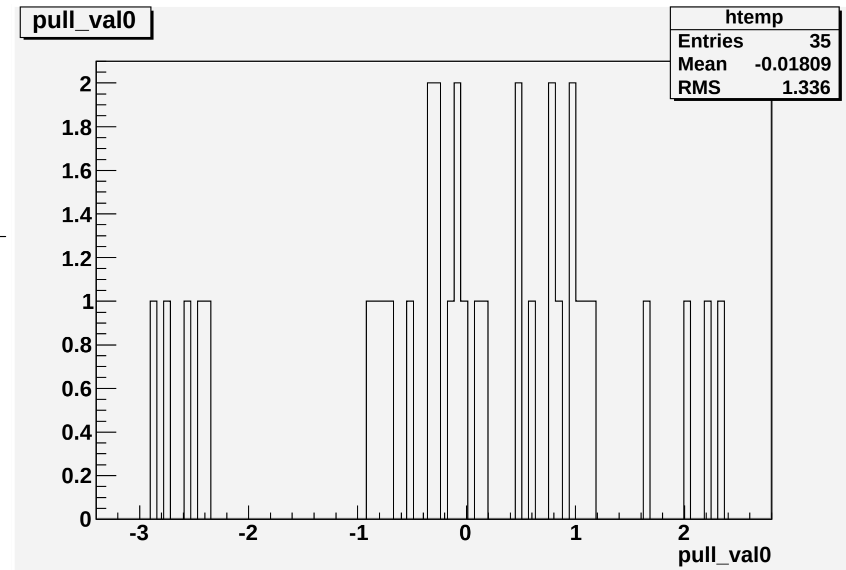
- Extended unbinned ML approach:
  - sum over all event types: $\{\texttt{signal}, \texttt{born}, \texttt{box}, \texttt{gamjet}, \texttt{qcd}\}$
  - multi-dimensional model: $\boldsymbol{m_{\gamma\gamma}, p_{T\gamma\gamma}, |\cos\theta^*|}$
  - splitted into 4 $\boldsymbol{R_9 - \eta}$ categories according to a signal purity

$$\mathcal{L} = \prod_{k=cat} \mathcal{M}_k(m_{\gamma\gamma}, cat(R_9^{min}, |\eta|^{max}))\cdot\mathcal{P}(p_{T\gamma\gamma})\cdot\mathcal{T}(\cos\theta^*)$$

- p.d.f.s for each component are fixed
  - measured in the "side-bands"
- floated parameters: N_s and N_b
  - compute a significance for each mass
  - floated mass is an option



- Toy Monte Carlo Validation of the fit model
  - model the signal according to full Monte Carlo model
    - produce 570 $H \to \gamma\gamma$ events per 10 fb^{-}
    - divide 20k of signal events into 35 independent samples
    - account real correlations
  - fit with proposed Likelihood model, i.e. no correlations between variables

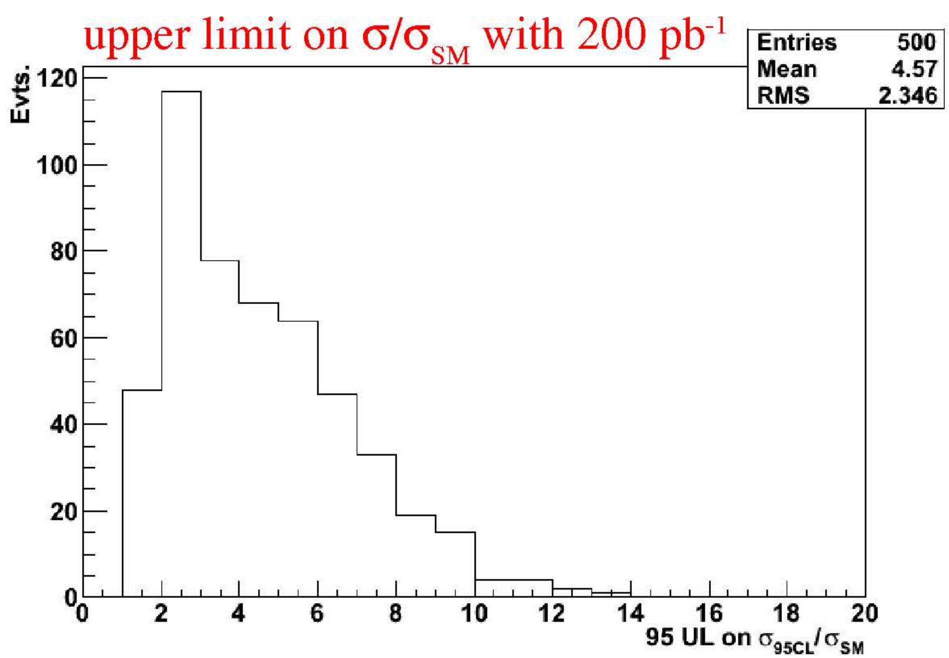


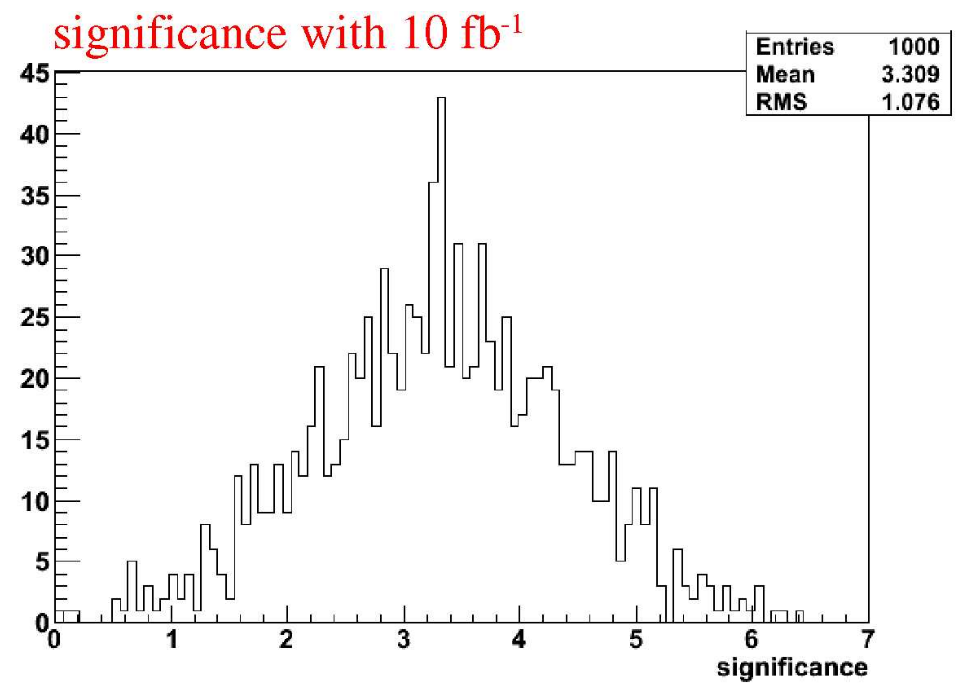
- No bias observed in the N_s pull distribution:

$$\langle N_s \rangle = 0.02 \pm 0.17; \quad r.m.s\langle N_s \rangle = 1.34 \pm 0.17$$

Large statistical MC sample of signal is strongly required for thorough validation test!

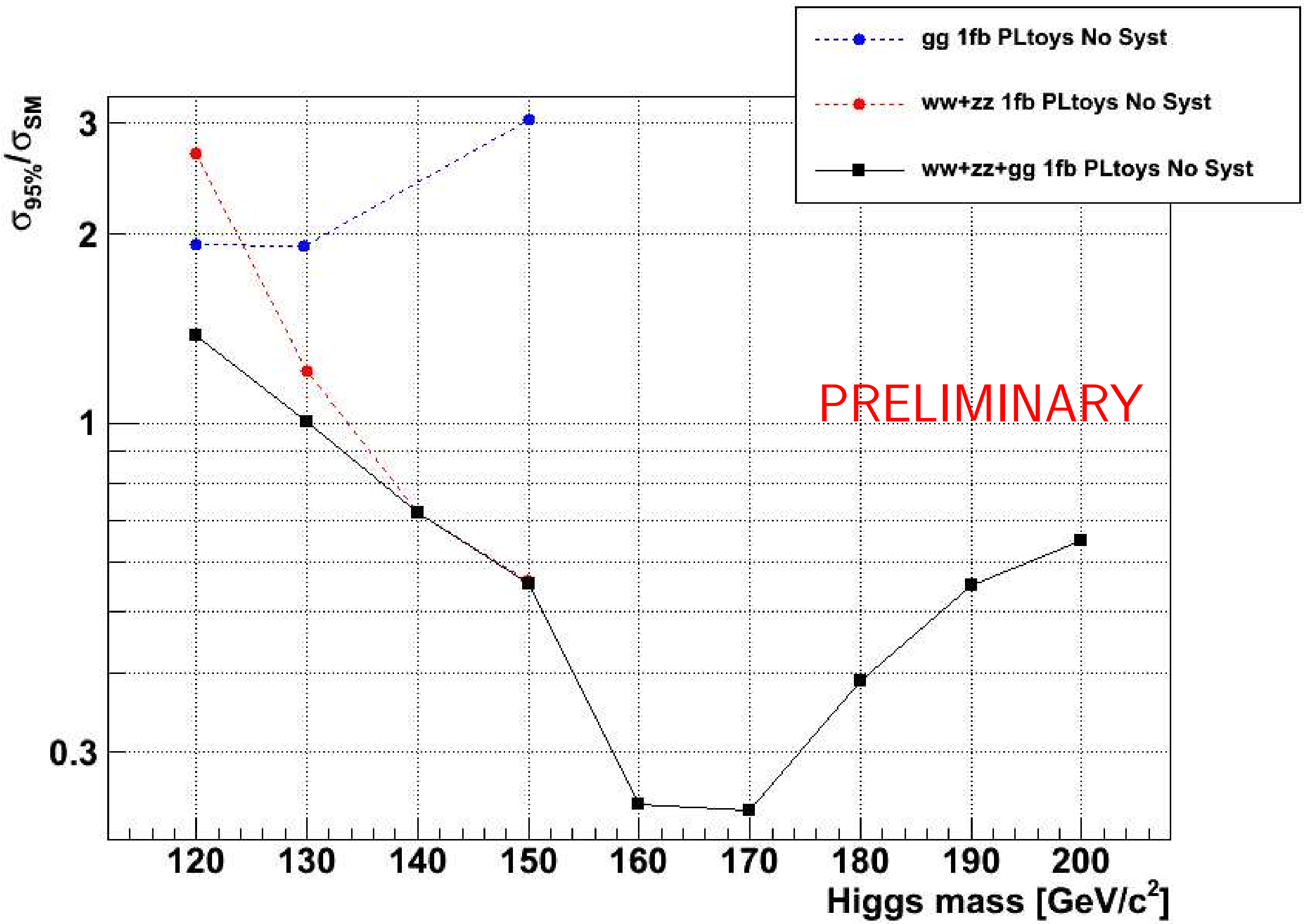
$$\boldsymbol{\sigma/\sigma_{\rm SM} < 4.6} \text{ at 95\% CL } (200\,\text{pb}^{-1}),\ \boldsymbol{\sqrt{s} = 10} \text{ TeV}$$
$$\boldsymbol{S_L = \sqrt{-2\Delta \ln \mathcal{L}} = 3.3 \pm 1.1} \ (10\,\text{fb}^{-1}),\ \boldsymbol{\sqrt{s} = 10} \text{ TeV}$$





ATLAS expects 3.6 σ with $10\,\text{fb}^{-1}$ at $\sqrt{s} = 14$ TeV

95% Upper Limit on $\sigma/\sigma_{\rm SM}$ with $\int Ldt$ = 1 fb^{-1} and $\sqrt{s} = 10$ TeV



Conclusions

☞ Carried out the $H \to \gamma\gamma$ analysis with new approach - unbinned ML

- analysis exploits a full CMS detector simulation (`CMSSW_2_2_9`);
- MCatNLO event generator complement standard `Pythia` signal samples;
- backgrounds (Born, γ+jets) includes high-p_T jets effects (`Madgraph`);

☞ Designed the Likelihood model with an eye towards early discovery

- built flexible model that can be easily extended once more data recorded;
- 3-d likelihood model $(m_{\gamma\gamma},\ p_T,\ |\cos\theta^*|)$ provides best performances;
- this model integrated into `RooStatsCms` to proceed with statistical tests;
- obtained signal significance for $\sqrt{s} = 10$ TeV for $m_H = 120\,\mathrm{GeV}$:

$$\boldsymbol{\sigma/\sigma_{\mathrm{SM}} < 4.6} \text{ at 95\% CL with } 200\,\mathrm{pb}^{-1}$$
$$\boldsymbol{S_L = 3.3 \pm 1.1} \text{ with } 10\,\mathrm{fb}^{-1}$$

- contributes significantly in combined Higgs mass constraint at low mass region $\leq$150 GeV

Backup slides

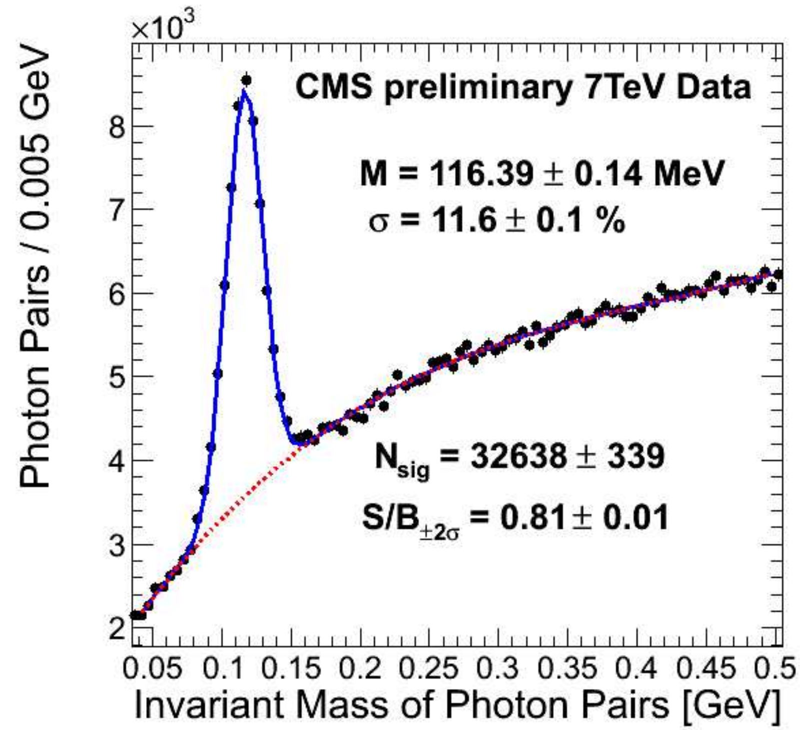
☞ Estimated energy in the ECAL:

$$E_{e,\gamma} = F \times \sum_{clusters} G\, c_i\, A_i$$

Corrections *Calibration*

☞ Energy correction scheme

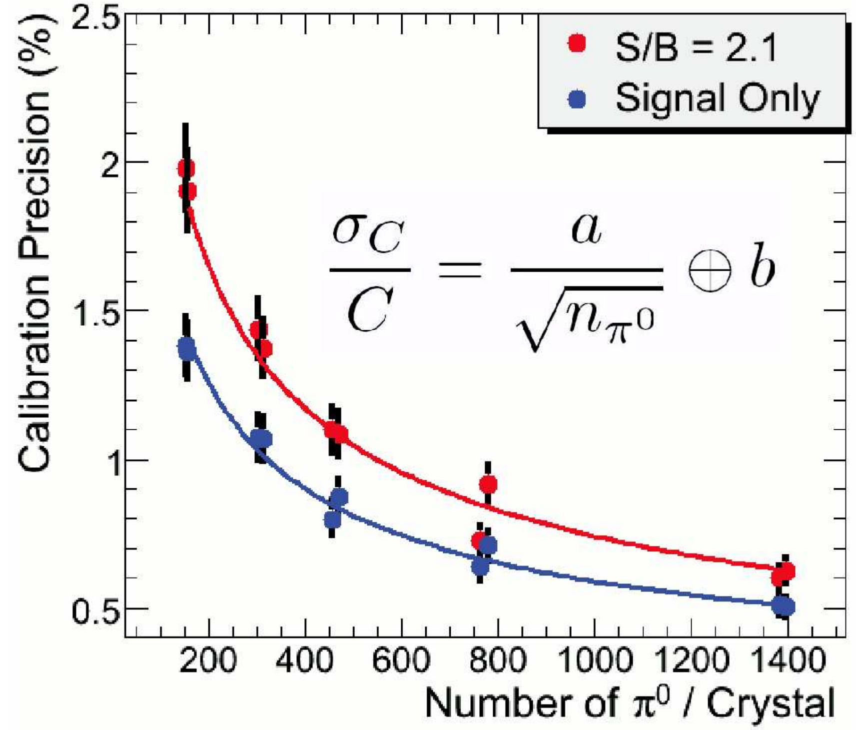
- $F = 1$ for 5x5 crystal sum for the energy of unconverted photons;
- overall containment factor;
- local containment and boundaries;
- correct for the bremsstrahlung;
- crystal transparency (laser monitoring)



☞ Calibration and alignment

- dedicated π^0 calibration
- physics events $W^+ \to e^+\nu,\ Z^0 \to e^+e^-$

- Start-up calibration precision
  - test beam calibration only for 9 SM for EB (500 Xtals for EE)
  - others have couple % calibration from cosmics for EB
  - about 10% lab calibration for EE
- Several paths for in-situ calibration



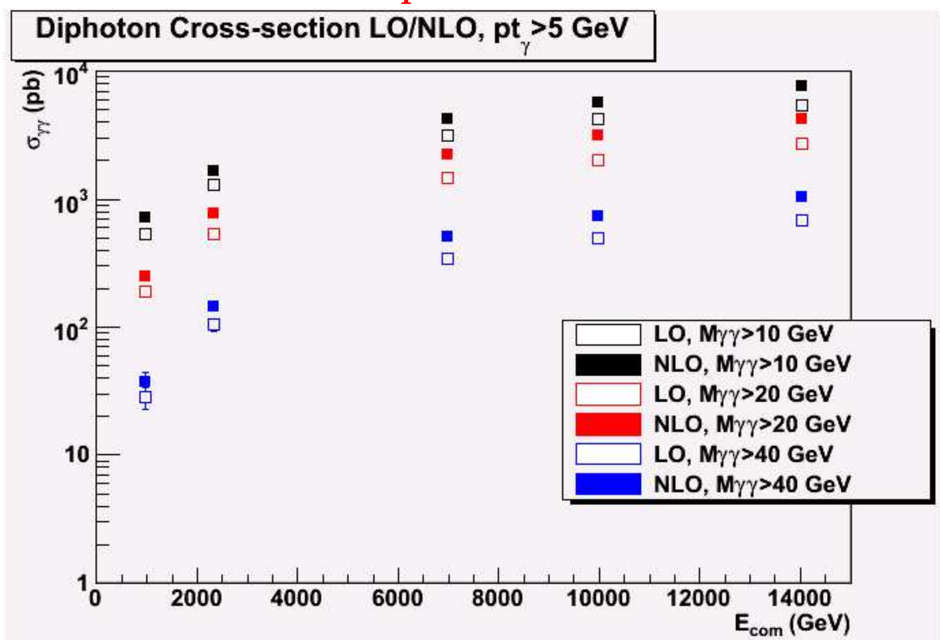
Strategy	Time	Precision
Mean energy deposited by jet triggers independent on ϕ at fixed η (correct for tracker material)	few hours	2-3%
π^0 mass peak ($L = 2 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$)	few days	$\leq$1%
$Z^0 \rightarrow e^+e^-$ absolute calibration	100 pb^{-1}	$\leq$1%
$W \rightarrow e\nu$ E/p measurement	5 fb^{-1}	$\leq$0.5%

N.Chanon et al.

DIPHOX, GAMMA2MC (NLO) pp$\to \gamma\gamma + X$ includes: born, box and one- and two- fragmentation photons

$\gamma\gamma + X$ σ_{NLO} (pb) at $\sqrt{s} = 7$ TeV

Process	> 10 GeV/c^2	> 20 GeV/c^2	> 40 GeV/c^2
Born	288	299	43
Box	416	213	50
OneFrag.	2060	915	181
TwoFrag.	1514	830	232
Total	4278	2257	506

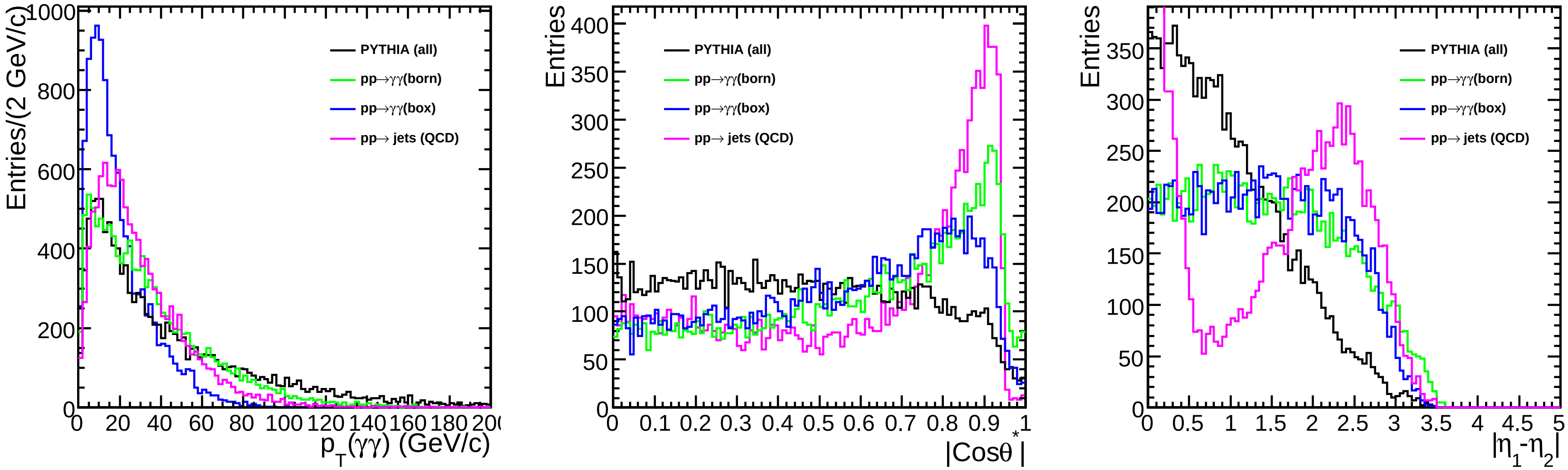


Di-photon σ_{NLO} (pb)

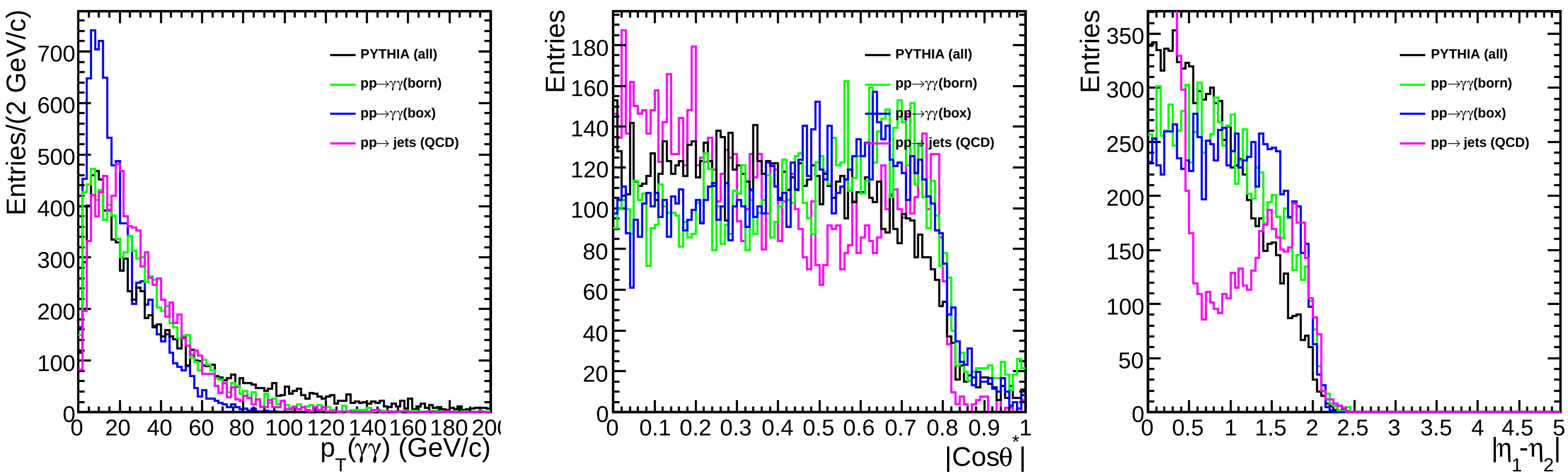
$m_{\gamma\gamma}^{min}$ (GeV/c^2)	0.9 TeV	2.36 TeV	7 TeV	10 TeV	14 TeV
10	716	1196	4278	5746	7627
20	251	768	2257	3155	4285
40	38	82	506	743	1051

Expect about 500 di-photon events above 40 GeV with 1 pb^{-1} at $\sqrt{s} = 7$ TeV. Obviously, 2010 publication!

High luminosity DoublePhoton HLT path $(20{,}20)$ GeV

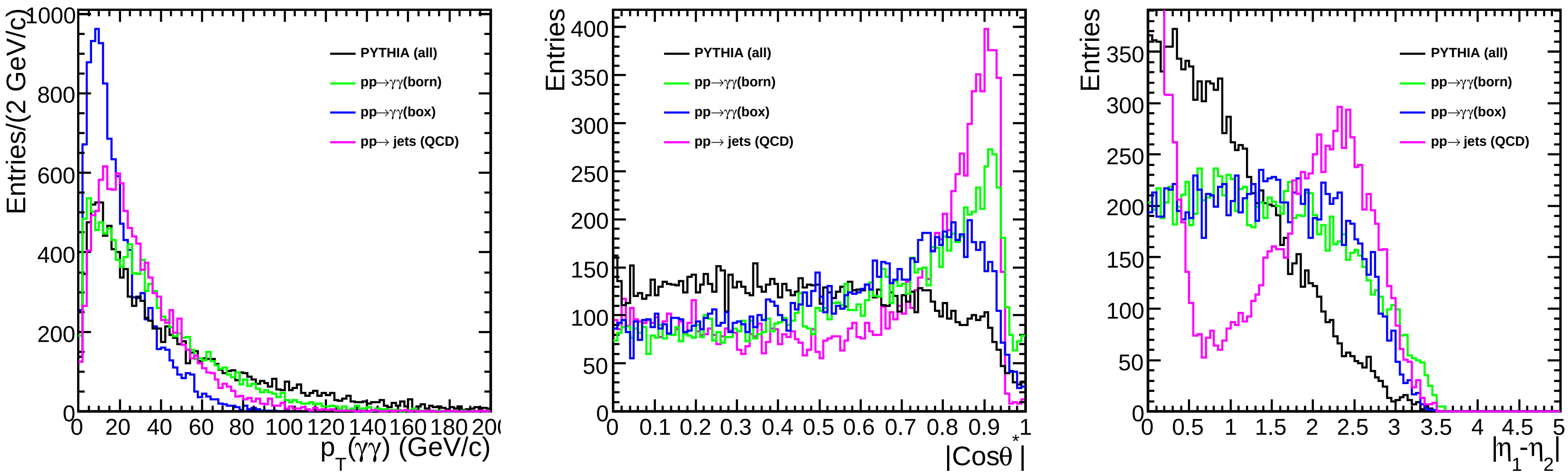


pTDR off-line analysis cuts $p_T^{\gamma 1,2} > (40{,}35)$ GeV

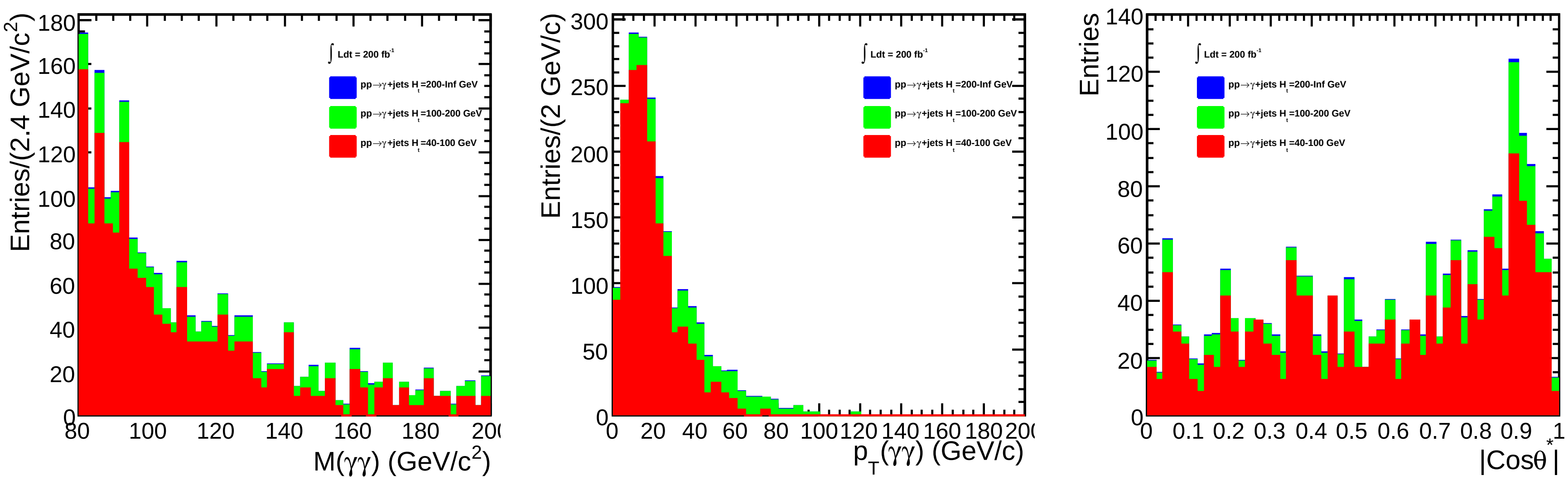


Reducible Background Kinematics

High luminosity DoublePhoton HLT path (20,20) GeV



γ+jets main contribution 40<H_t <100 GeV



- Two main aspects in the framework of discovery analysis: parameter estimation and statistical inference
- Likelihood function $\mathcal{L}$ provides a general approach
  - number counting experiment: $\mathcal{L}(N, n_s, n_b) = \frac{e^{-(n_s+n_b)}(n_s+n_b)^N}{N!}$
  - likelihood analysis based on p.d.f. functions $f^i \equiv f(x^i)$: $\mathcal{L}(\vec{\theta}) = \Pi_{i=1}^{N} f^i(\vec{\theta})$
  - p.d.f. function is a sum over different types of events: $f^i = \Sigma_{j=\{s,b,...\}} \epsilon_j f_j^i$

$$\mathcal{L}(\vec{\theta}) = \prod_{i=1}^{N} \epsilon_s f_s^i(\vec{\theta}) + (1-\epsilon_s) f_b^i(\vec{\theta})$$

  - extended ML accounts different statistical fluctuations $\epsilon_s << \epsilon_b$

$$\mathcal{L}(\vec{\theta}) = \frac{e^{-(n_s+n_b)}(n_s+n_b)^N}{N!} \prod_{i=1}^{N} n_s f_s^i(\vec{\theta}) + n_b f_b^i(\vec{\theta})$$

- Several observables $x \to \vec{x} = \{m_{\gamma\gamma}, p_{T\gamma\gamma}, |\cos\theta^*|, \ldots\}$ leads to the product p.d.f.

$$\boldsymbol{f^i = \mathcal{M}^i(m_{\gamma\gamma}) \cdot \mathcal{P}^i(p_{T\gamma\gamma}) \cdot \mathcal{T}^i(\cos\theta^*)}$$

- Possibilities to account correlations between observables:
  - Multi-Variate Analysis (MVA) technique: NN, BDT, Fisher,...

$$f^i = \mathcal{M}^i(m_{\gamma\gamma}) \cdot \mathcal{N}^i(NN_{out})$$

    - systematics uncertainties not trivial to extract
  - factorization of uncorrelated multi-dimensional p.d.f:

$$f^i = \mathcal{M}^i(m_{\gamma\gamma}) \cdot \mathcal{P}^i(p_{T\gamma\gamma}, \cos\theta^*)$$

    - complexity of multi-dimensional parameterization
- discrete variable or category split the p.d.f. as

$$\boldsymbol{\mathcal{M}^i(m_{\gamma\gamma}) = \prod_{k=cat} \mathcal{M}^i_k(m_{\gamma\gamma}, cat(R_9^{min}, |\eta|^{max}))}$$

- f^i are different for each event type {`signal`, `born`, `box`, `gamjet`, `qcd`}

☞ simultaneous fit to multiple data samples (control sample):

$$NLL = \sum_{i=1}^{n} -\log f_A^i(D_A) + \sum_{j=1}^{m} -\log f_B^j(D^j)$$

☞ further constraints are added in $-\log \mathcal{L}$ as additional terms

➠ suppose θ_s is affected by systematics described by $g(\theta_s)$ p.d.f: $f' = f(x,\theta) \cdot g(\theta_s)$

➠ penalty term appears in log-likelihood

$$NLL = -\log \mathcal{L} - \log g(\theta_s) = NLL_0 + NLL_P$$

☞ in case of Gaussian p.d.f. $\chi^2 = -2\log \mathcal{L}$:

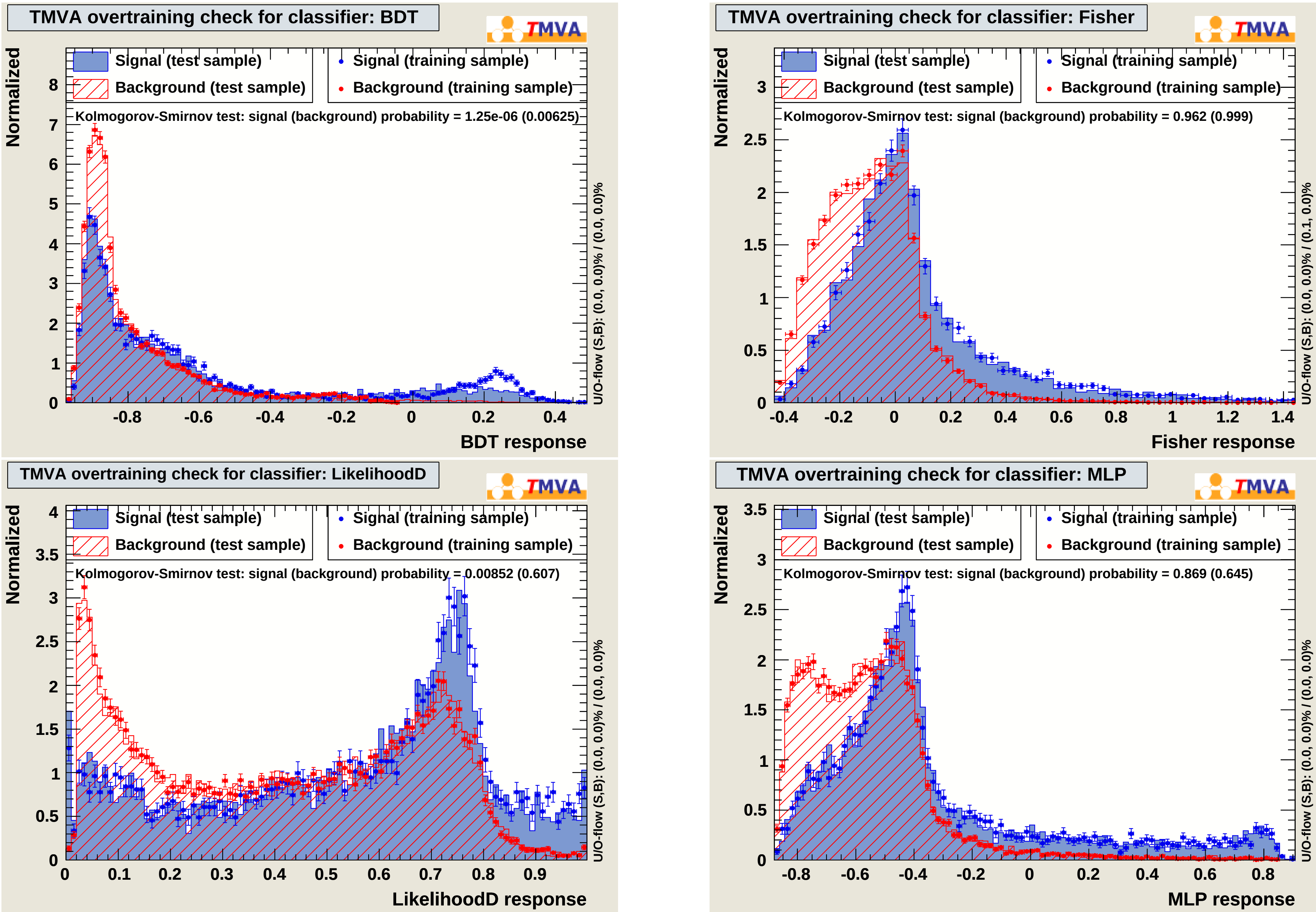
➠

$$NLL_P = \log(\sigma_\theta \sqrt{2\pi}) + \frac{1}{2}(\frac{\theta - \mu}{\sigma_\theta})^2$$

➠ account correlations between parameters and translate into multiple dimensions:

$$\boldsymbol{NLL_P = \log(|V|^{\frac{1}{2}} 2\pi^{\frac{N}{2}}) + \frac{1}{2}(\vec{\theta} - \vec{\mu})^T V^{-1} (\vec{\theta} - \vec{\mu})}$$

TMVA Discriminators



Correlation Matrix

lin. corr. coef. (%)		| cos θ* |	PT (H)	trkIso (γ1)	ecalIso (γ1)	hcalIso (γ1)	trkIso (γ2)	ecalIso (γ2)	hcalIso (γ2)
Signal	m (H)	3	0	2	4	-5	7	7	3
Box		3	4	1	1	4	-2	1	1
Born		2	5	1	2	2	4	1	4
γ+jets		4	5	1	1	2	1	-1	4
QCD γγ		6	2	2	-2	3	2	-2	3
Signal	hcalIso (γ2)	4	6	4	4	10	53	9	
Box		-4	10	1	3	2	13	8	
Born		2	26	22	6	26	51	6	
γ+jets		-8	33	-1	1	10	23	-2	
QCD γγ		-38	8	13	-10	33	15	-9	
Signal	ecalIso (γ2)	13	4	6	3	4	12		
Box		7	5	1	2	3	-6		
Born		4	5	-1	0	-2	-27		
γ+jets		-6	-7	-7	-3	-7	-68		
QCD γγ		8	-2	-5	3	-8	-84		
Signal	trkIso (γ2)	6	7	4	6	6			
Box		3	6	2	1	3			
Born		7	23	15	5	15			
γ+jets		10	18	-11	-2	-6			
QCD γγ		-7	-2	6	-4	11			
Signal	hcalIso (γ1)	2	5	41	3				
Box		3	8	11	18				
Born		-3	6	58	1				
γ+jets		-5	-7	55	1				
QCD γγ		-21	14	23	-15				
Signal	ecalIso (γ1)	-5	21	-7					
Box		-10	10	-4					
Born		-2	10	-20					
γ+jets		1	1	-28					
QCD γγ		6	-5	-84					
Signal	trkIso (γ1)	0	8						
Box		0	2						
Born		-2	3						
γ+jets		-2	-14						
QCD γγ		-7	3						
Signal	PT (H)	-1							
Box		-19							
Born		-2							
γ+jets		-3							
QCD γγ		-7							