

# Discrimination $\gamma / \pi^0$

pour la recherche du boson de Higgs dans le canal  $H \rightarrow \gamma\gamma$  dans l'expérience CMS

Nicolas Chanon (IPNL)

01/12/2009 – Journées de Rencontre Jeunes Chercheurs 2009

**I – Sélection des photons**

**II – Variables d'entrée**

**III – Résultats de l'analyse multivariée**

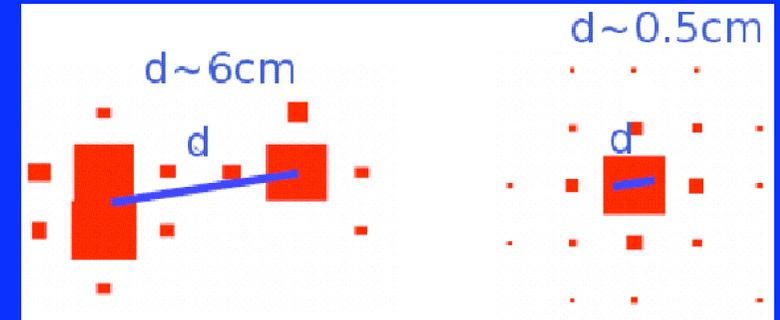
**IV – Les  $\pi^0$  dans les premières collisions de CMS**

Travail fait en collaboration avec l'IHEP-Beijing.

## Introduction

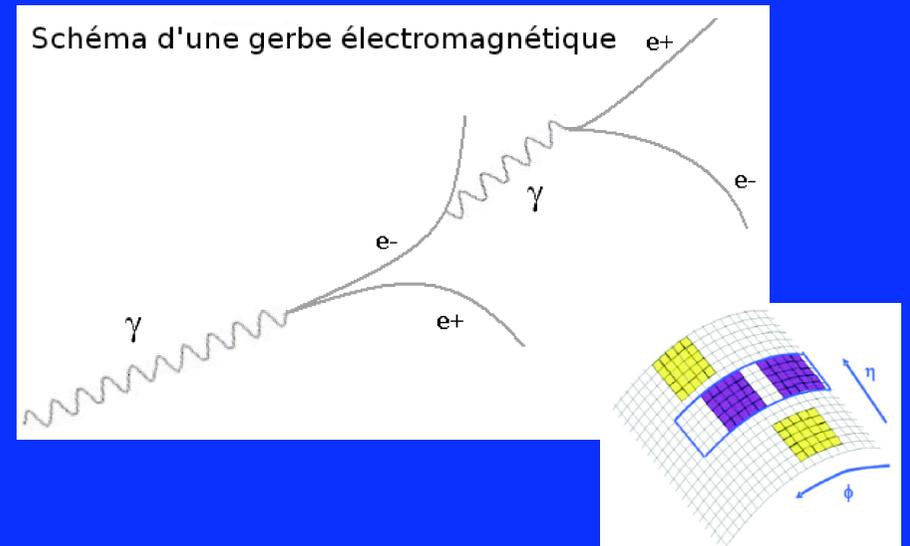
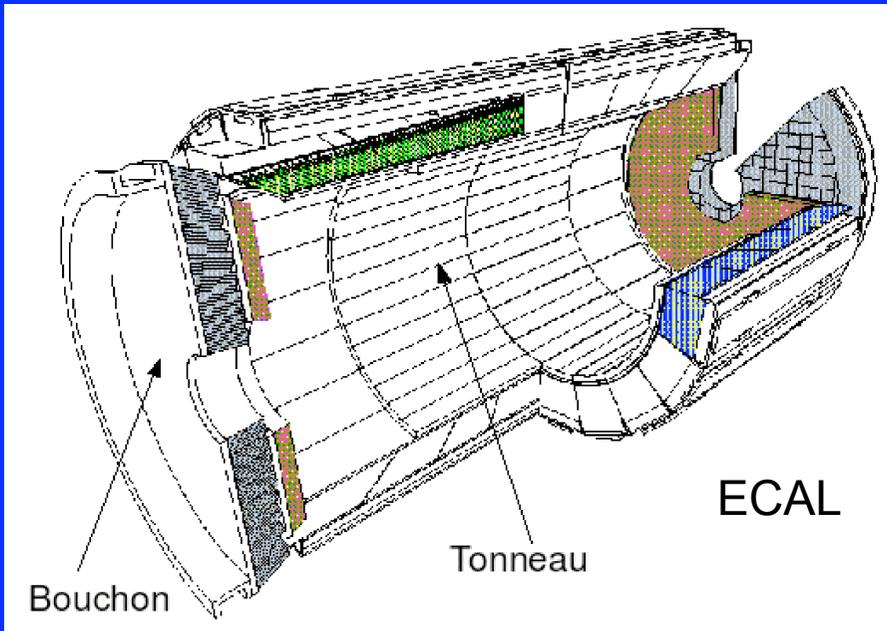


- **Motivation** : Les  $\pi^0$  très boostés émis par des jets se désintègrent généralement en paire  $\gamma\gamma$  très collimée. Leur énergie est récupérée dans un seul supercluster. Un  $\pi^0$  sera alors reconstruit comme un photon.



- **Problème** : la cinématique de ces événements est souvent très proche de celle des photons.

- **Idée** : utilisation d'une technique d'analyse multivariée pour discriminer les photons des  $\pi^0$ .



# Discrimination $\gamma / \pi^0$

## Sélection au niveau reconstruit :

- $|\eta| < 2.5$
- L'évènement est sélectionné si il y a **au moins un photon avec  $E_t > 40$  GeV**. Dans ces événements, on garde alors tous les photons ayant  **$E_t > 25$  GeV**
- **Tracker ISO** : Aucune trace ayant  $p_t > 1.5$  GeV dans un cône  $\Delta R < 0.3$  autour de la direction du photon candidat. On demande des traces ayant laissé des coups dans 2 couches du pixel.
- **Ecal ISO** : Somme de l'eT des basicclusters de l'ECAL dans un cône  $0.06 < \Delta R < 0.35$  doit être  $< 6$  GeV dans le tonneau (3 GeV dans les bouchons).
- **Hcal ISO** : Somme de l'eT dans les tours HCAL dans un cône  $\Delta R < 0.3$  doit être inférieure à 6 GeV dans le tonneau (5 GeV dans le bouchon)

## Note sur la conversion des photons en paire $e^+/e^-$

- Les photons peuvent se convertir en paire  $e^+/e^-$  dans le détecteur (env. 50% de probabilité). A l'heure actuelle, l'analyse pour les photons convertis est séparée de l'analyse pour les photons non convertis.
- Pour les photons qui convertissent, on apparie le supercluster aux traces laissées par la paire  $e^+/e^-$  dans le trajectographe. Parfois, une des traces est perdue.

# Discrimination $\gamma / \pi^0$

## I – Sélection des photons

### Pourcentage des photons passant les coupures

Cuts efficiency (%)	$H \rightarrow \gamma\gamma$ , $m_H=120$ GeV	QCD preselected
No Cut	100	100
% Photons in events having at least 1 photon with $p_T > 40$ GeV	91.34	76.60
1 PHOTON REQUIRED per event		
1 photon with $ \eta  < 2.5$	85.13	56.94
and $p_T > 25$ GeV	69.79	36.95
and Track Iso	55.23	3.74
and ECAL Iso	52.28	2.18
and HCAL Iso ( $\gamma/\pi^0$ disc selection)	50.81	1.41
2 PHOTONS REQUIRED per event		
2 photons with $ \eta  < 2.5$	81.05	50.22
and $p_T > 25$ GeV	62.85	25.93
and Track Iso	42.34	0.33
and ECAL Iso	38.13	0.09
and HCAL Iso	36.07	0.03
% Photons in events having 2 photons passing $\gamma/\pi^0$ disc selection	36.05	0.03
with $p_{T1} > 40$ GeV and $p_{T2} > 35$ GeV ( $H \rightarrow \gamma\gamma$ PTDR selection)	30.91	0.02

**Signal** : tous les photons du processus  $H \rightarrow \gamma\gamma$  PYTHIA passing cuts.

**Bruit de fond** : tous les photons des processus QCD passant les coupures

# Discrimination $\gamma / \pi^0$

## II – Variables d'entrée

**Les variables utilisées peuvent être classées en 3 catégories :**

- **Variables de forme du cluster** : cEP, cPP, R9, SigmaPhi, E3x3/(E3x3-E1st-E2nd)
- **Variables 'd'environnement'** : Pt\_phot/Pt\_jet, dR(photon, closest SC), dR(photon, closest track)
- **Variables de conversion** : Hcal energy around tracks

**L'étude va être divisée en 3 cas :**

**Photons convertis avec 1 trace** : cPP, R9, E3x3/(E3x3-E1st-E2nd),

Pt\_phot/Pt\_jet, dR(photon, closest SC), dR(photon, closest track) *6 (variables)*

**Photons convertis avec 2 traces** : cPP, SigmaPhi, E3x3/(E3x3-E1st-E2nd),

Pt\_phot/Pt\_jet, dR(photon, closest SC), dR(photon, closest track), HcalENearConvTrack  
*(7 variables)*

**Photons non convertis** : cPP, R9, SigmaPhi, E3x3/(E3x3-E1st-E2nd),

Pt\_phot/Pt\_jet, dR(photon, closest SC), dR(photon, closest track) *7 (variables)*

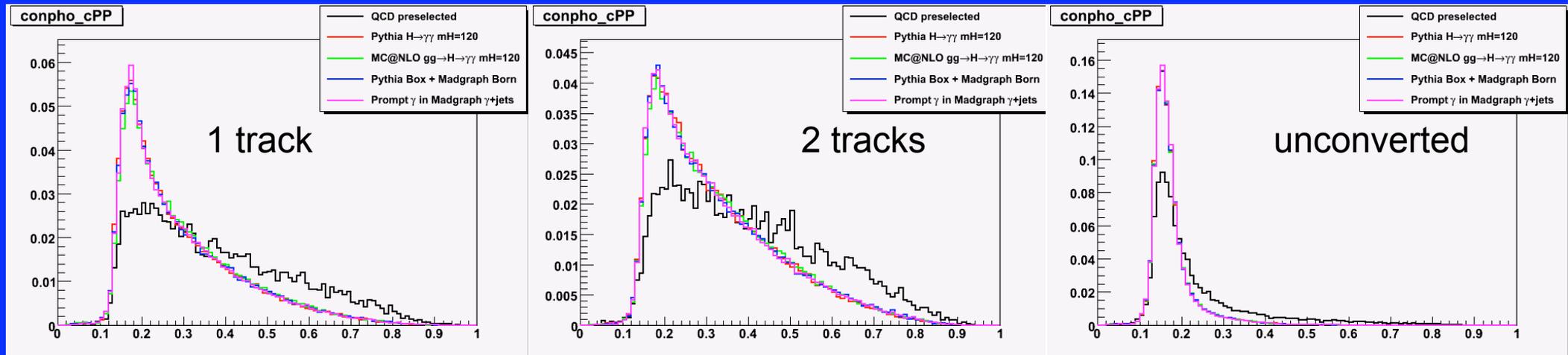
# Discrimination $\gamma / \pi^0$

## II – Variables d'entrée

### cPP (variable de forme du cluster)

cPP (correlation phi-phi) est une variable de forme du basiccluster. C'est l'élément diagonal du bas de la matrice de corrélation entre  $\phi_i$  et  $\phi_j$  de chaque cristal à l'intérieur du basiccluster de plus grande énergie

Pour les photons qui se convertissent, cPP est plus grand puisque le champ magnétique dévie les électrons dans la direction phi.



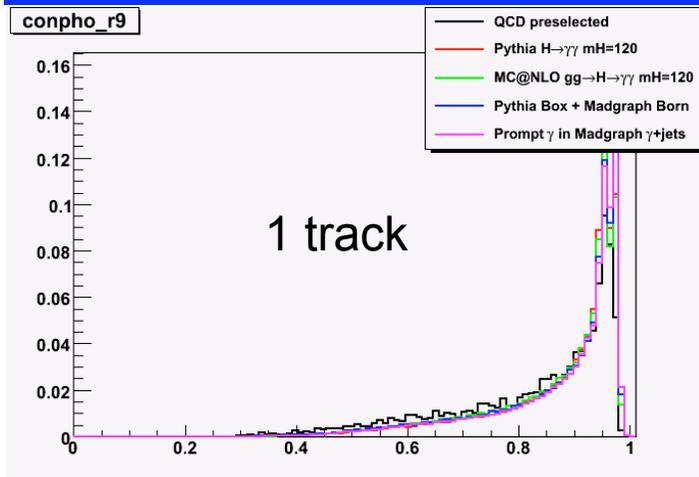
# Discrimination $\gamma / \pi^0$

## II – Variables d'entrée

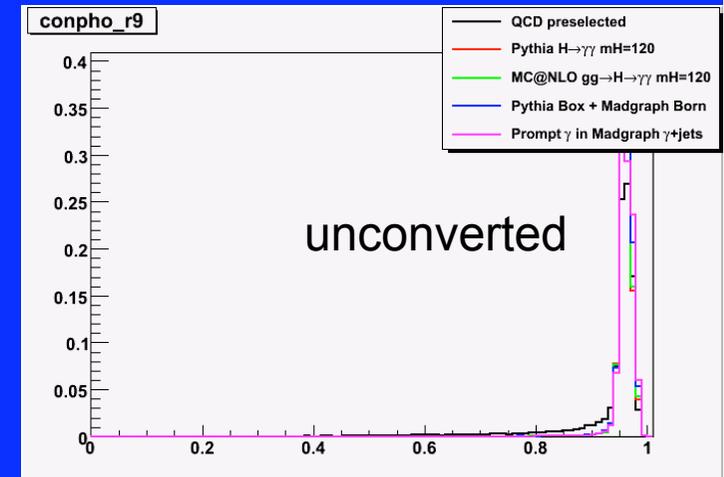
### R9 (variable de forme du cluster)

C'est le ratio de l'énergie collectée dans les cristaux 3x3, divisée par l'énergie du supercluster.

Quand R9 augmente, la plus grande partie de l'énergie se concentre dans dans 3x3 cristaux.  
Usuellement on considère que  $R9 > 0.93$  est le signe de photons non-convertis.



Not used in 2 tracks case



# Discrimination $\gamma / \pi^0$

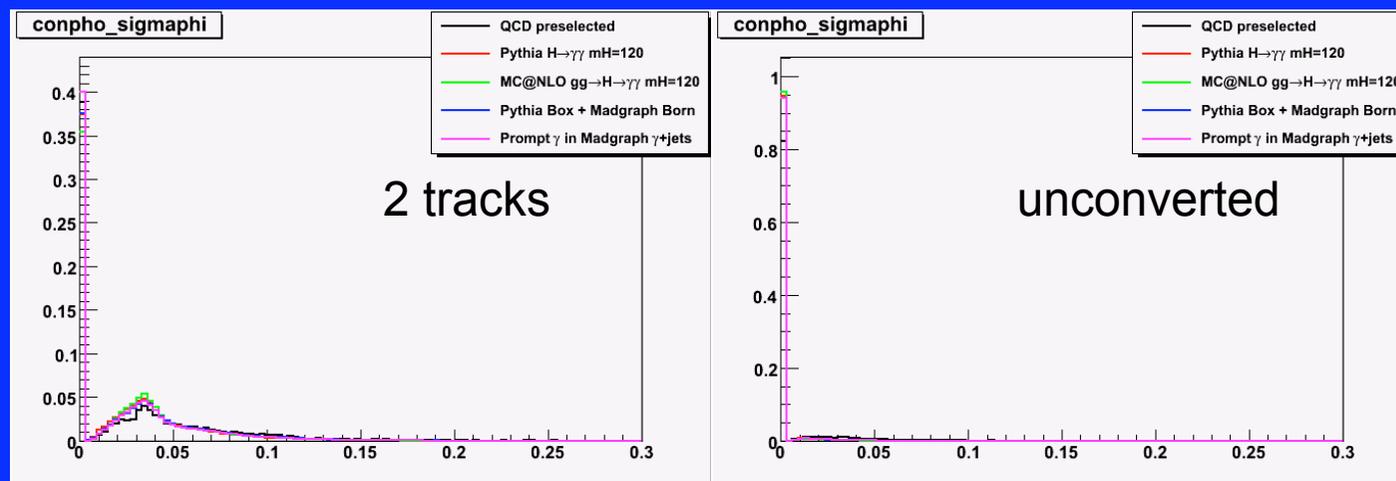
## II – Variables d'entrée

### Sigma\_Phi (variable de forme du cluster)

$$\sigma_{\phi} = \sum_i \sqrt{\frac{(\phi_i - \phi_{SC}) E_{BC,i}}{E_{SC}}}$$

SigmaPhi est la largeur naturelle du supercluster en phi, utilisée pour sa construction. Dans le cas d'un photon non-converti, où l'énergie se concentre souvent dans un seul basiccluster, on s'attend à ce que cette variable soit proche de 0.

Not used in 1 track case



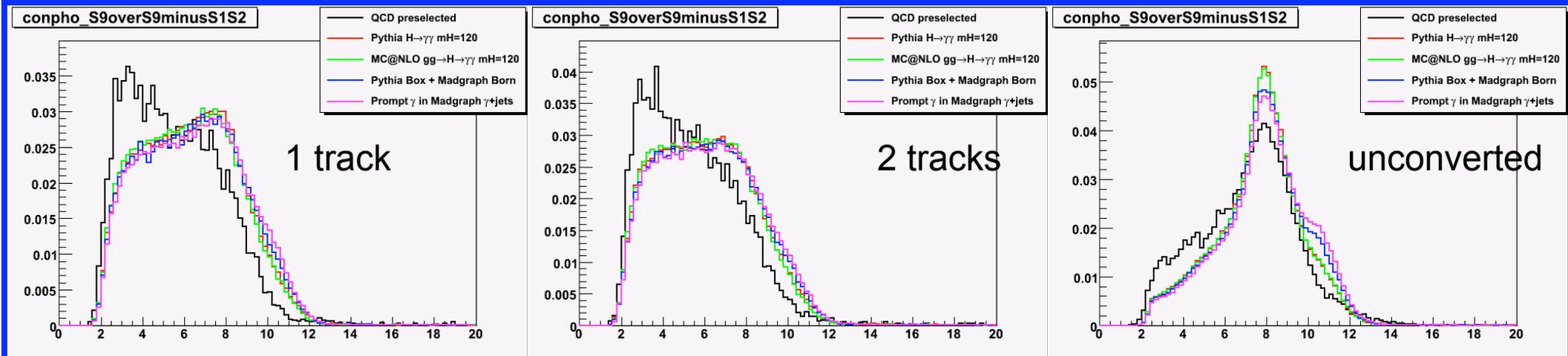
# Discrimination $\gamma / \pi^0$

## II – Variables d'entrée

### $E_{3 \times 3} / (E_{3 \times 3} - E_{1st} - E_{2nd})$ (variable de forme du cluster)

C'est une variable maison de forme du cluster, égale à l'énergie dans 3x3 cristaux divisée par cette même énergie moins l'énergie du premier et du deuxième cristal, dans le basiccluster de plus haute énergie.

$E_{3 \times 3} / (E_{3 \times 3} - E_{1st} - E_{2nd})$  caractérise l'asymétrie du dépôt d'énergie dans ce cluster.



# Discrimination $\gamma / \pi^0$

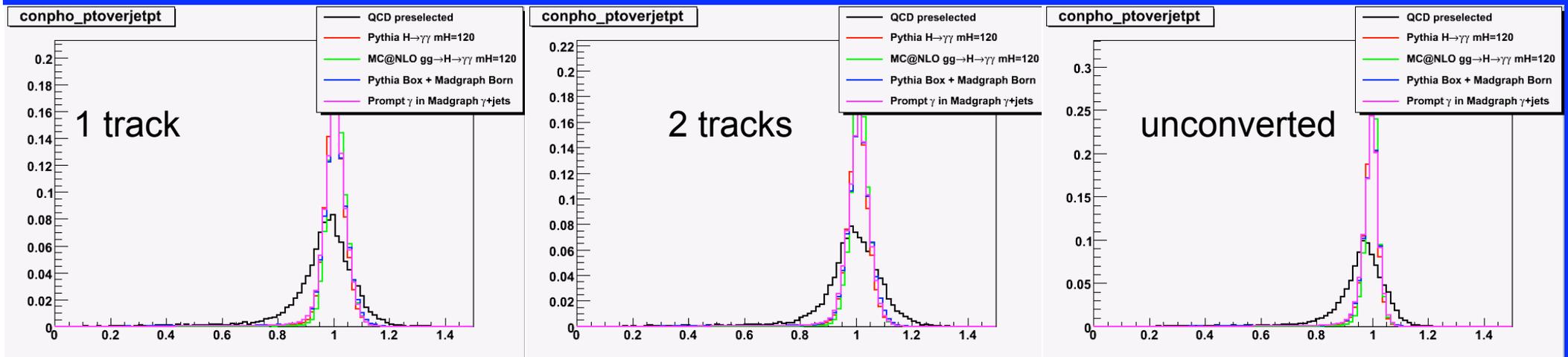
## II – Variables d'entrée

### $Pt_{\text{gamma}}/Pt_{\text{jet}}$ (variable d'environnement)

Cette variable est le pt du photon, divisé par le pt du jet calorimétrique le plus proche, avec  $pt_{\text{jet}} > 15$  GeV.

Pour le signal, le jet est toujours un faux jet (c'est +/- un photon, incluant quelques dépôts qu'il y a autour), c'est pourquoi la distribution est centrée en 1.

Pour le bruit de fond, ce peut être un jet réel incluant des  $\pi^0$  et d'autres particules. Le pic est donc plus étalé.



# Discrimination $\gamma / \pi^0$

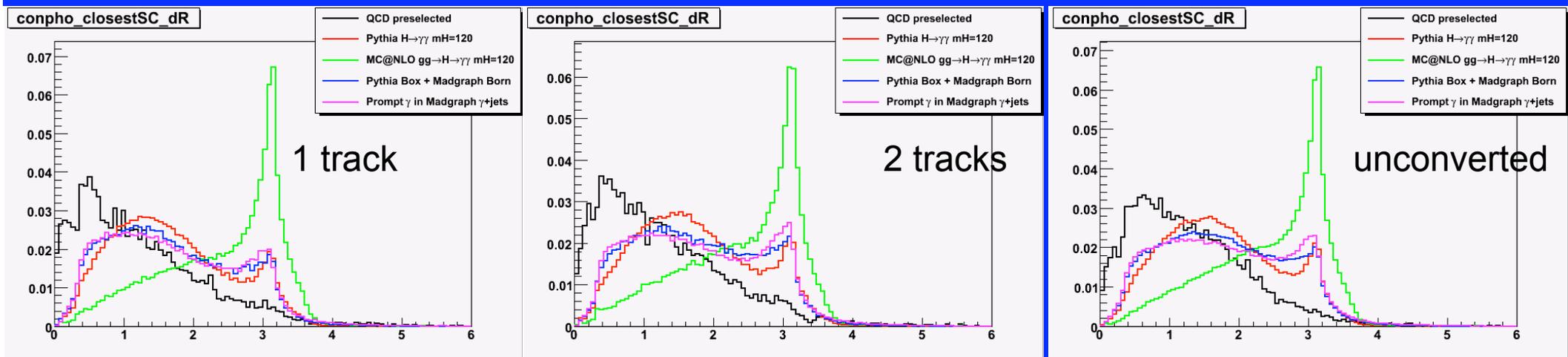
## II – Variables d'entrée

### $dR(\text{photon, closest SC})$ (variable d'environnement)

C'est le  $\Delta R$  entre le photon et le supercluster le plus proche dans le ECAL.

Pour le signal, ce SC peut être l'autre prompt photon (pic vers  $\sim 3$ ), alors que pour le bruit de fond, la distribution est beaucoup plus étalée

Si le photon vient de la désintégration d'un  $\pi^0$ , cette variable est signe de l'activité des jets autour du photon.



# Discrimination $\gamma / \pi^0$

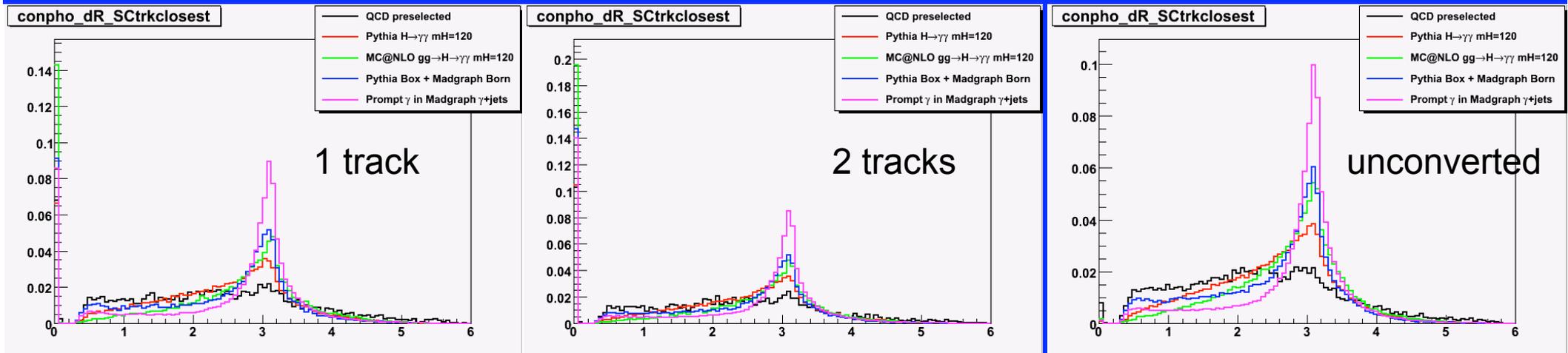
## II – Variables d'entrée

### $dR(\text{photon, closest track})$ (variable d'environnement)

C'est le  $\Delta R$  entre le photon et la trace la plus proche dans le Trajectographe (avec  $pt_{\text{trk}} > 10$  GeV).

Cette variable est indicatrice de l'activité des jets dans le trajectographe.

Le pic vers  $dR \sim 0$  est dû aux traces provenant de conversions précoces, qui ont 0 ou 1 coup dans le pixel.



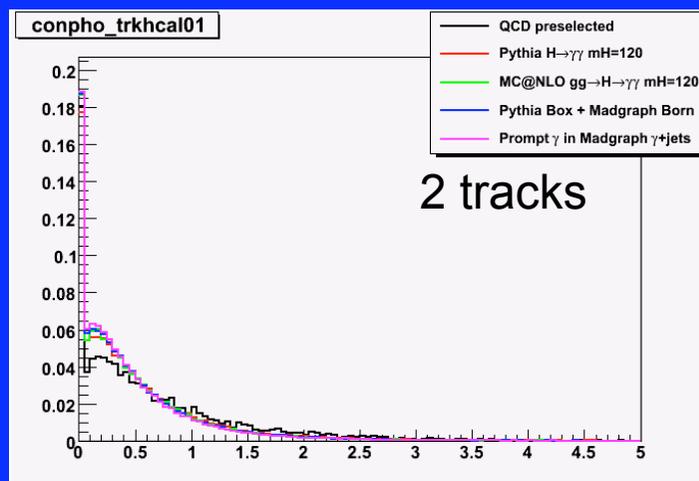
# Discrimination $\gamma / \pi^0$

## II – Variables d'entrée

### Energie dans le HCAL autour des traces (variable de conversion)

C'est la somme des coups dans le HCAL autour des traces des electrons issus de la conversion du photon, dans un cône  $dR < 0.1$ .

Si la trace vient d'un  $\pi^0$ , il devrait y avoir plus d'activité des jets autour d'elle.



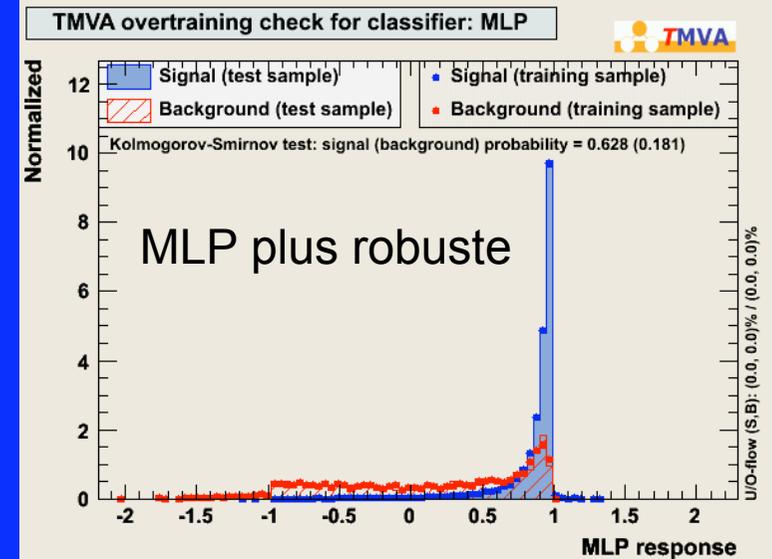
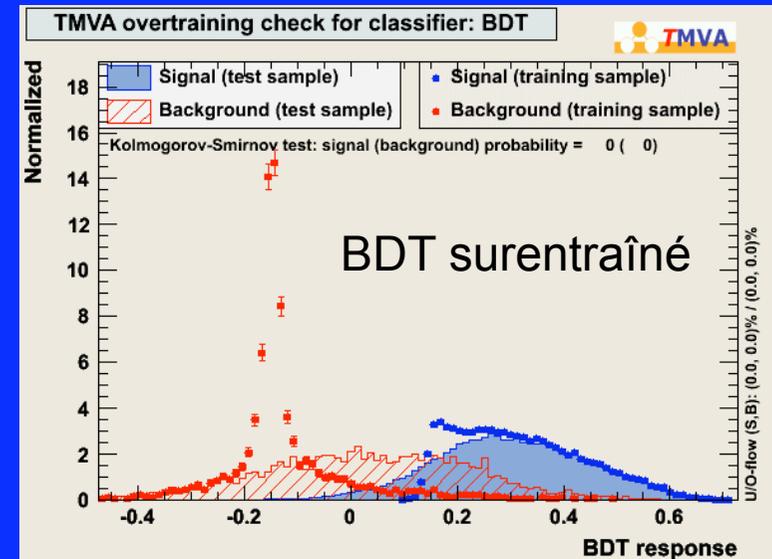
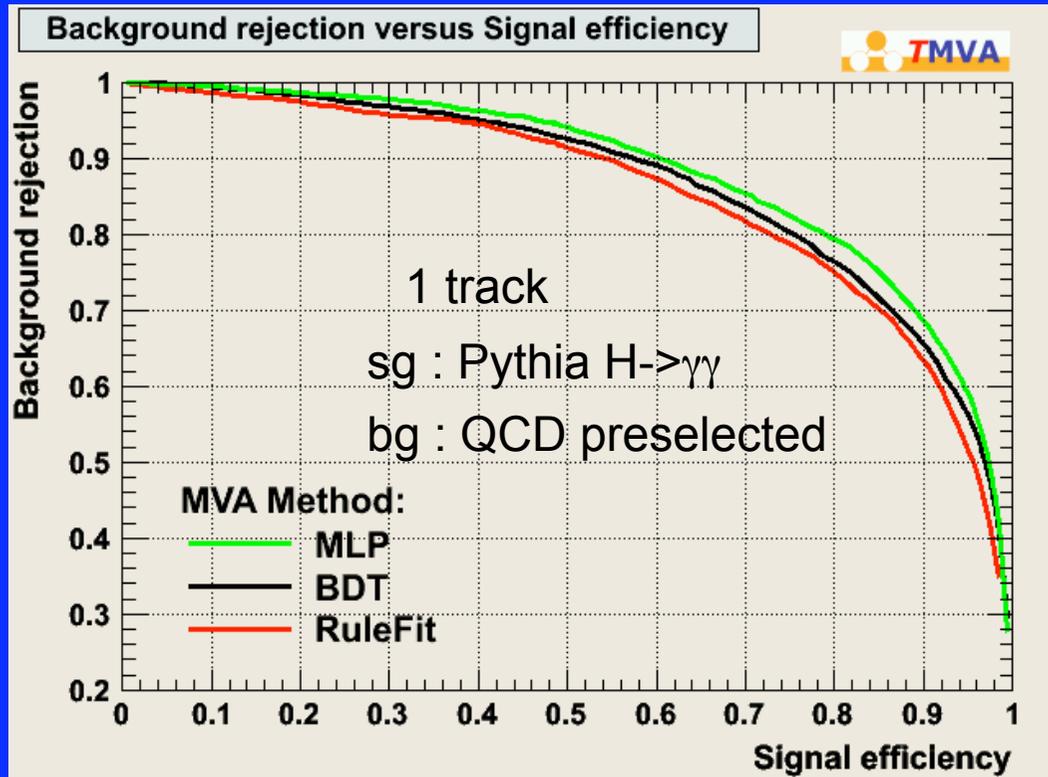
Not used in 1 track case

Not used in unconverted case

# Discrimination $\gamma / \pi^0$

## III – Résultats de l'analyse multivariée

Cas où seulement 1 electron issu de la conversion a été retrouvé

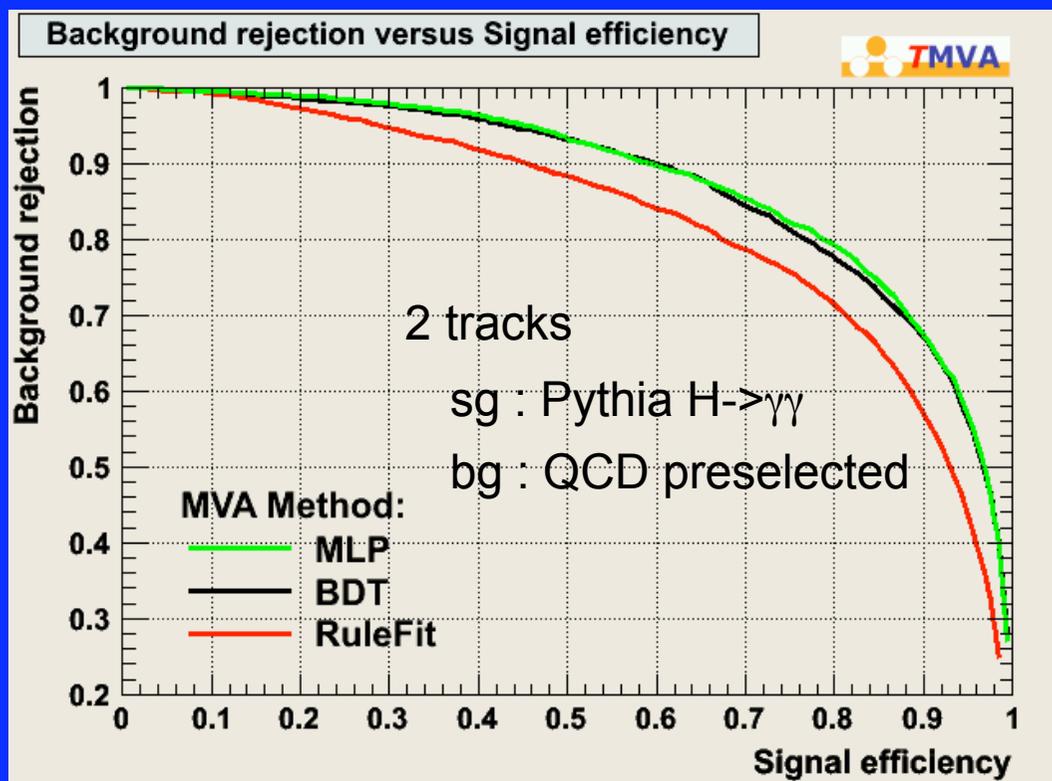


- Le meilleur résultat est obtenu avec le MLP (Multi Layer Perceptron, réseau de neurone)
- Dans le cas 1 trace, en gardant 90%

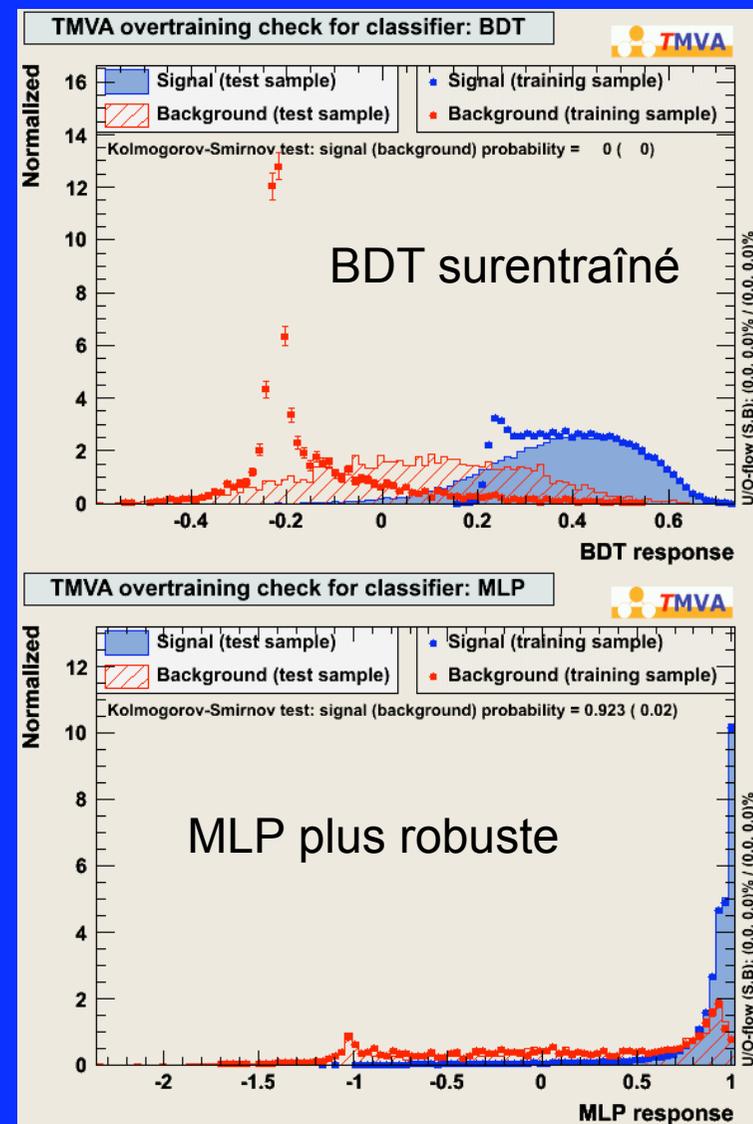
# Discrimination $\gamma / \pi^0$

## III – Résultats de l'analyse multivariée

Cas où la paire  $e^+e^-$  de la conversion a été retrouvé



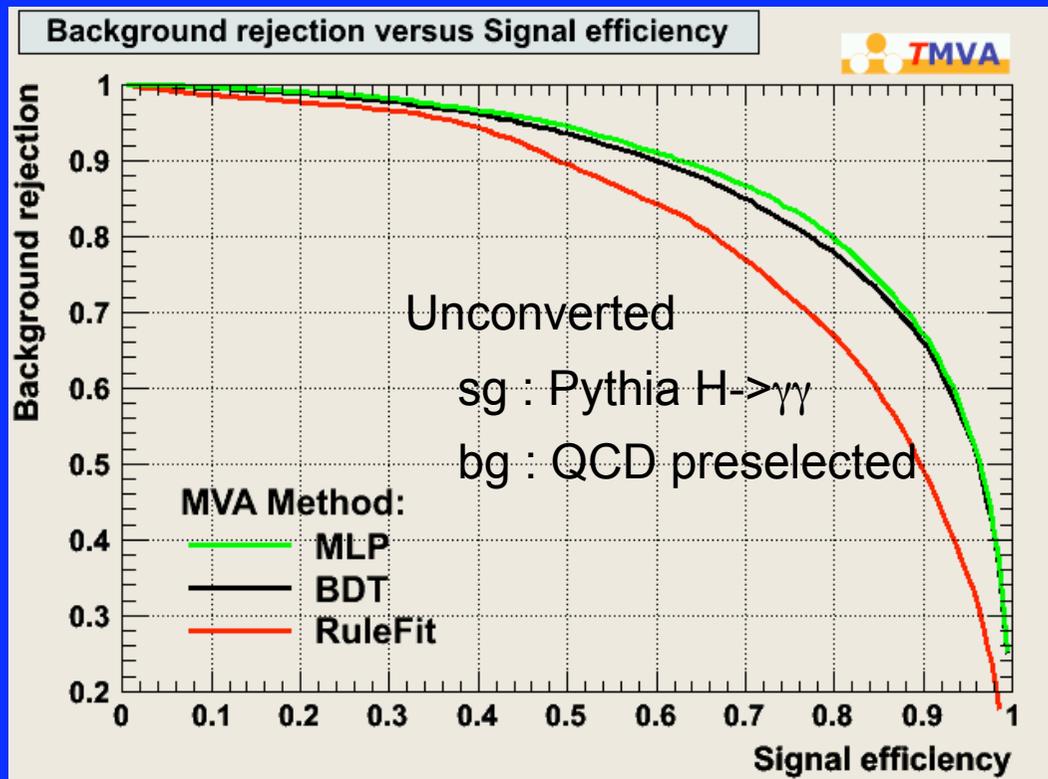
- Dans le cas 2 traces, en gardant 90% signal, 67% de rejection du bruit de fond QCD est obtenu.



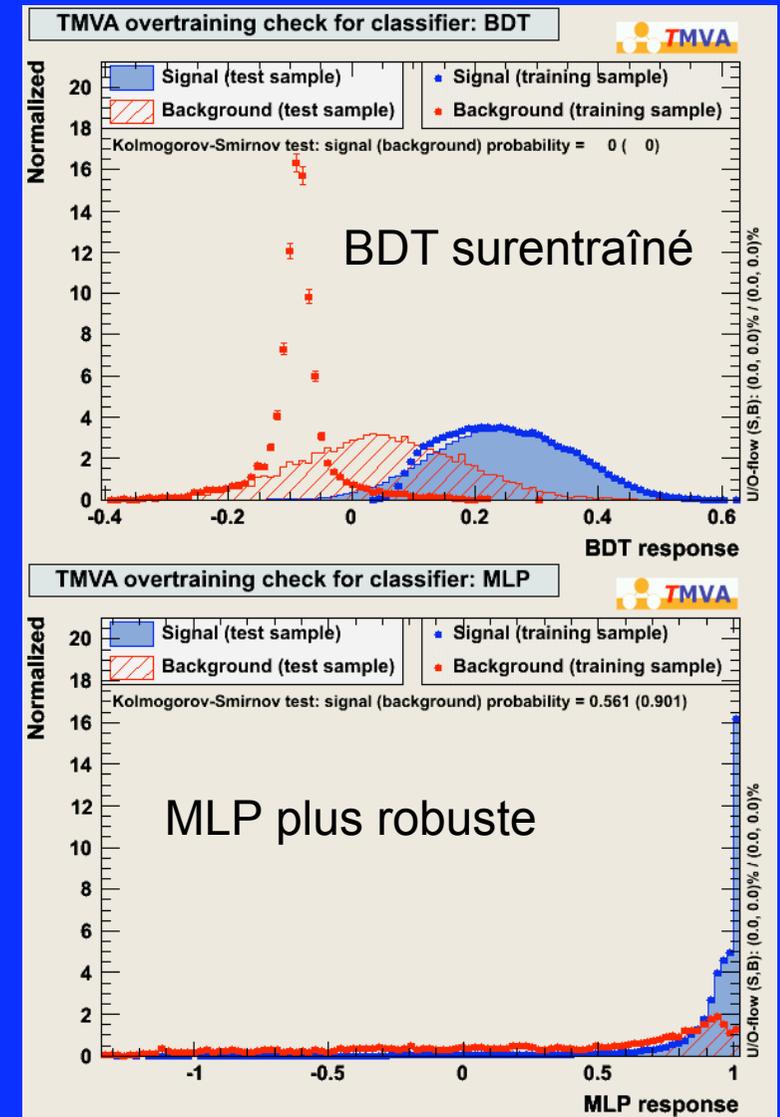
# Discrimination $\gamma / \pi^0$

## III – Résultats de l'analyse multivariée

Cas où le photon ne s'est pas converti

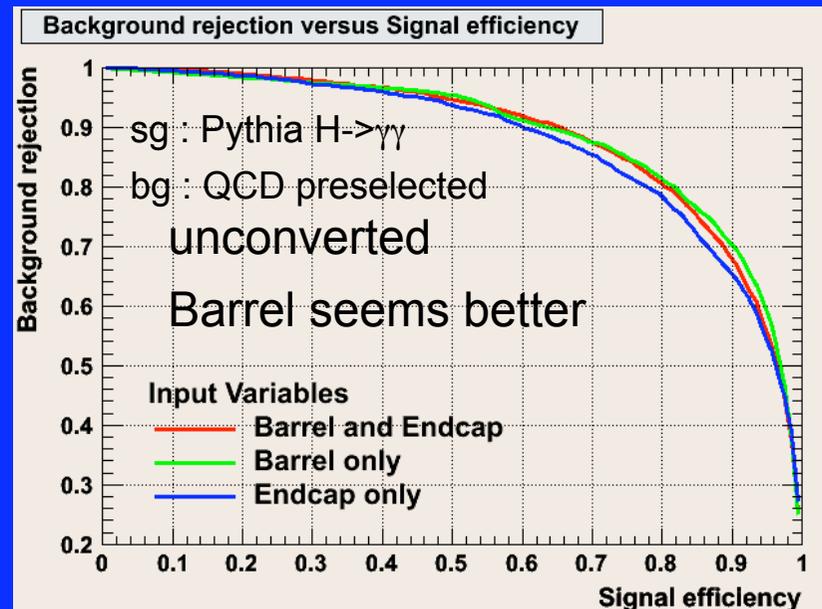
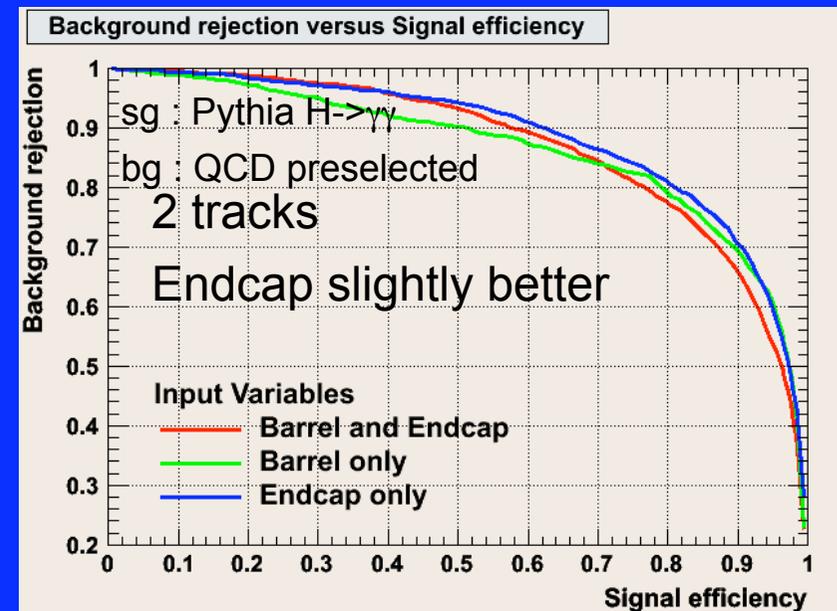
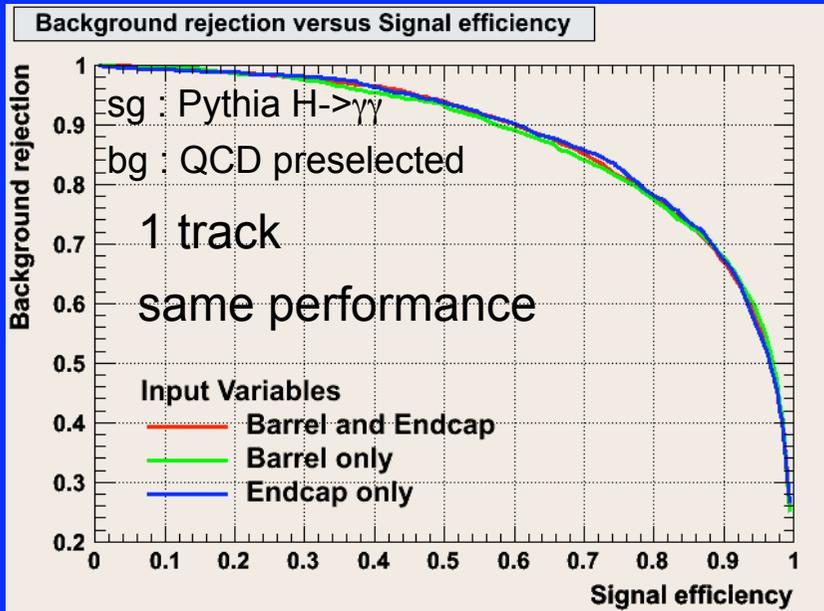


- Dans le cas non converti, en gardant 90% du signal, 67% de rejection du bruit de fond QCD est obtenu.



# Discrimination $\gamma / \pi^0$

## III – Résultats de l'analyse multivariée Comparaison tonneau/bouchon

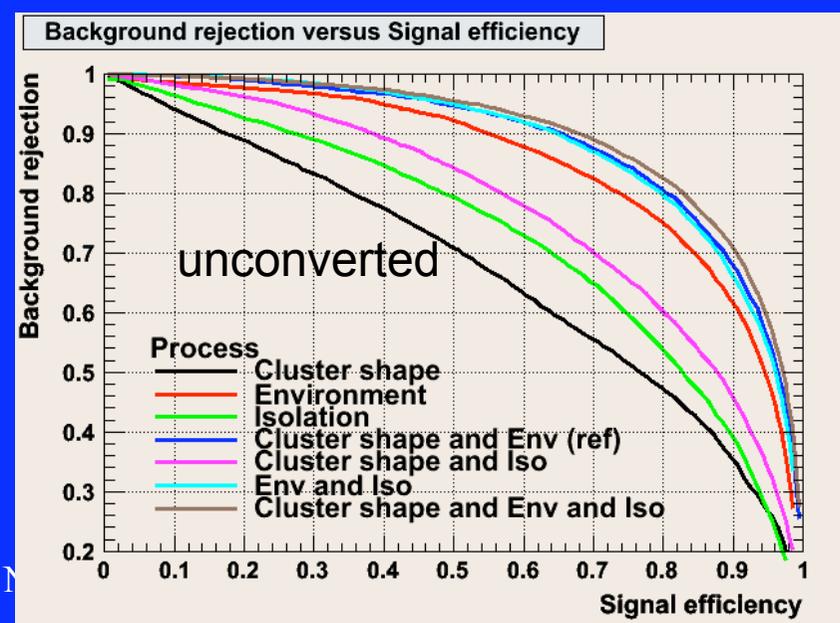
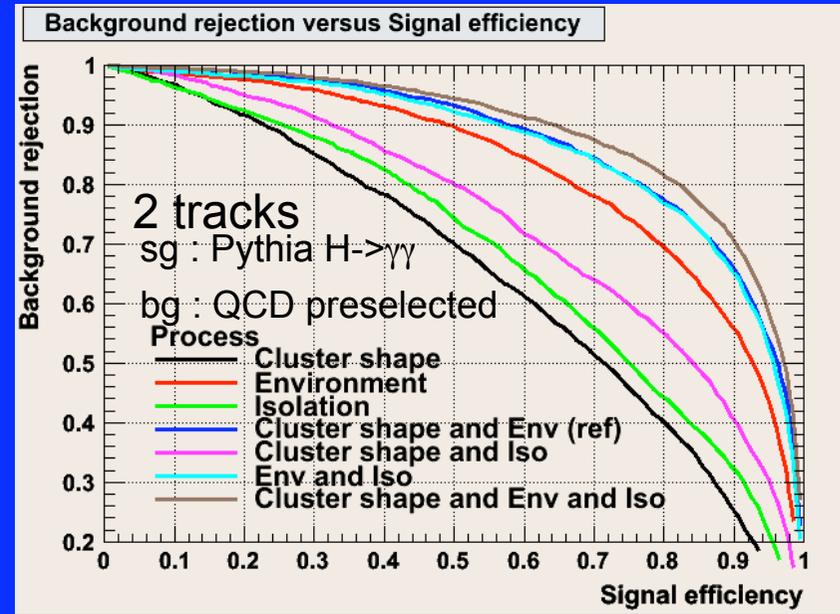
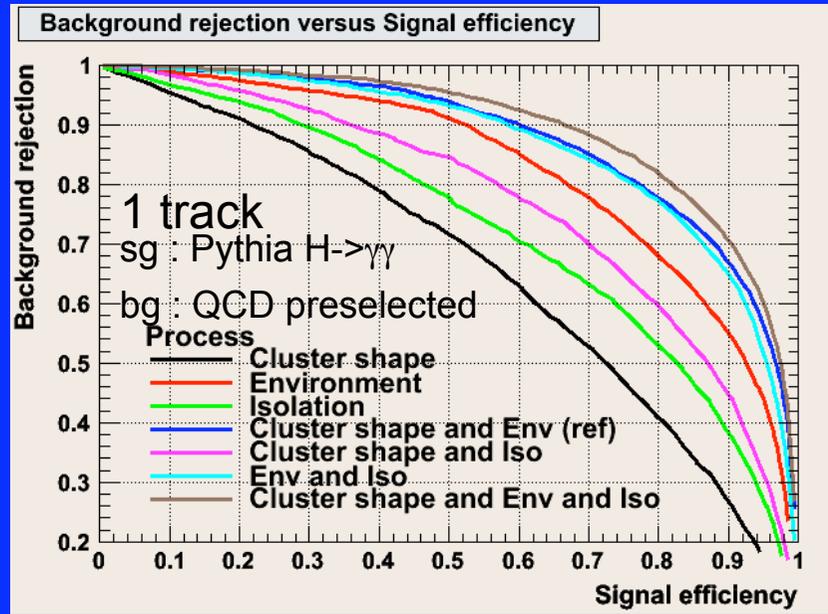


Tonneau et bouchon  
entraînés séparément

L'échantillon 'QCD  
preselected' contient les  
données du détecteur  
pied de gerbe

# Discrimination $\gamma / \pi^0$

## III – Résultats de l'analyse multivariée Effets des différents types de variables



Forme du cluster :

cPP

R9, s9/(s9-s1-s2), SigmaPhi

Environment : pt\_phot/

pt\_jet, dR(phot,

close

t

Pour chaque courbe, le NN a été réentraîné

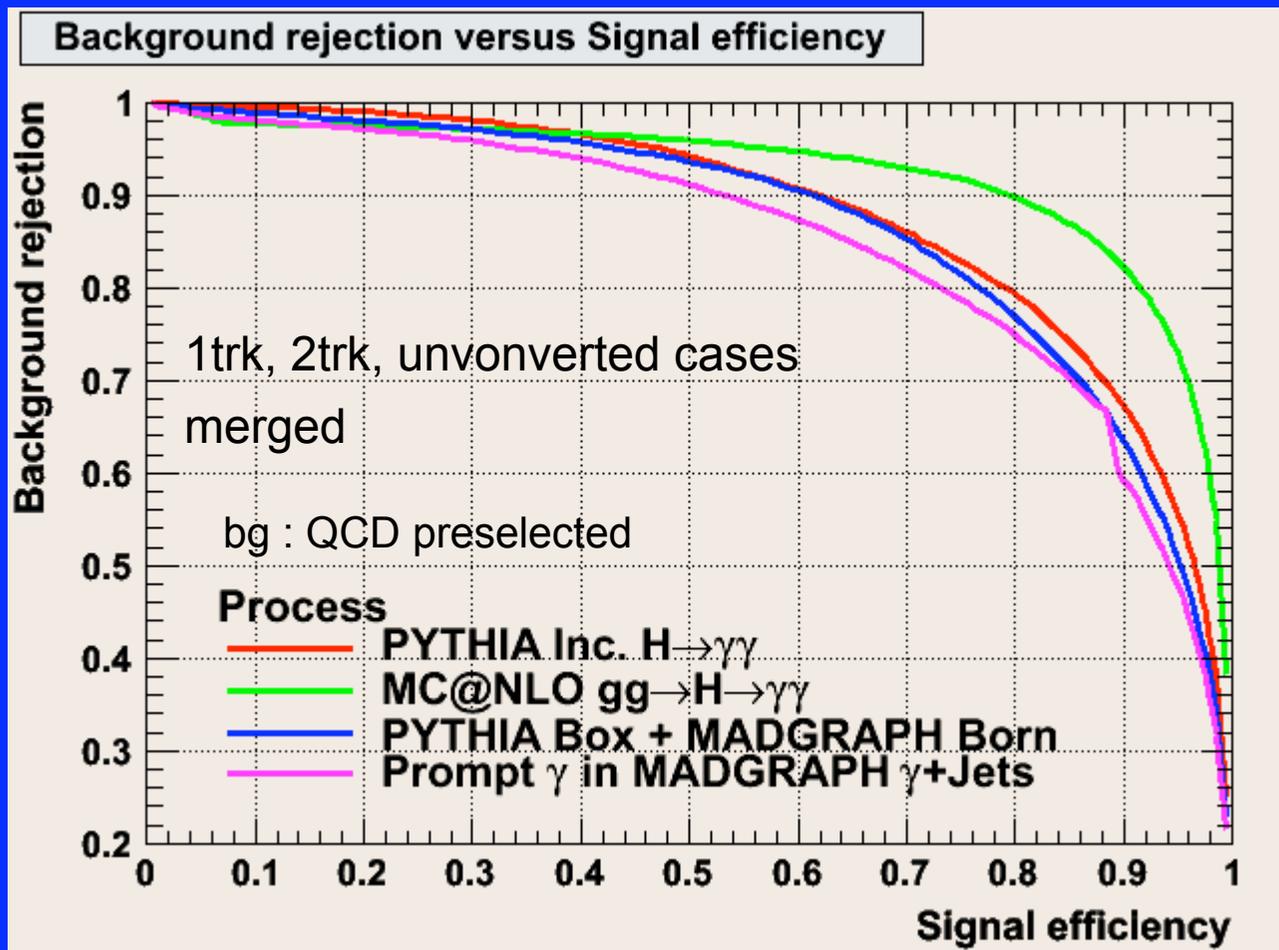
Les variables d'environnement donnent des meilleurs résultats que celles de forme du cluster ou d'iso

# Discrimination $\gamma / \pi^0$

## III – Résultats de l'analyse multivariée Rassemblement des 3 cas en un seul

- Rassemblement des cas 1 trace, 2 traces, non convertis en un seul cas. On utilise toutes les variables de la liste, qui ne concernent pas les photons convertis. *7 variables*.

La performance obtenue reste très proche : a-t-on vraiment besoin de diviser l'étude en 3 cas ?



- Signal **diphoton** (Born+Box) :  
64% de rejection pour 90%  
d'efficacité de signal

- **Prompt  $\gamma$  dans  $\gamma$ +jets** : 60%  
de rejection

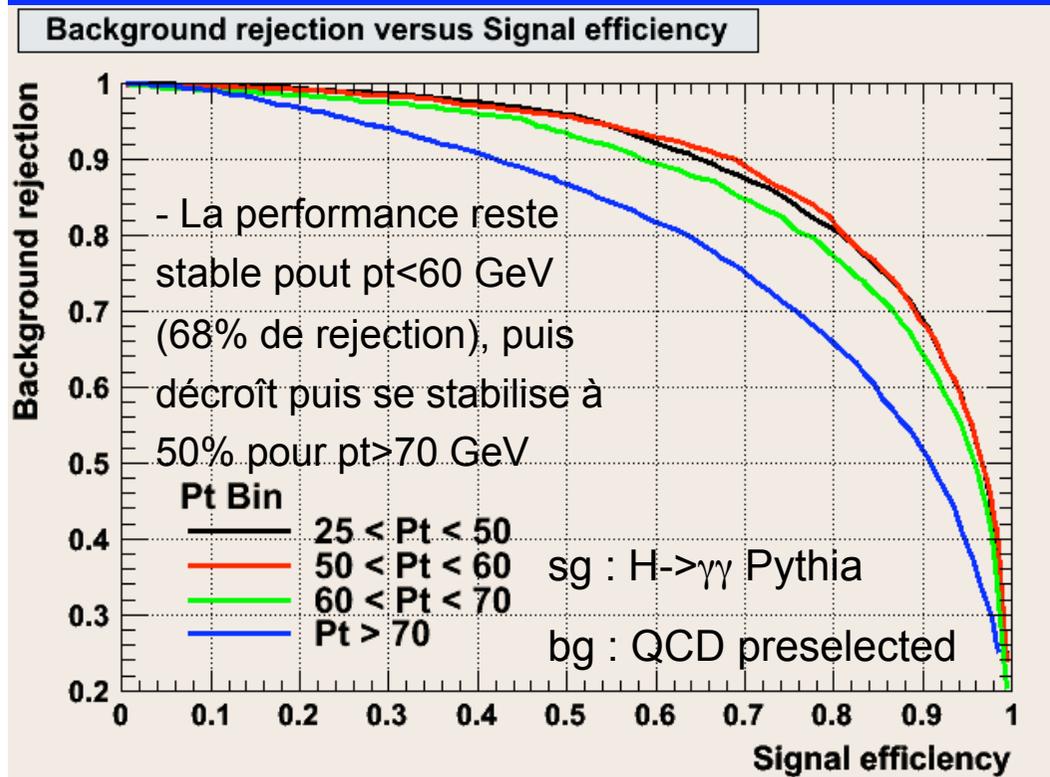
- **Pythia  $H \rightarrow \gamma\gamma$**  : 67% de  
rejection

- **MC@NLO  $gg \rightarrow H \rightarrow \gamma\gamma$**  : 83%  
de rejection

# Discrimination $\gamma / \pi^0$

## III – Résultats de l'analyse multivariée Performance en fonction de pt

Photons convertis et non-convertis :



### ANALYSES DEJA EXISTANTES :

Photons non-convertis :

K. Karafasoulis, A. Kyriakis, D. Maletic, **AN2008/063**

<b>barrel</b>	isolation and unconversion using " $N_{trk}^{ConvID} = 0$ "
$E_T$ [GeV]	$\pi^0$ rejection for 90% photon efficiency (%)
20-25	$70.0 \pm 1.4$
25-35	$62.3 \pm 0.9$
35-45	$43.4 \pm 1.0$
45-55	$33.2 \pm 0.7$
55-65	$29.3 \pm 0.6$

<b>endcap</b>	isolation and unconversion using " $N_{trk}^{ConvID} = 0$ "
$E_T$ [GeV]	$\pi^0$ rejection for 90% photon efficiency (%)
20-25	$64.1 \pm 2.0$
25-35	$62.1 \pm 1.3$
35-45	$57.1 \pm 1.2$
45-55	$54.4 \pm 1.2$
55-65	$51.8 \pm 1.0$

Photons convertis :

G. Anagnostou, Gamma conversion meeting, 28/04/08, ~57% de rejection

n

# Discrimination $\gamma / \pi^0$

## IV – Les Pi0 dans les premières collisions de CMS

**Recherche de Pi0** : 2 analyses indépendantes dans le tonneau

Le déclenchement n'est pas fait avec ECAL, mais avec le trigger faisceau. 191 évènements

### Première analyse :

- Utilisation de clusters 3x3 cristaux
- $\gamma$  :  $eT > 300$  MeV
- $\pi^0$  :  $eT > 900$  MeV
- $E_{2 \times 2} / E_{3 \times 3} > 0.85$

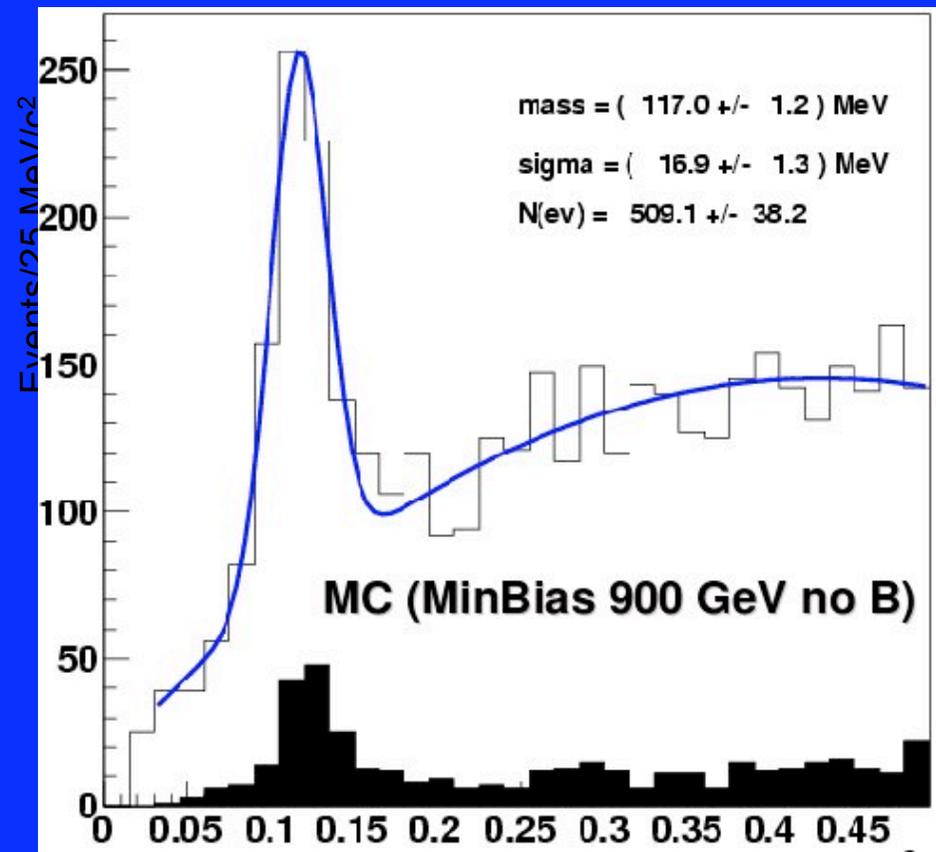
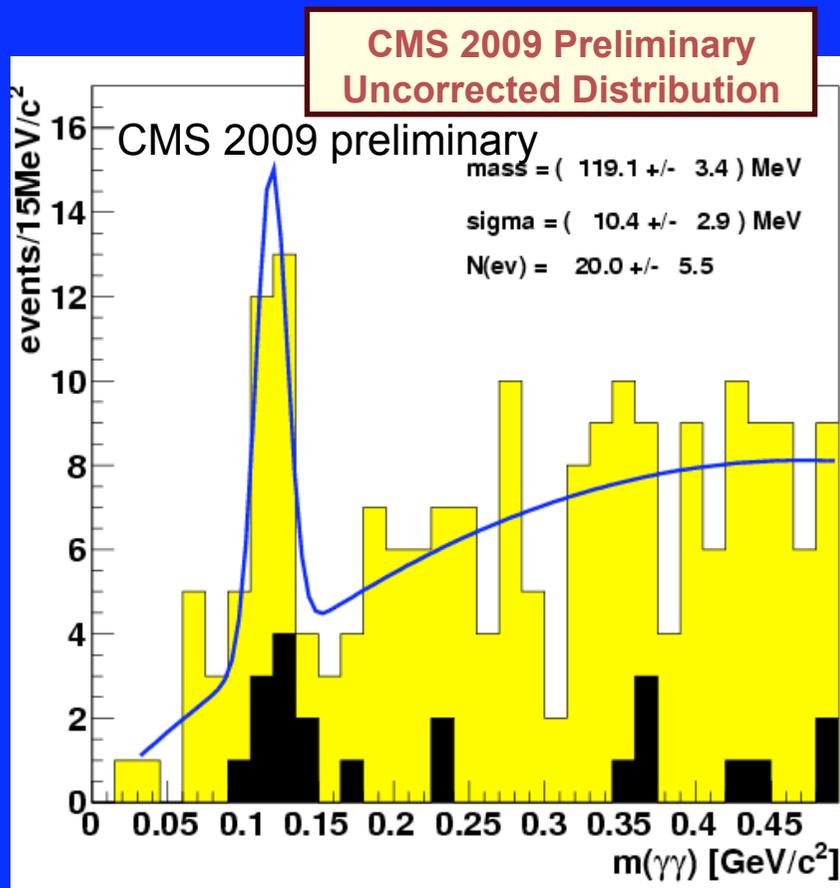
### Deuxième analyse :

- Superclusters « Island »
- $\gamma$  :  $eT > 400$  MeV
- $eT_1 + eT_2 > 900$  MeV
- $\cos(\theta) > 0.7$



# Discrimination $\gamma / \pi^0$

## IV – Les $\pi^0$ dans les premières collisions de CMS



- Première distribution diphoton dans CMS

- Le pic de la  
dis

t

# Discrimination $\gamma / \pi^0$

## Conclusions

- En gardant 90% du signal, la rejection du bruit de fond atteint ~67% pour H- $\rightarrow\gamma\gamma$  inclusif Pythia, ~83% pour gg- $\rightarrow$ H- $\rightarrow\gamma\gamma$  MC@NLO, ~64% pour Pythia Box + Madgraph Born, ~60% pour les prompt  $\gamma$  de Madgraph  $\gamma$ +jets.
- La performance est meilleure que lors des études précédentes dans CMS sur les photons convertis et non-convertis (avec une même technique et un nombre de variables plus petit).
- La rejection décroît puis reste stable lorsque le pt des photons augmente (encore 50% de rejection pour des photons ayant pt>70 GeV).
- La plus grande partie de la performance vient des variables d'environnement

## Perspectives

- Evaluer la performance de l'analyse à partir des premières données, sur des lots de Pi0 et photons avec des coupures relâchées.
- Essayer des MC-toy experiment pour évaluer la performance avec 10pb-1 d'évènements  $\gamma$ +jets
- Intégrer et vérifier la performance de cette analyse multivariée dans les analyses  $\gamma$ +jets,  $\gamma\gamma$  et H- $\rightarrow\gamma\gamma$

# BACK-UP SLIDES

# Discrimination $\gamma / \pi^0$

## I – Sélection des photons

### Qui sont les photons sélectionnés ?

Appariement avec les particules niveau générateur (après préselection) :

Dans un cône  $dR(\text{phot\_reco}, \text{part\_MC}) < 0.1$ , on apparie le photon reconstruit avec la particule niveau générateur qui a le pt le plus proche. On regarde de quelle désintégration elle est issue.

(No MC info in QCD preselected)

#### Signal Pythia inc. $H \rightarrow \gamma\gamma$

Prompt photon (issu du Higgs)	99.8 %
Photon non partonique d'un quark (différent de ci-dessous)	0.1 %
Photon d'un $\pi^0$	0.02 %
Photon d'un eta	0.007 %
Photon d'un rho	0 %
Photon d'un omega	0.0009 %

#### QCD madgraph (all HT bins)

Photon partonique	0 %
Photon non partonique d'un quark (différent de ci-dessous)	10.1 %
Photon d'un $\pi^0$	68.1 %
Photon d'un eta	17.1 %
Photon d'un rho	0.1 %
Photon d'un omega	4.6 %

## Sélection des traces pour les photons convertis

### Selection of the best tracks among all converted photon track candidates :

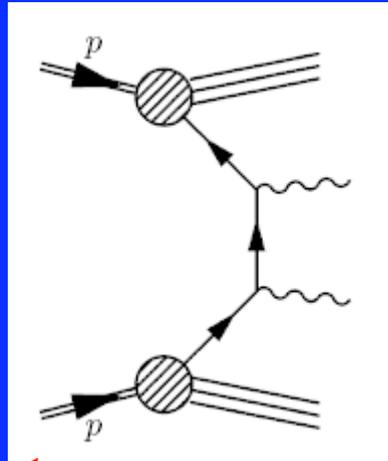
#### CMSSW\_2\_2\_9 (official selection) :

- 1 track case : EoverP closest to 1
- 2 track case : Use of a likelihood (T. Kolberg, N. Marinelli) which takes as input EoverP,  $\text{abs}(\cot_{\theta})$ ,  $\text{abs}(\Delta\phi)$ ,  $\text{chi2}_{\text{max\_pt}}$ ,  $\text{chi2}_{\text{min\_pt}}$

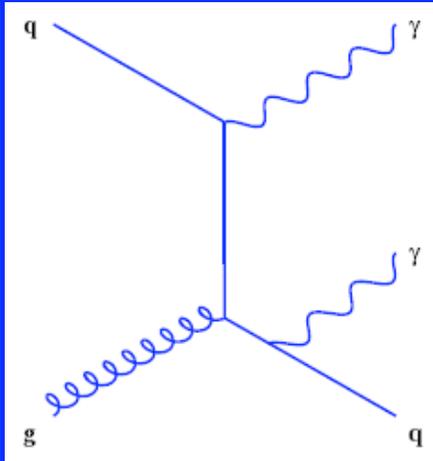
# Higher order $\gamma\gamma$ Generator Studies

## NLO generators of $\gamma\gamma$ processes

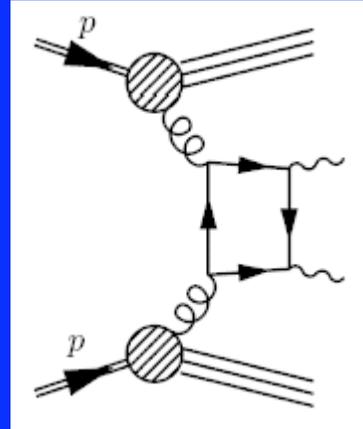
**BORN**



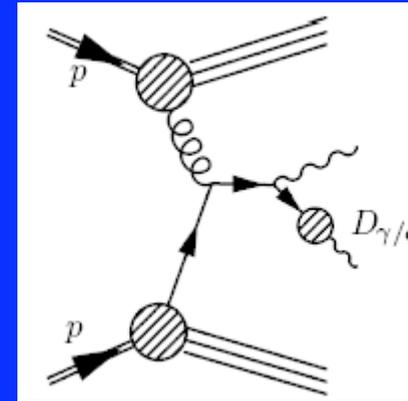
**BREM**



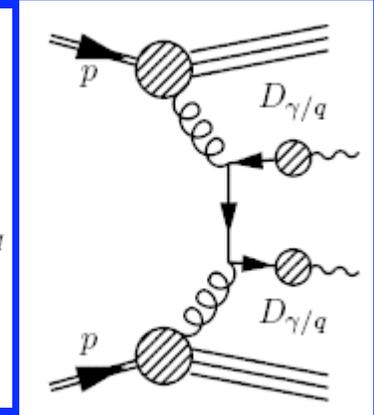
**BOX**



**ONE FRAG**



**TWO FRAG**



**DIRECT**

**BOX :**  
**gamma2MC & Resbos** [Csaba Balazs & C.-P. Yuan]: **LO and NLO**

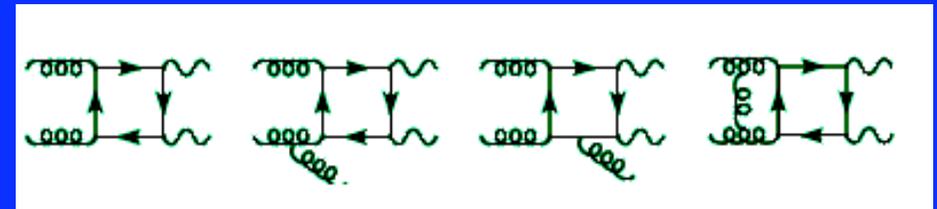
Box [already NLO] + corrections [up to N3LO]

NLO codes

	type of code	Direct	One Frag.	Two Frag.
Aurenche et al.	I/FO	NLO	LO	none
Owens et al.	G/FO	NLO	LO	none
DIPHOX (*)	G/FO	NLO	NLO	NLO
RESBOS	G/SGS	NLO	LO	none

I : Inclusive  
 G : Generator  
 FO : Fixed Order  
 SGS: Soft Gluon Summation

(\*) [http://wwwlapp.in2p3.fr/lapth/PHOX\\_FAMILY/main.html](http://wwwlapp.in2p3.fr/lapth/PHOX_FAMILY/main.html)



Corrections to box

=> The most complete description would seem to be possible with DIPHOX for all contributions but box, interfaced with Gamma2MC for Box