Recognition of Backgrounds Channels through Machine Learning Methods

U. Zarantonello

Ferrara - June 26, 2025

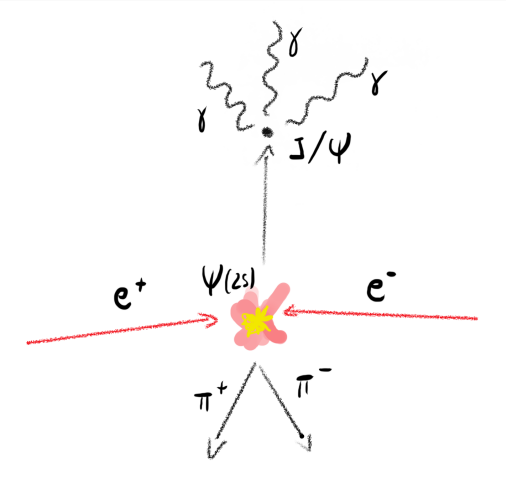
Table of Contents

Table of Contents

Decay Channel in Analysis

The scope of my analysis is to discriminate the decay channel:

$$\psi(2s) \to \pi^+\pi^- \, J/\Psi, \;\; J/\Psi \to \gamma\gamma\gamma \tag{1}$$



The study will use the $\approx 2.3 \cdot 10^9$ events of $\psi(2s)$ retrieved from the BESIII experiment (2009, 2012, 2021).

Present Measures

The previous measurements of the **branching fraction** $\mathbf{B}(J/\psi \to 3\gamma)$ are:

- CLEO Collaboration, 2008: $(1.2 \pm 0.3 \pm 0.2) \times 10^{-5}$
- BESIII, 2013: $(11.3 \pm 1.8 \pm 2.0) \times 10^{-6}$

(BESIII preceding analysis used only the 2009 datas)

The branching fraction theorized using **Lattice QCD** (2020) is:

- $(1.614 \pm 0.016 \pm 0.261) \times 10^{-5}$ with $a \simeq 0.085\,\text{fm}$;
- $(1.809 \pm 0.051 \pm 0.295) \times 10^{-5}$ with $a \simeq 0.067\,\text{fm}$;

where a is the lattice spacing used in the simulation.

Why do we study the process?

There are two main reasons why this analysis is interesting:

1. We have a much **larger dataset**; so the process can be measured with much higher precision (in order to build a complete picture of the decay of J/ψ);
2. In particular for the process $J/\psi \to 3\gamma$, we want to verify the **NRQCD** predictions with higher precision while also confirm the latter **lattice QCD** calculations;
3. In this energy region **glueballs** are predicted to exist, our analysis can therefore help narrow down the phase space where glueball contributions might still be hiding.

Table of Contents

Analysis Overview

The analysis revolve around the **recognition of the J/ψ events**, in order to do so we can recognize the following steps:

1. **Selection of the J/ψ events** using the $\pi^+\pi^-$ tracks as a recoil of $\Psi(2s)$;
2. **Selection of the $J/\psi \to n \cdot \gamma$ decays**.
3. **Selection of pure 3γ events** through the elimination of the intermediate particles events.

NOTICE: The dataset used in today presentation is the **Inclusive Monte Carlo** dataset based only on the 2021 datas.

Fit of $\pi^+\pi^-$ over the J/ψ mass

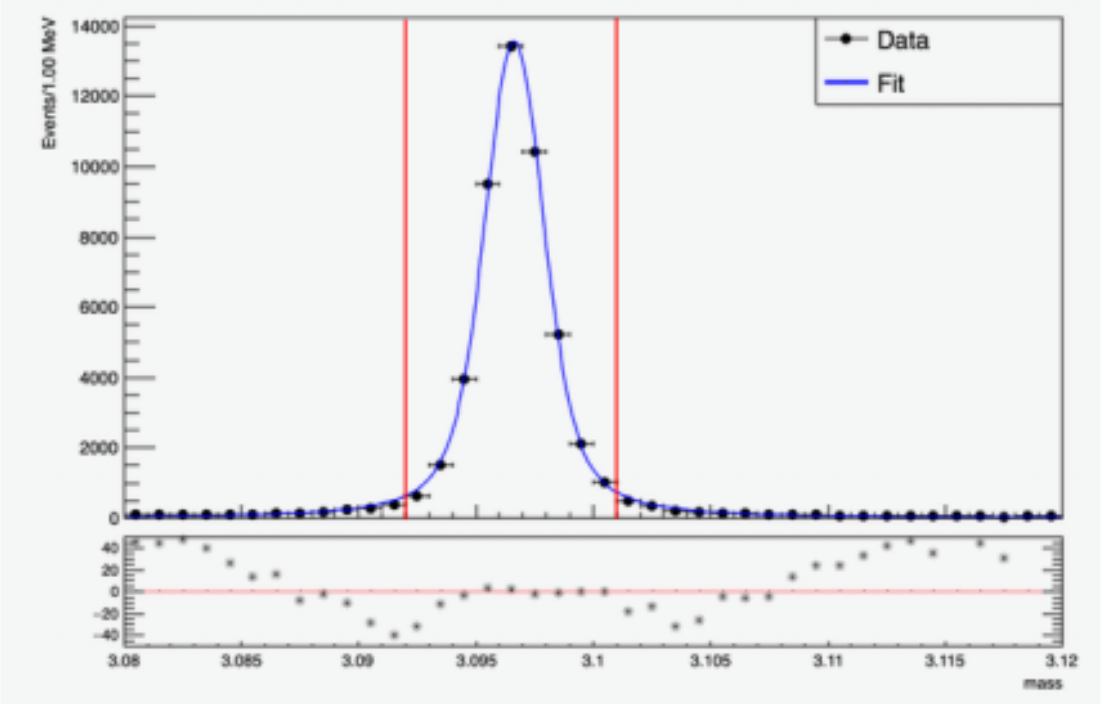


Figure 1: Fit of the $\Psi(2S)$ using the $\pi^+\pi^-$ mass.

We selected the best pion candidates for the J/ψ mass reconstruction optimizing this fit.

Recognition of the J/ψ events.

Fitting the J/ψ mass using the $\pi^+\pi^-$ tracks we were able to impose a cut to the combined mass (in order to reject every event that is not a J/ψ event):

$$3.092\,\mathrm{GeV}/c^2 \leq \mathrm{m}_{\pi\pi} \leq 3.101\,\mathrm{GeV}/c^2$$

.

γ's events selection

The first selection of γ's is composed by some **fiducial cuts** that are typical to gamma signals.

- Barrel cuts:

$$\text{rejected if: } |\cos(\theta)| < 0.80 \ \wedge \ \mathrm{E}_\gamma \leq 0.025\,\mathrm{GeV}$$

- Endcaps cuts:

$$\text{rejected if: } 0.86 < |\cos(\theta)| < 0.92 \ \wedge \ \mathrm{E}_\gamma \leq 0.05\,\mathrm{GeV}$$

- Timing cut:

$$0\,\mathrm{ns} \leq t \leq 700\,\mathrm{ns}$$

In our case we saved only the events that presented **at least three good candidates**.

Kalman Fit of γ signals

The selection of the neutral particles that best reproduce the J/ψ is done using a **Kalman Fit**.

$\Rightarrow$ The fit is done using: 3 γ tracks and the $\pi^+\pi^-$ tracks selected before.

$\Rightarrow$ The only constraint is to compose the mass of the $\Psi(2s)$.

The Kalman fit produced a cut based of the χ^2, each event is considered good if:

$$\chi^2 \leq 45 \ \wedge \ \chi^2 \neq 0\,.$$

Kalman Fit optimization

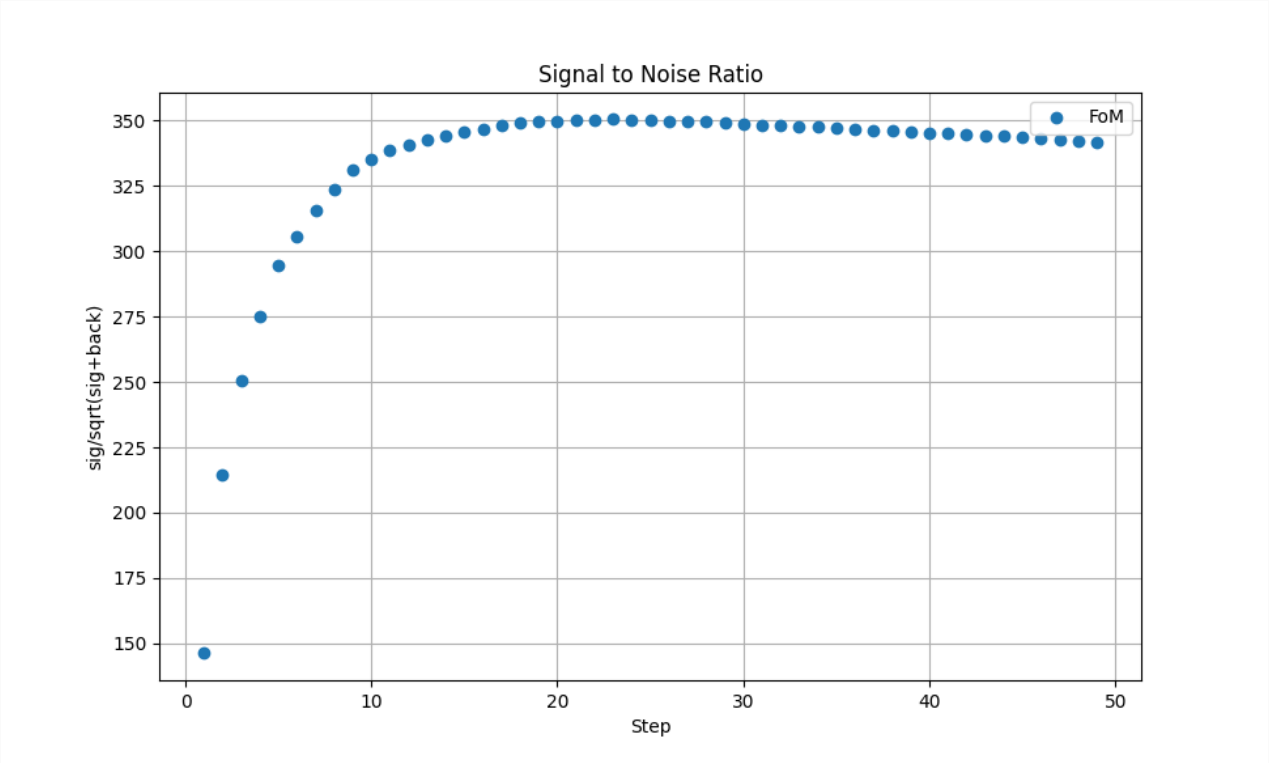


Figure 2: The optimization of this cut had been done by maximizing the significance $S/\sqrt{S+B}$. In this graphic each step correspond to 2 unit of χ^2.

Intermediate Particles Events Rejection

The principal problem of this type of process is the presence of many **intermediate events** that can produce at least 3 γ's. Two essential tools that can be used to recognize this events are: **Topology studies** and **Dalitz Plots**.

Topology table

rowNo	decay tree (decay initial-final states)	iDcyTr	nEtr	nCEtr
1	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \eta\gamma, \eta \to \gamma\gamma$ ($\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma$)	0	7419	7419
2	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \eta'\gamma, \eta' \to \gamma\gamma$ ($\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma$)	3	1865	9284
3	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \pi^0\gamma$ ($\psi' \dashrightarrow \pi^0\pi^+\pi^-\gamma$)	5	678	9962
4	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \pi^0\pi^0\gamma$ ($\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma$)	1	575	10537
5	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_4(2050)\gamma, f_4(2050) \to \pi^0\pi^0$ ($\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma$)	2	358	10895
6	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_2(1270)\gamma, f_2(1270) \to \pi^0\pi^0$ ($\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma$)	7	204	11099
7	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \gamma\gamma\gamma$ ($\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma$)	4	181	11280
8	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_0(1710)\gamma, f_0(1710) \to \pi^0\pi^0$ ($\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma$)	8	120	11400
9	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_0'\gamma, f_0' \to \gamma\gamma$ ($\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma$)	6	58	11458

Figure 3: Table of topology showing the 9 channels with higher count. The dataset used is the Inclusive Monte Carlo one after the Kalman selection.

Dalitz plot of Signal vs Inclusive MC

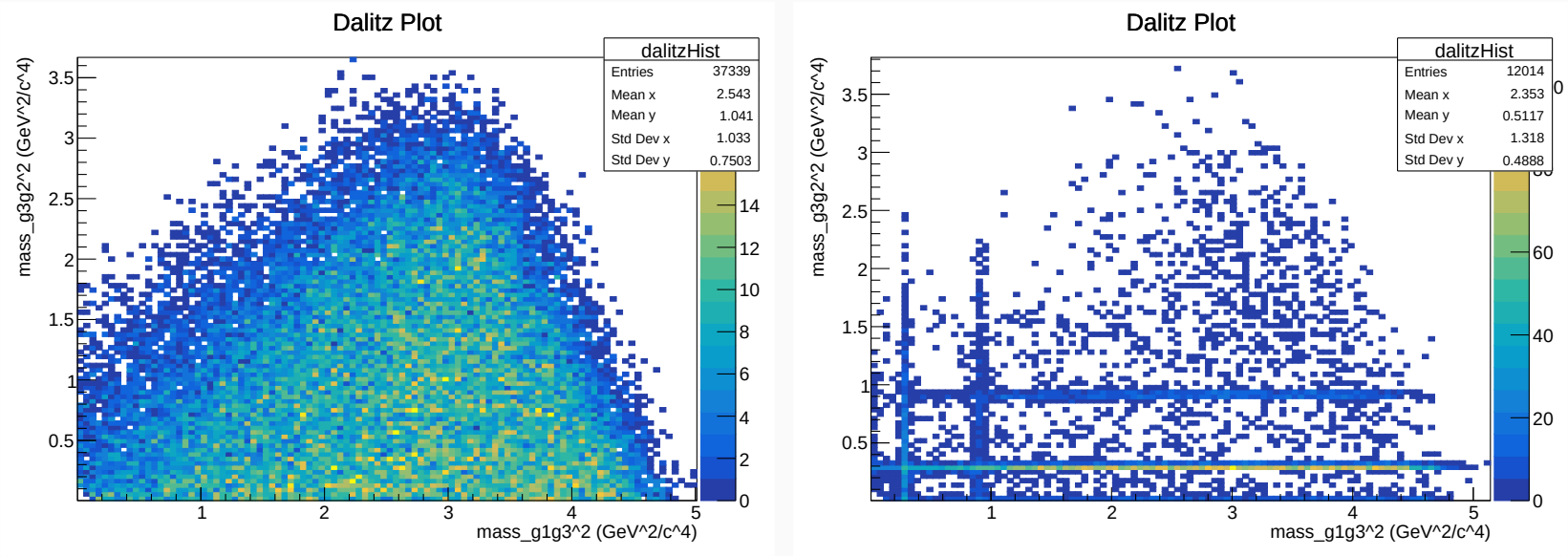


Figure 4: Here you can see a comparison of the Dalitz Plots of the combined mass $\gamma_1\gamma_3$ vs $\gamma_2\gamma_3$. On the left you can see a **pure signal** of 3γ while on the right we have the **inclusive Monte Carlo**; both are produced with a cut dataset (fiducial cuts + Kalman Fit).

Fit of resonances

From both the Topology and the Dalitz plot we see clearly that are present various **resonances**, the three most present are also the ones that was easiest to remove:

$$\eta,\ \eta',\ \pi^0.$$

To do so we did a **fit** of the masses obtainable combining $\gamma\gamma$ and we fitted them over the mass of the three particles using a **sum of a Breit-Wigner and a Crystal ball**.

Fit of Backgrounds

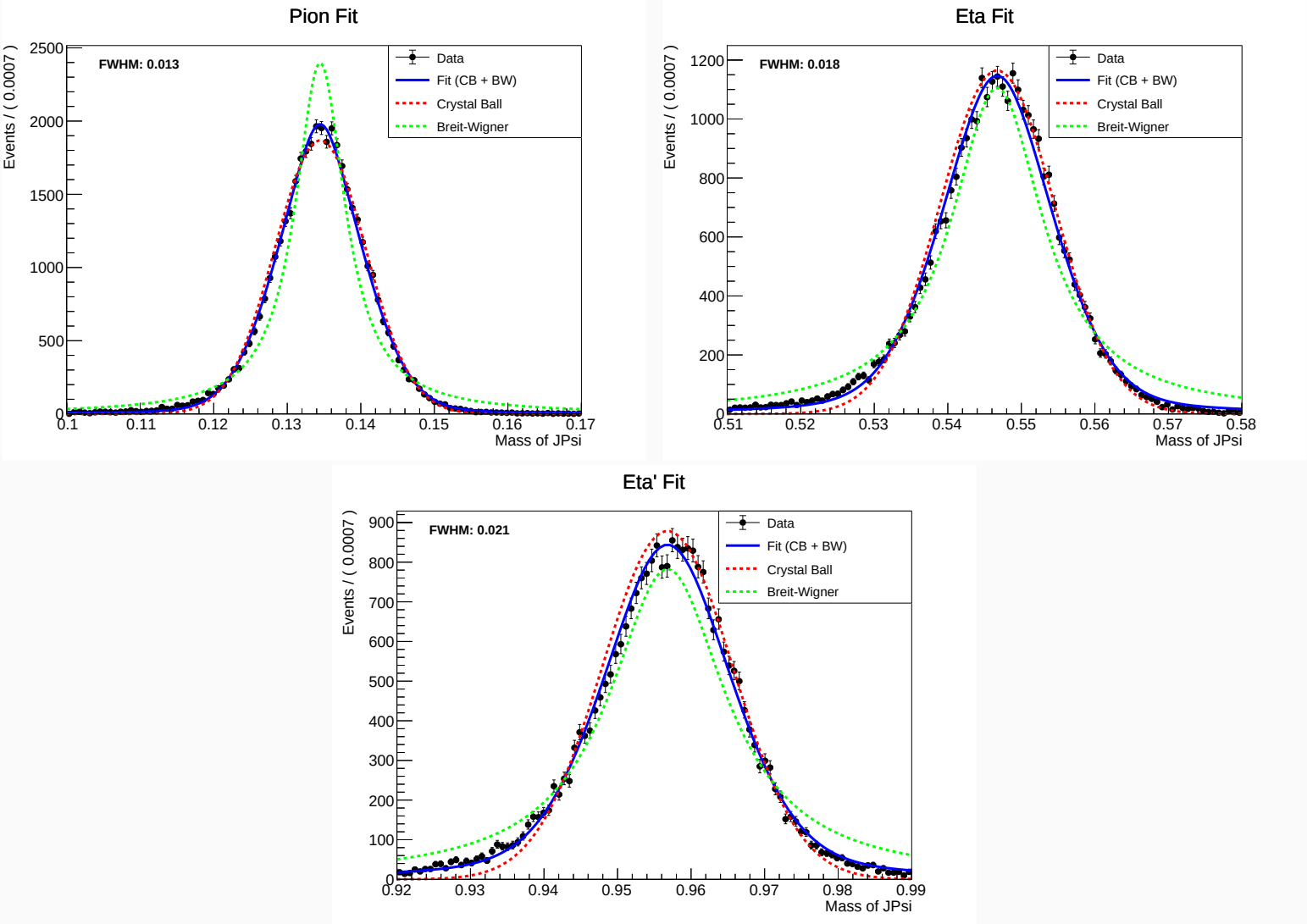


Figure 5: Data from an MC of each signal

Optimization of phase space cut

Found the FWHM of each fit, we used it to impose a cut, optimized using the significance. The cuts fixed are:

- Resonance η:

$$0.5103\,\mathrm{GeV}/c^2 \;\leq\; \mathrm{mass}_{\gamma\gamma} \;\leq\; 0.5854\,\mathrm{GeV}/c^2;$$

- Resonance η':

$$0.9277\,\mathrm{GeV}/c^2 \;\leq\; \mathrm{mass}_{\gamma\gamma} \;\leq\; 0.9878\,\mathrm{GeV}/c^2;$$

- Resonance π^0:

$$0.1149\,\mathrm{GeV}/c^2 \;\leq\; \mathrm{mass}_{\gamma\gamma} \;\leq\; 0.1550\,\mathrm{GeV}/c^2.$$

To reduce the presence of multiple π^0, we decided to include an additional Kalman fit made only when we have more than 3γ, and rejected each event that contained a fitted combined mass $\gamma\gamma$ inside the π^0 mass interval.

Resume of Cuts

Here you can see the number of events saved after each cut:

Cut	Counts
Total Events Analyzed	$2.30 \cdot 10^9$
Fiducial cuts for charged tracks	$2.15 \cdot 10^9$
At least 3 good γ tracks	$1.64 \cdot 10^9$
Good Charged Tracks	$6.35 \cdot 10^8$
Vertex Fit	$6.34 \cdot 10^8$
J/ψ Events as $\pi^+\pi^-$ recoil	$3.99 \cdot 10^7$
Kalman Fit with 3γ tracks	$7.66 \cdot 10^4$
Resonances Cut	$1.30 \cdot 10^4$
Kalman fit of pions	$1.12 \cdot 10^4$

Resume of Cuts: Efficiency

Cut	Efficiency
Total Events Analyzed	100,00%
Fiducial cuts for charged tracks	72,12%
At least 3 good γ tracks	48,20%
Good Charged Tracks	18,48%
Vertex Fit	18,45%
J/ψ Events as $\pi^+\pi^-$ recoil	15,95%
Kalman Fit with 3γ tracks	12,89%
Resonances Cut	11,31%
Kalman fit of pions	11,25%

Table 1: Here is shown the efficiency of each cut over a sample of 300000 events of signal simulated.

State of the Measurement

At the state of the art, using classical methods, We were able to cut out a dataset containing the following tracks:

rowNo	decay tree (decay initial-final states)	iDcyTr	nEtr	nCEtr
1	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \pi^0\pi^0\gamma$ $(\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma)$	4	3161	3161
2	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_4(2050)\gamma, f_4(2050) \to \pi^0\pi^0$ $(\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma)$	2	2074	5235
3	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_2(1270)\gamma, f_2(1270) \to \pi^0\pi^0$ $(\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma)$	5	1397	6632
4	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \gamma\gamma\gamma$ $(\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma)$	1	1019	7651
5	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_0(1710)\gamma, f_0(1710) \to \pi^0\pi^0$ $(\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma)$	3	879	8530
6	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \eta\gamma, \eta \to \gamma\gamma$ $(\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma)$	7	843	9373

Figure 6: Topology showing the 6 channels with the most events present in the MC dataset.

Backgrounds still to remove

As we can see, we have a persistent background of $\pi^0\pi^0$ events since they product a very similar signal to the one we are looking for. The backgrounds still present at this point are: f_0, f_2, f_4 and $\pi^0\pi^0$.

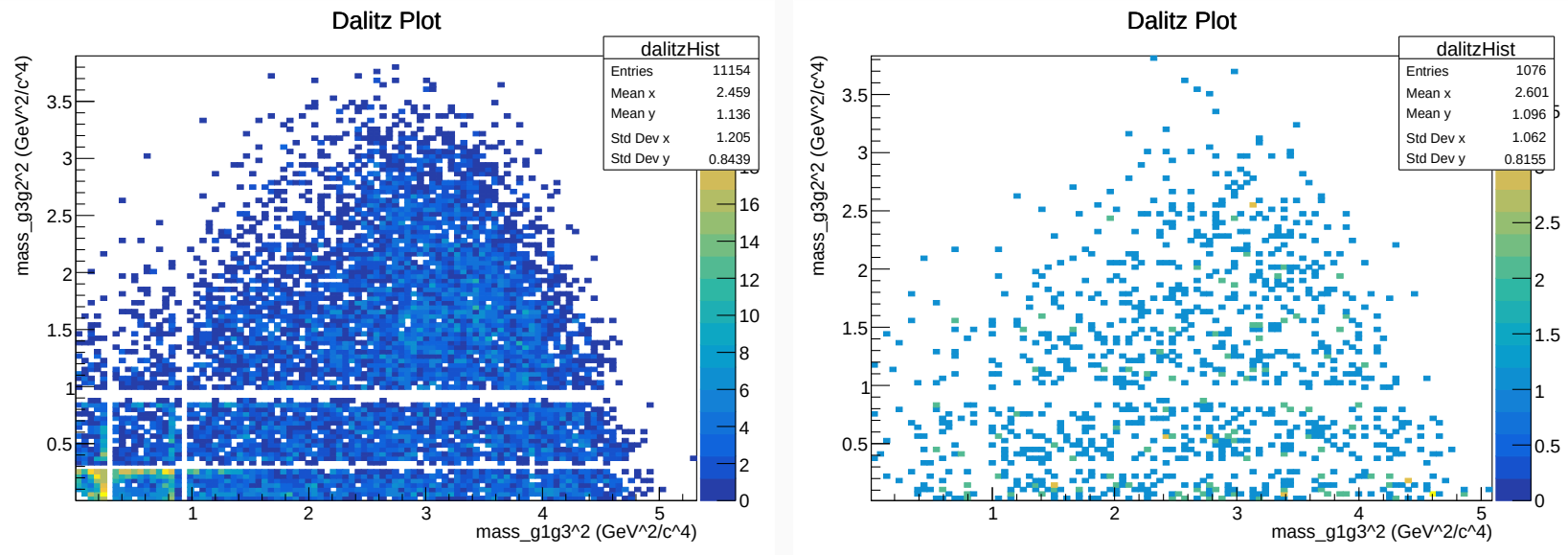


Figure 7: On the left, the Dalitz plot of the $\gamma\gamma\gamma$ signal, on the right the Dalitz plot of the $\pi^0\pi^0$ background.

Table of Contents

Why to use Machine Learning?

The past analysis took advantage of a partial wave analysis, we would like to avoid that by using Machine Learning methods.

In order to do so I tried to compare two different approaches:

- Boosted Decision Trees (BDT): a multivariate method that uses decision trees to classify events;
- Multi-Layer Perceptron (MLP): a neural network that uses multiple layers of neurons.

Preparation of the Dataset

I chose to use the **TMVA** package of the ROOT framework, which is a powerful tool for Machine Learning in High Energy Physics.

To train the Machine Learning algorithms I retrieved a MC dataset for each of the 6 channels:

- Backgrounds: $\pi^0\pi^0\gamma$, $f_4(2050)\gamma$, $f_2(1270)\gamma$, $f_0(1710)\gamma$, $\eta\gamma$;
- Signal: $\gamma\gamma\gamma$.

To each one of the channels has been assigned a weight based on the number of events counted in the topology. This is necessary to "normalize" the datasets that being created singularly have a different number of events with respect to the inclusive MC dataset.

variables used for the training

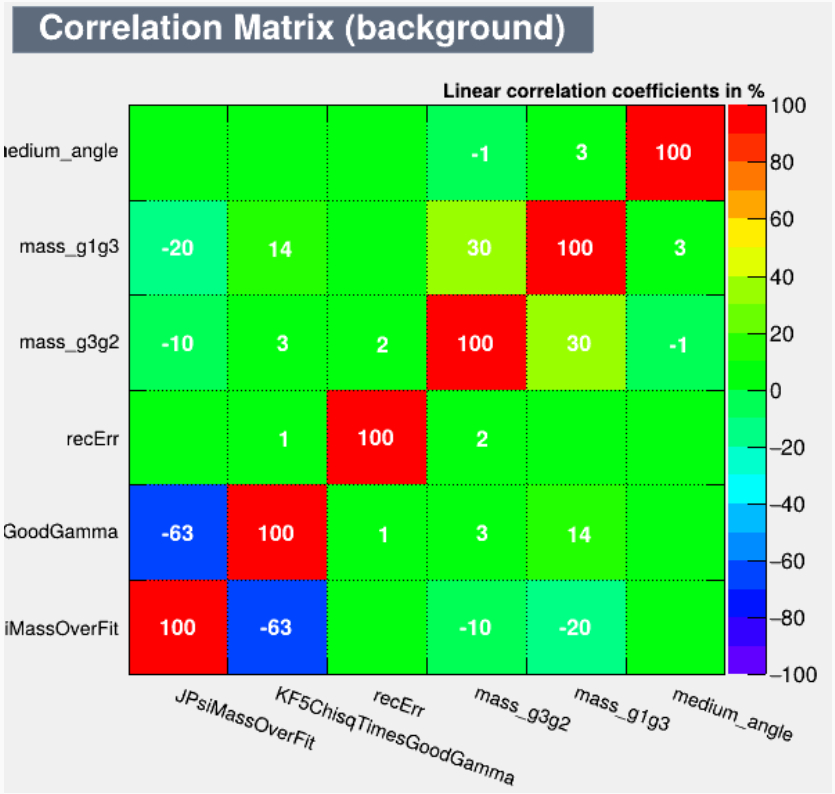
The variables used for the training are:

- $\frac{m_{J/\psi(rec)}}{m_{J/\psi(Fit)}}$;
- $\chi^2 \cdot n_{\mathrm{GoodGammas}}$ where χ^2 is for the fit of J/ψ mass.;
- $E(J/\psi)_{rec} - E(\pi^+) - E(\pi^-)$;
- invariant masses of $\gamma\gamma$ pairs: $m_{\gamma_1\gamma_3}$, $m_{\gamma_2\gamma_3}$;
- $(\theta_1 + \theta_2 + \theta_3)/3$;

To which has been applied a **Decorrelation** method to reduce the correlation between the variables.

In my case I chose to **split in half** the dataset, between the training and the testing set.

Correlation Matrices



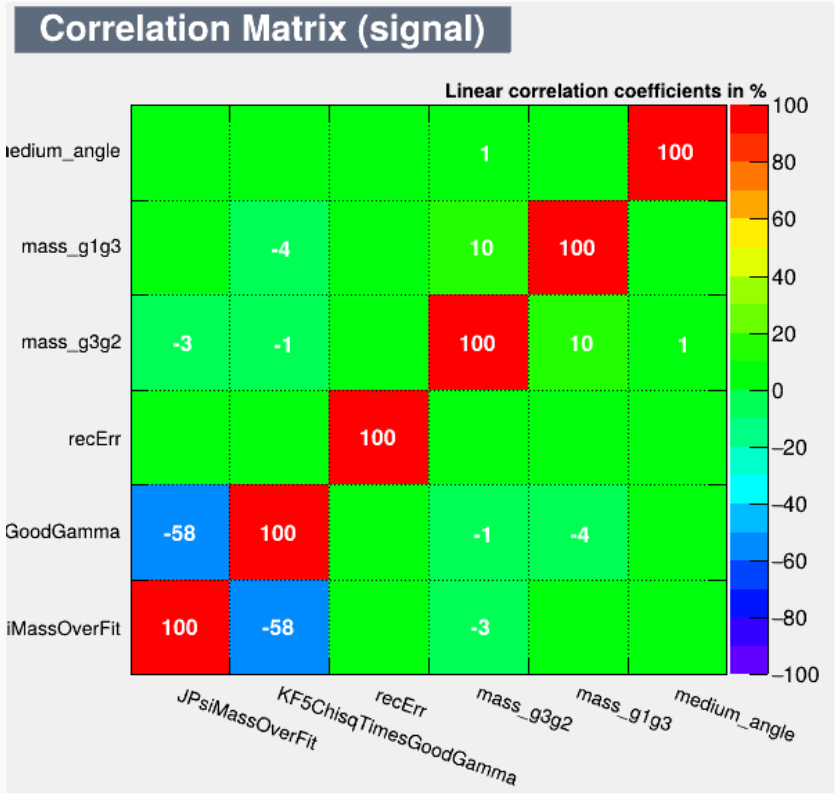


Figure 8: On the left, the correlation matrix of the background channels, on the right the correlation matrix of the signal channel.

Boosted Decision Trees

The Boosted Decision Trees (BDT) is a multivariate method that uses decision trees to classify events. It works by creating a series of decision trees that are trained on the dataset, each tree is trained on the residuals of the previous tree. The training is done using the following parameters:

- **Boosting Type**: AdaBoost;
- **Number of Trees**: 500;
- **Maximum Depth**: 3;
- **Minimum Number of Events in Leaf Node**: $2,5\%$;
- **Learning Rate**: 0.5.

NB: the BDT uses also the option `Use Bagging` that allows to use a random subset of the dataset for each tree.

Multi-Layer Perceptron

The Multi-Layer Perceptron (MLP) is a neural network that uses multiple layers of neurons to classify events. The parameters used for the training are:

- **Number of Layers**: 1;
- **Number of Neurons per Layer**: same as $n_{var} = 6$;
- **Activation Function**: Tanh;
- **Estimator Type**: Mean Square Error;
- **Training Method**: Back Propagation;
- **Cycles**: 700

Training results

Both methods gives us a **score** that evaluate the event classifying it as signal or background. It is very difficult to obtain a perfect separation so the discrimination is done by setting a **threshold** that separate at best the two classes.

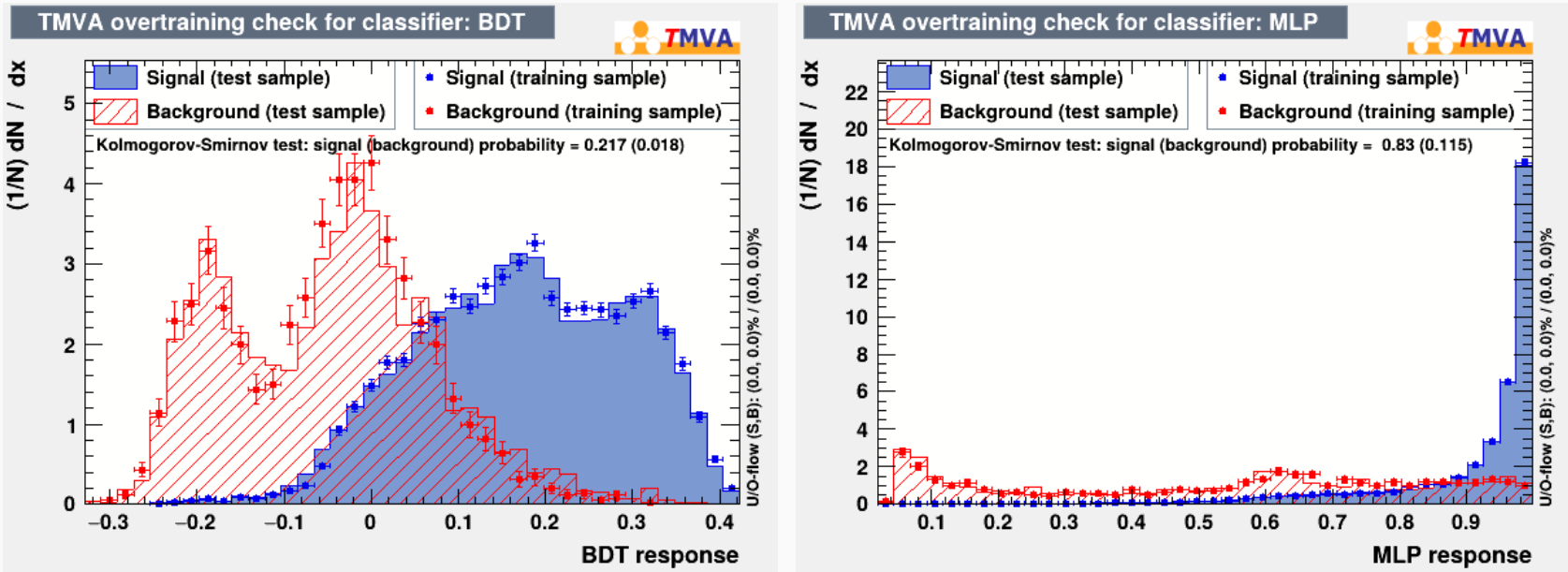


Figure 9: On the left, the separation using the Boosted Decision Trees, on the right the separation using the Multi-Layer Perceptron.

Roc Curve

The **Roc Curve** is a graphical representation of the performance of a binary classifier. One of the parameter to evaluate how good is a classifier is the **Area Under the Curve** (AUC), which is the integral of this curve.

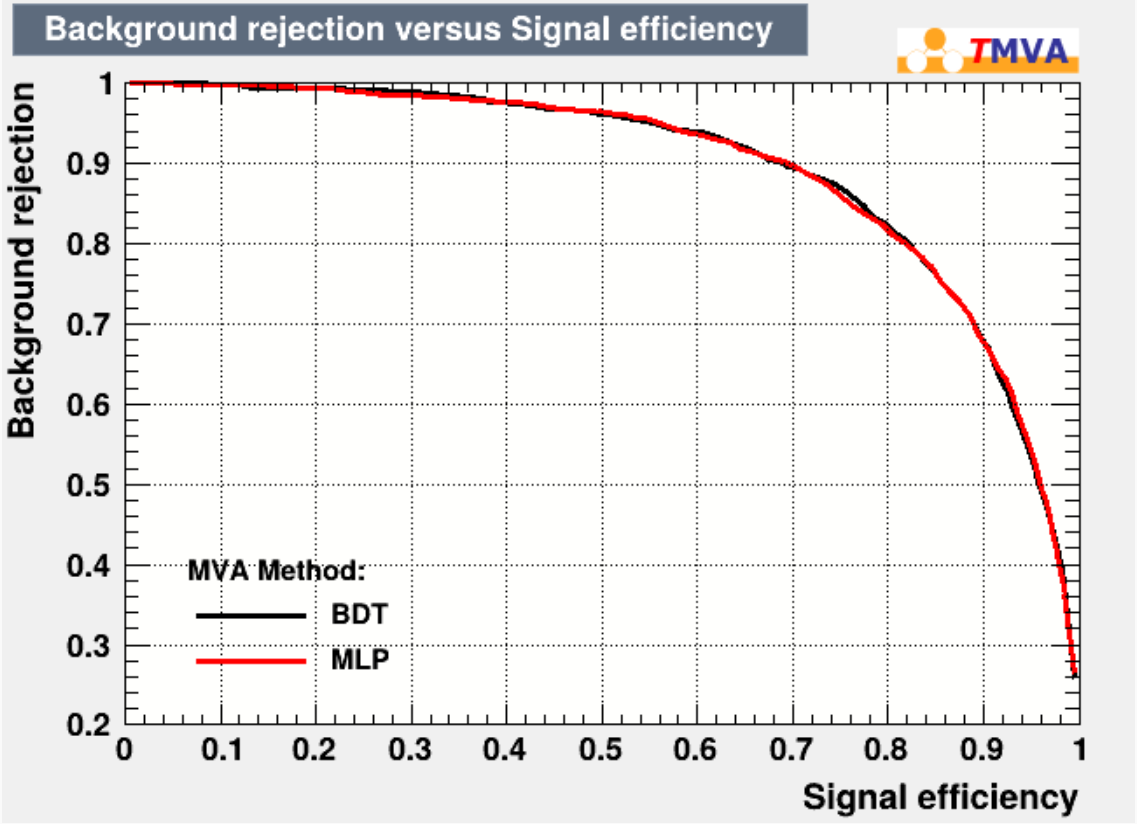


Table of Contents

Results

The results of the training are similar for both methods, and the AUC is around 0.89 for both methods.

I tried to apply the trained algorithms to the Inclusive MC and the results are quite good. I tried three different methods:

- a cut on **BDT** of: $score_{BDT} > 0.02$;
- a cut on **MLP** of $score_{MLP} > 0.76$;
- a combined cut of both values.

REMINDER: State of the Measurement

- **Goal:** Selection of the $J/\psi \to 3 \cdot \gamma$ decays.

rowNo	decay tree (decay initial-final states)	iDcyTr	nEtr	nCEtr
1	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \pi^0\pi^0\gamma$ $(\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma)$	4	3161	3161
2	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_4(2050)\gamma, f_4(2050) \to \pi^0\pi^0$ $(\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma)$	2	2074	5235
3	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_2(1270)\gamma, f_2(1270) \to \pi^0\pi^0$ $(\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma)$	5	1397	6632
4	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \gamma\gamma\gamma$ $(\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma)$	1	1019	7651
5	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_0(1710)\gamma, f_0(1710) \to \pi^0\pi^0$ $(\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma)$	3	879	8530
6	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \eta\gamma, \eta \to \gamma\gamma$ $(\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma)$	7	843	9373

Figure 10: Topology showing the 6 channels with the most events present in the MC dataset.

Topology result of BDT

rowNo	decay tree (decay initial-final states)	iDcyTr	nEtr	nCEtr
1	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \pi^0\pi^0\gamma$ $(\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma)$	1	960	960
2	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \gamma\gamma\gamma$ $(\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma)$	2	912	1872
3	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_4(2050)\gamma, f_4(2050) \to \pi^0\pi^0$ $(\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma)$	6	658	2530
4	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_2(1270)\gamma, f_2(1270) \to \pi^0\pi^0$ $(\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma)$	3	350	2880
5	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \eta'\gamma, \eta' \to \gamma\gamma$ $(\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma)$	0	347	3227
6	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_0'\gamma, f_0' \to \gamma\gamma$ $(\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma)$	5	332	3559

Figure 11: The 6 channels with the most events using BDT.

Topology result of MLP

rowNo	decay tree (decay initial-final states)	iDcyTr	nEtr	nCEtr
1	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \gamma\gamma\gamma$ $(\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma)$	1	910	910
2	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \pi^0\pi^0\gamma$ $(\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma)$	3	864	1774
3	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_4(2050)\gamma, f_4(2050) \to \pi^0\pi^0$ $(\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma)$	6	563	2337
4	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_2(1270)\gamma, f_2(1270) \to \pi^0\pi^0$ $(\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma)$	2	369	2706
5	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_0'\gamma, f_0' \to \gamma\gamma$ $(\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma)$	5	341	3047
6	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \eta'\gamma, \eta' \to \gamma\gamma$ $(\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma)$	0	316	3363

Figure 12: The 6 channels with the most events using MLP.

Topology result of Combined Method

rowNo	decay tree (decay initial-final states)	iDcyTr	nEtr	nCEtr
1	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \gamma\gamma\gamma$ $(\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma)$	1	893	893
2	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \pi^0\pi^0\gamma$ $(\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma)$	3	774	1667
3	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_4(2050)\gamma, f_4(2050) \to \pi^0\pi^0$ $(\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma)$	6	521	2188
4	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_0'\gamma, f_0' \to \gamma\gamma$ $(\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma)$	5	327	2515
5	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_2(1270)\gamma, f_2(1270) \to \pi^0\pi^0$ $(\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma)$	2	302	2817
6	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \eta'\gamma, \eta' \to \gamma\gamma$ $(\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma)$	0	301	3118

Figure 13: The 6 channels with the most events using both methods.

Comparison of the Fitted J/ψ mass

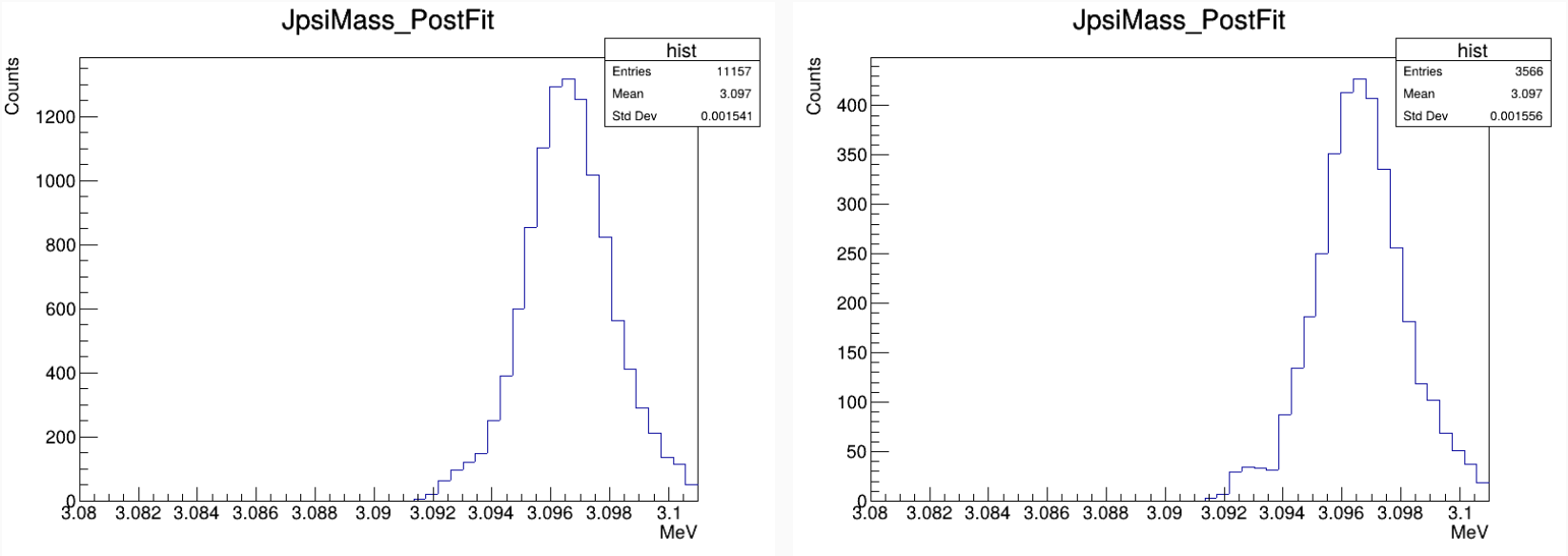


Figure 14: On the left we have the fitted J/ψ mass of the Inclusive MC dataset, on the right we have the fitted J/ψ mass after the application of the TMVA methods.

Table of Contents

Future Goals

- Surely the analysis is still in progress, so the main focus is to obtain a sharp separation of the Decay channel in analysis, using the complete dataset.
- The future step would be then to use what we've learned in this analysis as a basis to study the decay channel:

$$J/\psi \rightarrow \gamma\gamma\gamma\,,$$

using the same dataset.
This is more challenging to do due to the absence of strong tagging given by the $\pi^+\pi^-$, so the hope is that the TMVA methods will be useful.

Thank you

BACKUP: Distribution of the input variables

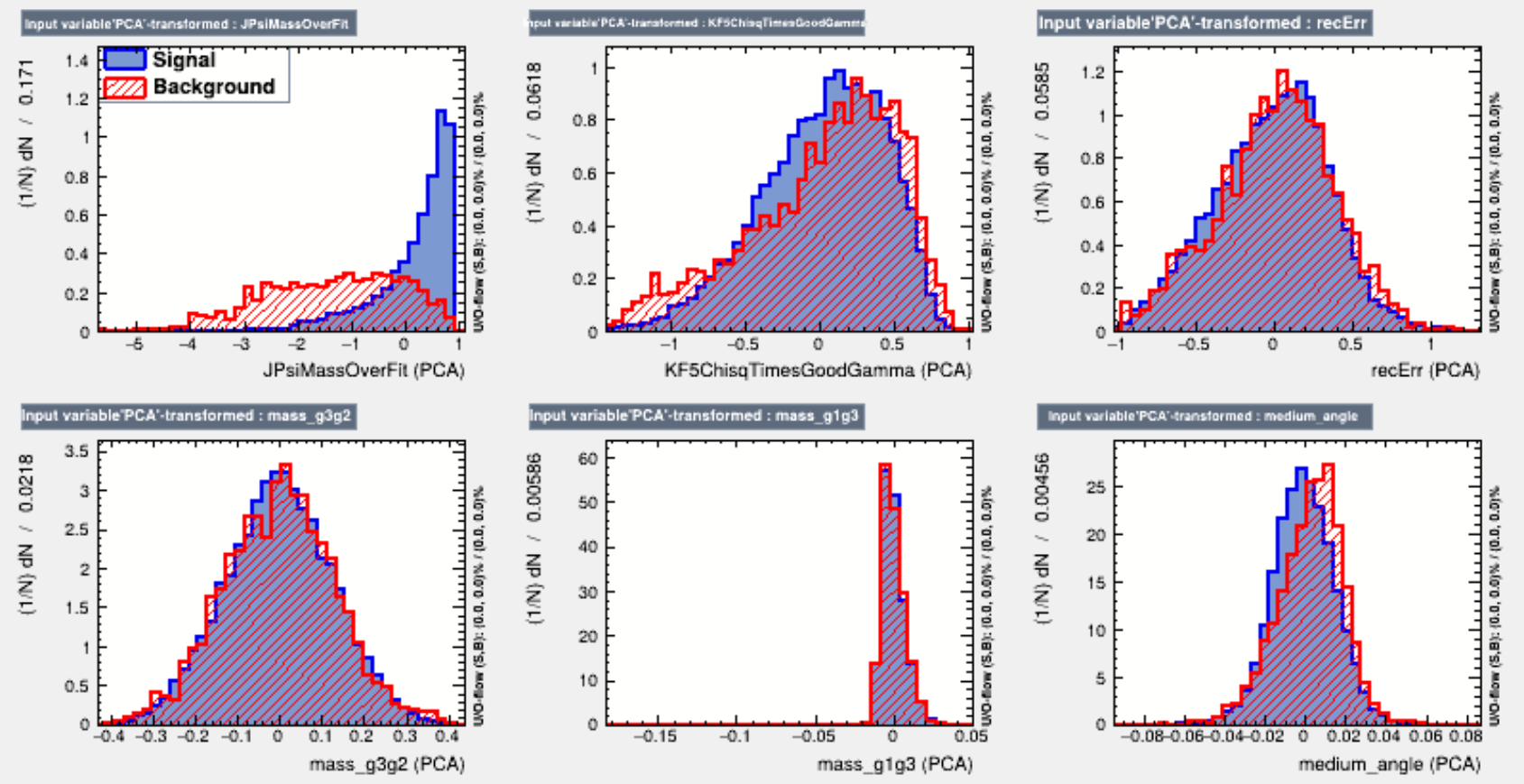


Figure 15: Distribution of the input variables used for the training of the BDT and MLP.

BACKUP: Comparison of the J/ψ mass

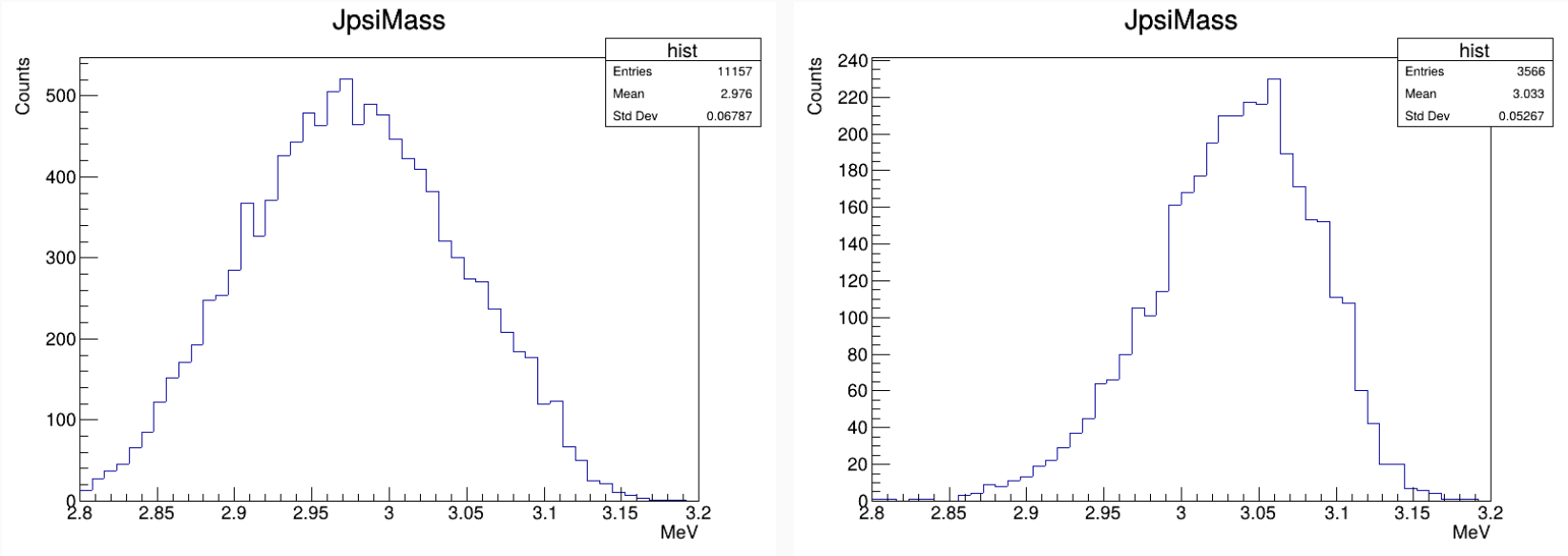


Figure 16: On the left we have the J/ψ mass of the Inclusive MC dataset, on the right we have the J/ψ mass after the application of the TMVA methods.

BACKUP: Branching fractions of intermediate Processes

The branching fractions for each intermediate decay involved are:

- $\mathrm{Br}(J/\psi \to \eta'\gamma) = (\ 5.13 \pm 0.17\ ) \times 10^{-3}$
- $\mathrm{Br}(J/\psi \to \eta\gamma) = (\ 1.104 \pm 0.034)\ \times 10^{-3}$
- $\mathrm{Br}(J/\psi \to \pi\gamma) = (\ 3.49 + 0.33 - 0.30\ ) \times 10^{-5}$
- $\mathrm{Br}(J/\psi \to \pi\pi\gamma) = (\ 1.15 \pm 0.05\ ) \times 10^{-3}$
- $\mathrm{Br}(J/\psi \to f_0\gamma \to \pi\pi\gamma) = (\ 3.8 \pm 0.5\ ) \times 10^{-4}$
- $\mathrm{Br}(J/\psi \to f_2\gamma \to \pi\pi\gamma) = (\ 1.64 \pm 0.12\ ) \times 10^{-3}$
- $\mathrm{Br}(J/\psi \to f_4\gamma \to \pi\pi\gamma) = (\ 2.7 \pm 0.7\ ) \times 10^{-3}$

BACKUP: fs Dalitz Plots

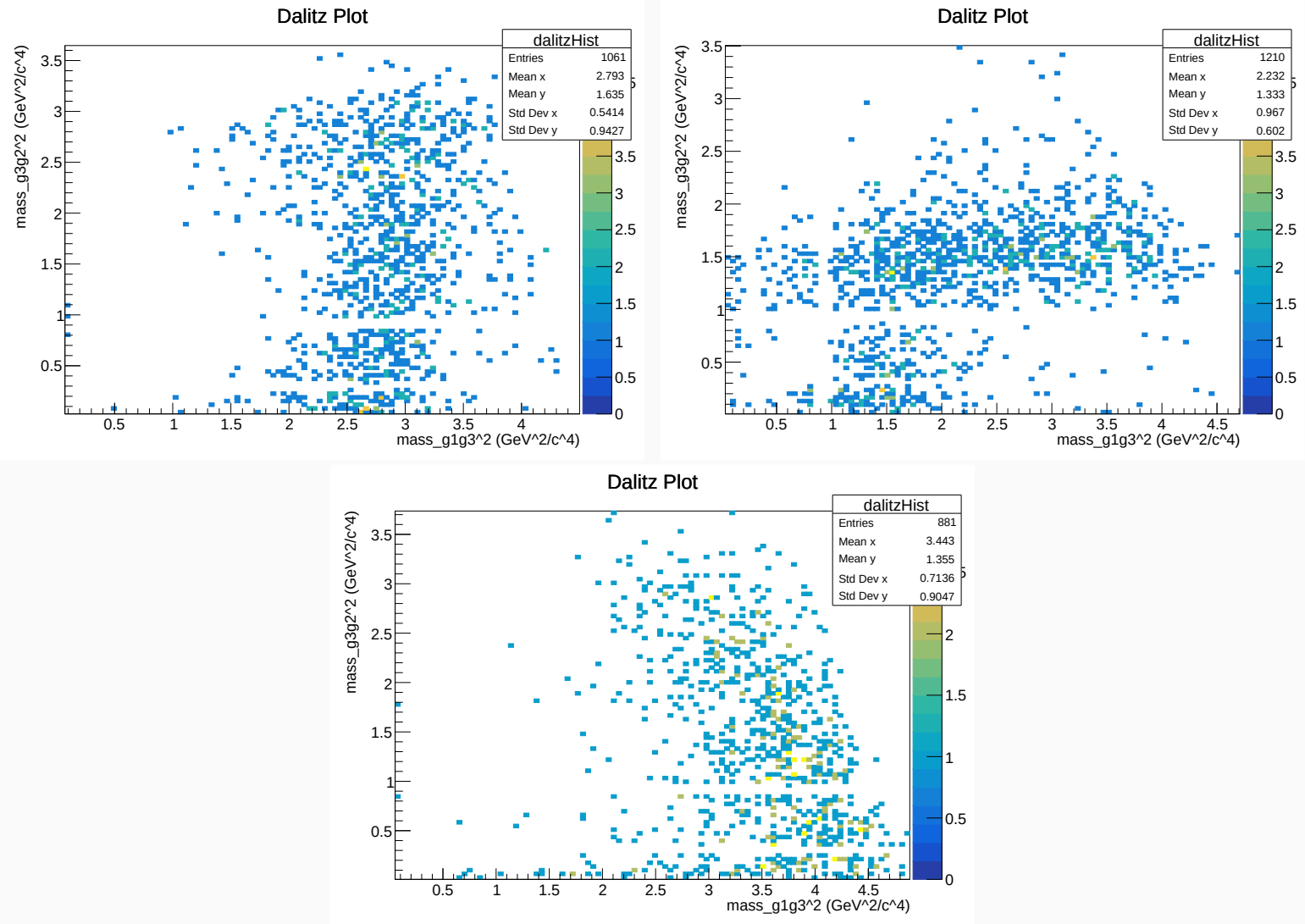


Figure 17: Dalitz plot respectively of f_0, f_2, f_4.

Discarded Methods: Kalman Fit instead of Cuts

We tried also another way of cutting the dataset, trying to use Kalman to **fit also the resonances**. In this way we thought that could be possible to avoid to cut out of the analysis entire sectors of the Phase Space. This method produced a **worse result** than the cut based one.

rowNo	decay tree (decay initial-final states)	iDcyTr	nEtr	nCEtr
1	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \pi^0\pi^0\gamma$ ($\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma$)	0	223	223
2	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_4(2050)\gamma, f_4(2050) \to \pi^0\pi^0$ ($\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma$)	1	141	364
3	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \gamma\gamma\gamma$ ($\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma$)	2	80	444
4	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \eta\gamma, \eta \to \gamma\gamma$ ($\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma$)	6	73	517

rowNo	decay tree (decay initial-final states)	iDcyTr	nEtr	nCEtr
1	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \pi^0\pi^0\gamma$ ($\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma$)	0	563	563
2	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_4(2050)\gamma, f_4(2050) \to \pi^0\pi^0$ ($\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma$)	1	352	915
3	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to f_2(1270)\gamma, f_2(1270) \to \pi^0\pi^0$ ($\psi' \dashrightarrow \pi^0\pi^0\pi^+\pi^-\gamma$)	6	238	1153
4	$\psi' \to \pi^+\pi^- J/\psi, J/\psi \to \gamma\gamma\gamma$ ($\psi' \dashrightarrow \pi^+\pi^-\gamma\gamma\gamma$)	3	161	1314

Figure 18: On the left we have the fit based topology, on the right we have the cut based one. As we can see the signal count in the fit version is halved wrt the cut version.

Discarded Methods: Kalman Fit of missing Tracks

Since the most difficult background to reject is the $\pi^0\pi^0\gamma$ one; we tried to reject it using the an option of the Kalman fit that is the AddMissTrack() function of Kalman Fit.

In such a way we wanted to test the possibility to have a t least **one missing photon** per π^0, due to the low energy of γ's coming from the decay of a π^0.

This method was also discarded because it produced empty dataset with no good data.

BACKUP: Boosted Decision Trees

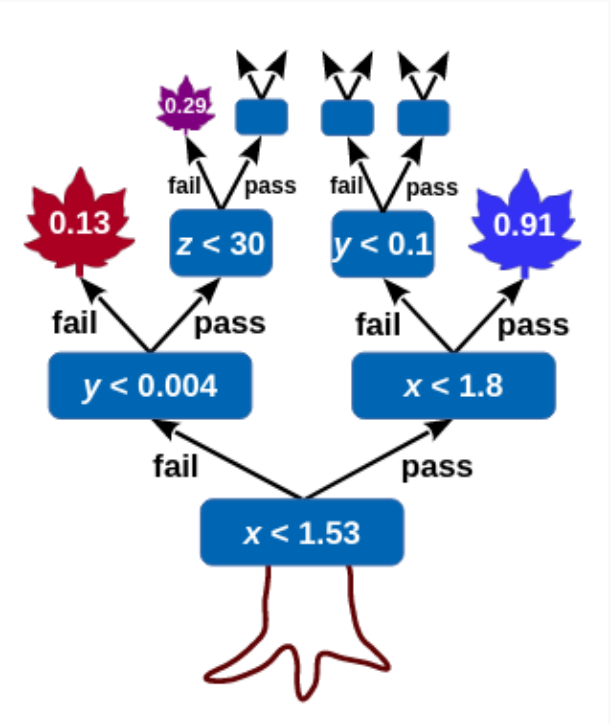
We are now trying to implement a **Boosted Decision Tree (BDT)** in the analysis in order to discriminate the Background events from the signal ones.

BDT is a **machine learning method**, included in the TMVA class of ROOT, and is based on the concept of Decision Trees

BACKUP: Decision Tree

Decision Tree

A decision tree is an algorithm that, using multiple variables, start by imposing a condition on one variable (e.g.: a rectangular cut over the energy), and verify if that condition bring to an improvement of the dataset, each branch than split into a secondary decision, that can be over the same or over another variable and so on until an objective is reached.



BACKUP: Boosting

The BDT take the concept of decision trees but **iterates** it by creating a **forest** of many small trees (which stops already at the third/forth branch). In such a way it is possible to evaluate many different picture of the same condition.

The **boosting** is given by the fact that, in the training phase, we give **different weights** to the trees based on how well that specific tree divide our dataset. In such a way we create a pattern that can then be applied to the dataset to separate background and signal