# Phenomenological studies of dijet azimuthal decorrelation at $\sqrt{s}$= 7 TeV in pp collisions at the LHC

Sandra Consuegra Rodríguez, Higher Institute of Technology and Applied Sciences, Cuba.

Supervisor(s): Daniela Domínguez Damiani
Hannes Jung

September 7, 2016

**Abstract**

The azimuthal angular decorrelation between the two jets with the largest transverse momentum in pp collisions at $\sqrt{s}$= 7 TeV is studied in this project. The comparison between the results obtained with Monte Carlo event generator PYTHIA and the pure NLO extracted from POWHEG using a RIVET analysis routine is presented. It was also analized the behaviour of the dijet azimuthal decorrelation varying the pTHatMin parameter of PYTHIA. From the technical point of view, running POWHEG only with the first two steps allowed us to obtain pure NLO calculations. Then the output of POWHEG was transformed in such a way that the data could be processed with RIVET framework and also a normalization was performed, being the main result the implementation of POWHEG without its parton shower as a first step to match PYTHIA and POWHEG.

# Contents

# 1 Introduction.

Quantum Chromodynamics (QCD) is the theory of the strong interactions that take place between quarks and gluons. High-energy proton-proton collisions with high momentum transfer are described within the framework of QCD as pointlike scatterings between the proton constituents, collectively referred to as partons. The outgoing partons manifest themselves, through quark and gluon soft radiation and hadronization processes, as localized streams of particles, identified as jets.
If only two jets are produced they will have equal transverse momenta pT with respect to the beam axis and will be back to back in the azimuthal angle ($\Delta\Phi_{dijet} = [\Phi_{jet1} - \Phi_{jet2}] = \pi$). Larger decorrelations from $\pi$ occur in the case of hard multijet production. If during the hard process itself, a bunch of jets were produced, the azimuthal angle between the two leading jets can approach zero, although very small angular separations are suppressed because of the finite jet sizes as well as the low probability to occur.
The measurement of the dijet azimuthal angular decorrelation is an interesting observable to probe the Standard Model. More general, and perhaps more important, in providing an abundant source of high momentum transfer events, the dijet production process acts as both a background to, and sensitive probe of, physics beyond the Standard Model.
From an experimental point of view, dijet events have a relevant practical role in jet measurements, because they are used in calibration of equipments, for example, in the determination of jet energy resolution. Jet production is the most common hard scattering processes taking place in hadronic collisions, therefore it is fundamental that it be thoroughly studied and understood.
Adding higher order QCD corrections could be really needed in order to obtain better predictions of $\Delta\Phi$ azimuthal decorrelation between the leading jets.

# 2 Theory.

When two hadrons collide at high energy, most of the collisions involve only soft interactions of the constituent quarks and gluons. Such interactions can not be treated using perturbative QCD, because $\alpha$ is large when the momentum transfer is small. In some collisions however, two quarks or gluons will exchange a large momentum. In this cases the elementary interaction takes place very rapidly compared to the internal time scale of the hadrons wave functions. So the lowest order(s) QCD predictions should accurately describe this process, considered as hard. The cross section for this kind of processes can be written as a factorized product of short and long distance terms:

$$\sigma(P_1, P_2) = \sum_{i,k} \int dx_1 dx_2 f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \sigma_{i,j}(\mu_R^2, \mu_F^2) \tag{1}$$

where $P_1$ and $P_2$ are the momentum of the incoming hadrons. The momentum of the partons that participate in the hard interaction are $p_1 = x_1 P_1$ and $p_2 = x_2 P_2$. The functions $f_i(x_i, \mu_F^2)$ are the usual QCD quark or gluon PDFs, defined at a factorized scale $\mu_F$, which takes into account the long distance effects. It is in this sense that $\mu_F$ can be thought of as the scale which separates long and short physics. $\sigma_{i,j}$ is the cross section of the hard process between the partons i and j. The following figure shows the general picture of this kind of events:

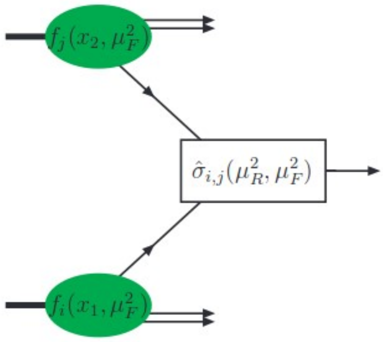



Figure 1: Hard process event.

The first terms in the perturbative QCD expansion are usually sufficient to describe successfully the hard interaction between two partons, because the scale of this process is large. However, in some regions of the phase space other terms are enhanced and can not be neglected. Furthermore, when a quark emits a gluon in a collinear way perturbation theory does not describe this process adequately and resumation theory is needed. Phase collinear emissions are described under the parton branching mechanism, occurring outside the hard process itself. Figure 2 shows a general view of the processes at the LHC that we are interested in:

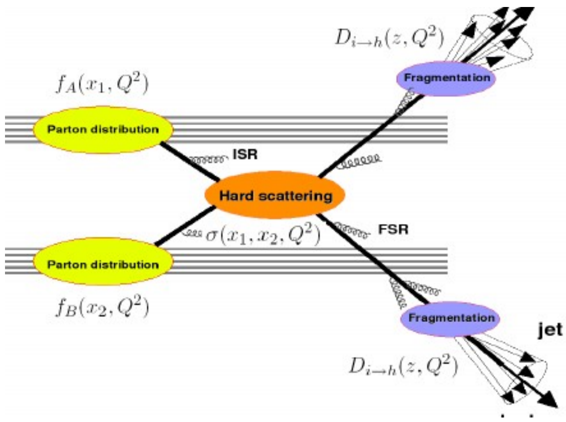



Figure 2: Hadron-Hadron collisions picture at High Energy Scale.

Parton branching typically happens for the ingoing and outgoing quarks and gluons of the hard interaction. The incoming quark, initially with low virtual mass-squared and carrying a fraction x of the hadrons momentum, moves to more virtual masses and lower momentum fraction by successive small-angle emissions. In a collision all outgoing partons will undergo parton showering and transform themselves into hadrons forming jets through hadronization. When more than two partons are involved in the collision (either by parton showering or in the hard process itself) the leading jets are not forced to be in the back-to-back region ($\Delta\Phi \approx \pi$) anymore and the distribution of $\Delta\Phi$ will be dependent on the parton showering and/or in the emission of a third jet during the hard process. Figure 3 shows the $\Delta\Phi$ decorrelation in an event where two hard jets were produced:

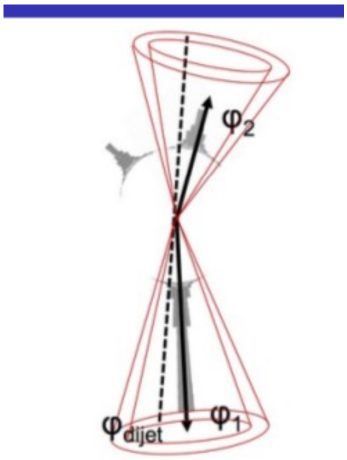



Figure 3: Dijet azimuthal decorrelation and general picture of LHC collisions.

In a high energy collision soft-gluon emissions will decorrelate the two highest pT (leading) jets and cause small deviations from $\pi$, as well as harder emissions will decorrelate them to higher values on $\Delta\Phi$. In this project we will focus on phenomenological studies of dijet azimuthal decorrelation between the two leading jets produced in pp collisions at $\sqrt{s} = 7$ TeV at the LHC.

# 3 Techniques.

## 3.1 Monte Carlo event generator PYTHIA and POWHEG BOX.

PYTHIA is a standard tool for the generation of events in high energy collisions, comprising a coherent set of physics models for the evolution from a few body hard process to a complex multiparticle final state. The physics involved is not understood well enough to give an exact description, instead the program has to be based on a combination of analytical results and various QCD-based models. The different steps in the simulation includes the hard subprocesses, the initial and final-state parton showers, the underlying events, beam remnants and finally hadronization and decays. Furthermore, you can

access extensive information: subroutines, functions, switches, parameters, particle and process data. This allows the user to tailor the generation task to the topic of interest. The problem of merging NLO calculations with parton shower simulations is basically that of avoiding overcounting, since the SMC programs do implement approximate NLO corrections already.
The MC@NLO proposal was the first one to give an acceptable solution to the overcounting problem. The basic idea is avoiding the overcounting by subtracting from the exact NLO cross section its approximation, as implemented in the SMC program to which the NLO computation is matched. Such approximated cross section is computed analytically, and is SMC dependent. In general, the exact NLO cross section minus the MC subtraction terms does not need to be positive. Therefore MC@NLO can generate events with negative weights, although their presence does not imply a negative cross section, since at the end physical distributions must turn out to be positive.
A method, to be called POWHEG in the following (for Positive Weight Hardest Emission Generator), was proposed that overcomes the problem of negative weighted events, and that is not SMC specific. The POWHEG BOX is a general computer framework for implementing NLO calculations in shower Monte Carlo programs according to the POWHEG method. The main purpose of the implementation of the POWHEG method is to generate hard events that can then be fed into a SMC program for subsequent showering. Therefore as an output from POWHEG we could get the NLO calculations plus an initial state of parton shower that later match with other SMC as PYTHIA.

## 3.2 RIVET Analysis and CMS_2011_S8950903 routine.

The RIVET toolkit (Robust Independent Validation of Experiment and Theory) is a system for validation of Monte Carlo event generators. It provides a large (and ever growing) set of experimental analyses useful for MC generator development, validation, and tuning, as well as a convenient infrastructure for adding your own analyses. RIVET is one of the most widespread way by which analysis code from the LHC and other high-energy collider experiments is preserved for comparison and development of future theory models. It is used by phenomenologists, MC generator developers, and experimentalists on the LHC and other facilities.
The study of dijet azimuthal decorrelation done in this contribution was performed using the CMS_2011_S8950903 RIVET routine. The .top file obtained with POWHEG was transformed into a .yoda file in order to be able to make a comparison between PYTHIA, POWHEG and experimental data using the rivet-mkhtml command.
The analysis routine data is based on an inclusive dijet event sample corresponding to an integrated luminosity of 2.9 pb. Jets are anti-kt with R=0.5, $pT > 80(30)$ GeV and

$|\eta| < 1.1$. The routine is divided into 5 different pT regions:

- $80GeV < pT < 110GeV$
- $110GeV < pT < 140GeV$
- $140GeV < pT < 200GeV$
- $200GeV < pT < 300GeV$
- $pT > 300GeV$

## 3.3 Need of adding NLO calculations into SMC programs.

Although LO calculations with PYTHIA generally describe broad features of a particular process and provide the first estimate of its cross section, in many cases this approximation is insufficient. The inherent uncertainty in a lowest-order calculation derives from its dependence on the non-physical renormalization and factorization scales, which is often large.

That is why it would be great to have in our study additionally to PYTHIA, some MC program that goes further, at least up to NLO matrix elements. For this we will implement an intermediate step with POWHEG to obtain NLO pure computations (without Parton Showering) to get the dijet $\Delta\Phi$ distributions between the leading jets in hard scattering processes. This will be our main technical work.

To understand better what is the difference between PYTHIA and POWHEG let's see how the total cross section is estimated in both cases. In PYTHIA the computation starts from a kinematic configuration (hard) which is generated according to an exact LO computation. Usually such configuration is that of a $2 \rightarrow 2$ partonic process. The final-state multiplicity is then iteratively increased, by letting each initial and final state parton branch into a couple of partons with a probability related to a Sudakov form factor. Thus, if at a given stage of the shower, the scattering process is described by m partons, the algorithm decides with a certain probability whether branching is over at this stage, or further branchings will take place.

On the other hand in the POWHEG formalism, the generation of the hardest emission is performed first, obtaining full NLO accuracy, and using the SMC to generate subsequent radiation. At the same time the POWHEG formula can be used to feed a SMC program that will perform all subsequent (softer) showers and hadronization. If the SMC is ordered in pT, like PYTHIA, we simply require that the shower is started with an upper limit on the scale equal to the kt (minimum transversal momentum generated) of the POWHEG event.

There is an intermediate step between NLO computations themselves and the parton shower where we could get some kind of files that already contain histograms filled with needed information of the analysis using only NLO computations. These files are called: pwg-NLO.top.

It was necessary to create a mechanism to convert these .top files into .yoda files. To proceed with this aim we had to first extract from the RIVET routine the basic histogram information (i.e width of the binning) to give them to the POWHEG analyzer in order to

get coherent results to compare and then converting these .top files in a txt flat format. After this a tool from RIVET was used (flat2yoda) to convert flat formats to yoda files. Summarizing, the technical part or the project consisted of three steps:
• Incorporating the binning and pT ranges of the RIVET routine that take care of the $\Delta\Phi$ decorrelation at 7 TeV to the POWHEG input file.
• Transforming the POWHEG output .top file into a .yoda file to process it with RIVET.
• Normalizing the histograms obtained with POWHEG.

# 4 Results.

## 4.1 Effects of increasing the pTHatMin in dijet azimuthal decorrelations.

The pTHatMin parameter is the minimum invariant pT of jets produced in a strong interaction with PYTHIA. As a first task we proposed to study the effects of increasing the pTHatMin parameter in dijet azimuthal decorrelations for pTHatMin=5 GeV, 10 GeV, 15 GeV and 18 GeV.
Here the reported experimental dijet azimuthal decorrelation in *pp* collisions at $\sqrt{s} = 7$ TeV stored at CMS_2011_S8950903 RIVET analysis routine is compared with PYTHIA results for three different ranges of leading pT jet and varying the pTHatMin parameter. It can be observed the improvement of statistics with the increase of the pTHatMin parameter. This is due to the fact that we are looking at high pT jets, so if we put a higher pTHatMin value PYTHIA will produce more events in the range of the leading jets analized in the routine.

## 4.2 Comparison of experimental dijet azimuthal decorrelations with PYTHIA and POWHEG simulations.

Finally, a comparison of experimental data, PYTHIA and pure NLO POWHEG results is presented in the 5 pT ranges of the RIVET routine. It can be observed that at low pT region PYTHIA behaves better than POWHEG because the phase space is smaller and the parton shower is enough to describe the $\Delta\Phi$ behaviour. Once we increase the pT region, the contribution from NLO terms becomes larger and then we can see a better accuracy of POWHEG than PYTHIA. In the $pT > 300GeV$ region we already see the improvement of the results from POWHEG, even though it is without any parton shower.

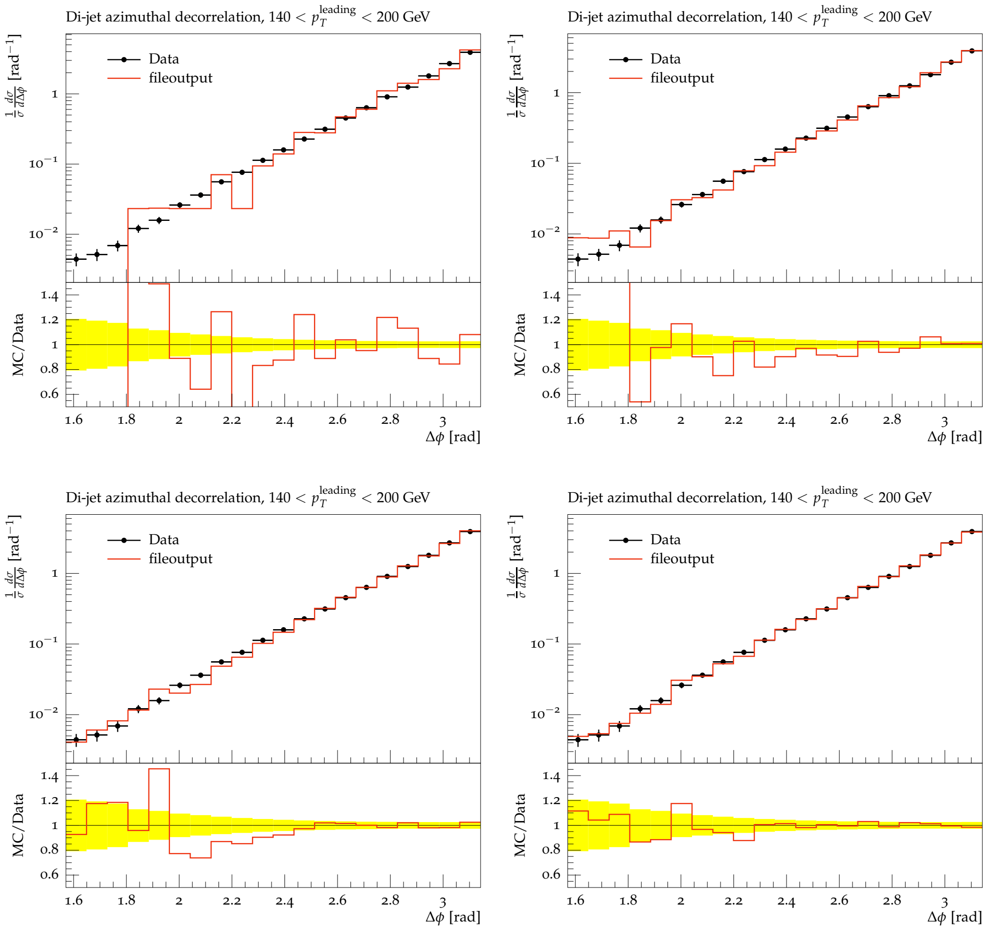
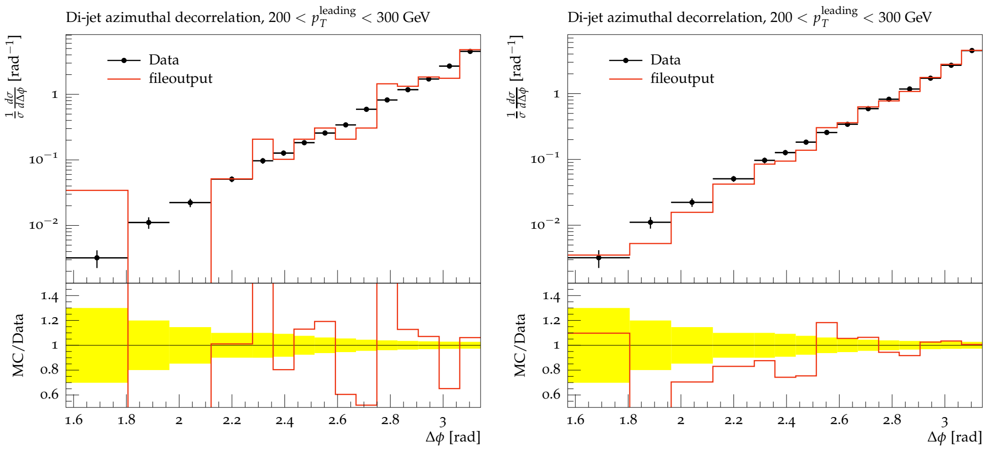



Figure 4: Dijet azimuthal decorrelation for pTHatMin=5, 10, 15 and 18 GeV for $140GeV < p_T^{leading} < 200GeV$.



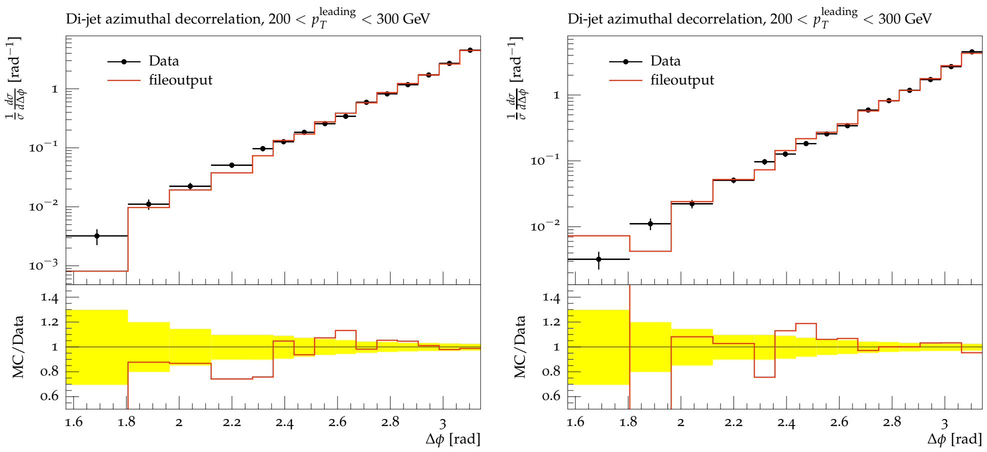



Figure 5: Dijet azimuthal decorrelation for pTHatMin=5, 10, 15 and 18 GeV for $200GeV < p_T^{leading} < 300GeV$.

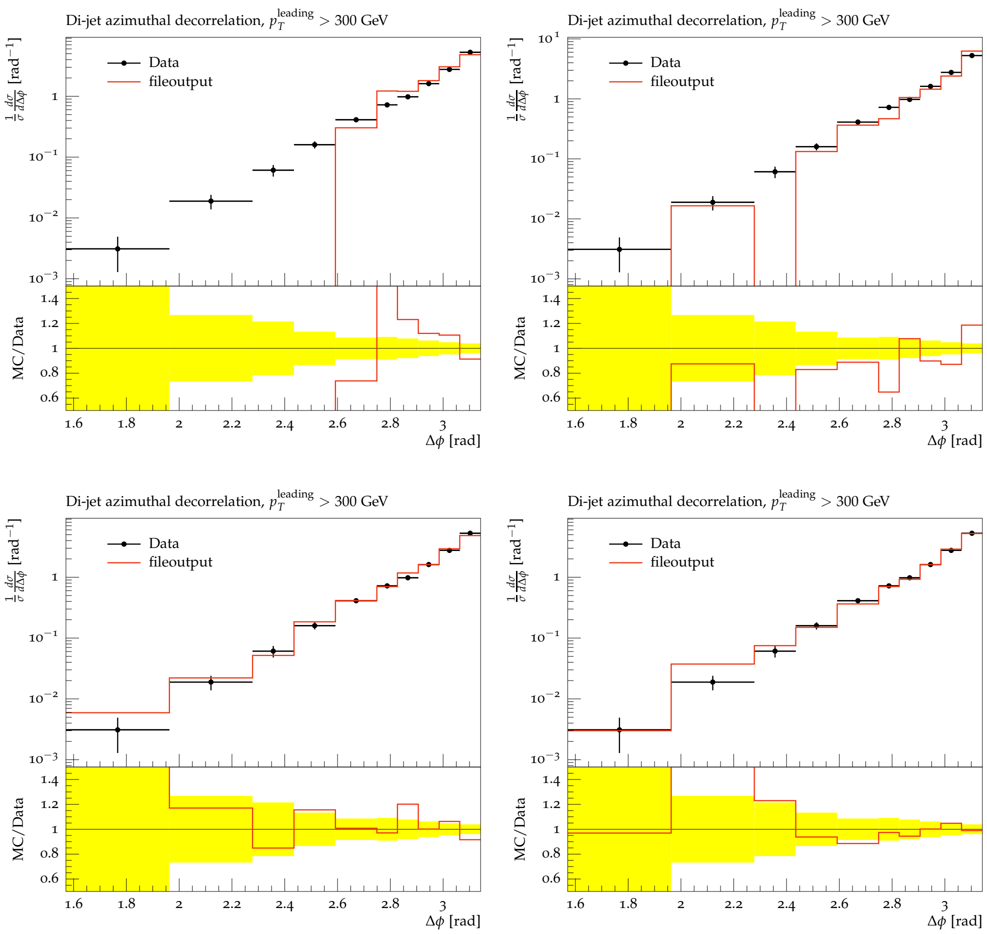



Figure 6: Dijet azimuthal decorrelation for pTHatMin=5, 10, 15 and 18 GeV for $p_T^{leading} > 300GeV$.

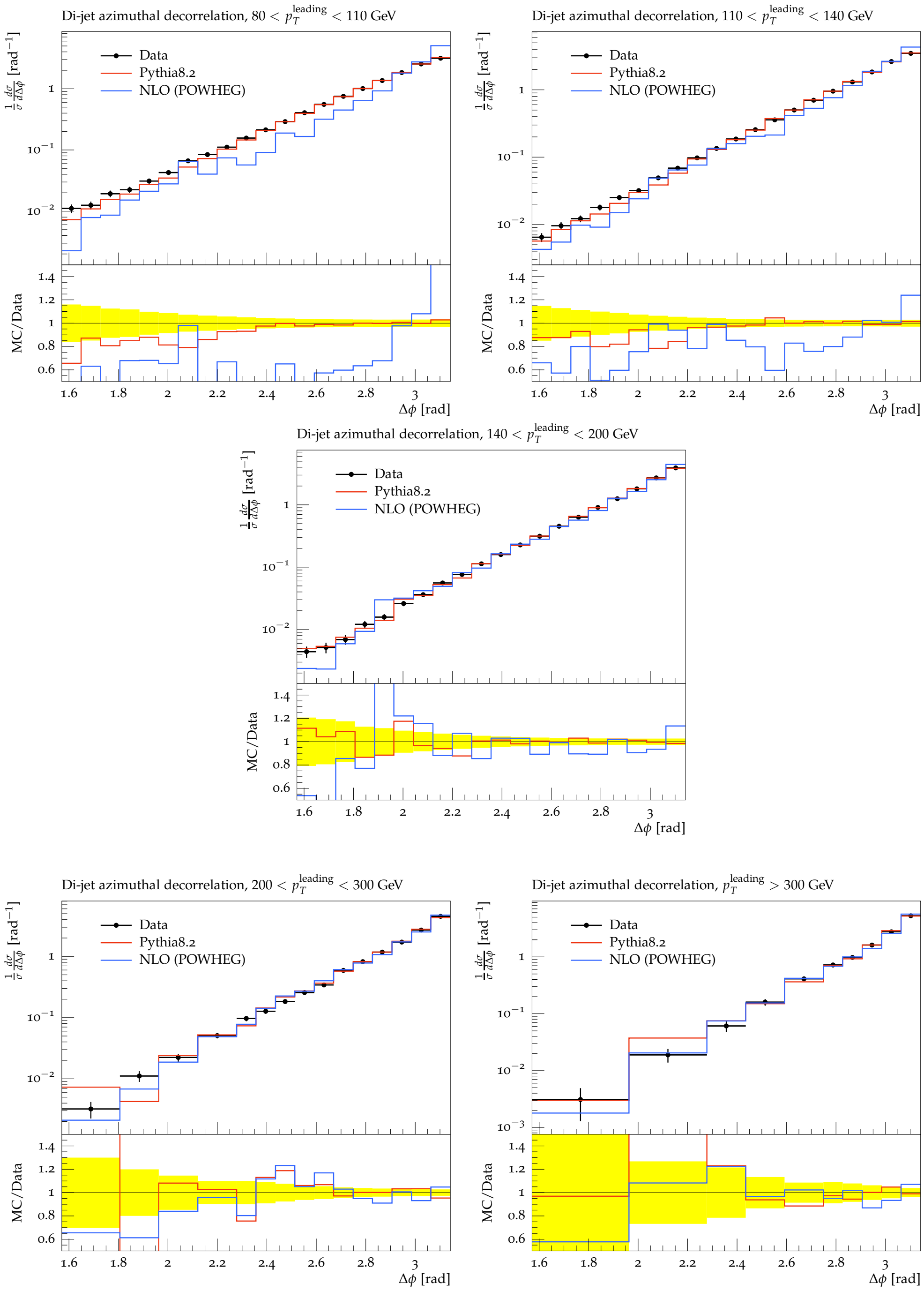



Figure 7: Dijet azimuthal decorrelation comparison of PYTHIA, pure NLO (POWHEG) and experimental results.

## 5 Conclusions.

The study of dijet azimuthal decorrelations in pp collisions at $\sqrt{s}$= 7 TeV has been carried out using a RIVET routine with emphasis in the observable: $(\Delta\Phi)$ distribution between the leading and sub leading jets at 5 diferent pT ranges.
As a result POWHEG was used to get pure NLO matrix elements. This allowed us to explore the influence of the POWHEG method through its modified NLO accuracy in solving the matching-merging problem.

## 6 Acknowledgements.

I would like to thank the QCD DESY-CMS group for their support during the summer student program. I am specially grateful to Hannes Jung, Daniela Domínguez and Armando Bermúdez. Also, I would like to thank my professors at Higher Institute of Technology and Applied Sciences (InSTEC), in Cuba.

## References

[1] Measurement of dijet azimuthal decorrelation in pp Collisions at $\sqrt{s} = 8$ TeV, *CMS Collaboration, arXiv:1602.04384v1 [hep-ex] 13 Feb 2016.*

[2] Dijet Azimuthal Decorrelations in *pp* Collisions at $\sqrt{s} = 7$ TeV, *CMS Collaboration, Phys. Rev. Lett.* **106**, *122003 (2011).*

[3] Matching NLO QCD computations with parton shower simulations: the POWHEG method, *Stefano Frixione and Paolo Nason and Carlo Oleari, Journal of High Energy Physics,***11***,(2007).*

[4] PYTHIA 6.4 Physics and Manual, *T. Sjstrand, S. Mrenna and P. Skands, JHEP05 (2006) 026.*

[5] A Brief Introduction to PYTHIA 8.1, *T. Sjstrand, S. Mrenna and P. Skands, Comput. Phys. Comm. 178 (2008) 852 [arXiv:0710.3820].*

[6] RIVET user manual version 2.5.1, *Andy Buckley, Jonathan Butterworth, David Grellscheid, Hendrik Hoeth, Leif Lonnblad, James Monk, Chris Pollard, Holger Schulz and Frank Siegert, (2013) [arXiv:1003.0694v8].*

[7] The POWHEG BOX user manual: common features, *Simone Alioli, Paolo Nason, Carlo Oleari and Emanuele Re. (2011)*

[8] Jet pair production in POWHEG, *Simone Alioli, Keith Hamilton, Paolo Nason, Carlo Oleari and Emanuele Re, Journal of High Energy Physics,***4***,(2011).*