Matching TMD Parton Shower with POWHEG (NLO) on Z boson production with pp collisions.

Authors: R. Ramos Blazquez, M. Fernández Moreira, J. Gironés Domínguez
Supervisors: Dr. Armando Bermudez Martínez & Dr. Hannes Jung

Higher Institute of Technology and Applied Sciences, University of Havana, InSTEC.
Deutches Elektronen Synchrotron, DESY.

September, 2021

Table of Contents

1 Introduction

In the past, parton shower generators and next-to-leading-order (NLO) calculations were seen as complementary approaches to computing hadronic interactions. The latter reported, after many tests of QCD carried out at lepton and hadron colliders, convincing evidence that showed that per-perturbative QCD at the NLO level works well, and improves the agreement of theoretical prediction with data. The interest in precise NLO calculations has then shifted in the direction of predicting cross-sections and backgrounds for collider processes. On the other hand, an effort has been made to improve the shower generators under the use of a TMD (transverse momentum distribution) dependence or the use of more precise matrix elements. These efforts have created the way for two new approaches that have impacted the collider phenomenology field. On one side, the (ME+PS) (matrix element and shower matching) while on the other (NLO+PS) NLO calculations interfaced with showers, the latter one would be the essence behind the simulations performed along with this project. The main goal of our work was to get some answers to whether the usage of POWHEG (NLO) + TMDs PS is a good tool, at least, when simulating the Z boson production in proton-proton collisions. Before the idea and the interest came to us we noticed that there was an important background on this type of study by using instead of POWHEG + TMD PS the MC@NLO approach which had been quite successful, but it was lacking a couple of advantages that POWHEG might give to us if we were able to get good enough results that could compete with the results coming from MC@NLO. From the spectrums that we obtained we noticed that we couldn't do the mixing of POWHEG and TMD PS directly, which is by the usage of default parameters of POWHEG, so we needed to understand the dependence of the results on parameters like h_{damp} and p_{tmin} being the last one fundamental on solving the scale-up problem. We used a couple of observables also, to get a closer look inside of the theoretical models that we have used when chasing our main goal. All of these will be better detailed in the next sections of this report.

2 Theory

2.1 The PB (Parton Branching) method

There is a logical question that needs to be answered when moving forward on this work which is, why TMD? The answer is that the inclusion of a tied dependence of PS with TMD (transverse momentum distribution function) results in an improvement to conventional PS methods. It provides, for example, the strong convenience of fixing the small transverse momentum phenomena. The way to include TMD in our analyses was through the Parton Branching method, through which it's being defined a TMD parton density function which evolution will be very tied to the Parton Shower behavior, let's take a deeper look inside this method.
The PB method describes the TMD parton density as:

$$xA_a(x,k_t,\mu) = \Delta_a(\mu)xA_a(x,k_t,\mu_0) + \sum_b \int \frac{dq^2}{q^2}\frac{d\emptyset}{2\pi}\frac{\Delta_a(\mu)}{\Delta_a(q)}\theta(\mu-q)\theta(q-\mu_0)$$

$$\times \quad \int_x^{z_M} dz\, P_{ab}{}^{(R)}\left(\alpha_s\left(f(z,q)\right),z\right)\frac{x}{z}A_b\left(\frac{x}{z},k'_t,q\right) \qquad (1)$$

with $z_M < 1$ defining resolvable branchings, $k_t(q_c)$ being the transverse momentum vector from the propagating (emitted) parton, respectively. The transverse momentum of the parton before branch-ing is defined as $k_{t'} = |k + (1-z)q|$ with $q = q_c/1-z$ being the rescaled transverse momentum vector of the emitted parton, and ϕ being the azimuthal angle between $\vec{q}$ and $\vec{k}$. The argument in α_s is in general a function of the evolution scale q. Higher-order calculations indicate the transverse momentum of the emitted parton as the preferred scale. The real emission branching probability is denoted by $P_{ab}^{(R)} = (\alpha_s(f(z,q)),z)$.
The Sudakov form factor is given by:

$$\Delta_a(z_M,\mu,\mu_0) = \exp\left(-\sum_b \int_{\mu_0{}^2}^{\mu}\frac{dq^2}{q^2}\int_0^{z_M} dz\, z\, P_{ba}{}^{(R)}\right) \qquad (2)$$

Dividing the TMD parton density over $\triangle_a(\mu^2)$ and then differentiating the fraction over μ^2 gives the differential form of the evolution equation describing the probability for resolving a parton with transverse momentum k' and momentum fraction x/z into a parton with momentum fraction x and emitting another parton during a small decrease of μ, if it's being further normalized it would be obtained:

$$\frac{\Delta_a(\mu)}{xA_a(x,k_t,\mu)}d\left(\frac{xA_a(x,k_t,\mu)}{\Delta_a(\mu)}\right) = \sum_b \frac{d\mu^2}{\mu^2}\int_x^{zM} dz\,\frac{d\emptyset}{2\pi}P_{ab}{}^{(R)}\frac{\frac{x}{z}A_b\left(\frac{x}{z},k'_t,\mu\right)}{xA_a(x,k_t,\mu)} \qquad (3)$$

This equation can be integrated between μ_{i-1}^2 and μ^2 to give the no-branching probability (Sudakov form factor) for the backward evolution Δ_{bw}

$$\log\Delta_{bw}(x,k_t,\mu,\mu_i-1) = \log\left(\frac{\Delta_a(\mu)}{\Delta_a(\mu_i-1)}\frac{xA_a(x,k_t,\mu_i-1)}{xA_a(x,k_t,\mu)}\right) \qquad (4)$$

$$= -\sum_b \int_{\mu_{i-1}{}^2}^{\mu^2}\frac{dq'^2}{q'^2}\frac{d\emptyset}{2\pi}\int_x^{zM} dz\, P_{ab}{}^{(R)}\frac{x'A_b(x',k'_t,q')}{xA_a(x,k_t,q')}$$

with $x' = x/z$. This Sudakov form factor is very similar to the Sudakov form factor in ordinary par-ton shower approaches, with the difference that for the PB TMD shower the ratio of PB TMD densities $[x'A_b(x',k'_t,q')]/[xA_a(x,k'_t,q')]$ is applied, which includes a dependence on k_t.
It is important to remark that, when describing conventional showers with standard collinear pdfs; the z_M limit was identified as a source of systematic uncertainties. In this particular parton shower

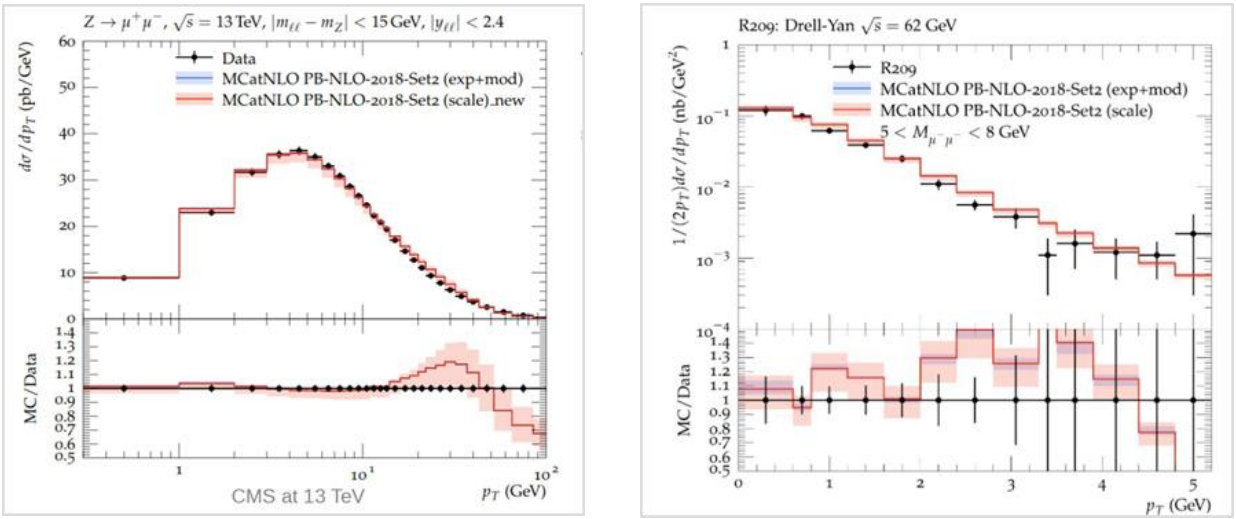
through the PB approach, the same z_M limit is present. However, The PB approach allows a consistent formulation of the parton shower with the PB TMDs, as in both Sudakov form factors Δ_a and Δ_{bw} the same value of z_M is used. The splitting functions $P_{ab}^{(R)}$ contains the coupling:

$$P_{ab}(\alpha_s, z) = \sum_{n=1}^{\infty} \left(\frac{\alpha_s(f(z,q))}{2\pi}\right)^n P_{ab}{}^{(n-1)}(z) \qquad (5)$$

This method brings a couple of advantages, such as:
The fact that the solutions associated with it are valid for LO, NLO, NNLO calculations. On the other side, and one of the biggest improvements is that as long as the initial parameters of the PS stay so tied to the TMDs the uncertainties of the PS process can be relocated with the TMDs uncertainties which can be fitted through experimental data. [1].

2.2 Background support

Before trying POWHEG with TMD PS there was a previous work on mixing the NLO calculations with TMD PS. It was performed through the usage of MC@NLO. The main idea in this particular approach is to subtract the PS approximation for the first emission from the NLO and add it back to the LO plus unresolved contributions. It is important to remark the good results obtained through this approach on the following graphs:



(Fig. 1,2. Drell Yan p_t spectrum for the real (left) and virtual (right) regions)

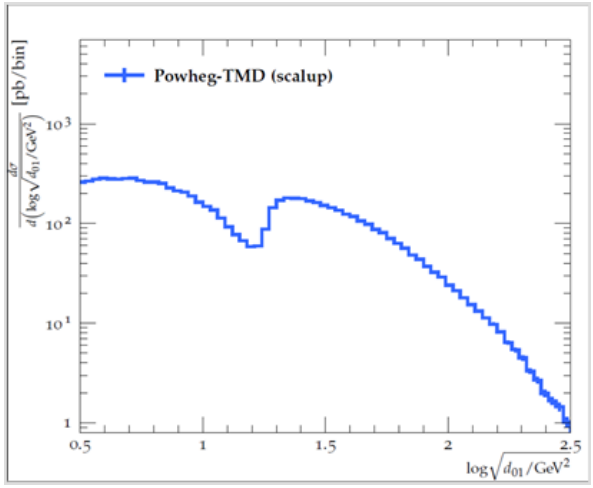
These two results are from a study of PB parton showers applied to Drell Yan (which describes the lepton and anti-lepton productions in high energy hadrons collisions) + jets productions. In both curves, the lepton and anti-lepton were produced through the previous productions of a Z boson as a result of the interactions of a quark and an antiquark from the two initial hadrons.

We can observe the excellent fitting of the MC@NLO calculations with the experimental data of the Drell Yan p_t spectrum for two very different situations, in the left curve, for higher lepton masses (which is the real region, located close to the Z boson peak) and on the right one for low lepton masses (on the virtual region, located far from the Z boson peak). As it can be seen there were obtained very good results for both low and medium p_t values, which are the kind of results that we are looking for in our pursuit.
Now, another logical question could be asked, which is why would anyone be interested on trying to look for a new matching between NLO and TMD PS through POWHEG + PS if it is possible to get such good results with MC@NLO? There are several reasons but the main ones are two, the first, and

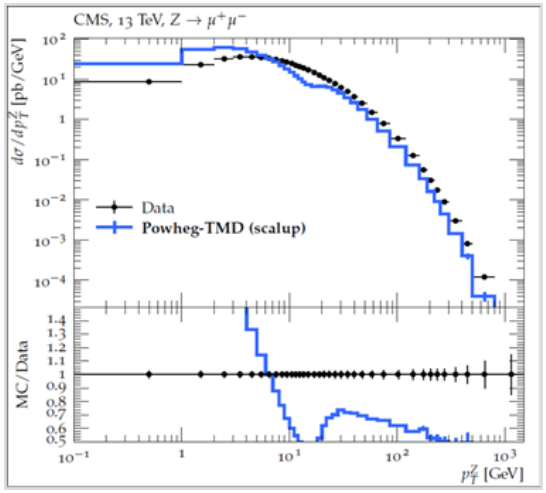
which can be considered the most important one, is the fact that MC@NLO brings negative weights which means nonphysical process probabilities, this can be resolved, otherwise such results couldn't be obtained, but to solve this we need to run longer calculations times, and the second reason is the strong dependence on the method as the whole machinery over the PS method that is being used. And that is why we are interested in using and testing the POWHEG (NLO) + TMDs Parton Shower.

2.3 The scale up problem

POWHEG vouches for (Positive Weights Hardest Emission Generator) and the fundamental idea behind it is to generate on the first moment the hardest radiation by modifying the NLO cross-section and then to give place and to "turn on" the PS generator. The main advantage is that it solves both of the MC@NLO deficiencies; it doesn't bring negative weights into the simulation so it is possible to get equally good fittings with fewer statistics, and it doesn't depend so directly on the PS method that is being used. However, would it work just by mixing POWHEG (NLO) and TMD PS directly? The answer is no, the why, is related to the scale-up problem. One of the most important parameters when simulating non-real radiation in the POWHEG method is the p_{tmin}, if the set default value for this parameter by POWHEG is the one that gets used we would probably get similar results to the following ones:



(Fig. 3: behavior of the d_{01} observable)



(Fig. 4: Transverse momentum of the Z boson)

The graph on the left shows the behavior of the differential jet rates d_{01} observable which usage will be better detailed along this report's following section. The one on the right shows the transverse momentum of the Z boson and as it can be observed both of these curves are not giving place to a good simulation of the non-real radiation, by using the default p_{tmin}. In this case, the number of events affected by the PS are dramatically reduced and this is why it can't produce a soft behavior when merging the real emission region (NLO) with the virtual one (PS). However, it can be solved. Now it would be good to take a deeper look inside how cross sections can be described.

2.4 NLO, PS and pure POWHEG

We can write the cross section for the emission of a parton according to NLO as:

$$\frac{d\sigma^R}{dp_t{}^2} = R(p_t{}^2) \qquad (6)$$

The expression in Eq. 6 is divergent when the radiation becomes soft or collinear to any of the incoming or outgoing partons. In other words, when we try to have the exclusive cross section in the radiation variables we obtain unphysical results in regions of phase space in which a small scale

is resolved. On the other hand, the PS formulation is able to resolve small scales by the subsequent emission of partons (resolvable and non-resolvable) [2]. The PS approximation for the radiation of one resolvable parton from the Born configuration is:

$$\frac{d\sigma^{PS}}{dp_t{}^2} = R^{PS}(p_t{}^2)exp[-\int_{p_t{}^2} dp'_t{}^2 \frac{R\left(p'_t{}^2\right)}{B}] \quad (7)$$

where the Sudakov form factor enforces no further resolvable emission above p_t^2. One should notice that the LO cross section (B) is recovered after integrating over the radiation phase space. This should be the case from the probabilistic interpretation of the Sudakov factor, in other words, the PS is by construction, conserving unitarity. The matching/merging of a PS to a fixed-order relies on the relation between Eq. 6 and Eq. 7. It is interesting to notice that both R and R^{PS} are divergent in the soft and collinear limit, however in the case of Eq. 7 the Sudakov factor overtakes the divergency as it tends to zero more rapidly in this limit.

The POWHEG solution corresponds to define a cross section for the real emission which mimics the PS form in Eq. 7:

$$\frac{d\sigma^{PH}}{dp_t{}^2} = \frac{\bar{B}}{B} R(p_t{}^2)exp[-\int_{p_t{}^2} dp'_t{}^2 \frac{R\left(p'_t{}^2\right)}{B}] \quad (8)$$

where the superscript PH stands for POWHEG emission, which corresponds to the real emission weighted by the Sudakov factor contained in Eq. 8. The expression in Eq. 8 is a mixture of Eq. 6 and Eq. 7. It contains a Sudakov factor, however this factor carries the real matrix elements emission kernel R. The ratio between the Born matrix elements $\overline{B}/B$ is included to ensure that after integrating Eq. 8 over the radiation phase space one obtains the NLO cross section in the Born variables $\overline{B}$. In addition, the inclusion of the Sudakov factor formally does not diminish, at the given order and large P_t, the accuracy of the cross section:

$$R(p_t{}^2)exp[-\int_{p_t{}^2} dp'_t{}^2 \frac{R\left(p'_t{}^2\right)}{B}] \approx R(p_t{}^2)(1 + O(\alpha_S)) \quad (9)$$

3 Results

3.1 Transverse momentum spectrum for NLO and POWHEG. The parameter h_{damp}

Before we show the results and in order to explain them it is primordial to remark some of the mathematical advantages that POWHEG provides. For this it is defined a function $D(p_t; h)$ as:

$$D(p_t; h) = \frac{h^2}{p_t{}^2 + h^2} \quad (10)$$

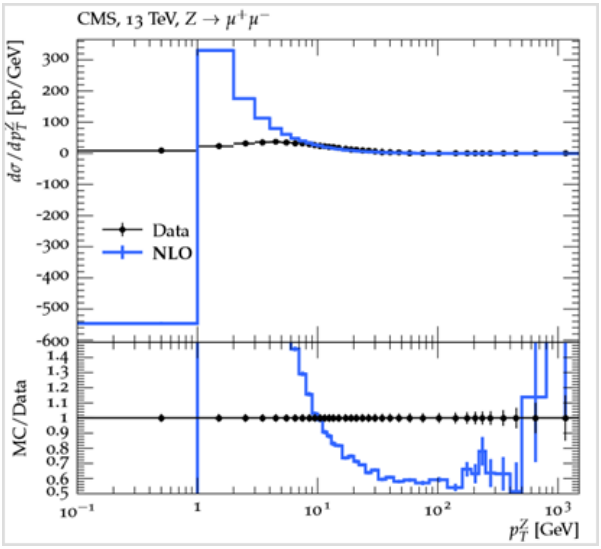
The parameter h is called h_{damp} on the POWHEGBOX library.

Using the function D one can then rewrite Eq. 8 as:

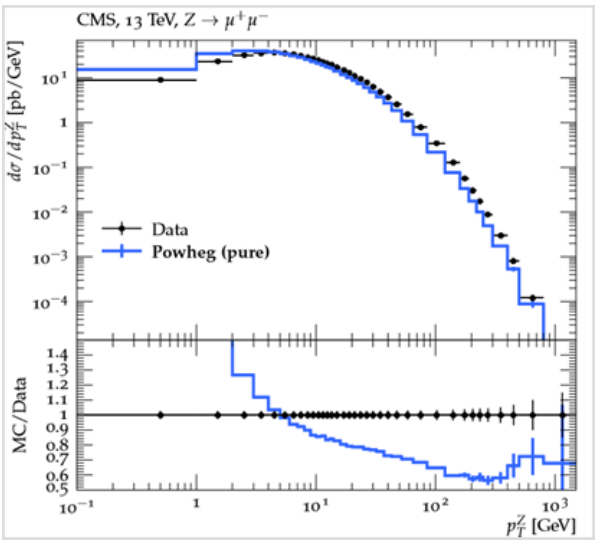
$$\frac{d\sigma^{PH}}{dp_t{}^2} = D\frac{\bar{B}}{B} R(p_t{}^2)exp[-\int_{p_t{}^2} dp'_t{}^2 \frac{R\left(p'_t{}^2\right)}{B}] + (1 - D)R(p_t{}^2) \quad (11)$$

If $D \approx 0$ then Eq. 11 approaches Eq. 2 and no Sudakov weight is applied to the event whereas if $D \approx 1$ the expression Eq. 8 is recovered. So POWHEG provide the possibility to analyze the behaviour of the real contribution and the pure POWHEG.

In order to analyze the NLO transverse momentum spectrum of the Z boson corresponding to a center of mass energy of $13TeV$ we chose $h = 0{,}001$ GeV (see figure 5). For this we use the routine rivet CMS_2019_I1753680.



(Fig.5. p_t spectrum for NLO)

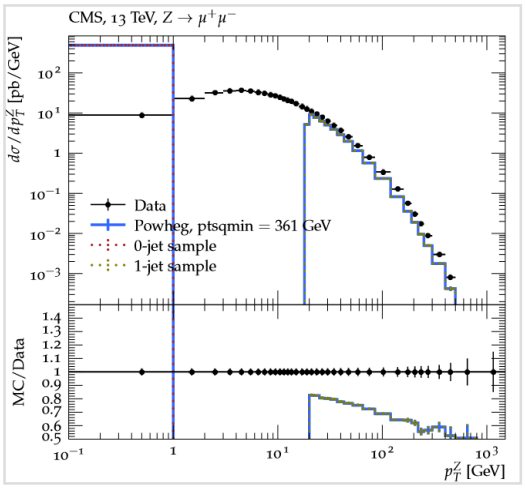


(Fig.6. p_t spectrum for pure POWHEG)

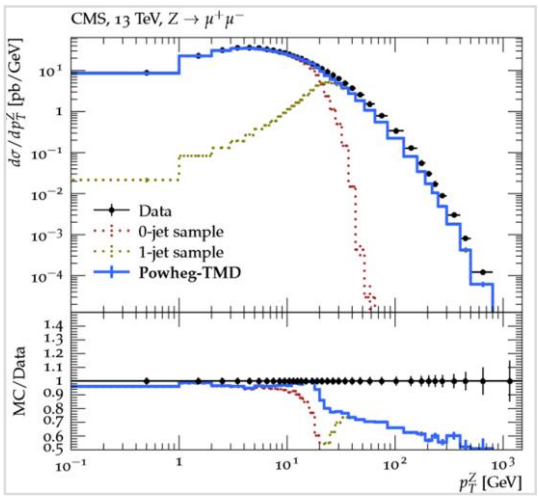
There is a divergence for low p_t in Figure 5 that corresponds to events with non-radiation from exact calculations and the behaviour for higher p_t values is given by events with exact radiation. If we want to analyze the transverse momentum spectrum obtained for pure POWHEG we would have to tend $h \sim \infty$ (see Figure 6). In this case there is not divergences for low p_t because of the influence of the Sudakov factor, but as it was previously treated and as it can be seen there is no accomplish of a good correspondence with the experimental data.

3.2 POWHEG (NLO) +TMD PS. The parameter p_{tmin}

As it was introduced, in order to match TMD PS and NLO in POWHEG properly, so a good agreement with the experimental data could be obtained, a p_{tmin} is defined. Below this value it won't be possible to have exact radiation. We used a $p_{tmin} = 19GeV$ (see Figure 7).



(Fig.7. The p_t spectrum without PS)



(Fig.8. The p_t spectrum with PS)

In the region of low p_t (Figure 7) through conservation of the momentum we obtain a delta in $p_t = 0$. This region below p_{tmin} is suppose to be filled by the PS. When we turned on the PS (see Figure 8), this region of phase space is filled by the PS according to the TMD evolution. At low p_t we have approximate radiation from the PS and at high p_t exact radiation. *As can be seen there is a good agreement with the experimental data and we achieved a good matching between TMD PS and NLO with POWHEG.*

3.3 Observables

The transverse of the z-boson is sensitive to the sum of the momentum of the partons that were emitted, it is no sensitive to each one separately. Taking this into account, let's analyze the matching between TMD and NLO in POWHEG studying the jet production rates at different resolution scales. To this end, splitting scales of jets are constructed using an infrared safe clustering algorithm based on sequential combination of the input momenta. In this analysis the k_t algorithm is used, with distance measures defined for every iteration as follows [3]:

$$d_{ij} = min\left(p_{T,i}^2, p_{T,j}^2\right) x \frac{\Delta R_{ij}^2}{R^2} \quad (12)$$

$$d_{ib} = p_{T,i}^2 \quad (13)$$

where the transverse momentum p_t carries an index corresponding to the i^{th} and j^{th} constituent momentum in the input list, for all possible permutations of i and j in the given clustering step.

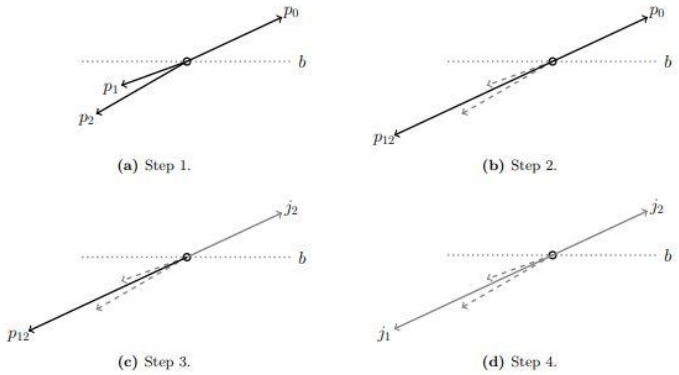
d_{ib} is a measure of the distance in energy between particle i and the initials, the index b denotes the beam line. The input momenta separation ΔR_{ij} is being defined in rapidity and azimuthal angle (η-ϕ) space as $(\Delta R_{ij})^2 = (\eta_i - \eta_j)^2 + (\varphi_i - \varphi_j)^2$. The parameter R governs the average cone size in (η-ϕ) the space around the jet axis.

For a given iteration of the algorithm in which the number of input momenta drops from $k+1$ to k, the associated squared splitting scale d_k is given by the minimum in d_{ij} and d_{ib} scales that are being defined for that iteration step:

- If this minimum is a d_{ij}, the i^{th} and j^{th} momenta in the input list are replaced by their vectorial sum.
- If the minimum is a d_{ib}, the i^{th} momenta is removed from the input collection and is declared a jet.

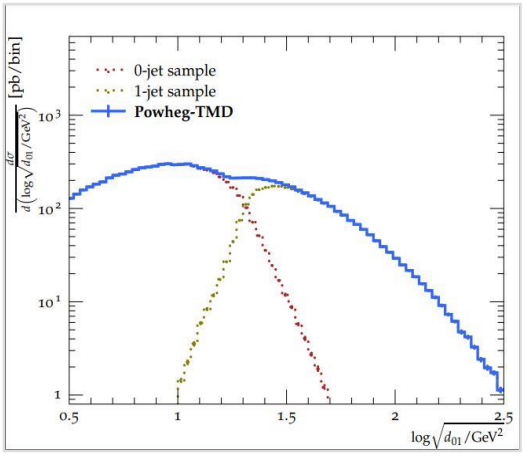
The index k defines the order of the splitting scale, with k = 0 being the last iteration step before the algorithm terminates. Hence the zeroth-order splitting scale, $\sqrt{d_0}$, corresponds to the p_t of the leading k_t-jet, and one can regard the N^{th} splitting scale, $\sqrt{d_N}$, as the distance measure at which an N-jet event it's being resolved as an $(N+1)$-jet event.

The steps of a k_t clustering sequence using three input momenta are illustrated in the following figure:



(Fig. 9 Simplified illustration of the k_t clustering algorithm, starting with three input momenta P_0, P_1 and P_2 (step 1). The dotted line labelled b represents the beam line. In step 2, the minimum distance measured is the one between two input momenta p_1 and P_2, so that the two input momenta are replaced by their vector combination. In step 3, the minimum distance measure is between the P_0 and the beam line, so that P_0 is declared a jet (j_2) and removed from the input list. Finally in step 4, there is only the combined input momentum P_{12} left and so it will be declared a jet (j_1) and the algorithm terminates).

The following graph shows that the approximate radiation of the parton shower is precisely adjusted with the 0 jet sample for low values of d_{01}, as well as for high values of d_{01} where the 1 jet sample can be observed with a very good approximation, which is radiation associated with the exact calculation, proving the effectiveness of the used method:



(Fig. 10 Matching TMD PS and POWHEG (NLO) using the observable d_{01})

4 Conclusions

Along this report we have performed a successful test when matching POWHEG (NLO) with TMD PS. In order to do this several important things were done along the way, so that, they can be integrated in a timeline:
The potentiality of an inclusion of TMD on the PS method was recognized, due to the detection of possible several advantages of tiding the PS method to a TMD parton density through the Parton Branching method. Then it was obtained through MC@NLO, a successful merging between NLO and TMD PS, but the analysis didn't stop there, in fact the deficiencies of MC@NLO (the inclusion of negative weights and the strong dependence over the PS method that was used) were good reasons to keep digging in the possible improvement of the calculations with POWHEG. Although, good results were obtained at last, it wasn't a direct line to success. When merging POWHEG with TMD PS directly, it was crucial to get a deeper understanding of the p_{tmin} parameter in order to get a smooth spectrum without scale up issues. After that it became significant how POWHEG provides an easy way to study NLO and pure POWHEG transverse momentum spectrum by means of the variation of the parameter h_{damp}. In both cases a good agreement with the experimental data at low p_t couldn't be achieved, However, it was obtained with the appropriate combination between TMD PS and NLO. In the last section of this report were showed some of the most important observables that were used in order to get a deeper look inside of the used models. After all of this it was confirmed that there was *a good agreement with the experimental data and a good matching between TMD PS and NLO with POWHEG was obtained.*

5 Acknowledgements

We would like to use the following words to communicate our sincere acknowledgment and gratefulness towards the QCD DESY-CMS group among which, in special manner, to Dr. Hannnes Jung for his desire to include us in this wonderful summer student program at DESY and share through lectures and discussion part of the knowledge he has acquire through his scientific career, to Dr. Armando Bermudez Martinez, for his patience with us, his dedication as our direct supervisor, and the sharpness in every answer to any of our questions and for his good explanations regarding the theory and the computational aspects behind this project. Last, but certainly not least to our professor at our home institute (Higher Institute of Technology and Applied Sciences) Dr. Fernando Guzman for his intention of linking us with this program at DESY and his work on making physics even more attractive for us.

6 Bibliography

[1] S. Baranov, A. Bermudez Martinez, L.I. Estevez Banos, F. Guzman, F. Hautmann, H. Jung, A. Lelek, J. Lidrych, A. Lipatov, M. Malyshev, M. Mendizabal, S. Taheri Monfared, A.M. van Kampen, Q. Wang, H. Yang. *CASCADE3 A Monte Carlo event generator based on TMDs*. 2021. arXiv:2101.10221v2 [hep-ph].

[2] A. Bermudez Martinez. *Measurements and phenomenology of azimuthal correlations in high transeverse momentum multi-jet. Topologies in CMS at the center of mass energy of* $13TeV$. 2019.

[3] ATLAS Collaboration. *Measurement of the kt splitting scales in* $Z \rightarrow$ *events in pp collisions at* $\sqrt{S} = 8TeV$ *with the ATLAS detector*. JHEP08026 (2017).

[4] P. Nason, B. Webber. *Next-to-Leading-Order Event Generators*. Cavendish-HEP-2012-02, CERN-PH-TH-2012-028. 2012. https://arxiv.org/abs/1202.1251v1.

[5] *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].