Proceedings of the LHCP2015 Conference, St. Petersburg, Russia, August 31 - September 5, 2015

Editors: V.T. Kim and D.E. Sosnov

Soft QCD models and general-purpose Monte Carlo simulation

ANDRZEJ SIÓDMOK[1,2]

[1]*Theoretical Physics Department, CERN, Geneva, Switzerland.*
[2]*The Henryk Niewodniczański Institute of Nuclear Physics Polish Academy of Sciences Radzikowskiego 152, 31-342 Kraków, Poland.*

andrzej.siodmok@cern.ch

Abstract. We begin with a brief description of the main building blocks of Monte Carlo event generators (MCEG) for the full simulation of hadron–hadron collisions. Next, we focus on soft QCD models and in particular on Multiple Partonic Interaction (MPI) models that are implemented in MCEG. Finally, we present a comparison of three main Monte Carlo event generators, Herwig++, PYTHIA and SHERPA, and also a cosmic-rays model EPOS to a range of LHC data sets which are sensitive to the MPI activity.

INTRODUCTION

Monte Carlo event generators (MCEG) are of crucial importance for particle physics. They provide fully exclusive simulations of high-energy collisions, therefore they are used by almost all experimental collaborations to plan their experiments and analyze their data, and by theorists to simulate the complex final states of the fundamental interactions which are essential in the search for Beyond the Standard Model physics. Currently there are three families of MCEG available which all offer suitable frameworks for LHC physics: Herwig(6/++/7) [1, 2, 3], Pythia(6/8)[4, 5] and Sherpa[6, 7]. In this short contribution we briefly discuss the main building blocks of a Monte Carlo event generation[1] and focus on soft QCD models and in particular Multiple Parton Interaction (MPI) models. MPI models are essential for a proper description of the minimum-bias (MB) and underlying event (UE) data from hadron colliders. The amount of UE activity at the LHC is measured, so it is tempting to think that the contribution of the UE is known. However, there are observables that depend on correlations or fluctuations away from average value of the UE, including, to varying extents, any measurement relying on jets or isolation criteria. In fact, almost every observable[2] that will be used for beyond the standard model searches or precision measurements falls into this class, so the correction must be represented by a model tuned to data, rather than by a single number measured from data.

Event Generation

The main building blocks of MCEG needed for a simulation of a proton-proton collision at the LHC are:

1. Hard process
2. Parton shower
3. Hadronization
4. Multiple Interactions
5. Decays of unstable hadrons

[1]For a detailed description of MCEG with much more information on the physics background we refer the reader to the MCnet review paper on Monte Carlo event generators [8] and the much shorter but more recent documents [9, 10].
[2]Some measurements based on boosted jet substructure techniques are examples of an exception in that respect, see for example [11, 12, 13].

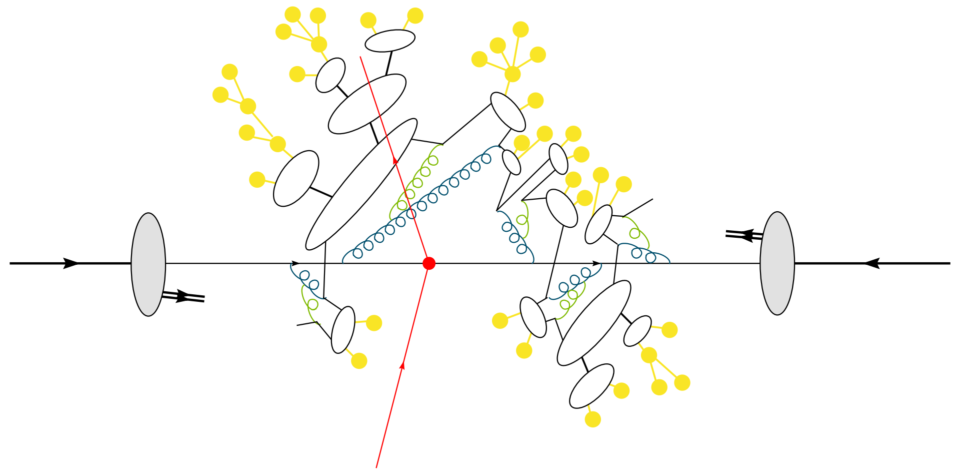


FIGURE 1. Drawing of the simulation of a pp collision shown in [14]. The MPI are not included in this figure.

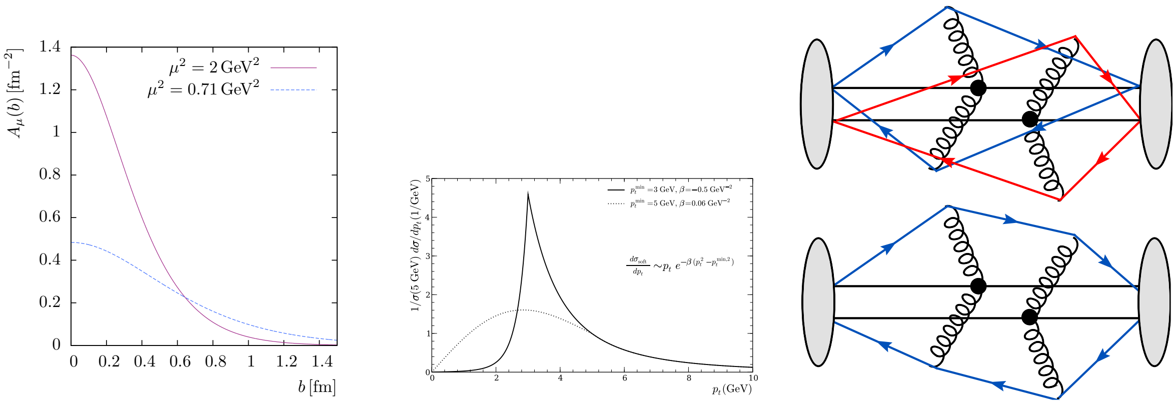
In Fig. 1 we show a cartoon of the event generation (excluding MPI). The generation begins with a hard signal process, in the figure this is lepton pair production (red). Then the parton shower evolution (marked with the dark and light green gluon curly lines) starts from the hard process and evolves downwards in momentum scale to a point $\sim$ 1 GeV where perturbation theory breaks down. At this scale the partonic degrees of freedom are converted into hadrons (yellow circles) via a hadronization model. In the case of Fig. 1 the cluster hadronization model (white blobs) is presented. The cluster model [15] was introduced in Herwig and also implemented in Sherpa[16]. While Pythia's hadronization model is based on the famous Lund string model [17]. The last step of event generation is based on the fact that many of these hadrons (yellow blobs) are not stable particles and therefore decay. In addition to this sequence of steps, all initiated by the hard subprocess, there may be additional semi-hard processes, called multiple partonic interactions, which are the main subject of this note. These are mostly fairly soft QCD interactions that also undergo all of the steps described above for the hard process and produce additional particles in all the available phase space.

Multiple parton interactions

The first detailed Monte Carlo model for perturbative MPI was proposed in [18] and was the main model in Pythia for a long time. The models implemented in Herwig [19, 20] and the first model in Sherpa are based on a similar physical picture. There are however important details where the approach deviates from the formalism in Pythia. For example in the recent Pythia versions the additional hard scatters are interleaved with the parton shower [21, 22] which is not the case in Sherpa or Herwig++. This approach allows a picture where MPI and the parton shower radiation are interleaved in one common sequence of decreasing $p_\perp$ values. Before we highlight some other differences between the models let us first briefly describe the MPI model implemented in Herwig++ [23, 24]. The model is formulated in impact parameter space. At fixed impact parameter $\vec{b}$, multiple parton scatterings are assumed to be independent which leads to an expression for the average number of hard interactions:

$$\bar{n}(\vec{b}, s) = A(\vec{b}; \mu^2)\sigma^{\text{inc}}(s; p_t^{\text{min}}) , \tag{1}$$

where $A(\vec{b}; \mu^2)$ is a so-called overlap function describing the two colliding protons as a function of the impact parameter $\vec{b}$. In Herwig++ $A(\vec{b}; \mu^2)$ is modelled by the electromagnetic form factor, see Fig. 2(a). The parameter μ^2 appearing in the function is one of the main tuning parameters and can be interpreted as an effective inverse proton radius.In Pythia 8 there are five different options for the overlap function, including a double Gaussian matter distribution which is very similar to the one obtained from the electromagnetic form factor. In Herwig++, parton-parton scatterings ($\sigma^{\text{inc}}(s; p_t^{\text{min}})$) are divided into soft and hard by a parameter p_t^{min}, which is another main tuning parameter in the model. Above p_t^{min}, scatters are assumed to be perturbative and take place according to leading order QCD matrix elements. Below p_t^{min}, scatters are assumed to be non-perturbative, with "Gaussian" transverse momentum distribution and valence-like longitudinal momentum distribution, see Fig. 2(b). In Pythia in order to damp the cross



(a) The overlap function used in Herwig++ (two values of μ^2 parameter).

(b) Division of parton-parton scatterings into soft and hard in Herwig++.

(c) An example of two different colour connection of MPI.

FIGURE 2. Main building blocks of MPI model in Herwig++.

section at $p_T \to 0$, the cross section is multiplied by a regularization factor $(p_T^2/(p_T^2 + p_{T0}^2))^2$, and simultaneously the α_s is evaluated at a scale $p_T^2 + p_{T0}^2$, where p_{T0} is a free parameter of the model. It is interesting to mention that there is a CMS measurement [25] which gives some hints towards the mechanism by which the parton-parton cross sections are unitarised when approaching low-p_T. In both Herwig++ and Pythia p_{T0} and $p_t^{\min}$ vary with energy. In Herwig++ according to:

$$p_t^{\min}(s) = p_{\perp,0}^{\min}\left(\frac{\sqrt{s}}{E_0}\right)^b, \tag{2}$$

and, in principle, $p_{\perp,0}^{\min}$ and b are fitted to data, with $E_0 = 7$ TeV. A similar evolution of p_{T0} is implemented in Pythia. Finally, the last and the least understood building block of MPI models is so-called colour reconnection (CR), see Fig. 2(a). The idea behind CR is based on colour preconfinement [26], which implies that parton jets emerging from different partonic interactions are colour-connected (clustered in Herwig++) if they are located closely in phase space. As the MPI model does not take that into account, those colour connections have to be adapted afterwards by means of a CR procedure. In Herwig++ the CR model defines the distance between two partons based on their invariant mass, i.e the distance is small when their invariant mass (cluster mass) is small. Therefore, the aim of the CR model is to reduce the colour length $\lambda \equiv \sum_{i=1}^{N_{cl}} m_i^2$, where N_{cl} is the number of clusters in an event and m_i is the invariant mass of cluster i. A similar model of CR was implemented some time ago in Pythia [4] and recently there were new ideas of how to improve the model for example by going beyond leading colour approximation [27, 28]. In Fig. 3 we show an observable that is sensitive to CR: $\langle p_t \rangle$ vs N_{ch}. We expect that it should be almost flat when the MPI system hadronizes independently (i.e. without additional correlations due, to for example CR). In Fig. 3 (left panel) we see that the measured distribution is not flat and that all generators (except Sherpa which does not have a CR model) are able to describe $\langle p_t \rangle$ vs N_{ch} MB data collected at 7 TeV. In Fig. 3 (middle panel) we show charged-particle multiplicities as a function of the pseudorapidity, ATLAS MB data was collected at the new energy frontier 13 TeV [29] which means that it was not used for the tuning of MCEG. Pythia 8 seems to describe the data reasonably well, however the cosmic-rays model EPOS [30, 31] seems to describe the MB data even better. On the other hand we will see, that it fails to describe the UE data. Herwig++ fails to describe the MB data, one of the reasons for this is that it was only tuned to the UE data. However it is worth noticing that the older UE tune of Herwig++ gives a much better description of MB data, Fig. 3 (right panel) shows the potential of the model. This observation brings us to the point that in general MPI models have a problem describing both UE and MB data sets simultaneously (using the same tune). Similar tension is also visible in the disagreement of MPI models with the effective cross section for double-parton scattering σ_{eff} [32, 33]. However, in the recent publication [34] the authors provided a Herwig++ tune which is able to describe UE data over the range of energies in addition to σ_{eff} measured by the CDF and more modern experiments. The other type of measurements which are very sensitive to MPI activity are Underlying Event measurements which are made relative to a leading object (the hardest charged track or a jet). Then, the transverse

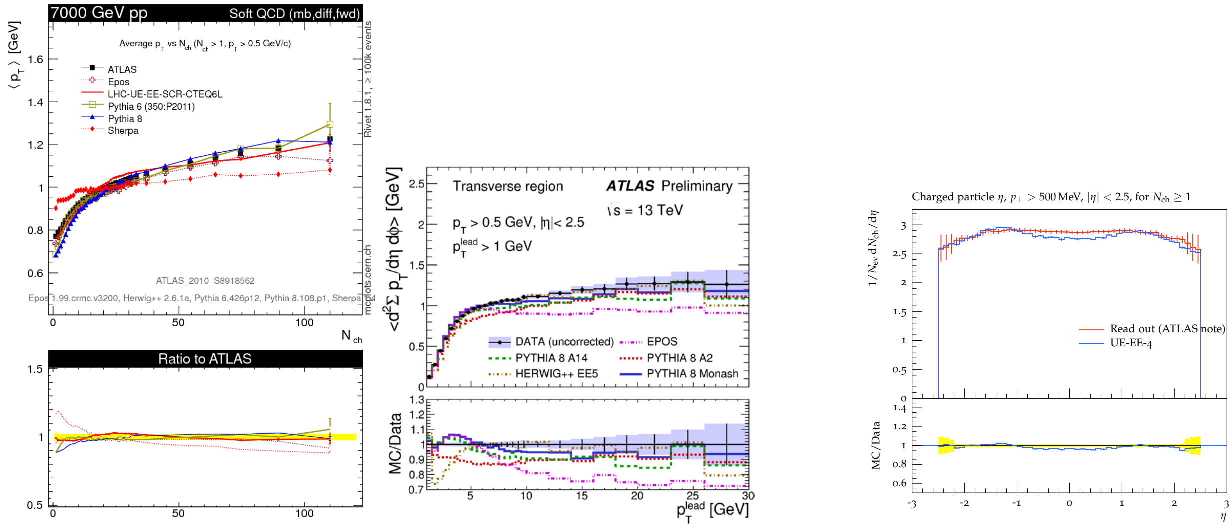


FIGURE 3. ATLAS MB data showing the average transverse momentum as a function of the number of charged particles at 7 TeV— [35] (left panel), charged-particle multiplicities as a function of the pseudorapidity at 13 TeV [29] (middle panel) and the same as the middle panel but with data compared to an older tune EE-4 of Herwig++. The dots represent the data and the curves represent the predictions from different MC models.

plane is subdivided in azimuthal angle ϕ relative to this leading object at $\phi = 0$. The region around the leading object, $|\phi| < \pi/3$, is called the "towards" region and the opposite region, where we usually find a recoiling hard jet, $|\phi| > 2\pi/3$, is called the "away" region. The remaining region, transverse to the leading object and its recoil, where the underlying event is expected to be least 'contaminated' by activity from the hard subprocess, is called the "transverse" region. Therefore, in Fig. 4 we show the mean scalar $p_\perp$ sum of stable particles as a function of $p_\perp^{\mathrm{lead}}$ in the transverse

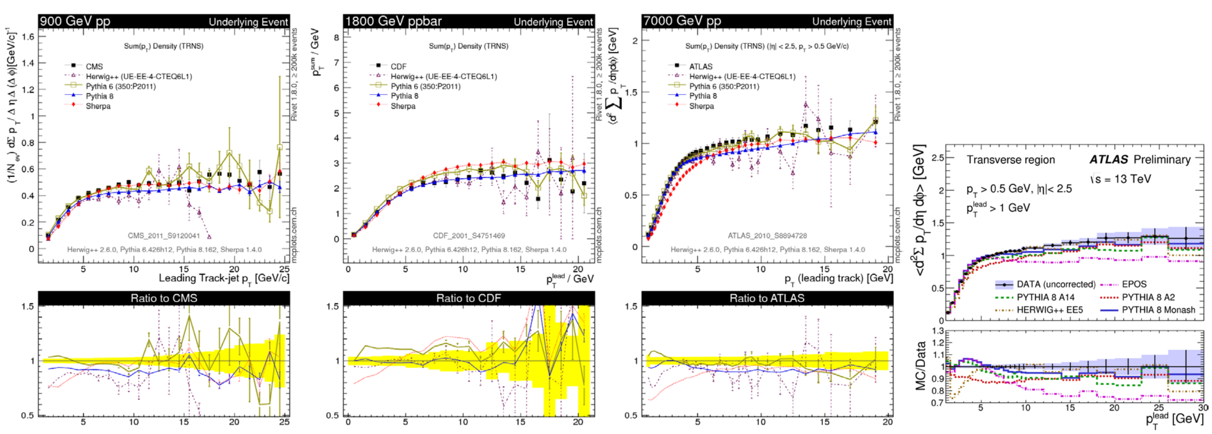


FIGURE 4. ATLAS data at 900GeV (1st column), CDF data at 1800GeV (2nd column), ATLAS data at 7TeV (3rd column) and 13TeV (4th column) showing the mean scalar $p_\perp$ sum of the stable charged particles in the "transverse" area as a function of $p_\perp^{\mathrm{lead}}$.

region measured at different energies (from 900 GeV to 13 TeV). The plateau at high p_T in Fig. 4 shows nicely that the transverse activity decouples from the leading object momentum for large momenta, hence the interpretation as underlying event activity is correct. We can also see that all MCEG describe the data collected at different collider energies reasonably well, which would not be possible without good modeling of MPI. In the last column we show new data collected at 13 TeV, this is not used in the tuning procedure and therefore can be used to test the predictions of the models. We can see that the models are in good agreement with the data, except, as mentioned before, EPOS. This is because EPOS currently has no hard component in the model. Finally, despite the significant success of MPI models in describing a range of MB and UE there are still open problems. Let me show just one which is clearly visible in Fig. 5, taken from a CMS measurement of strange particle production in UE [36]. We can see that if we start to ask more detailed questions about the nature of the UE (for example asking about strange particles) some models have significant problems producing a correct answer. However, there is a constant effort to improve the MPI models and already there has been a first attempt [37] to solve the problem in Fig. 5. Finally, let me just mention that there is a new approach to MPI in Sherpa called SHRiMPS. It aims at a smooth inclusion of diffractive and soft interactions into the multiple iteraction picture, based on a Gribov–Regge formalism [38]. However it is not yet fully developed and tuned, therefore we don't show its results.

Conclusion

Amazing progress has been made in the development of MCEG over the previous decades. We have seen that a wide range of LHC data sensitive to soft QCD is well described by MCEG. The other LHC data sets triggered new developments of the MC event generators. Finally, there are data sets which show that there is still need for further development. As the LHC studies more subtle effects, generators must keep increasing their precision. This is only possible with an appropriate input from the experimental community which gives a solid basis for future developments of MCEG.

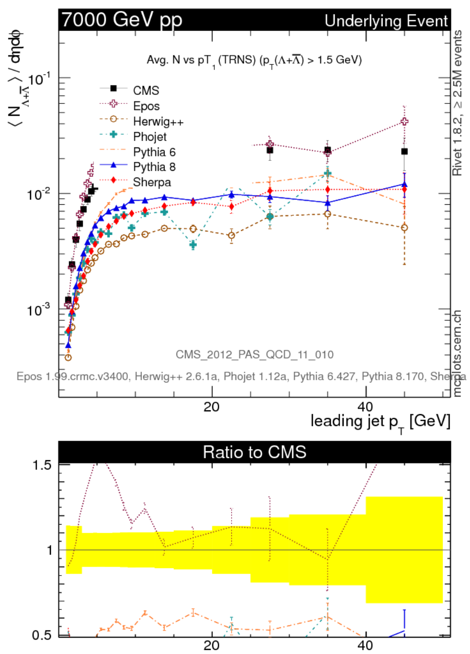


FIGURE 5. CMS data at at 7TeV showing the multiplicity density of the $\Lambda + \bar{\Lambda}$ particles in the "transverse" area as a function of $p_{\perp}^{\text{lead}}$.

ACKNOWLEDGMENTS

We thank the organizers for the very pleasant and fruitful conference. We wish to acknowledge Stephen Webster for his critical reading of these proceedings.

REFERENCES

[1] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, JHEP **01**, p. 010 (2001), arXiv:hep-ph/0011363 [hep-ph] .

[2] M. Bahr *et al.*, Eur. Phys. J. **C58**, 639–707 (2008), arXiv:0803.0883 [hep-ph] .
[3] J. Bellm *et al.*, (2015), arXiv:1512.01178 [hep-ph] .
[4] T. Sjostrand, S. Mrenna, and P. Z. Skands, JHEP **05**, p. 026 (2006), arXiv:hep-ph/0603175 [hep-ph] .
[5] T. Sjstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, Comput. Phys. Commun. **191**, 159–177 (2015), arXiv:1410.3012 [hep-ph] .
[6] T. Gleisberg, S. Hoeche, F. Krauss, A. Schalicke, S. Schumann, and J.-C. Winter, JHEP **02**, p. 056 (2004), arXiv:hep-ph/0311263 [hep-ph] .
[7] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, JHEP **02**, p. 007 (2009), arXiv:0811.4622 [hep-ph] .
[8] A. Buckley *et al.*, Phys. Rept. **504**, 145–233 (2011), arXiv:1101.2599 [hep-ph] .
[9] M. H. Seymour and M. Marx, "Monte Carlo Event Generators," (2013) arXiv:1304.6677 [hep-ph] .
[10] A. Sidmok, Acta Phys. Polon. **B44**, 1587–1601 (2013).
[11] M. H. Seymour, Z. Phys. **C62**, 127–138 (1994).
[12] A. Altheimer *et al.*, Eur. Phys. J. **C74**, p. 2792 (2014), arXiv:1311.2708 [hep-ex] .
[13] M. Dasgupta, A. Powling, and A. Siodmok, JHEP **08**, p. 079 (2015), arXiv:1503.01088 [hep-ph] .
[14] S. Gieseke, Nucl. Phys. Proc. Suppl. **222-224**, 174–186 (2012).
[15] B. R. Webber, Nucl. Phys. **B238**, p. 492 (1984).
[16] J.-C. Winter, F. Krauss, and G. Soff, Eur. Phys. J. **C36**, 381–395 (2004), arXiv:hep-ph/0311085 [hep-ph] .
[17] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, Phys. Rept. **97**, 31–145 (1983).
[18] T. Sjostrand and M. van Zijl, Phys. Rev. **D36**, p. 2019 (1987).
[19] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, Z. Phys. **C72**, 637–646 (1996), arXiv:hep-ph/9601371 [hep-ph] .
[20] I. Borozan and M. H. Seymour, JHEP **09**, p. 015 (2002), arXiv:hep-ph/0207283 [hep-ph] .
[21] T. Sjostrand and P. Z. Skands, JHEP **03**, p. 053 (2004), arXiv:hep-ph/0402078 [hep-ph] .
[22] T. Sjostrand and P. Z. Skands, Eur. Phys. J. **C39**, 129–154 (2005), arXiv:hep-ph/0408302 [hep-ph] .
[23] M. Bahr, S. Gieseke, and M. H. Seymour, JHEP **07**, p. 076 (2008), arXiv:0803.3633 [hep-ph] .
[24] M. Bahr, J. M. Butterworth, S. Gieseke, and M. H. Seymour, "Soft interactions in Herwig++," in *Proceedings, 1st International Workshop on Multiple Partonic Interactions at the LHC (MPI08)* (2009), pp. 239–248, arXiv:0905.4671 [hep-ph] .
[25] C. Collaboration (CMS), (2014).
[26] D. Amati and G. Veneziano, Phys. Lett. **B83**, p. 87 (1979).
[27] J. R. Christiansen and P. Z. Skands, JHEP **08**, p. 003 (2015), arXiv:1505.01681 [hep-ph] .
[28] J. R. Christiansen and T. Sjostrand, Eur. Phys. J. **C75**, p. 441 (2015), arXiv:1506.09085 [hep-ph] .
[29] "Charged-particle distributions in sqrt(s)=13 TeV p p interactions measured with the ATLAS detector at the LHC," Tech. Rep. ATLAS-CONF-2015-028 (CERN, Geneva, 2015).
[30] H. J. Drescher, M. Hladik, S. Ostapchenko, T. Pierog, and K. Werner, Phys. Rept. **350**, 93–289 (2001), arXiv:hep-ph/0007198 [hep-ph] .
[31] K. Werner, I. Karpenko, T. Pierog, M. Bleicher, and K. Mikhailov, Phys. Rev. **C82**, p. 044904 (2010), arXiv:1004.0805 [nucl-th] .
[32] F. Abe *et al.* (CDF), Phys. Rev. **D56**, 3811–3832 (1997).
[33] M. Bahr, M. Myska, M. H. Seymour, and A. Siodmok, JHEP **03**, p. 129 (2013), arXiv:1302.4325 [hep-ph] .
[34] M. H. Seymour and A. Siodmok, JHEP **10**, p. 113 (2013), arXiv:1307.5015 [hep-ph] .
[35] G. Aad *et al.* (ATLAS), New J. Phys. **13**, p. 053033 (2011), arXiv:1012.5104 [hep-ex] .
[36] "Measurement of Strange Particle Production in Underlying Events in proton-proton collisions at sqrt(s)," Tech. Rep. CMS-PAS-QCD-11-010 (CERN, Geneva, 2011).
[37] C. Bierlich and J. R. Christiansen, Phys. Rev. **D92**, p. 094010 (2015), arXiv:1507.02091 [hep-ph] .
[38] V. A. Khoze, F. Krauss, A. D. Martin, M. G. Ryskin, and K. C. Zapp, Eur. Phys. J. **C69**, 85–93 (2010), arXiv:1005.4839 [hep-ph] .
[39] "Leading Track Underlying Event at 13 TeV," Tech. Rep. ATL-PHYS-PUB-2015-019 (CERN, Geneva, 2015).