WHY DO WE NEED DETAILED ANALYSIS OF DATA TO UNDERSTAND THE MECHANISMS OF *pp* COLLISIONS AT THE LHC?*

G. Paić [a], L. Serkin [b], G.G. Barnaföldi [c], G. Bíró [c,d]

[a]Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México
Apartado Postal 70-543, 04510 Ciudad de México, Mexico
[b]Facultad de Ciencias, Universidad Nacional Autónoma de México
Circuito Exterior s/n, Ciudad Universitaria
04510, Coyoacán, Ciudad de México, Mexico
[c]HUN-REN Wigner Research Centre for Physics
29–33 Konkoly-Thege Miklós Rd., 1121 Budapest, Hungary
[d]Eötvös Loránd University, Institute of Physics
Pázmány Péter Sétány 1/A, 1117 Budapest, Hungary

Received 9 July 2025, accepted 3 December 2025,
published online 19 December 2025

In this paper, we investigate the transverse momentum spectra of charged particles in proton–proton collisions at the LHC using detailed event-by-event simulations with the PYTHIA 8 event generator. Rather than relying solely on average observables such as the mean transverse momentum, we analyse the full shape of the spectra across different charged-particle multiplicity intervals. Our results show that significant variations in spectral shape occur with increasing multiplicity, revealing a transition from soft to hard processes that is not captured by global mean values. Comparisons with other commonly used Monte Carlo models, including Herwig 7 and EPOS 4, demonstrate that while the mean values are broadly consistent across generators, the underlying spectral shapes differ.

DOI:10.5506/APhysPolBSupp.18.6-A15

1. Introduction

The way data from proton–proton (*pp*) or heavy-ion collisions are commonly analysed often involves presenting averages of relevant observables — such as the mean transverse momentum, $\langle p_{\mathrm{T}} \rangle$, or mean charged multiplicity, $\langle N_{\mathrm{ch}} \rangle$, see *e.g.* Refs. [1–3]. However, this practice can be problematic or

* Presented at the V4-HEP 4 — Theory and Experiment in High Energy Physics Workshop, Bratislava, Slovakia, 26–28 July, 2023.

even dangerous, as it may obscure or blur important features of the collisions. Scientific progress often comes from examining the details, not just the averages.

Let us consider a well-known example: the Landau curve representing the energy loss of particles [4]. Quoting only the mean offers little information, whereas examining the entire distribution provides important insight into the different underlying processes contributing to it.

Bringing the discussion closer to our field, we know that the transverse momentum spectra observed in nucleus–nucleus collisions can, at first approximation, be viewed as consisting of two contributions: a low-p_{T} component associated with soft thermal processes, and a high-p_{T} tail resulting from jet fragmentation [5–7]. Analyses often extract $\langle p_{\mathrm{T}} \rangle$ as a function of the event charged multiplicity N_{ch}. However, these two competing mechanisms are not cleanly separated in the p_{T} distribution. Instead, they mix significantly — especially in high-multiplicity events, where fragmentation contributes to many low-p_{T} particles, effectively "contaminating" the soft component [8–10].

The LHC at CERN is a powerful microscope that allows us to perform a detailed exploration of the outcome of the collisions. But the question we would like to raise is: are its possibilities being fully exploited? This paper is a call to investigate interactions in more detail than it is currently done, with the firm conviction that novel approaches may lead to new discoveries.

2. Predictions of PYTHIA Monte Carlo event generator on an event-by-event basis

In order to explore both the potential and the necessity of adopting new methods of analysis, we make use of the well-known and widely employed for studying LHC physics PYTHIA 8 (version 8.309) Monte Carlo (MC) simulation [11]. 500 million inelastic pp collisions were generated at a centre-of-mass energy of $\sqrt{s} = 13$ TeV using the ATLAS A14 global LHC data tune [12]. The results are presented for primary charged particles in the kinematic range of $0.15\ \mathrm{GeV} \leq p_{\mathrm{T}} \leq 20\ \mathrm{GeV}$ [1], and pseudorapidity interval of $|\eta| \leq 4$, following the definition provided in Ref. [13]. The outcome of the simulation for the event-by-event mean transverse momentum at a given multiplicity is shown in Fig. 1.

Figure 1 shows an increase of $\langle p_{\mathrm{T}} \rangle$ with multiplicity, reflecting the interplay between soft and hard processes in the event. At low multiplicities, events are dominated by soft, low-p_{T} particle production. As multiplicity increases, the contribution from semi-hard and hard processes, such as mini-jets and jet fragmentation, becomes more significant, and hence the value

[1] Unless otherwise specified, all quantities are expressed in natural units ($c = 1$).

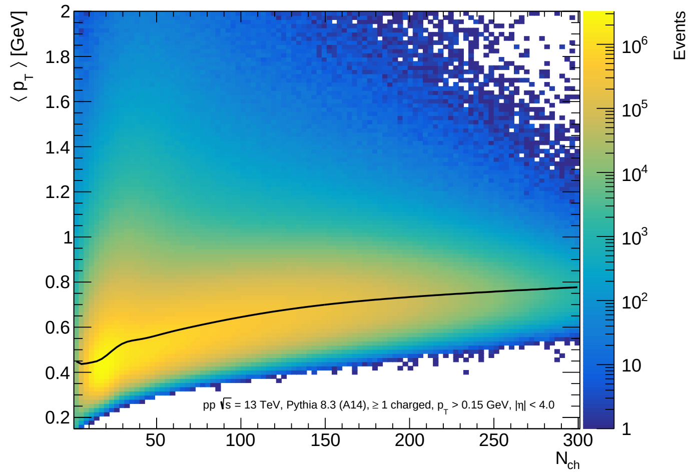

Fig. 1. Event-by-event mean transverse momentum as a function of the event charged particle multiplicity for pp collisions at $\sqrt{s} = 13$ TeV simulated with PYTHIA 8. The colour scale indicates the number of events on a logarithmic scale; the black line represents the $\langle p_{\mathrm{T}} \rangle$ trend.

of $\langle p_{\mathrm{T}} \rangle$ is higher. The black line in Fig. 1 represents the $\langle p_{\mathrm{T}} \rangle$ value in each multiplicity bin. While it captures the central trend of the distribution, a wide range of p_{T} values is created in the collisions.

Figure 2 presents the p_{T} distributions of charged particles for several event multiplicity intervals in pp collisions at 13 TeV, as simulated with PYTHIA 8. While the $\langle p_{\mathrm{T}} \rangle$ values increase only gradually with multiplicity, from approximately 0.44 GeV in low-multiplicity events to 0.69 GeV in the highest bin, the shape of the spectra varies significantly across multiplicity bins. This indicates that changes in particle production dynamics are not fully captured by the mean alone.

Figure 3 shows the ratio of the normalized p_{T} spectra in each charged-particle multiplicity bin to the (minimum bias) inclusive spectrum for pp collisions at 13 TeV. Each subplot corresponds to one multiplicity interval, matching those defined in Fig. 2.

At low multiplicities (top row of Fig. 3), the spectra are slightly softer than the inclusive reference, which leads to ratios above one at low-p_{T} and close to unity at higher-p_{T} once the spectra are normalized to equal area. As the charged-particle multiplicity increases (middle and bottom rows), the spectral shape becomes progressively harder. The excess at low-p_{T} is reduced, and the ratios tend to rise gradually at higher-p_{T}. In the highest multiplicity intervals (bottom right), this hardening becomes most pronounced, with the ratio exceeding unity across a wide p_{T} range relative to the inclusive spectrum.

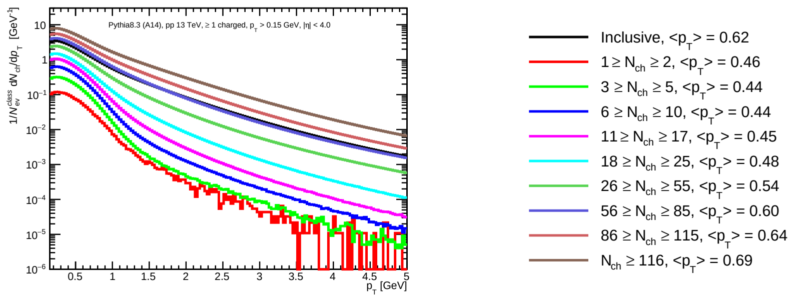

Fig. 2. The transverse momentum spectra of primary charged particles in the kinematic range 0.15 GeV $\leq p_{\mathrm{T}} \leq$ 20 GeV, and $|\eta| \leq 4$ for different charged-particle multiplicity intervals in *pp* collisions at 13 TeV, simulated with PYTHIA 8 (tune A14). Each curve corresponds to a specific multiplicity range as indicated in the legend on the right-hand side. Each spectrum is normalized to the number of events in that multiplicity range.

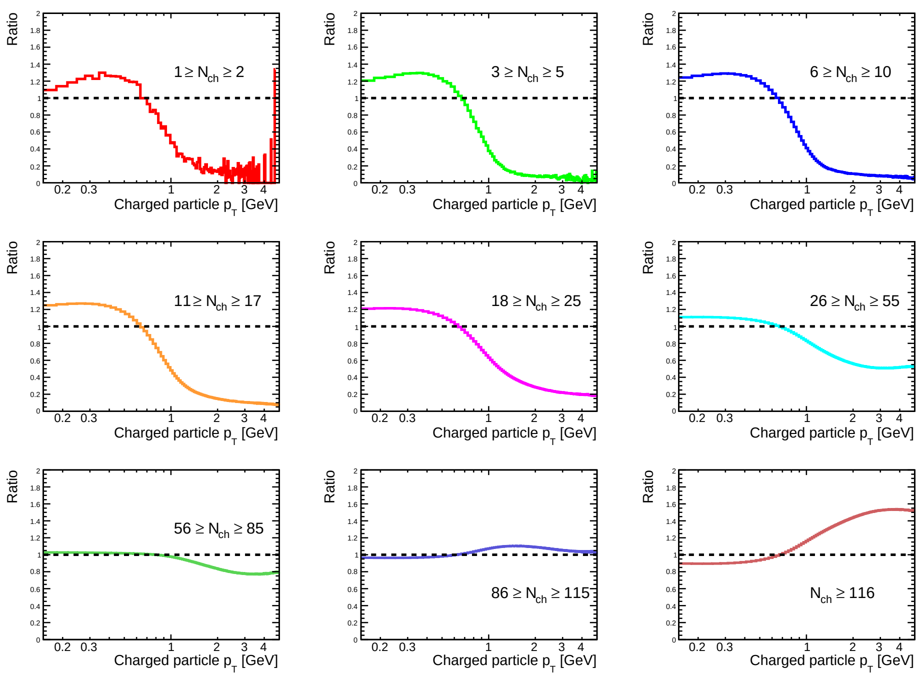

Fig. 3. Ratio of the normalized transverse momentum spectra in different charged-particle multiplicity intervals to the inclusive spectrum, for *pp* collisions at 13 TeV simulated with PYTHIA 8. Each panel corresponds to a specific multiplicity range, matching those defined in Fig. 2. The dashed horizontal line at unity represents the inclusive reference.

These trends indicate that the contribution from hard processes becomes increasingly important with event multiplicity and dominates the spectrum at high-p_{T}, while soft production remains significant at low-p_{T}.

3. Comparison with the predictions of other models

In order to understand a little bit better the underlying mechanisms behind the observed p_{T} spectra and their evolution with multiplicity, it is essential to compare the PYTHIA 8 results with predictions from other MC models. While PYTHIA 8 provides a well-tuned baseline for minimum bias pp collisions, alternative event generators, such as Herwig and EPOS, implement different hadronisation models, multiple parton interaction schemes, and collective effects.

Figure 4 compares the charged-particle p_{T} spectra as predicted by three event generators, PYTHIA 8.3, Herwig 7.2 [14], and EPOS 4 [15], for pp collisions at 13 TeV. Spectra are shown for different charged-particle multiplicity

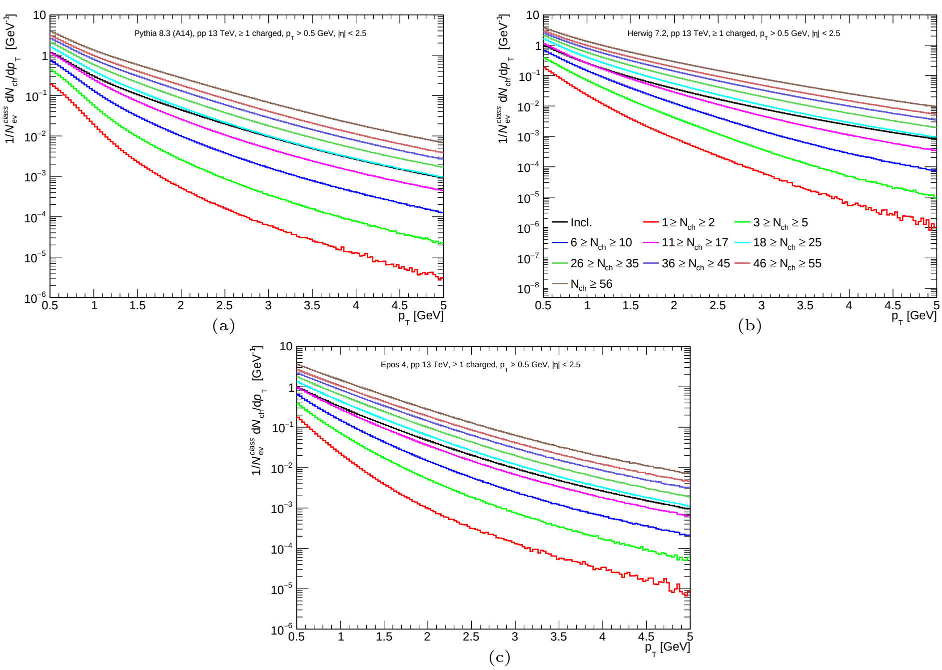

Fig. 4. Comparison of charged-particle p_{T} spectra simulated with (a) PYTHIA 8, (b) Herwig 7.2 and (c) EPOS 4 in pp collisions at 13 TeV. Each spectrum is shown for a range of charged-particle multiplicities, with the corresponding $\langle p_{\mathrm{T}} \rangle$ values indicated in the legend shown in Table 1. Each spectrum is normalized to the number of events in that multiplicity range.

intervals, and the corresponding $\langle p_{\rm T}\rangle$ values are listed in Table 1. In this case, the results are presented for primary charged particles in the kinematic range of 0.5 GeV $\leq p_{\rm T} \leq$ 20 GeV, and a pseudorapidity interval of $|\eta| \leq 2.5$.

Table 1. Mean transverse momentum $\langle p_{\rm T}\rangle$ in GeV for different multiplicity ranges in PYTHIA 8, Herwig 7, and EPOS 4.

Multiplicity range	PYTHIA 8	Herwig 7	EPOS 4
Inclusive	0.98	0.99	1.02
$1 \leq N_{\rm ch} \leq 2$	0.72	0.76	0.77
$3 \leq N_{\rm ch} \leq 5$	0.77	0.82	0.85
$6 \leq N_{\rm ch} \leq 10$	0.84	0.87	0.91
$11 \leq N_{\rm ch} \leq 17$	0.90	0.92	0.98
$18 \leq N_{\rm ch} \leq 25$	0.95	0.97	1.01
$26 \leq N_{\rm ch} \leq 35$	1.00	1.02	1.05
$36 \leq N_{\rm ch} \leq 45$	1.03	1.06	1.07
$46 \leq N_{\rm ch} \leq 55$	1.07	1.10	1.09
$N_{\rm ch} \geq 56$	1.11	1.15	1.11

A common trend is observed in Fig. 4: the $p_{\rm T}$ spectra is getting harder when multiplicity increases, and the $\langle p_{\rm T}\rangle$ increases correspondingly. Although the inclusive $\langle p_{\rm T}\rangle$ values differ by only a few percent between models (0.98 GeV for PYTHIA 8, 0.99 GeV for Herwig 7, and 1.02 GeV for EPOS 4), the shapes show clear differences, particularly in the high-$p_{\rm T}$ region. This reinforces the idea that comparisons based solely on average quantities, such as $\langle p_{\rm T}\rangle$, may obscure crucial differences in event dynamics and modeling approaches.

These results confirm that a more detailed approach, focusing on distributions rather than global quantities, can significantly deepen our understanding of the soft and hard mechanisms in transverse momentum spectra in hadronic collisions [10], and provide more stringent tests for tuning and validating event generators [16].

4. Conclusion

In this work, we have presented a case for moving beyond mean values and emphasising the importance of studying the full shape of transverse momentum spectra in proton–proton collisions. Through detailed event-by-event simulations with PYTHIA 8, and comparisons with other commonly used MC generators such as Herwig 7 and EPOS 4, we have shown that while the mean transverse momentum evolves only modestly with multiplicity, the shapes themselves undergo substantial changes. The complex interplay between soft and hard processes, as seen in the p_{T} spectra shape variations, cannot be captured by averages alone, and the three model comparison reveals that different MC generators may produce similar mean values while differing significantly in the underlying dynamics.

We wish to acknowledge the Wigner Scientific Computing Laboratory for providing the necessary resources and support for the calculations. This work was supported by the research grants No. NKFIH NEMZ_KI-2022-00058, 2025-1.1.5-NEMZ_KI-2025-00005, 2025-1.1.5-NEMZ_KI2025-00013, NKKP ADVANCED_25 153456, 2024-1.2.9-NETWORKING-2024-00047, 2024-1.2.5-TET-2024-00022 for which we are deeply grateful. Support for this work has been received from the Mexican National Council of Humanities, Sciences and Technologies CONAHCYT under grant No. CF-2042.

REFERENCES

[1] ALICE Collaboration (K. Aamodt *et al.*), *Eur. Phys. J. C* **68**, 89 (2010).

[2] ALICE Collaboration (K. Aamodt *et al.*), *Phys. Lett. B* **693**, 53 (2010).

[3] ALICE Collaboration (S. Acharya *et al.*), *J. High Energy Phys.* **2018**, 013 (2018).

[4] L. Landau, *J. Phys. (USSR)* **8**, 201 (1944).

[5] ALICE Collaboration (B. Abelev *et al.*), *Phys. Lett. B* **727**, 371 (2013).

[6] A.N. Mishra *et al.*, *Eur. Phys. J. A* **57**, 245 (2021).

[7] E. Wang, X.-N. Wang, *Phys. Rev. C* **64**, 034901 (2001).

[8] A.N. Mishra, G.G. Barnaföldi, G. Paić, *J. Phys. G: Nucl. Part. Phys.* **50**, 095004 (2023).

[9] A.N. Mishra, A. Ortiz, G. Paić, *Phys. Rev. C* **99**, 034911 (2019).

[10] G. Bíró, L. Serkin, G. Paić, G.G. Barnaföldi, *Eur. Phys. J. Spec. Top.* **234**, 2931 (2025).

[11] S. Torbjörn *et al.*, *Comput. Phys. Commun.* **191**, 159 (2015).

[12] ATLAS Collaboration, ATL-PHYS-PUB-2014-021, 2014.

[13] ALICE Collaboration (S. Acharya *et al.*), *Eur. Phys. J. C* **79**, 857 (2019).

[14] J. Bellm *et al.*, *Eur. Phys. J. C* **76**, 196 (2016).

[15] K. Werner, *Phys. Rev. C* **108**, 064903 (2023).

[16] ATLAS Collaboration, ATL-PHYS-PUB-2016-001, 2016.