

# PYTHIA and the preoccupied proton

Thomas A. Trainor<sup>1\*</sup>

<sup>1</sup>University of Washington, Seattle, Washington, USA

**Abstract.** The PYTHIA Monte Carlo (PMC) has been applied broadly to simulations of high-energy  $p$ - $p$  and  $p$ - $\bar{p}$  collisions. The PMC is based on several assumptions, such as that most hadrons result from jet production (multiple parton interactions or MPIs), that  $p$ - $p$  centrality is relevant and that color reconnection (CR) strongly influences fragmentation to jets. An alternative description is provided by the two-component (soft + hard) model (TCM) of hadron production. TCM analysis of  $p$ -Pb ensemble-mean- $p_t$  data reveals centrality trends quite different from those estimated via a geometric Glauber model based on the eikonal approximation. Glauber estimates of binary-collision number are three times TCM estimates. Detailed study of  $p$ -Pb data conflicts with a basic Glauber assumption – that a projectile proton may interact simultaneously with multiple target nucleons. Instead, in both  $p$ - $p$  and  $p$ -A collisions, a  $p$ -N collision once initiated is exclusive of other possible interactions (during that collision), and within the collision any pair of participant partons may interact – a  $p$ -N collision is thus “all or nothing.” In this presentation the PMC is challenged by an assortment of contradictory data, and evidence for  $p$ -N exclusivity is reviewed to make a case for the “preoccupied proton” of the title.

## 1 Introduction

This presentation confronts the PYTHIA Monte Carlo (PMC) with an alternative two-component (soft + hard) model (TCM) of hadron production near midrapidity. Modeling of  $p$ - $p$  centrality is a major issue. Centrality modeling for  $p$ -A collisions provides important evidence against the relevance of centrality in elementary  $N$ - $N$  collisions. Results suggest that a new principle of *exclusivity* governs binary  $N$ - $N$  collisions within composite A-B collisions, leading to the “preoccupied proton” of the title.

Within the PMC multiple parton interactions (MPIs) are assumed to be the dominant or exclusive mechanism for hadron production. The PMC thus contains no feature comparable to the TCM soft component. Non-diffractive inelastic scattering is modeled within the PMC by extending the pQCD parton scattering cross section down to  $p_t = 0$  with collision-energy-dependent soft cutoff parameter  $p_{\perp 0}$  as a main tuning parameter.  $p$ - $p$  centrality described by the Glauber model with eikonal approximation is a major feature of the model. Color reconnection (CR) is assumed so as to minimize the total string length (i.e. fragment number) resulting from multiple hard parton scatters (MPIs). The PMC is tuned to accommodate a specific subset of currently available data volumes and analysis methods [1].

---

\*e-mail: ttrainor99@gmail.com

In contrast, the TCM is interpreted such that for hadron production near midrapidity the soft component represents the majority of hadrons, those arising from longitudinal dissociation or fragmentation of projectile nucleons, while the minority hard component represents hadrons from large-angle-scattered partons (gluons) fragmenting to minimum-bias (MB) dijets [2–4]. The observation in  $p$ - $p$  collisions of a quadratic relation between soft and hard components [2] is a major result with its own implications for  $p$ - $p$  centrality as modeled in the PMC. The TCM provides an accurate description of a broad array of data including  $\bar{p}_t$  data for  $p$ -Pb collisions as described in this presentation [5].

## 2 Two-component model of hadron production

The TCM was first applied to hadron production in  $p$ - $p$  ( $p$ - $\bar{p}$ ) collisions near midrapidity [2, 6]. Mean charge density  $\bar{\rho}_0 = n_{ch}/\Delta\eta = \bar{\rho}_s + \bar{\rho}_h$  averaged over some  $\eta$  acceptance  $\Delta\eta$  is decomposed into soft and hard components. Soft-component charge density  $\bar{\rho}_s = n_s/\Delta\eta$  is a product of projectile nucleon dissociation along  $z$  and corresponds to participant low- $x$  gluons, with  $\bar{\rho}_s \propto \log(\sqrt{s}/10 \text{ GeV})$ . Hard-component density  $\bar{\rho}_h$  consists of fragments from MB dijets. Those interpretations are supported by comparing TCM results for a variety of collision systems to measured jet properties [7–9]. An essential feature of the  $p$ - $p$  TCM is the observed quadratic relation  $\bar{\rho}_h \approx \alpha \bar{\rho}_s^2$  with  $\alpha \approx O(0.01)$ . Generalized to composite A-B systems (A-B geometry parameters and  $N$ - $N$  densities are factorized) the spectrum TCM is

$$\begin{aligned} \bar{\rho}_0(y_t) &= (N_{part}/2)\bar{\rho}_{sNN}\hat{S}_0(y_t) + N_{bin}\bar{\rho}_{hNN}\hat{H}_0(y_t) \\ \frac{\bar{\rho}_0(y_t)}{\bar{\rho}_s} &= \hat{S}_0(y_t) + x(n_s)\nu(n_s)\hat{H}_0(y_t), \end{aligned} \quad (1)$$

where  $y_t \equiv \ln[(m_t + p_t)/m_\pi]$ ,  $x(n_s) = \bar{\rho}_{hNN}/\bar{\rho}_{sNN} \approx \alpha \bar{\rho}_{sNN}$  and  $\nu \equiv 2N_{bin}/N_{part}$ . The hatted soft and hard model functions on  $y_t$  are unit normal.

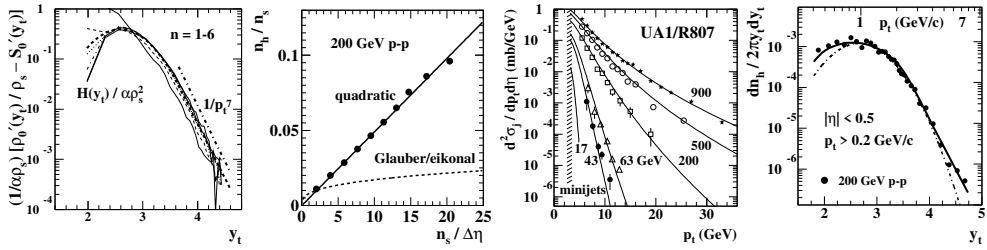
## 3 Minimum-bias dijets, the TCM hard component and MPIs

The PMC is based in part on the assumption that most hadrons from a  $p$ - $p$  collision emerge from multiparton interactions (per event) or MPIs (i.e. jets) described by pQCD. The PMC is then in effect a one-component model or OCM (hard only). However, experimental evidence conflicts with that assumption [2, 3, 7–9].

Figure 1 (first) shows hadron spectrum hard components for six multiplicity  $n_{ch}$  classes of 200 GeV  $p$ - $p$  collisions (thin curves of several line styles). The bold dashed curve is hard-component model  $\hat{H}_0(y_t)$ , a Gaussian with exponential tail. The systematic deviations from the model have been interpreted to result from bias of the underlying jet spectrum by the imposed  $n_{ch}$  condition [7]. The second panel shows the ratio  $n_h/n_s$  of soft and hard yields integrated within acceptance  $\Delta\eta = 1$  vs  $\bar{\rho}_s$ . The data follow a linear trend accurately corresponding to  $\bar{\rho}_h \approx \alpha \bar{\rho}_s^2$  over a ten-fold increase in  $\bar{\rho}_s$ , corresponding therefore to a 100-fold increase in dijet production [3]. The first panel indicates no significant alteration of jet formation over that interval, in conflict with color reconnection as described below.

Figure 1 (third) shows jet  $p_t$  spectra (points) from  $p$ - $p$  ( $p$ - $\bar{p}$ ) collisions at five collision energies. The solid curves are determined by the model function [10]

$$p_t \frac{d^2\sigma_j}{dp_t d\eta} = \frac{d^2\sigma_j}{dy_{max} d\eta} = 0.052 \Delta y_b^2 \frac{1}{2\pi\sigma_u^2} e^{-u^2/2\sigma_u^2}, \quad \sigma_u \approx 1/7 \quad (2)$$



**Figure 1.** First: Spectrum hard components for six multiplicity classes of 200 GeV  $p$ - $p$  collisions. Second: Relation between soft and hard charge yields. Third: Jet spectra for five  $p$ - $p$  collision energies (points) and associated curves from Eq. (2). Fourth: Hard component for NSD  $p$ - $p$  collisions (points) and pQCD prediction from jet measurements (solid curve).

where  $y_{max} = \ln(p_t/m_\pi)$ ,  $\Delta y_{max} = \ln(\sqrt{s}/2E_{cut})$ ,  $u = \ln(p_t/E_{cut})/\Delta y_{max} \in [0, 1]$  and  $\Delta y_b = \ln(\sqrt{s}/10 \text{ GeV}) \propto \bar{\rho}_s$ , with  $E_{cut} \approx 3 \text{ GeV}$  being the effective lower bound of a jet spectrum. Factor  $\Delta y_b^2$  corresponds exactly to the quadratic relation between soft and hard components in Eq. (1). The exponential is a Gaussian on normalized rapidity  $u$ . Jet spectra for all collision energies are described by that simple function.

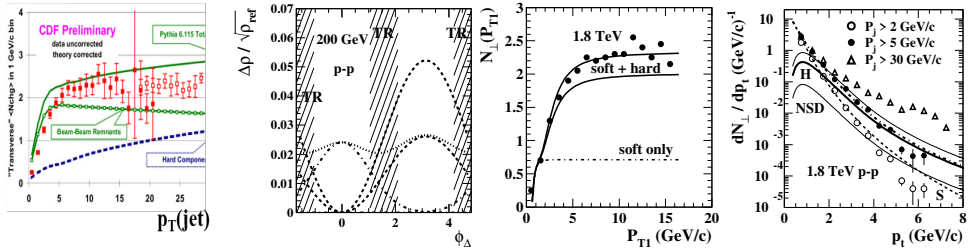
Figure 1 (fourth) shows the spectrum hard component for 200 GeV NSD  $p$ - $p$  collisions (points) as in the first panel. The solid curve is a convolution of the 200 GeV jet spectrum in the third panel and CDF-measured  $p$ - $\bar{p}$  fragmentation functions [8]. A lower bound of 3 GeV on the jet spectrum is required by the hadron spectrum data. Given those results the integral of Eq. (2) over  $p_t$  and  $\eta$  estimates a jet total cross section  $\sigma_{j0} \approx 4 \text{ mb}$  which can be compared with the NSD cross section 36 mb for that collision energy to infer that the frequency per event of MB dijets in  $4\pi$  is  $O(0.1)$  at 200 GeV.

## 4 PYTHIA and the underlying event – UE

In the context of the PMC a triggered dijet is assumed to be confined to two regions on azimuth centered at 0 and  $\pi$  relative to the trigger. Complementary regions centered at  $\pm\pi/2$  are described as the transverse region or TR. The integrated charge within the TR is denoted by  $N_\perp$ , and the trend of  $N_\perp$  with trigger  $p_t$  condition is said to reveal properties of the *underlying event* or UE. In particular, the rise of  $N_\perp$  with  $p_t(\text{jet})$  to a saturation value (pedestal) should reflect increase of the MPI rate with increasing  $p$ - $p$  centrality in response to the trigger condition [1, 11].

Figure 2 (first) shows CDF UE data for  $N_\perp$  vs  $p_t(\text{jet})$  (points) in a conventional plot format [12]. The curves are from PYTHIA. The second panel shows the fragment azimuth distribution for MB dijets from 200 GeV  $p$ - $p$  collisions (dotted curve) derived from model fits to 2D angular correlations [3]. The relevant conclusion from this plot is that any dijet, triggered or not, must make a substantial contribution to the TR.

Figure 2 (third) shows  $N_\perp$  data from the first panel (points) compared to running integrals of the TCM for 1.8 TeV  $p$ - $\bar{p}$   $p_t$  spectra (dash-dotted and lower solid) [13]. Those curves are TCM predictions based on the spectrum TCM reported in Ref. [7], not fits to data. The panel demonstrates that the UE pedestal effect does not result from increasing  $p$ - $p$  centrality, instead reflects running integration of the single-particle  $p_t$  spectrum in response to the jet  $p_t$  trigger condition. It is notable that because of the 0.5 GeV/c  $p_t$  acceptance lower limit for the CDF data only 25% of the spectrum soft component is accepted. With full  $p_t$  acceptance the soft component would dominate the  $N_\perp$  data.

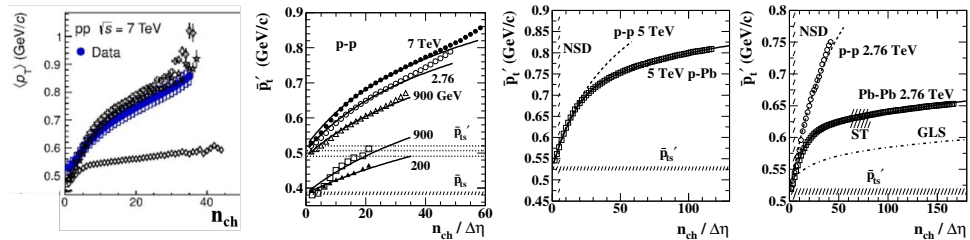


**Figure 2.** First: CDF data associated with the underlying event (points) compared to PYTHIA predictions (curves). Second: Model functions derived from 2D angular correlations. Third: Data from first panel (points) compared to running integrals of single-particle spectra (solid curves). Fourth:  $N_{\perp}$  spectra (points) compared to TCM predictions (curves).

Figure 2 (fourth) shows  $N_{\perp} p_t$  spectra for three trigger conditions (points). The curves represent a TCM prediction for 1.8 TeV derived from the spectrum TCM in Ref. [7], the dashed curve being the fixed soft component. The condition  $P_j > 2$  GeV/c completely suppresses jet production; data are consistent with the soft component alone. Condition  $P_j > 5$  GeV/c corresponds to 50% of events with a dijet (bold solid), and  $P_j > 30$  GeV/c corresponds to 100% (upper thin solid). For any trigger condition the data at low  $p_t$  are consistent with the fixed soft component alone, arguably representing the “underlying event.” There is no variation corresponding to changing  $p$ - $p$  centrality. Data at higher  $p_t$  are biased by the trigger condition as one might expect if the triggered jet makes a strong contribution to the TR. The same trigger bias corresponds to the upper solid curve in the third panel (with 20% increase of the TCM hard component to accommodate the data).

## 5 PYTHIA and color reconnection – CR

The default PMC is unable to reproduce  $p$ - $p$   $\bar{p}_t$  vs  $n_{ch}$  data; to accommodate such data a *color reconnection* or CR mechanism was introduced [1]. Color connections (e.g. strings) said to result from multiple parton interactions (MPIs) are rearranged or reconnected to minimize total string length. A consequence of such rearrangement should be a resulting strong dependence of jet formation on jet number or density.



**Figure 3.** First: ALICE  $\bar{p}_t$  data (solid points) compared to PYTHIA predictions (open points). Second:  $\bar{p}_t$  data (points) compared to TCM trends (curves). Third: ALICE  $\bar{p}_t$  data for 5 TeV  $p$ -Pb collisions (points) compared to a TCM trend (solid curve). Fourth: ALICE data for 2.76 TeV Pb-Pb collisions (open boxes) compared to a TCM trend (solid curve).

Figure 3 (first) shows  $\bar{p}_t$  vs  $n_{ch}$  ( $n_{ch}$  integrated within  $\Delta\eta = 0.6$ ) data (solid points) for 7 TeV  $p$ - $p$  collisions from Ref. [14]. The lower open points from the default PMC can be compared with the curve marked ‘‘Glauber/eikonal’’ in Fig. 1 (second). The upper open points are from the PMC with CR mechanism adjusted to accommodate the data. The second panel includes those 7 TeV  $p$ - $p$  data and  $\bar{p}_t$  data from other collision energies (points). The curves are derived from a comprehensive  $p_t$  spectrum TCM that accurately describes spectrum data from SPS to top LHC energies [4, 7]. Data deviations from the TCM curves result from small shifts of hard components on  $y_t$  with varying  $n_{ch}$  [7]. The accurate TCM representation of  $p$ - $p$   $\bar{p}_t$  data without a CR mechanism arises from the quadratic relation  $\bar{\rho}_h \propto \bar{\rho}_s^2$  also reflected in the jet spectrum description of Eq. 2.

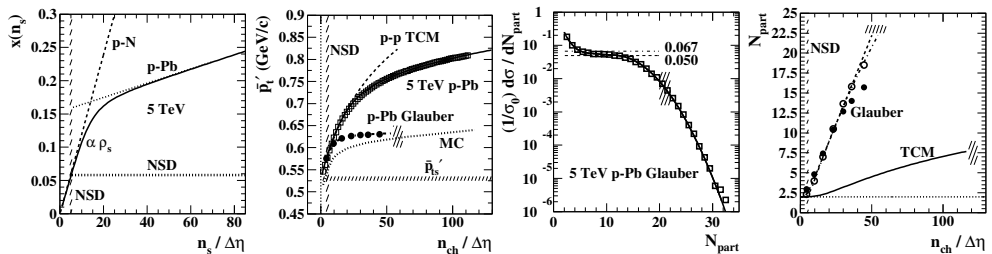
Figure 3 (third) shows  $\bar{p}_t$  data for 5 TeV  $p$ -Pb collisions (points) compared to the TCM reported in Ref. [4] (solid). The dashed curve represents the  $p$ - $p$  TCM as in the second panel but with interpolation to 5 TeV. The data are described within their uncertainties. It is notable that  $p$ - $p$  and  $p$ -Pb data coincide precisely up to  $\bar{\rho}_0 \approx 20$ . The fourth panel shows a comparable TCM description of 2.76 TeV Pb-Pb  $\bar{p}_t$  data.

## 6 $p$ -Pb centrality analysis: TCM vs Glauber model

$p$ -Pb centrality determination can be inferred from the TCM for  $\bar{p}_t$  data as described in Refs. [4, 15]. The TCM for extensive mean *total*  $p_t$  integrated over all  $y_t$  and some  $\Delta\eta$  acceptance is derived from Eqs. 1

$$\begin{aligned} \bar{P}_t &= (N_{part}/2)n_{sNN}\bar{p}_{tsNN} + N_{bin}n_{hNN}\bar{p}_{thNN}(n_s) \\ \frac{\bar{P}_t}{n_s} &= \bar{p}_{ts} + x(n_s)v(n_s)\bar{p}_{th0}, \end{aligned} \quad (3)$$

where universal  $\bar{p}_{ts} \approx 0.40$  GeV/c corresponds to universal slope parameter  $T \approx 145$  MeV within  $\hat{S}_0(y_t)$ . The second line assumes  $\bar{p}_{tsNN}(n_s) \rightarrow \bar{p}_{th0}$  fixed, i.e. no jet modification in  $p$ -Pb collisions. The two  $\bar{p}_{tx}$  values are determined by model functions  $\hat{S}_0(y_t)$  and  $\hat{H}_0(y_t)$ . Hard/soft ratio  $x(n_s) = n_{hNN}/n_{sNN} \approx \alpha\bar{\rho}_{sNN}$  is assumed as in  $p$ - $p$  collisions (Fig. 1, second). If  $x(n_s)$  is defined other parts of the TCM are also determined as a consequence.



**Figure 4.** First: Model for TCM  $x(n_s)$  parameter (solid curve). Second: 5 TeV  $p$ -Pb  $\bar{p}_t$  data (open points) compared to TCM (solid curve) and predictions from a Glauber analysis (solid points). Third: Glauber-model cross-section distribution on  $N_{part}$ . Fourth: Comparison of Glauber (points) and TCM (solid curve) predictions for  $N_{part}$  vs charge density  $\bar{\rho}_0$ .

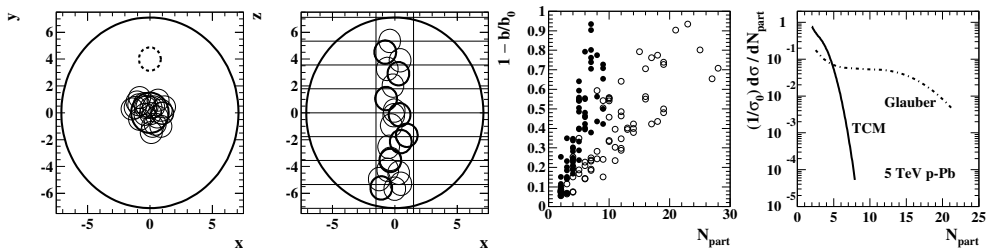
Figure 4 (first) shows  $x(n_s)$  vs  $\bar{\rho}_s$  for  $p$ -Pb collisions (solid curve) based on agreement between  $p$ - $p$  and  $p$ -Pb  $\bar{p}_t$  data below  $\bar{\rho}_0 \approx 20$  (dashed line) and the simplest possible modification above that point, a linear trend with reduced slope (dotted line). The second panel shows the resulting TCM (solid curve) compared to  $\bar{p}_t$  data (open squares). The data are

described within their uncertainties. Transition point  $\bar{\rho}_{s0} \approx 15$  and slope reduction factor  $m_0 \approx 0.1$  are the only new parameters required to extend the  $p$ - $p$  TCM to  $p$ -Pb data [4]. Centrality parameters  $N_{part} = N_{bin} + 1$  and  $v(n_s)$  are then also determined. The solid points are implied by the Glauber analysis of Ref. [16]

Figure 4 (third) shows the differential cross section  $d\sigma/dN_{part}$  (points) determined by a Glauber analysis of  $p$ -Pb centrality [16]. That analysis is based in part on the assumption that  $n_x \propto N_{part}$  for charge multiplicity  $n_x$  within a V0A detector. The measured probability distribution  $dP/dn_x$  on  $n_x$  is assumed to be equivalent to the differential cross section in the form  $(1/\sigma_0)d\sigma/dn_x$  to assign fractional cross sections to event classes on  $n_x$  as in Fig. 1 of Ref. [16]. The fourth panel shows the consequent  $N_{part}$  vs  $\bar{\rho}_0$  relation (points). The TCM equivalent derived from  $\bar{\rho}_i$  data is the solid curve. There is clearly a fundamental disagreement between the two methods, prompting a reexamination of the Glauber model as it is applied to  $p$ - $A$  collisions.

### 7 $p$ - $N$ exclusivity within $p$ -Pb collisions

Reference [17] reports a possible resolution of the conflict between Glauber model and TCM noted above. The critical issue lies with the eikonal approximation, assumed within the Glauber model, as it applies to  $p$ - $N$  collisions within a  $p$ -Pb collision. The eikonal approximation applied to  $p$ - $p$  collisions assumes that each participant parton in a projectile proton flies freely and may interact with any participant parton within a target proton that is inside its *eikonal corridor* defined by a parton-parton cross section. As a consequence of that assumption the number of binary parton-parton collisions (i.e. dijet production) in a simulated  $p$ - $p$  collision goes as the  $4/3$  power of the number of participant partons, equivalent to  $A$ - $A$  collisions with  $N_{bin} \approx (N_{part}/2)^{4/3}$ . The eikonal result is shown as the dashed curve in Fig. 1 (second) with  $n_h/n_s \propto \bar{\rho}_s^{-1/3}$ . That dependence arises because in each event participants are restricted to a  $p$ - $p$  overlap region determined by an impact parameter: i.e.  $p$ - $p$  centrality matters. In contrast,  $p$ - $p$  spectrum data follow a quadratic  $N_{bin} \propto N_{part}^2$  relation with  $n_h/n_s \propto \bar{\rho}_s$  characteristic of full overlap between  $p$ - $p$  collision partners in each event. To resolve the conflict of centrality modeling for  $p$ -Pb collisions an *exclusion-time constraint* may be imposed: a projectile proton already fully engaged in a  $p$ - $N$  collision is excluded from interacting with the next nucleon for a time comparable to a nucleon diameter –  $\delta t \approx 1.5$  fm/c.



**Figure 5.** First: Nucleons within eikonal corridor for  $b = 0$  projectile. Second:  $p$ - $N$  encounters (22 total circles) and actual collisions with exclusion-time constraint (8 dark circles). Third: Glauber-model relation of relative impact parameter  $b/b_0$  to participant number  $N_{part}$  with (solid) and without (open) an exclusion-time constraint. Fourth: Differential cross section on  $N_{part}$  with (solid) and without (dash-dotted) an exclusion-time constraint.

Figure 5 (first) shows a Glauber-model simulation of a  $p$ -Pb collision viewed along the trajectory of a projectile proton with  $b = 0$ . The bold circle is the eikonal corridor defined by

a nucleon-nucleon cross section. Any target nucleon with center inside the bold circle marks a  $p$ - $N$  *geometric encounter*, of which there are 22 in this event (light circles). The second panel shows a transverse view of the collision system. The 22 encounters are again indicated by light circles. Imposing an exclusion-time constraint ( $\delta t > 1.5$  fm/c between collisions) limits the actual  $p$ - $N$  *collisions* to 8 (bold circles).

Figure 5 (third) shows number of participants vs relative impact parameter for an ensemble of simulated  $p$ -Pb collisions with (solid points) and without (open circles) an exclusion-time constraint. The first and second panels represent one element of such an ensemble. The fourth panel shows the resulting cross-section distribution with exclusion-time constraint (solid curve) whereas the dash-dotted curve represents the Glauber data in Fig. 4 (third) with no constraint. Without the constraint a projectile nucleon may collide simultaneously with several overlapping target nucleons, leading to large overestimates of  $N_{part}$  as reported in Ref. [15]. The need for an exclusion-time constraint within  $p$ - $A$  simulations is also consistent with the quadratic  $\bar{\rho}_h \propto \bar{\rho}_s^2$  relation for isolated  $p$ - $p$  collisions. That relation implies centrality is not relevant for such collisions – any  $p$ - $N$  *collision* exhibits 100% overlap. Thus, individual  $p$ - $N$  collisions are “all or nothing” and multiple *simultaneous* collisions are forbidden.

## 8 $p$ -Pb TCM and PID spectra

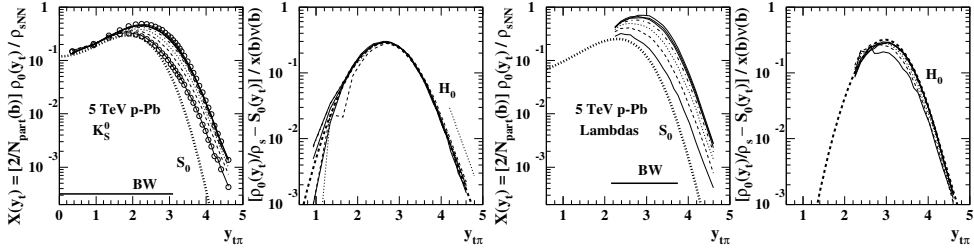
The  $p$ -Pb TCM for collision geometry, spectra and  $\bar{p}_t$  data may be tested by application to spectrum data for identified (PID) hadrons reported in Ref. [18]. The TCM centrality derived from  $\bar{p}_t$  data as described above is retained unchanged, as is the description of spectra and yields for unidentified hadrons. To describe data for identified hadrons new parameters  $z_{si}$  and  $z_{hi}$  represent soft- and hard-component yields of hadron species  $i$  as *fractions* of the yields for unidentified hadrons. The TCM for normalized spectra from identified hadrons then follows from Eq. (1) (second line)

$$\begin{aligned} \frac{\bar{\rho}_{0i}(y_t)}{\bar{\rho}_{si}} &= \hat{S}_{0i}(y_t) + (z_{hi}/z_{si})x(n_s)v(n_s)\hat{H}_{0i}(y_t) \\ \frac{1}{\bar{\rho}_{si}} &= \frac{1 + (z_{hi}/z_{si})x(n_s)v(n_s)}{1 + x(n_s)v(n_s)} \frac{1}{z_{0i} \bar{\rho}_s}, \end{aligned} \quad (4)$$

where the second line defines the normalization for PID spectra in terms of the equivalent  $1/\bar{\rho}_s$  for unidentified hadrons already determined in previous analysis. New parameters  $z_h/z_s$  and  $z_0$  are determined for each hadron species by the spectrum structure below  $y_t = 2$  ( $p_t \approx 0.5$  GeV/c) and are therefore unaffected by the spectrum hard component.

Figure 6 shows full spectra (first and third) and extracted hard components (second and fourth) for  $K_S^0$  and  $\Lambda + \bar{\Lambda}$  from 5 TeV  $p$ -Pb collisions. Similar results for pions, protons and charged kaons are obtained. The  $K_S^0$  data below  $y_t = 2$  ( $p_t < 0.5$  GeV/c) confirm that the soft component, including slope parameter  $T$ , is independent of centrality. For each hadron species parameter  $z_h/z_s$  is adjusted to bring spectra for seven centralities into coincidence for  $y_t < 2$ , and parameter  $z_0$  is then adjusted to bring all spectra into coincidence with the unit-normal  $\hat{S}_0(y_t)$  model function (bold dotted). Model parameters for charged and neutral kaons are assumed to be identical (consistent with spectrum data), and those for protons and Lambdas are similar.

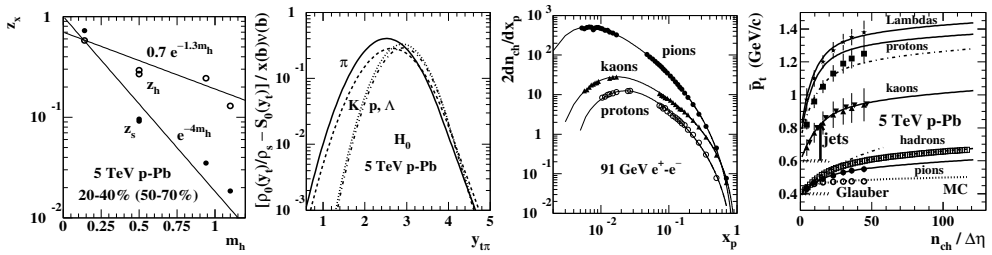
Figure 7 (first) shows  $z_s$  and  $z_h$  parameter values vs hadron mass for a specific  $p$ -Pb centrality. The centrality numbers in parentheses are from Ref. [15] based on the  $p$ -Pb TCM whereas the others are from Ref. [16] based on a Glauber analysis. The lines show exponential dependences on hadron mass. Whereas  $z_s$  decreases strongly with mass as expected



**Figure 6.** Spectra from seven centrality classes of 5 TeV  $p$ -Pb collisions: Left: Full spectra (first) and isolated hard components (second) for  $K_S^0$ . Right: Full spectra (first) and isolated hard components (second) for  $\Lambda + \bar{\Lambda}$ . The bold dotted and dashed curves are TCM models.

from the statistical model  $z_h$  decreases much less quickly so that for baryons there is a large excess of jet fragments. The effect is most notable in Fig. 6 (third) where the Lambda hard component peaks near  $y_t = 3$  ( $p_t \approx 1.5$  GeV/c) and for more-central collisions dominates the soft component (bold dotted). Thus, the origin of the baryon/meson “puzzle” encountered in A-A collisions is due to jet production according to these data.

Figure 7 (second) shows TCM hard-component models  $\hat{H}_0(y_t)$  for pions, kaons, protons and Lambdas [see bold dashed curves in Fig. 6 (second and fourth)]. The curves tend to coincide on the high- $y_t$  side but have a strong mass dependence on the low- $y_t$  side. It is interesting to compare those TCM hard components with corresponding fragmentation functions from LEP  $e^+e^-$  collisions in the third panel (ALEPH data). The same mass trend is observed, buttressing the conclusion that the PID TCM hard components represent fragments from a common underlying jet spectrum.



**Figure 7.** First: TCM soft  $z_s$  and hard  $z_h$  fractions for identified hadrons relative to unidentified hadrons vs hadron mass. Second: Spectrum hard components for five hadron species. Third: Fragmentation functions for three hadron species. Fourth: PID  $\bar{p}_t$  data (points) from 5 TeV  $p$ -Pb collisions for seven multiplicity classes vs TCM (solid curves).

Figure 7 (fourth) shows PID  $\bar{p}_t$  vs  $n_{ch}$  data for seven centrality classes of 5 TeV  $p$ -Pb collisions from Ref. [18] (solid points) and corresponding TCM derived from the PID spectrum TCM described above (solid curves). The open circles are implied by results of the Glauber analysis of Ref. [16], and the dotted curve (MC) represents default PYTHIA.  $\bar{p}_t$  data for unidentified hadrons from Ref. [14] (open squares) are included for comparison. Whereas the most central (0-5%) point for PID data is reported to be at  $\bar{p}_0 \approx 45$  the unidentified-hadron data from the same collaboration extend to 2.6 times that number, further evidence that the  $p$ -Pb Glauber centralities reported in Ref. [16] can be questioned [15]. The TCM description of pion, kaon and Lambda  $\bar{p}_t$  data is quite accurate, reflecting the quality of the



TCM description of spectrum data. However, the proton data deviate substantially from the TCM expectation (proton solid curve). The difference in data is an apparent suppression of the proton hard component near its mode. The dash-dotted curve corresponds to inclusion of a fixed suppression factor to model the proton spectrum data. The remaining systematic deviations correspond to shifts of the proton hard component to higher  $p_t$  with increasing  $n_{ch}$  or  $p$ -Pb centrality. Similar but smaller shifts are observed for Lambda data as in Fig. 6 (fourth panel) but not for kaons (second panel) or pions.

## 9 Summary

In this presentation the PYTHIA Monte Carlo (PMC) is confronted with several results from two-component model (TCM) analysis of particle data that challenge certain of its fundamental assumptions. The TCM accurately describes a broad array of data from  $p$ - $p$ ,  $p$ - $A$  and  $A$ - $A$  collision systems, and from SPS to top LHC collision energies, with only a few parameters. Physical interpretation of TCM soft and hard components is consistent with basic QCD energy trends and jet measurements.

The PMC is based in part on the assumption that almost all hadron production in  $p$ - $p$  collisions arises from multiple parton interactions (MPIs) in each collision event that are described by pQCD. The underlying jet spectrum is extended down to  $p_t = 0$  with a soft cutoff adjusted to accommodate data. In contrast, the TCM describes the same particle data via a soft component representing the majority of hadrons and a minority hard component representing jet production that is quantitatively compatible with measured jet spectra and fragmentation functions. The rate of MB dijet production per NSD  $p$ - $p$  collision is  $O(0.1)$ , not multiple jets per collision. Measured  $p_t$  spectrum hard components are consistent with a jet spectrum cutoff near 3 GeV.

The *underlying event* (UE) is said to be accessed via charge multiplicity  $N_{\perp}$  integrated within a limited azimuth interval (TR) relative to an imposed jet trigger. Increase of  $N_{\perp}$  to a saturation value (pedestal) with increasing trigger  $p_t$  condition reportedly arises from increasing  $p$ - $p$  centrality in response to requirement of more jets at higher  $p_t$ . But the pedestal effect is actually predicted by a simple running integral of the single-particle  $p_t$  spectrum, and the  $N_{\perp}$   $p_t$  spectrum indicates no change in the lower- $p_t$  soft component (the actual UE) with trigger but substantial changes at higher  $p_t$  consistent with the triggered jet making a strong contribution to the TR.

$p$ - $p$  centrality is treated within the PMC by a Glauber model based on the eikonal approximation. As in  $A$ - $A$  collisions dijet production then varies as the  $4/3$  power of the participants (i.e. low- $x$  gluons), as reflected in default-PMC predictions for  $\bar{p}_t$  vs  $n_{ch}$ . However, the TCM description of  $p$ - $p$  spectra demonstrates a quadratic relation between MB jet production and participant gluons implying that centrality is irrelevant for  $p$ - $p$  collisions, that each  $p$ - $p$  collision involves full overlap of the collision partners and no restriction to an eikonal corridor for participant partons. The Glauber model applied to  $p$ -Pb collisions also fails and for closely-related reasons: a  $p$ - $N$  collision is “all-or-nothing” – simultaneous collisions via partial overlaps are forbidden.

In response to failure of the default PMC to describe  $\bar{p}_t$  vs  $n_{ch}$  trends a color-reconnection or CR mechanism was introduced that minimizes string lengths (i.e. jet fragments) within MPI production. But that mechanism must be equivalent to strong changes in parton fragmentation with increasing jet number which are not observed in data. For  $n_{ch}$  variation equivalent to 100-fold increase in dijet production spectrum hard components, equivalent to a convolution of jet spectrum and fragmentation-function ensemble, show little variation.

Successful TCM descriptions of  $\bar{p}_t$  vs  $n_{ch}$  data for  $p$ - $p$  and  $p$ -Pb collisions argue against recent claims for “collectivity” or flows in small collision systems. Strong increases in  $\bar{p}_t$  are

fully accounted for by the systematics of the spectrum hard component that is quantitatively related to MB dijets. Given a flow hypothesis it is difficult to account for the strong *decrease* in  $\bar{p}_t$  with increasing system size from  $p$ - $p$  to  $p$ - $A$  to  $A$ - $A$ , whereas in a TCM context dijet production per  $N$ - $N$  pair depends quadratically on  $N$ - $N$   $n_{ch}$ , which is largest in high-multiplicity  $p$ - $p$  collisions and much smaller in central  $p$ -Pb collisions. Given the accurate and complete description of  $p_t$  spectra in terms of a *fixed* soft component and a hard component quantitatively linked to jet production it is unlikely that radial flow (for example) plays a significant role in  $p$ - $p$  or  $p$ - $A$  collisions.

The ability of the TCM to accommodate a variety of A-B data is demonstrated by its extension to describe identified-hadron (PID) spectra from  $p$ -Pb collisions. The TCM for unidentified hadrons from  $p$ -Pb collisions, itself a very simple extension from  $p$ - $p$  collisions, is maintained unchanged, and only two additional parameters (soft and hard fractions of unidentified-hadron yields) are required to describe PID data accurately. The same centrality estimates, based on  $N$ - $N$  exclusivity, are also retained. The TCM description of PID data demonstrates that strong  $\bar{p}_t$  increase with hadron mass, interpreted by some to be a manifestation of radial flow, is a consequence of MB dijet production: jet formation favors baryon production by a large factor, a result directly related to the so-called baryon/meson puzzle emerging from  $A$ - $A$  collision data.

In conclusion, accurate TCM descriptions of a variety of A-B collision data emphasize the central role of MB dijets for hadron production near midrapidity, the absence of centrality dependence for  $N$ - $N$  collisions and the importance of  $p$ - $N$  “exclusivity” within  $p$ - $A$  collisions. TCM results present strong challenges to several assumptions underlying the PYTHIA Monte Carlo, especially the prevalence of MPIs, relevance of the CR mechanism and conventional interpretations of underlying-event trends.

## References

- [1] T. Sjöstrand, arXiv:1706.02166.
- [2] J. Adams *et al.* (STAR Collaboration), Phys. Rev. D **74**, 032006 (2006).
- [3] T. A. Trainor and D. J. Prindle, Phys. Rev. D **93**, 014031 (2016).
- [4] T. A. Trainor, arXiv:1708.09412.
- [5] T. A. Trainor, arXiv:1805.09681.
- [6] G. Pancheri and Y. Srivastava, Conf. Proc. C **850313**, 28 (1985) [Phys. Lett. B **159**, 69 (1985)].
- [7] T. A. Trainor, J. Phys. G **44**, no. 7, 075008 (2017).
- [8] T. A. Trainor, Phys. Rev. C **80**, 044901 (2009).
- [9] T. A. Trainor, Int. J. Mod. Phys. E **17**, 1499 (2008).
- [10] T. A. Trainor, Phys. Rev. D **89**, 094011 (2014).
- [11] L. Frankfurt, M. Strikman and C. Weiss, Phys. Rev. D **83**, 054012 (2011).
- [12] T. Affolder *et al.* (CDF Collaboration), Phys. Rev. D **65**, 092002 (2002).
- [13] T. A. Trainor, Phys. Rev. D **87**, 054005 (2013).
- [14] B. B. Abelev *et al.* (ALICE Collaboration), Phys. Lett. B **727**, 371 (2013).
- [15] T. A. Trainor, arXiv:1801.05862.
- [16] J. Adam *et al.* (ALICE Collaboration), Phys. Rev. C **91**, no. 6, 064905 (2015).
- [17] T. A. Trainor, arXiv:1801.06579.
- [18] B. B. Abelev *et al.* (ALICE Collaboration), Phys. Lett. B **728**, 25 (2014).