A two-component (soft + hard) model (TCM) of hadron production near mid-rapidity was derived from the charge-multiplicity $n_{ch}$ dependence of $p_t$ spectra from 200 GeV $p$-$p$ collisions [1]. The TCM, interpreted to represent longitudinal projectile-nucleon dissociation and minimum-bias (MB) transverse dijet production near midrapidity, has been extended recently to consider the $n_{ch}$ dependence of $p_t$-integral angular correlations [2].

Analysis of new high-statistics $p_t$ spectra in the latter study gave results quantitatively consistent with the earlier spectrum analysis. The $n_{ch}$ dependence of 2D angular correlations revealed a significant nonjet quadrupole ($v_2$) component as a novel aspect of $p$-$p$ collisions. The $n_{ch}$ dependence of yields, spectra and correlations has played a key role in establishing (a) the nature of hadron production mechanisms in $p$-$p$ collisions and (b) that the dijet contribution to $p_t$ spectra is quantitatively consistent with predictions based on reconstructed jets [3-5].

The phenomenology of $p$-$p$ collisions as data serves as an essential reference for high-energy $p$-$A$ and $A$-$A$ collisions, specifically regarding claims of novel physical mechanisms such as formation of a quark-gluon plasma [6] or possible manifestations of hydrodynamic flows (“collectivity”) even in small collision systems [7-8]. It is important therefore to extend the TCM for $p$-$p$ collisions formulated with data from the relativistic heavy ion collider (RHIC) to higher energies with analysis of $p$-$p$ $p_t$ spectra from the large hadron collider (LHC). The question of recently-claimed collectivity or flows in small ($p$-$p$ and $p$-$A$) systems may be addressed in terms of the resolved TCM soft and hard components and evidence (or not) for radial flow in differential study of $p_t$ spectra [9].

In Ref. [10] analysis of the $n_{ch}$ dependence of 13 TeV $p_t$ spectra was presented with indirect reference (via a spectrum ratio) to a 7 TeV spectrum. The relation to previou

**I. INTRODUCTION**

A two-component (soft + hard) model (TCM) of hadron production in yields and spectra derived from the charge-multiplicity $n_{ch}$ dependence of $p_t$ spectra from 200 GeV $p$-$p$ collisions [1]. The TCM, interpreted to represent longitudinal projectile-nucleon dissociation and minimum-bias (MB) transverse dijet production near midrapidity, has been extended recently to consider the $n_{ch}$ dependence of $p_t$-integral angular correlations [2]. Analysis of new high-statistics $p_t$ spectra in the latter study gave results quantitatively consistent with the earlier spectrum analysis. The $n_{ch}$ dependence of 2D angular correlations revealed a significant nonjet quadrupole ($v_2$) component as a novel aspect of $p$-$p$ collisions. The $n_{ch}$ dependence of yields, spectra and correlations has played a key role in establishing (a) the nature of hadron production mechanisms in $p$-$p$ collisions and (b) that the dijet contribution to $p_t$ spectra is quantitatively consistent with predictions based on reconstructed jets [3-5].

The phenomenology of $p$-$p$ collision data serves as an essential reference for high-energy $p$-$A$ and $A$-$A$ collisions, specifically regarding claims of novel physical mechanisms such as formation of a quark-gluon plasma [6] or possible manifestations of hydrodynamic flows (“collectivity”) even in small collision systems [7-8]. It is important therefore to extend the TCM for $p$-$p$ collisions formulated with data from the relativistic heavy ion collider (RHIC) to higher energies with analysis of $p$-$p$ $p_t$ spectra from the large hadron collider (LHC). The question of recently-claimed collectivity or flows in small ($p$-$p$ and $p$-$A$) systems may be addressed in terms of the resolved TCM soft and hard components and evidence (or not) for radial flow in differential study of $p_t$ spectra [9].

In Ref. [10] analysis of the $n_{ch}$ dependence of 13 TeV $p_t$ spectra was presented with indirect reference (via a spectrum ratio) to a 7 TeV spectrum. The relation to previous TCM results is not clear due to the spectrum-ratio strategy adopted. Comparisons with Monte Carlo (MC) models seem inconclusive. Those results raise significant questions about spectrum analysis: what information is conveyed by $p_t$ spectra, what is the best method to extract all significant information, how should spectrum information be used to test theoretical models?

An initial study of the $n_{ch}$ dependence of $p_t$ spectra from 200 GeV $p$-$p$ collisions established that based on their $n_{ch}$ dependence $p_t$ spectra can be decomposed into two distinct components subsequently identified as “soft” (associated with projectile-nucleon dissociation) and “hard” (associated with minimum-bias large-angle parton scattering and jet formation) [1]. The decomposition does not rely on imposed model functions. The shapes of the isolated data components were found to be approximately independent of multiplicity and could be modeled by simple functions if the soft component was expressed on transverse mass $m_t$ and the hard component on transverse rapidity $y_t$. The 200 GeV TCM decomposition has been confirmed with higher-statistics data [2]. The present study shows that the TCM is required by data from a broad array of $p$-$p$ collision systems.

The present study begins with a review of TCM methods as applied to $p_t$ spectra in Refs. [1,2]. Spectrum data from the RHIC and LHC as used in this analysis are introduced, and the analysis strategy to be employed is described. In App. A spectrum-ratio methods from Ref. [10] are applied to 200 GeV $p$-$p$ spectra from Ref. [2] to illustrate the consequences for a data system with established properties. Results reveal that the TCM hard component has a substantial $n_{ch}$ dependence not previously accommodated. The hard-component $n_{ch}$ dependence is determined first for 200 GeV data and then 13 TeV data. Combining those results with spectra at 17.2 GeV and 0.9 TeV an accurate description of the energy dependence of $p$-$p$ $p_t$ spectra extending over all presently-
accessible p-p collision energies is established. The resulting TCM arguably represents all information carried by the \( p_t \) spectra of unidentified hadrons. The hard component is quantitatively related to the properties of isolated jets, and Monte-Carlo-simulated spectrum ratios are interpreted physically in relation to data systematics.

This article is arranged as follows: Section II introduces a TCM description of spectra and yields from p-p collisions. Section [11] summarizes \( p_t \) spectra from RHIC p-p collisions based on the TCM in Refs. [1] [2] and from LHC p-p collisions featuring spectrum ratios in Ref. [10]. Section [IV] presents a revised hard-component model that describes the full \( n_{ch} \) dependence of 200 GeV spectrum data. Section [V] applies the TCM to LHC spectrum-ratio data to determine the \( n_{ch} \) dependence of those spectra. Section [VI] describes the collision-energy evolution of p-p \( p_t \) spectra via the TCM. Section [VII] discusses systematic uncertainties. Sections [VIII] and [IX] present discussion and summary. Appendix [A] introduces TCM analysis of 200 GeV p-p spectra applying spectrum-ratio methods as in Ref. [10]. Appendix [B] compares a TCM-predicted 0.9 TeV \( p_t \) spectrum with data as a quantitative test of TCM energy dependence. Appendix [C] reviews 13 TeV probability distributions on event multiplicity and multiplicity collision-energy dependence relevant to Ref. [10].

II. A TCM FOR p-p COLLISIONS

Final-state hadrons from high-energy nuclear collisions are distributed within a momentum space approximated near midrapidity (\( \eta = 0 \)) by the cylindrical space \( (p_t, \eta, \phi) \), where \( p_t \) is transverse momentum, \( \eta \) is pseudorapidity and \( \phi \) is azimuth angle. Transverse mass \( m_t = \sqrt{p_t^2 + m_h^2} \) with hadron mass \( m_h \). Pseudorapidity is \( \eta = - \ln[\tan(\theta/2)] \) (\( \theta \) is polar angle relative to collision axis \( z \)), and \( \eta \approx \cos(\theta) \) near \( \eta = 0 \). To improve visual access to low-\( p_t \) structure and simplify description of the \( p_t \) spectrum hard component (defined below) spectra may be presented on transverse rapidity \( y_t = \ln[(m + p_t)/m_h] \) with \( p_t = m_h \sinh(y_t) \) and \( m_t = m_h \cosh(y_t) \). For unidentified hadrons \( y_t \) with pion mass assumed (80% of hadrons) serves as a regularized logarithmic \( p_t \) measure \( y_t = \ln(2p_t/m_h) \). A typical acceptance limit \( p_t > 0.15 \) GeV/c corresponds to \( y_t > 1 \).

The p-p spectrum soft component is most efficiently described on transverse mass \( m_t \) whereas the spectrum hard component is most efficiently described on transverse rapidity \( y_t \). The spectrum TCM thus requires a heterogeneous set of variables for its simplest definition. The components can be easily transformed from one variable to the other by a Jacobian factor defined below.

A. TCM context

The two-component model of hadron production in high-energy nuclear collisions has been reviewed in Refs. [1] [11] for p-p collisions and Refs. [3] [9] [12] for A-A collisions. The TCM represents both a mathematical data model and a system of inferred data components isolated via their systematic properties. The TCM model functions may serve as a reference for interpretation of data properties based on comparisons with theory (e.g. Ref. [3]). Differences between data and TCM reference functions may reveal novel physical mechanisms.

The TCM applied to elementary collisions has been interpreted to represent two principal sources of final-state hadrons near midrapidity: longitudinal projectile-nucleon dissociation (soft) and large-angle-scattered (transverse) parton fragmentation (hard). In A-A collisions the two processes scale respectively proportional to \( N_{part} \) (participant nucleons \( N \) and \( N_{bin} \) (\( N \)-\( N \) binary encounters) as determined by a Glauker Monte Carlo model. Analogous scalings for p-p collisions in terms of charge multiplicities, as described in Sec. [TC], have been described in Refs. [2] [11]. The TCM accurately represents hadron yield and spectrum systematics [11] [9] and related aspects of angular correlations [2] [12] [14].

B. TCM for p-p single-particle \( p_t \) or \( y_t \) spectra

The joint single-charged-particle (SP) 2D (azimuth integral) density on \( y_t \) and \( \eta \) is denoted by \( \rho_0(y_t, \eta) = d^2n_{ch}/y_t dy_t d\eta \). The \( \eta \)-averaged (over \( \Delta \eta \)) \( y_t \) spectrum is \( \bar{\rho}_0(y_t; \Delta \eta) \). The \( y_t \)-integral mean angular density is \( \bar{\rho}_0(\Delta \eta) = \int dy_t \bar{\rho}_0(y_t; \Delta \eta) = \Delta n_{ch}/\Delta \eta \) averaged over acceptance \( \Delta \eta \). According to the \( p_t \) spectrum TCM hadron density \( \bar{\rho}_0 \) has soft \( \bar{\rho}_s \) and hard \( \bar{\rho}_h \) components related by \( \bar{\rho}_h = \alpha \bar{\rho}_s^2 \) for \( \alpha = O(0.01) \) and \( \bar{\rho}_0 = \bar{\rho}_s + \bar{\rho}_h \). Given some hypothesis \( \alpha \) determined by spectrum analysis (see below) \( \bar{\rho}_s \) and \( \bar{\rho}_h \) can be obtained for any \( \bar{\rho}_0 \) as the solutions to a quadratic equation. For 200 GeV the value determined by spectrum analysis is \( \alpha \approx 0.006 \) [12].

The \( n_{ch} \) dependence of p-p SP \( p_t \) spectra over a large \( n_{ch} \) interval (10-fold increase) was established in Ref. [1]. Systematic variation of spectrum shapes leads to identification of two approximately fixed forms whose relative amplitudes vary smoothly with \( n_{ch} \) (approximately linear and quadratic respectively for soft and hard). The TCM was not assumed initially, emerged instead from an inductive study. The relation of the hard component to isolated-jet properties was established in Ref. [3]. The soft component in Au-Au collisions retains its fixed form but the hard-component form changes substantially with centrality, reflecting quantitative jet modification [3].

A TCM for p-p \( y_t \) spectra conditional on uncorrected \( n_{ch} \) integrated over \( 2\pi \) azimuth and averaged over some \( \eta \) acceptance \( \Delta \eta \) was defined in Ref. [1] by

\[
\bar{\rho}_0(y_t; n_{ch}) = S(y_t; n_{ch}) + H(y_t; n_{ch})
\]

\[
\approx \bar{\rho}_s(n_{ch}) \bar{S}_0(y_t) + \bar{\rho}_h(n_{ch}) \bar{H}_0(y_t),
\] (1)

where \( \bar{\rho}_s = n_s/\Delta \eta \) and \( \bar{\rho}_h = n_h/\Delta \eta \) are corresponding \( \eta \)-averaged soft and hard hadron densities. The soft and
hard \( y_t \) spectrum shapes [unit normal \( \hat{S}_0(y_t) \) and \( \hat{H}_0(y_t) \)] inferred from data were assumed to be independent of \( n'_h \), with parametrized forms defined in Refs. [1] [3] [9]. Conversion from densities on \( p_t \) or \( m_t \) to densities on \( y_t \) is easily accomplished via the Jacobian factor \( p_t m_t / y_t \).

The fixed unit-normal soft component is most efficiently approximated by a Lévy distribution on \( m_t \)

\[
\hat{S}_0(m_t) = \frac{A(T,n)}{[1 + (m_t - m_t^\prime)/(nT)]^n},
\]

with hadron mass \( m_h \), slope parameter \( T \) and Lévy exponent \( n \), that goes to a Maxwell-Boltzmann exponential on \( m_t \) in the limit \( 1/n \to 0 \). \( \hat{S}_0(m_t) \) describes spectra normalized as ratios \( \hat{\rho}_0(y_t; n'_h)/\hat{\rho}_s \) in the limit \( n'_h \to 0 \). The Lévy distribution may represent a system near equilibrium represented by parameter \( T \), with additional parameter \( 1/n \) measuring deviations from full equilibrium [15].

The fixed unit-normal hard-component model \( \hat{H}_0(y_t) \) is most efficiently approximated by a Gaussian plus exponential tail on \( y_t \) determined by Gaussian centroid \( \hat{y}_t \), Gaussian width \( \sigma_{y_t} \) and “power-law” parameter \( q \). The slope is required to be continuous at the transition point from Gaussian to exponential. An algorithm for computing \( \hat{H}_0(y_t) \) is provided in Ref. [9] (App. A). Whereas \( \hat{H}_0(y_t) \) on \( y_t \) has an exponential form \( \propto e^{-q y_t} \) at larger \( y_t \) the corresponding \( \hat{H}_0(p_t) \) approximates the power-law form \( \propto 1/p_t^{q+2} \) at larger \( p_t \) (hence “power-law” tail). Hard component \( H(y_t) = \alpha \hat{\rho}_s^2 \hat{H}_0(y_t) \) is well-approximated by measured \( p_t \)-fragmentation functions convoluted with a measured 200 GeV minimum-bias (MB) jet spectrum [3] [5]. Note that the two TCM “power-law” exponents \( n \) and \( q \) represent distinct soft and hard hadron production mechanisms near \( y_t = 0 \).

The fixed hard-component model developed in previous studies as described above is revised twice in the present study in response to higher-statistics 200 GeV spectrum data, first to accommodate data above the hard-component mode in Sec. IV A and then to accommodate data below the mode in Sec. IV B. Hard-component parameters \( \sigma_{y_t} \) and \( q \) are allowed to vary smoothly with \( n'_h \) as required by the spectrum data. Centroid \( \hat{y}_t \) does not vary significantly according to data.

C. \( p-p \) hard components and dijet production

Based on a dijet interpretation for the spectrum hard component [3] [4] [9] the corresponding yield should be \( \hat{\rho}_h = \epsilon(\Delta \eta) f(n'_h) 2n'_{ch,j} \), where \( f(n'_h) \) is the dijet frequency per collision and per unit \( \eta, \epsilon(\Delta \eta) \in [0.5,1] \) is the average fraction of a dijet appearing in acceptance \( \Delta \eta \) and \( 2n'_{ch,j} \) is the mean dijet fragment multiplicity. For 200 GeV non-single-diffractive (NSD) \( p-p \) collisions with \( \hat{\rho}_s \approx 2.5 \) and mean fragment multiplicity \( 2n'_{ch,j} \approx 2.5 \pm 0.5 \) inferred from measured jet systematics [1] frequency \( f_{NSD} = 0.006 \times 2.5^2/(0.55 \times 2.5) \approx 0.027 \) is inferred from \( p-p \) spectra integrated within \( \Delta \eta = 2 \).

That value can be compared with results from isolated-jet measurements in the form \( f_{NSD} = (1/\sigma_{NSD}) d\sigma_{dijet}/d\eta \approx (1/36.5 \text{ mb}) \times 1 \text{ mb} \approx 0.028 \) for 200 GeV \( p-p \) collisions [5] based on a measured jet spectrum [16] and NSD \( p-p \) cross section [17]. Measured NSD hard-component density \( \hat{\rho}_h \) is thus quantitatively consistent with dijet systematics derived from eventwise-reconstructed jets [5] [16] [15]. If a non-NSD \( p-p \) event sample with arbitrary mean \( n'_{ch} \) is selected the dijet frequency should vary with soft hadron density \( \hat{\rho}_s \) as

\[
f(n'_h) \approx 0.027 \left[ \frac{\hat{\rho}_s(n'_h)}{\hat{\rho}_{s,NSD}} \right]^2
\]

with \( \hat{\rho}_{s,NSD} = 2.5 \) for 200 GeV \( p-p \) collisions according to spectrum results from Ref. [1]. The same quadratic production trend is observed for dijet manifestations in 2D angular correlations [2], further supporting the relation.

III. \( p-p \) \( p_t \) SPECTRUM DATA

\( p_t \) spectra for the present study are obtained from SPS data at 17.2 GeV, RHIC data at 200 GeV and LHC data at several energies. The RHIC data are in the form of isolated spectra over an extended range of collision multiplicities whereas the LHC data are in the form of a few isolated spectra and spectrum ratios over a limited \( n'_h \) range. The goal of the study is to obtain an accurate and self-consistent TCM parametrization for a broad range of event multiplicities and collision energies.

A. RHIC \( p_t \) spectrum analysis

The multiplicity dependence of \( p_t \) spectrum structure from 200 GeV \( p-p \) collisions was reported in Ref. [1], and the trend for angular correlation structure was reported more recently in Ref. [2]. The latter also included an updated TCM spectrum analysis of high-statistics data.

| Class | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-------|---|---|---|---|---|---|---|
| \( n_h' \) | 2.3 | 4.6 | 7.9 | 10.2 | 13.7 | 18.4-24 | 25-50 |
| \( \langle n'_h \rangle \) | 2.52 | 4.87 | 7.81 | 10.8 | 14.3 | 19.6 | 26.8 |
| \( \hat{\rho}_h(n'_h) \) | 1.90 | 3.65 | 5.82 | 8.00 | 10.6 | 14.3 | 19.3 |
| \( \hat{\rho}_s(n'_h) \) | 1.88 | 3.57 | 5.63 | 7.65 | 9.96 | 13.3 | 17.5 |
| Events (M) | 2.31 | 2.21 | 0.91 | 0.33 | 0.14 | 0.02 | 0.001 |

Table I defines seven multiplicity classes for the study in Ref. [2] that apply to the spectrum data considered below. \( n'_h \) is an uncorrected multiplicity within \( \Delta \eta = 2 \).
related to corrected multiplicity \( n_{ch} \) by \( n'_{ch} = \xi n_{ch} \). The \( \langle n'_{ch} \rangle \) are distribution-weighted bin mean values. \( \bar{\rho}_s \) and \( \bar{\rho}_0 \) are corrected for efficiencies and \( p_t \) acceptance. \( n'_{ch} \) in this text replaces symbol \( \hat{n}_{ch} \) from Ref. 1.

Figure 1 (left) shows \( y_t \) spectra for six multiplicity classes. The spectra (uncorrected for tracking inefficiencies) are normalized by corrected soft component \( \bar{\rho}_s(n_{ch}') \). A common \( y_t \)-dependent inefficiency function is introduced for comparison of this analysis with corrected spectra in Ref. 1, indicated below \( y_t = 2 \) by the ratio of two bold dotted curves representing uncorrected \( S_0(y_t) \) and corrected (unit-normal) \( \hat{S}_0(y_t) \) model functions. Data spectra are represented by spline curves rather than individual points to emphasize systematic variation with \( n'_{ch} \).

The bold dashed curve labeled \( \alpha \bar{\rho}_s \hat{S}_0(y_t) \) (for \( \bar{\rho}_s = 2.5 \)) estimates a fixed hard component for 200 GeV \( p-p \) collisions in relation to corresponding soft component \( \hat{S}_0(y_t) \) (bold dotted curve). The NSD curves cross near \( y_t = 3.75 \) \( (p_t \approx 3 \text{ GeV/c}) \) where \( S(y_t) = H(y_t) \) in Eq. (1).

Figure 1 (right) shows normalized spectra from the left panel for six multiplicity classes in the form of spectrum ratios. It is acknowledged that both hard and soft QCD processes may play a role in hadron production (the two elements of the TCM) but they are not isolated in that study. The study is based on less than 1.5 million (M) events (vs 3M accepted events for Ref. 1 and 6M events for Ref. 2). Reference 1 is cited but its detailed spectrum analysis is not acknowledged.

The evolution of \( p-p \) \( p_t \) spectra with \( n_{ch} \) and collision energy is studied via spectrum ratios that discard some of the information in the individual spectra as demonstrated below. \( p-p \) data are said to provide a reference for \( A-A \) data and particularly for spectrum ratio \( R_{AA} \) intended to study jet modification in \( A-A \) collisions. But as a ratio \( R_{AA} \) also discards essential information: in particular it conceals most of the jet contribution (whatever appears below 4 GeV/c) that is essential to understand QCD processes in high-energy nuclear collisions 13–19.

Measurements include extension of charge \( n \) density to 13 TeV (Fig. 2), a \( p_t \) spectrum extending to 20 GeV/c (Fig 3), a spectrum ratio comparing 13 and 7 TeV spectra (Fig 4) and spectrum ratios comparing 13 TeV \( p_t \) spectra from three multiplicity classes to a common reference (Fig. 5). The spectrum-ratio data are said to show “...rich features when correlated with the charged-particle multiplicity...” It is concluded that spectrum ratios in Fig. 5 demonstrate stronger correlation of spectra with \( n_{ch} \) at higher \( p_t \), but the structure of individual spectra varies most rapidly with \( n_{ch} \) at the spectrum hard-component mode near 1 GeV/c as shown in Ref. 1.

It is acknowledged that jets may play a role in \( p_t \) spectra, but no previous analysis addressing that subject is considered (e.g. Refs. 1 3 9). Qualitative comments are offered about Monte Carlo comparisons with data.

In App. A a revised TCM defined in terms of spectrum ratios is applied to these 200 GeV data in preparation for analysis of LHC spectrum ratios. A fixed hard-component model is assumed in App. A but variations to accommodate spectrum data are explored in Sec. IV.

B. LHC \( p_t \) spectrum analysis

Reference 10 (ALICE collaboration) reports \( p_t \) spectra from 13 TeV \( p-p \) collisions for several event-multiplicity classes in the form of spectrum ratios. It is acknowledged that both hard and soft QCD processes may play a role in hadron production (the two elements of the TCM) but they are not isolated in that study. The study is based on less than 1.5 million (M) events (vs 3M accepted events for Ref. 1 and 6M events for Ref. 2). Reference 1 is cited but its detailed spectrum analysis is not acknowledged.

Figure 2 (left) shows ratio data from Ref. 10 (points) for three \( n_{ch} \) conditions (multiplicity bins A, B and C in App. C) relative to an INEL > 0 (inelastic collisions with at least one detected particle in \( \Delta n \)) reference and ensemble-mean accepted \( \hat{n}_{ch} \rightarrow \hat{\rho}_0 = \hat{\rho}_0 + \rho_0 \). The statistical uncertainties are typically smaller than the data points. Tracking efficiencies are assumed to cancel in ratios, and the \( p_t \) lower limit affects \( \hat{\rho}_s \) but not \( \hat{\rho}_h \) because the latter is localized on \( p_t \). Limiting values at small \( p_t \) are \( R \approx 1.12 \) for bin A, 1.02 for bin B and 0.91 for bin C. The dash-dotted curves approximate MC results (e.g. PYTHIA 20) in Fig. 5 of Ref. 10 as discussed in Sec. V.

Figure 2 (right) shows a spectrum ratio comparing
\[ R(p;E) = \frac{\rho(p|n)}{\rho(p|n_0)} \]

The soft-component model \( S_0(p_t) \) for the higher energy is then determined from a single 13 TeV spectrum fit, and the multiplicity-dependent TCM hard components \( H_0(p_t;n'_{ch}) \) are finally isolated from inferred data ratios \( T(p_t;n'_{ch}) \).

**Energy dependence:** With \( n_{ch} \)-dependent TCMs defined at 200 GeV and 13 TeV it is possible to establish an energy-dependent TCM covering the full available range of \( p-p \) collision energies relevant to dijet production near midrapidity. In Sec. [VIIA] an energy parametrization of soft-component Lévy exponent \( n \) is defined based on the 200 GeV and 13 TeV results and a simple \( \log(s/s_0) \) QCD trend for \( 1/q \). The energy dependence of \( y_t \) and \( \sigma_{yt} \) are supplemented by analysis of a 13 vs 7 TeV spectrum ratio. A key element is the energy dependence of TCM parameter \( \alpha \) that relates soft and hard yields. Its energy dependence is predicted in Sec. [VIIA] based on measured jet-related QCD quantities and a revised trend for jet fragment multiplicities \( 2n_{ch,J} \).

A major point of this exercise is a demonstration that the \( n_{ch} \) and energy dependence of available spectrum data require a specific model inferred via an inductive study of data properties. The resulting TCM is not an arbitrary model based on hypotheses, is intimately related to measured jet properties and QCD expectations.

**IV. 200 GeV \( p_t \) SPECTRUM TCM vs \( n_{ch} \)**

A main achievement of Ref. [1] was isolation of two spectrum components based on dramatically different scalings with \( n_{ch} \) variation. The “hard” component was later interpreted as jet-related after comparison with measured jet properties (FFs and jet spectrum). It was noted that significant deviations from the TCM hard component occurred for the two lowest \( n_{ch} \) classes.

In App. [A] TCM analysis of 200 GeV spectrum ratios reveals a substantial systematic \( n_{ch} \) dependence of the hard component at higher \( y_t \). In this section variations of hard-component parameters \( \sigma_{yt} \) and \( q \) are found to accommodate spectrum data above the hard-component mode. But significant deviations from the fixed model below the mode are closely related: the entire hard-component shape is biased by a changing \( n_{ch} \) condition. It is desirable to develop a complete TCM description accommodating all aspects of \( n_{ch} \) evolution. The term “data” below refers to normalized spectra \( \bar{\rho}(y_t;n'_{ch})/\bar{\rho}_n \).

**A. Hard-component \( n_{ch} \) evolution above the mode**

Figure. [18] (right) of App. [A] shows ratios of spectrum data from Fig. [1] (left) to the corresponding TCM expression in Eq. [1] with fixed hard component. Above \( y_t = 3 \) significant systematic variation (10% increase per multiplicity class at 4 GeV/c) suggests that power-law...
exponent $q$ should decrease with increasing $n_{ch}$, as might be expected if demand for larger multiplicities biases to more jet fragments by distorting the jet spectrum.

Figure 3 (left) shows data/TCM ratios based on a revised TCM with two hard-component $\hat{H}_0(p_t)$ parameters varying. Whereas $\hat{H}_0(y_t)$ was previously held fixed with power-law index $q = 5$ and width $\sigma_{y_t} = 0.465$ those parameters are now varied to accommodate individual spectra as described below. The modified TCM describes data above the mode within statistics (bold solid curves). Substantial spectrum deviations from the fixed TCM correlated with $n_{ch}$ appear in Fig. 1 (right) below the mode, but the corresponding manifestations in Fig. 3 (left) below $y_t = 3$ are strongly suppressed by the ratio format. Further modification could also accommodate those deviations (with the exception of the $n = 1$ high solid curve).

Figure 3 (right) repeats Figure 17 (left) with a revised TCM including varying hard-component model (curves through points) and log-log plotting format.

The data-model deviations at larger $y_t = 3$ are strongly suppressed by the ratio format.

Further modification could also accommodate those deviations (with the exception of the $n = 1$ high solid curve). The updated TCM ratios (new solid and dashed curves) pass through all data points (modulo statistical fluctuations), but the asymptotic limit in Eq. (A2) is no longer $\hat{\rho}_{s1}/\hat{\rho}_{s2}$ (dotted lines), is instead $\hat{\rho}_{s1}\hat{H}_0(p_t; q_1, \sigma_{y_1})/\hat{\rho}_{s2}\hat{H}_0(p_t; q_2, \sigma_{y_2})$ confirming that dijet production changes significantly with increasing event multiplicity. The log-log format reveals the reciprocal relation of the two ratios whereas the linear format in Fig. 17 gives the misleading impression that a ratio and its reciprocal carry different information.

Figure 4 (right) provides a check on the overall consistency of the TCM description and corresponds in format to Fig. 5 of Ref. [10] for comparison. The reference spectrum in this case is the TCM spectrum for $n_0 = 3$, with $\hat{H}_0(y_t)$ parameters that happen to coincide with the fixed model from Refs. [11, 12]. Following Ref. [10] each spectrum is normalized by its integral $\int p_t^2 \hat{H}_0(y_t) dy_t$.

The nominal parameter values for the 200 GeV fixed $\hat{H}_0(y_t)$ model (bold dashed curves in Fig. 1) are represented by the dotted and dash-dotted lines (corresponding to parameter values for $\hat{\rho}_s/\hat{\rho}_{s,ref} = 2$). Variation of two parameters in combination serves to broaden the hard-component model above the mode toward higher $p_t$. The saturation of $\sigma_{y_t}$ larger $\hat{\rho}_s$ is a consequence of increasing $2/q$. The transition point on $\hat{H}_0(y_t)$ from Gaussian to exponential form then moves back toward the mode and the exponential/power-law tail increasingly dominates the higher-$p_t$ structure. $\hat{H}_0(p_t)$ broadening could represent hardening of the underlying jet spectrum and/or modified jet formation. The corresponding parameter trends for 13 TeV are discussed in Sec. V C.

FIG. 3: Left: Ratios of data $y_t$ spectra to TCM equivalents with a varying hard component for six multiplicity classes. The data-model deviations at larger $y_t$ are consistent with statistical uncertainties (bold solid curves symmetric about unity) and tracking errors. The overall unweighted r.m.s. deviation for curves 2-6 is 3%. Right: Figure 17 (left) replotted with revised TCM including varying hard-component model (curves through points) and log-log plotting format.

FIG. 4: Left: Hard-component parameters varying with $n_{ch}$ or $\hat{\rho}_s$ for a revised TCM. 200 GeV solid and dashed curves through parameter data are defined by Eqs. (4). The $13$ TeV points and curves are discussed in Sec. V C. $\hat{\rho}_{s,ref} = 2.5$ for 200 GeV NSD $p-p$ collisions and 6 for 13 TeV INEL $> 0$ (inelastic events with at least one charged particle accepted) collisions. The factor 2 in $2/q$ permits greater plot sensitivity. The mean-value energy trend for $\sigma_{y_t}$ is shown in Fig. 19 (left) and for $1/q$ is shown in Fig. 13 (left). Right: Spectrum ratios for data (points) and TCM with varying hard component (solid curves) relative to an $n_0 = 3$ TCM reference, emulating Fig. 5 of Ref. [10] and demonstrating the accuracy of the 200 GeV TCM. The dashed curves represent the fixed hard-component TCM of Refs. [11, 12].
note that whereas comparison of classes 2 and 6 provides an example spectrum ratio the TCM for class 3 above serves as a reference to which all data are compared.

The present study demonstrates that high-statistics data from Ref. [2] convey substantial new information about the dijet contribution to p-p hadron spectra that is simply represented by smooth variation of existing $H_0(y_t)$ parameters. It also illustrates the utility of the TCM as a reference relative to which novel data properties can be detected and interpreted. Data/TCM spectrum ratios are apparently more easily described and interpreted than data/data ratios presented without a reference.

### B. Optimized hard component below the mode

Figure 3 (left) shows high-statistics 200 GeV p-p spectra for six multiplicity classes from Ref. [2] compared in ratio to the TCM with hard component varying as in Fig. 2 (left). One could conclude that the TCM data description is good except for the lowest ($n = 1$) multiplicity class (high solid curve). However, two bold solid curves symmetric about unity indicate one-sigma statistical errors that become very small in ratio at lower $y_t$. Spectrum ratios tend to strongly suppress statistically-significant information below the hard-component mode.

To progress further requires a differential data-TCM comparison relative to bin-wise statistical errors.

Figure 5 (left) is equivalent in principle to the ratios in Fig. 2 (left) but instead of the $\rho/\rho_{ref} + 1$ form of the latter this comparison has the per-particle form $\Delta \rho/\sqrt{\rho_{ref}}$ discussed in Sec. VII B comparable to Fig. 6 of Ref. [1]. The hatched band about zero indicates one-sigma statistical errors uniform on $y_t$ in this plot format. The variable hard component reduces residuals above the mode to the statistical level, but residuals below the mode remain very large compared to statistical errors and represent substantial spectrum information not accommodated within the revised TCM derived in Sec. IV B.

Figure 5 (right) shows the result of further modification of the hard-component model. Gaussian widths $\sigma_{y_t+}$ above and $\sigma_{y_t-}$ below the mode are varied separately. The values given in Fig. 4 (left) and Eq. 4 (lower) are retained for $\sigma_{y_t+}$ above the mode, but the values for $\sigma_{y_t-}$ are varied independently (in the form $1/\sigma_{y_t-}^2$) to accommodate data below the mode (except for $n = 1$). The resulting residuals for $n = 2-6$ are consistent with point-to-point systematic errors (about 1 permil of data values).

Figure 6 (left) shows variation of the hard-component width required to accommodate data below the mode in the form $1/\sigma_{y_t-}^2$ (solid points). The solid curve through points is $13.5\tan \{(\rho_s - 3.1)/5\}$. Also included is the trend for the width above the mode from Fig. 4 (left) and Eq. 4 (lower) as $1/\sigma_{y_t+}^2$ (open points and dashed curve respectively) demonstrating correlation of the two trends. The two widths become equal near $\rho_s \approx 5$, or $\rho_s/\rho_{ref,S} \approx 2$ in Fig. 3 (left), where the hard-component peak model is then approximately symmetric as in Refs. [1, 2].

### C. 200 GeV TCM $n_{ch}$ dependence summary

Figure 7 (left) summarizes the revised 200 GeV TCM hard-component model for seven multiplicity classes. The model for class $n = 1$ cannot accommodate data below the mode: the trend for $1/\sigma_{y_t-}^2$ in Fig. 6 (left) requires a negative entry for $n = 1$, and the shape of the model function would remain very different from the
data. The value of $\sigma_{y_1}$ for $n = 2$ is retained for $n = 1$. This model can be compared with data in Fig. 1 (right).

![Graph of $\alpha H(y_t)$](image)

**FIG. 7:** Left: Evolution of the hard-component model over seven multiplicity classes that exhausts all information in high-statistics spectrum data from Ref. 2. The model for $n = 1$ is undefined below the hard-component mode $y_t \approx 2.6$. Right: Difference between spectrum data and the full TCM relative to the hard-component model demonstrating that the latter is accurate to a few percent over the relevant $y_t$ interval.

Figure 4 (right) shows the scaled differential $\Delta \rho_0(y_t)/\rho_s$ denoted by expression (data – TCM) divided by the TCM hard component in the form $\alpha \rho_s H_0(y_t; n_{ch})$. Whereas full-spectrum ratios [e.g. Fig. 3 (left)] may be misleading this ratio to the hard-component model alone is informative, revealing that residuals between full (soft + hard) data spectra and a full revised TCM are less than 10% of hard-component values for $n = 2-6$ over the entire $y_t$ range relevant to jet-related spectrum structure.

V. 13 TeV p_t SPECTRUM RATIOS vs $n_{ch}$

In this section the spectrum TCM is applied to 13 TeV LHC $p_t$ spectrum ratios from Ref. 10 to infer separate soft- and hard-component spectrum models from those data. Relations derived for 200 GeV data in App. A are based on normalized spectra as defined in Ref. 11 over a large $n_{ch}$ interval and include a known soft component and fixed hard-component model. Relations defined in this section are based on ratios of spectra as normalized in Ref. 11 over a more-limited $n_{ch}$ interval with no soft-component model previously established but include an $n_{ch}$-varying hard component as in the previous section. Expressions below including an “$\approx$” symbol that define data quantities (points) are followed by corresponding TCM expressions including an “$\approx$” symbol (curves).

A. 13 TeV spectrum-analysis strategy

The 13 TeV spectrum study in Ref. 10 defined three multiplicity classes A, B and C as described in App. C1. Ratios were formed in which the $p_t$ spectrum from each class is divided by the INEL $> 0$ ensemble-mean spectrum adopted as a reference. The results are presented in Fig. 5 of Ref. 10 repeated in Fig. 2 (left) of the present study. It is concluded that “correlation of the spectrum with multiplicity...is stronger at high $p_T$...with jets presumably dominating the high-multiplicity domain.” It is further noted that “The general features...are similar to those first seen at $\sqrt{s} = 0.9$ TeV 21.” There is no reference to previous spectrum studies e.g. in Refs. 11 12 13 14 9 where differential analysis of $p$-$p$ spectrum $n_{ch}$ dependence and quantitative theoretical interpretations were first established, as reviewed in Sec. III and App. A.

Appendix A demonstrates that for high-statistics 200 GeV $p$-$p$ spectra a TCM with fixed hard component deviates significantly from ratio data at higher $p_t$. A hard-component model with two parameters varying with $n_{ch}$ is required by those data. For analysis of spectrum-ratio data from Ref. 10 a varying hard component is assumed from the beginning to establish parameter $n_{ch}$ trends.

The spectrum ratios in Fig. 2 (left) provide only indirect information on TCM elements in the form of ratios $T(p_t; n_{ch}') = H_0(p_t; n_{ch}')/S_0(p_t)$ per Eq. (6) below. By suitable transformation of ratio data $(R \rightarrow X \rightarrow Y \rightarrow T$ below) ratios $T(p_t; n_{ch}')$ for multiplicity bins A and C can be isolated as in Fig. 10 (left) (the bin-B ratio data provide no significant information).

An intermediate parametrization to describe 13 TeV $T(p_t; n_{ch}')$ ratio data is defined in terms of 200 GeV $S_0(p_t; T, n)$ by adjusting $H_0(p_t; n_{ch}')$ model parameters to fit the inferred ratio data. That intermediate parametrization is not intended as a final description of 13 TeV spectra. From that parametrization a model $T_0(p_t)$ for the 13 TeV INEL $> 0$ reference spectrum is determined by interpolation. To obtain the correct soft-component model for 13 TeV data the full reference spectrum in Fig. 3 of Ref. 10 is then fitted by varying only parameter $n$ of $S_0(p_t, n)$ in the modified TCM expression

$$\tilde{\rho}_0(p_t; n_{ch}') = \tilde{\rho}_s(n_{ch}') S_0(p_t, n) [1 + \alpha \tilde{\rho}_s(n_{ch}') T_0(p_t)],$$

where $n_{ch}'$ here corresponds to INEL $> 0$.

Given a correct 13 TeV $S_0(p_t, n)$ model the 13 TeV hard-component model is recovered from 13 TeV spectrum-ratio data by adjusting hard-component parameters to fit ratios $T(p_t; n_{ch}') = H_0(p_t; n_{ch}')/S_0(p_t, n)$ for bins A and C and the INEL $> 0$ spectrum, thus establishing a full $n_{ch}$-dependent TCM for 13 TeV $p_t$ spectra.

B. Spectrum ratios vs $n_{ch}$ and hard/soft ratio $T(p_t)$

Spectrum ratios $R(p_t; n_{ch}')$ are formed relative to a reference spectrum $\tilde{\rho}_{00}(p_t; \Delta \eta)$ (e.g. INEL $> 0$) for each of three multiplicity bins indexed by $n_{ch}'$. Spectra averaged over acceptance $\Delta \eta = 1.6$ are first normalized by their $p_t$ integrals as $\rho_0(p_t; n_{ch}', \Delta \eta)/\rho_{00}(n_{ch}', \Delta \eta)$. The spectrum-
ratio data are then represented by the first line of
\[
R(p_t; n'_{ch}) \equiv \frac{\bar{\rho}_{00}(\Delta \eta)}{\bar{\rho}_{00}(n'_{ch}, \Delta \eta)} \frac{\bar{\rho}_{0}(p_t; n'_{ch})}{\bar{\rho}_{0}(p_t; \Delta \eta)} \frac{\bar{\rho}_{0}(p_t; n_{ch})}{\bar{\rho}_{0}(p_t; \Delta \eta)} \approx \frac{\bar{\rho}_{00}(\Delta \eta) + \bar{\rho}_{0}(p_t; n_{ch})}{\bar{\rho}_{00}(\Delta \eta) + \bar{\rho}_{0}(p_t; \Delta \eta)}
\]
(6)

The TCM for unit-normal spectrum ratios based on Eq. (1) is described by the last three lines, where the \(\bar{\rho}_{00}\)s refer to the reference INEL > 0 event class and primess refer to incomplete \(p_t\) acceptance. The middle factor in the third line is assumed to be unity based on efficiency cancellations. The ratio limiting value at small \(p_t\) (first factor) has the approximate form \([1 + \alpha'\bar{\rho}_{00}(1 - \bar{\rho}_s/\bar{\rho}_{00})].\)

Figure 8 (left) shows ratio data from Fig. 2 (left) (points) for three \(n'_{ch}\) conditions (multiplicity bins A, B and C) relative to the INEL > 0 reference and ensemble-mean accepted \(n_{ch} \rightarrow \bar{\rho}_{00} = \bar{\rho}_{00} + \bar{\rho}_{0} \). Tracking efficiencies are assumed to cancel in ratios, and the \(p_t\) lower limit affects \(\bar{\rho}_s \rightarrow \bar{\rho}_s'\) but not \(\bar{\rho}_{00}\) because of its localization on \(p_t\). Limiting values at small \(p_t\) are R ≈ 1.12 for bin A, 1.02 for B and 0.91 for C. TCM curves are defined by Eq. (9) with reference ratio \(T_0(p_t)\) (dashed) and \(n_{ch}\)-dependent \(T(p_t; n'_{ch})\) (solid) defined below. The solid curve for case B does not appear (in this and later plots) because the \(\bar{\rho}_s/\bar{\rho}_{00}\) ratio is close to unity and spectrum ratio B is relatively insensitive to hard-component structure. The dash-dotted curves represent MC results (e.g. PYTHIA [20]) in Fig. 5 of Ref. [10].

The general form of the ratio data is a clear manifestation of the spectrum TCM, as anticipated by Fig. 3 (left) of Ref. [11] with spectra normalized by \(\bar{\rho}_0\) \((n_{ch}\) within \(\Delta \eta = 1)\) rather than \(\bar{\rho}_s\) as in Fig. 1 (left) of the present study. The excursions about unity at lower \(p_t\) defined by \(\alpha'\bar{\rho}_{00}(\bar{\rho}_{s}/\bar{\rho}_{00} - 1)\) ≈ 0.1 with \(\alpha'\bar{\rho}_{00} \approx 0.15\), can be compared with similar excursions at 200 GeV in Fig. 4 (right) where \(\alpha'\bar{\rho}_{00} \approx 0.025\) but the \(\bar{\rho}_{s}/\bar{\rho}_{00}\) range is larger.

Figure 8 (right) shows the first intermediate quantity \(X(p_t; n'_{ch})\) extracted from R(p_t; n'_{ch}) data (points) for conditions n'_{ch} represented in Eq. (7) for data (first line) and defined for the TCM (second line)
\[
X(p_t; n'_{ch}) \equiv R(p_t; n'_{ch}) \left( \frac{1 + \alpha'\bar{\rho}_s}{1 + \alpha'\bar{\rho}_{00}} \right) \rightarrow \frac{\bar{\rho}_s}{\bar{\rho}_{00}} \frac{T(p_t; n'_{ch})}{T_0(p_t)} \quad (at \ large \ p_t)
\]
(7)

In App. A1 ratio \(X(p_t)\) emerges directly from the spectrum ratio in Eq. (A2), thus bypassing quantity \(R(p_t)\) because of the choice of normalization in Ref. [11] and Fig. 4. Reference \(X_0(p_t)\) (dashed) results if \(T(p_t; n'_{ch}) \rightarrow T_0(p_t)\) in the numerator. The asymptotic limits at right (dotted lines) are \(\bar{\rho}_s/\bar{\rho}_{00} \approx 0.25\) for bin A, 0.85 for bin B and 1.75 for bin C. Given corrected \(\bar{\rho}_{00} \approx 6\) for 13 GeV INEL > 0 collisions the mean values 1.5, 5 and 10.5 for three bins do not match values 3, 9 and 15 estimated in App. C1.

From the combination of limiting values for \(R(p_t; n'_{ch})\) and \(X(p_t; n'_{ch})\) at small and large \(p_t\) respectively the product \(\alpha'\bar{\rho}_{00}\) for NSD collisions can be inferred uniquely from the spectrum-ratio data. \(\bar{\rho}_s/\bar{\rho}_{00}\) is estimated from limiting cases of Eq. (7) at larger \(p_t\) and combined with Eq. 9 limiting cases at smaller \(p_t\) to provide the estimate \(\alpha'\bar{\rho}_{00} \approx 0.15 \approx \alpha\bar{\rho}_{00}/\xi\). Since \(\xi \approx 0.6\) for these data \(\alpha\bar{\rho}_{00} \approx 0.09\) and (for \(\bar{\rho}_{00} \approx 6\) \(\alpha \approx 0.015 \pm 0.0015\) at 13 TeV vs \(\alpha \approx 0.006 \pm 0.001\) at 200 GeV [11, 2]. The jet-fragment yield per participant \(\sim \alpha\bar{\rho}_{00}\) is 6 times larger at 13 TeV than at 200 GeV and the total jet-fragment yield \(\sim \bar{\rho}_h \approx \alpha\bar{\rho}_{00}\) is 14 times larger. The energy variation of TCM soft-hard parameter \(\alpha\) is discussed in Sec. VII E in connection with Fig. 10.

Figure 9 (left) converts Fig. 8 (right) to a log-log format to demonstrate that the change in \(H_0(p_t; n'_{ch})\) with \(n_{ch}\) relative to the fixed reference in \(V_0(p_t)\) is larger for Bin A data (open points) than for Bin C data (solid points) explaining why bin A data in Fig. 10 (left) are further from the reference (upper dashed curve) than Bin C data. Compare that panel with the symmetry of Fig. 3 (right).
from that panel. The TCM is described by the second line. The first factor contains information on \( T(p_t; n'_{ch}) \) variation with \( n_{ch} \) class and leads to the solid curves representing the revised TCM. The factor in parentheses defines \( \hat{Y}_0(p_t) \) (dashed) that would result if \( T(p_t; n'_{ch}) \rightarrow T_0(p_t) \). \( \hat{Y}_0(p_t) = 0.5 \) when \( H(p_t) = S(p_t) \) or \( \alpha_\rho S_0 T_0(p_t) = 1 \). Compare with Fig. 17 (right) for 200 GeV data.

Figure 10 (left) shows the hard/soft ratio \( T(p_t; n'_{ch}) \) (solid and open points) obtained from \( Y(p_t; n'_{ch}) \) data as

\[
\alpha \hat{\rho}_s T(p_t; n'_{ch}) = \alpha \hat{\rho}_s T_0(p_t) \left[ \left( \frac{\bar{p}_s}{\hat{\rho}_s} - 1 \right) \frac{Y(p_t; n'_{ch})}{\hat{Y}_0(p_t)} + 1 \right] \tag{9}
\]

with \( T_0(p_t) \) (upper dashed) obtained from a 13 TeV spectrum fit as described below. Data and TCM transform equivalently. Dotted curves show the result if data were transformed as in Eq. (9) assuming a fixed \( T_0(p_t) \). The lower dashed curve is \( T_0(p_t) \) for 200 GeV from Fig. 18 (left). The change in ratio \( H(p_t)/S(p_t) \) indicates a larger role played by jets at lower \( p_t \) for 13 TeV \( p-p \) collisions.

The dash-dotted curve is the MC curve(s) in Fig. 8 (left) transformed as fixed \( T(p_t) \rightarrow T_0(p_t) \) indicating that MCs assume a fixed hard component, larger in amplitude and skewed to lower \( p_t \) than data, similar to results presented in Sec. IX of Ref. [1] and consistent with an assumed parton/jet spectrum extending well below 3 GeV, therefore predicting a large excess of low-energy jets.\(^1\)

A temporary parametrization for \( T_0(p_t) \) (upper dashed curve) can be constructed by retaining the 200 GeV \( S_0(p_t) \) model and adjusting \( \hat{H}_0(p_t; n'_{ch}) \) parameters to fit the 13 TeV \( T(p_t; n'_{ch}) \) data in Fig. 10 (left). Resulting parameter multiplicity trends can be interpolated to define \( T_0(p_t) \) describing the INEL > 0 reference spectrum.

\(^1\) The conclusion in Ref. [1] that the hard-component yield from PYTHIA is less than that inferred from data is incorrect. Plotted PYTHIA curves are comparable to \( \alpha_\rho H_0(p_t) \) with \( \alpha_\rho \approx 0.015 \) for 200 GeV NSD \( p-p \) collisions. The PYTHIA hard-component density \( \hat{\rho}_h \) is thus 2-3 times larger than that inferred from data.
from Fig. 3 of Ref. [10]. The fitted TCM (solid) is
\[ \rho_{00}(p_t) = \rho_{00}\tilde{S}_0(p_t; T, n)\left[1 + \alpha\rho_{00}T_0(p_t)\right] \]
(10)
with \(\alpha\rho_{00}T_0(p_t)\) represented by the upper dashed curve in Fig. [10] (left). The only TCM adjustment is variation of exponent \(n\) in the Lévy form of \(\tilde{S}_0(m_t; T, n)\) in Eq. (10) to fit the spectrum data, with slope parameter \(T = 145\) MeV held fixed. The fitted Lévy exponent \(n \approx 7.8\) at 13 TeV can be compared with \(n \approx 12.5\) at 200 GeV [2].

![Graph](image)

**FIG. 11:** Left: Data obtained from the INEL \(> 0\) 13 TeV \(p_t\) spectrum in Fig. 3 of Ref. [10] (points). The solid curve is the model described by Eq. (10) with \(n\) optimized. The three dashed curves represent \(H(p_t; n_{ch}')\) for \(T_0(p_t)\) (middle) and for \(T(p_t; n_{ch}')\) from bins A and C (lower and upper respectively). The 13 TeV soft component is represented by the dotted curve. The dash-dotted curve represents \(H(p_t)\) for 200 GeV. Right: Fit residuals for the left panel (points). Spectrum ratios with varying hard component for bins A and C (dashed) can be compared with 200 GeV ratio data in Fig. [18] (right). The hatched band indicates \(\pm 5\%\) deviations.

Given optimized soft-component model \(\tilde{S}_0(p_t; T, n)\) the 13 TeV hard-component model parameters \(y_t, \sigma_t, \) and \(q\) are determined by fitting the TCM for \(T(p_t; n_{ch}') = H_0(p_t; n_{ch}')/\tilde{S}_0(p_t; T, n)\) to data from Fig. [10] (left) (bins A and C) and the reference spectrum in Fig. [11] (left) (INEL \(> 0\) reference). That parametrization back transformed generates solid model curves that pass through data for bins A and C in all previous figures of Sec. V.

Figure [11] (right) demonstrates the quality of the TCM reference description with fit residuals (points). There is no systematic deviation of data from the optimized TCM. Deviations are substantially less than \(5\%\) (hatched band) over a momentum interval from 0.15 GeV/c to 16 GeV/c. That result can be compared with claims of “log-periodic” oscillations [22] in deviations of a so-called Tsallis distribution [23] (equivalent to Eq. [B1] of the present study) from a spectrum for 7 TeV \(p-p\) collisions that are similar to the residuals at 0.9 TeV in Fig. [20] (b).

The soft and hard spectrum components in Fig. [11] (left) are \(S(p_t) = \rho_{00}\tilde{S}_0(p_t)\) (dotted curve) and \(H(p_t) = \alpha\rho_{00}T_0(p_t)\) (dashed curve so labeled). The 200 GeV hard component (dashed-dotted curve) is shown for comparison. Two other dashed curves in Figure [11] (left) correspond to Eq. (10) with \(T_0(p_t) \rightarrow T(p_t; n_{ch}')\) represented by solid curves A and C in Fig. [10] (left). The dashed curves in Figure [11] (right) are Eq. (10) including \(H_0(p_t; n_{ch}')\) (for bins A and C) and the INEL \(> 0\) reference that can be compared with ratio data in Fig. [18] (right).

The fitted \(H_0(p_t; n_{ch}')\) parameter values for bins A and C and the INEL \(> 0\) reference are included in Fig. [4] (left) (upper points, labeled 13 TeV) along with simple parametrizations (upper solid and dashed curves) that show variations with \(\rho_0/\rho_{00}\) interval. The 13 TeV INEL \(> 0\) reference values are included in Table III.

These results plus previous figures in this section indicate that the 13 TeV TCM with \(n_{ch}\)-varying hard component provides an accurate \(p_t\) spectrum description over a significant range of event multiplicities. It also buttresses results from Ref. [1] that revealed a 200 GeV spectrum hard component with mode near \(p_t = 1\) GeV/c and approximate power-law tail at higher \(p_t\) compatible with QCD predictions derived from reconstructed jets.

**VI. \(p_t\) SPECTRUM ENERGY EVOLUTION**

Given \(n_{ch}\)-dependent TCMs inferred from 200 GeV and 13 TeV spectrum data an energy-dependent TCM continuously covering the interval from 17 GeV to 13 TeV is derived using supplementary spectrum data.

### A. Energy evolution of spectrum soft exponent \(n\)

The spectrum soft component at 13 TeV is substantially different from that at 200 GeV from Ref. [1]. The Lévy exponent changes from \(n \approx 12.5\) at 200 GeV to \(n \approx 7.8\) at 13 TeV (harder spectrum). Results from CERN super proton synchrotron (SPS) \(p-p\) spectra extend the energy trend over a larger energy interval.

Figure [12] (left) shows an \(m_t\) spectrum for identified pions (points) from 17.2 GeV inelastic \(p-p\) collisions well described by a Lévy distribution (solid curve) with exponent \(n \approx 27\) and the universal slope parameter \(T = 145\) MeV. Also shown is the corresponding Maxwell-Boltzmann (M-B) distribution (dashed-dotted curve) with \(1/n \rightarrow 0\) and the same slope parameter. Given the known energy dependence of the jet contribution [2] the spectrum hard component should produce at most a slight deviation from the soft component in that \(p_t\) interval, especially for inelastic \(p-p\) collisions (the dashed curve represents a TCM predicted sum of soft + hard components – see Sec. VIB). The spectrum data at 17.2 GeV then constrain only soft component \(\tilde{S}_0(m_t; T, n)\).

Figure [12] (right) shows soft-component exponents in the form \(1/n\) inferred from spectrum data for three collision energies (solid points) at the SPS, RHIC and LHC. The solid curve is an algebraic hypothesis based on variation of the soft component due to Gribov diffusion [26]. Low-\(x\) gluons result from a virtual parton...
splitting cascade within projectile nucleons whose mean depth on \( x \) is determined by the collision energy. Each step of the cascade adds transverse-momentum components in a random-walk process. The depth of the cascade is proportional to \( \ln(s/s_0) \), and \( \sqrt{s_0} \approx 10 \text{ GeV} \) is inferred from dijet systematics \[12\]. Given the properties of a random walk and with \( 1/n \) as a measure of transverse-momentum excursions \[15\] its trend is estimated as \( \propto \ln(\sqrt{s}/10 \text{ GeV}) \) (solid curve). The open circles at 0.9, 2.76 and 7 TeV are interpolations of the Lévy exponent to \( n = 9.82, 8.83 \) and 8.16 respectively.

### B. Energy evolution of spectrum hard exponent \( q \)

Figure 13 (left) shows inverse values (solid points) of exponents \( q = 5.15 \) for 200 GeV as in Fig. 2 and \( q = 3.65 \) for 13 TeV as in Fig. 11 (left) plotted vs quantity \( \Delta y_{\text{max}} = \ln(\sqrt{s}/6 \text{ GeV}) \) observed to describe the energy trend for jet spectrum widths \( \propto \Delta y_{\text{max}} \) from NSD \( p-p \) collisions assuming a jet spectrum low-energy cutoff near 3 GeV \[5\] [see Fig. 5 of Ref. 5 for a direct comparison with measured jet spectra]. The inverse \( 1/q \) effectively measures the hard-component peak width at larger \( \Delta y \). The relation \( 1/q \propto \Delta y_{\text{max}} \) (solid line) is expected given that the \( p-p \) \( p_T \)-spectrum hard component can be expressed as the convolution of a fixed \( p-p \) FF ensemble with a collision-energy-dependent jet spectrum \[3\], and the jet-spectrum width trend has the same dependence \[3\]. The vertical hatched band indicates an inferred cutoff to dijet production from low-\( x \) gluon collisions near 10 GeV. That the same relation applies to the ensemble-mean-\( p_T \) hard component has been established in a separate study \[27\]. The inverse values of \( q = 3.80 \) for 7 TeV, \( q = 4.05 \) for 2.76 TeV and \( q = 4.45 \) for 0.9 TeV (open circles) are obtained by interpolation.

### C. \( p_T \) spectrum ratio for two LHC energies

Figure 4 of Ref. 10 [Fig. 2 (right) of this paper] provides partial information on the energy variation of \( p_T \) spectra between 7 and 13 TeV. Supplementary information must be introduced to obtain a full spectrum description. Soft and hard spectrum components are considered separately.

The data spectrum ratio in Fig. 2 (right) is represented by the first line of

\[
R(p_T; E_1, E_2) = \frac{\tilde{\rho}_0(p_T; E_2)}{\tilde{\rho}_0(p_T; E_1)} \equiv \frac{\tilde{\rho}_{s2} S_0(p_T; E_2) + \tilde{\rho}_{s2} H_0(p_T; E_2)}{\tilde{\rho}_{s1} S_0(p_T; E_1) + \tilde{\rho}_{s1} H_0(p_T; E_1)} \approx \frac{\tilde{\rho}_{s2}}{\tilde{\rho}_{s1}} \frac{\tilde{S}_0(p_T; E_2)}{\tilde{S}_0(p_T; E_1)} \left[ 1 + \alpha \tilde{\rho}_{s2} T(p_T; E_2) \right] \frac{1}{\tilde{S}_0(p_T; E_1)} \left[ 1 + \alpha \tilde{\rho}_{s1} T(p_T; E_1) \right].
\]

The second line defines the TCM for this case and the third line indicates a factorization similar to that in Eq. 6. However, the \( \tilde{S}_0 \) ratio does not cancel, is determined instead by the exponent-\( n \) trend in Fig. 12 (right).

Figure 13 (right) shows spectrum-ratio data (points) as in Fig. 2 (right). The low-\( p_T \) limit of \( R(p_T) \) is density ratio \( \tilde{\rho}_{s2}/\tilde{\rho}_{s1} \approx 1.12 \) (hatched band) compared with expected ratio 1.10 derived from the soft-component trend (dotted curve) in Fig. 22 of App. C. This ratio of full spectra confuses soft and hard TCM components, is insensitive to
energy-dependent jet physics obscured at lower $p_t$ by the spectrum soft component, and the changes between 7 and 13 TeV are small as shown below. Improved sensitivity to jet physics could be obtained by analyzing spectra from higher-multiplicity $p$-$p$ collisions.

Given energy trends for the soft- and hard-component exponents derived above, the 7 TeV spectrum hard component can be isolated. Quantity $R_0(p_t)$ (dashed curve) is Eq. (11) with both hard-component forms $\bar{H}_0(p_t; E)$ fixed at $E_2 = 13$ TeV so that $T(p_t; E_1) \to T(p_t; E_1) \equiv \bar{H}_0(p_t; E_2)/\bar{S}_0(p_t; E_1)$ with Lévy $\hat{S}_0$ index $n = 8.16$ at 7 TeV derived from Fig. 12 (right). $R_0(p_t; E_1; E_2)$ then represents only the soft-component contribution to $R(p_t; E_1; E_2)$ variation with energy, which is known.

Figure 14 (left) shows data for quantity $X(p_t; E_1, E_2)$ (points) defined by the first line of

$$X(p_t; E_1, E_2) = \frac{R(p_t; E_1, E_2)}{\rho(p_t; E_1; E_2)}$$

$$\approx \frac{1 + \alpha \rho_1 T_0(p_t; E_1)}{1 + \rho_0 T(p_t; E_1)}.$$

The second line defines a TCM expression having asymptotic form $\bar{H}_0(p_t; E_2)/\bar{H}_0(p_t; E_1)$ at larger $p_t$. From Fig. 10 (left) $S(p_t) \approx H(p_t)$ near $p_t \approx 1.4$ GeV/c (vertical dotted line). To the right of that point ratio $X(p_t)$ is dominated by the spectrum hard components. With $T_0(p_t; E_1) \equiv T(p_t; E_2) \times S_0(p_t; E_2)/S_0(p_t; E_1)$ and $\alpha \rho_1$ already determined by data, either directly or by interpolation, data ratio $T(p_t; E_1)$ can be obtained via

$$T(p_t; E_1) = \frac{1}{\alpha \rho_1} \left\{ \frac{1 + \alpha \rho_1 T_0(p_t; E_1)}{X(p_t)} - 1 \right\},$$

yielding $H(p_t; E_1) = \alpha \rho_1^2 \bar{S}_0(p_t; E_1)/T(p_t; E_1)$.

![Figure 14: Left: Ratio $X(p_t)$ derived from data in the left panel with Eq. (12) (first line) (points). The solid curve is the second line with 7 TeV $\bar{H}_0(p_t)$ adjusted to accommodate the data by varying $\bar{y}_0$ and $\sigma_{p_t}$ to obtain the values appearing in Table 1. Right: Hard-component ratio comparing 13 and 7 TeV (points) derived from points in the left panel via Eq. (13) and $H(p_t; E) = \alpha(E)\rho_0^2(E)\bar{S}[p_t; n(E)]T(p_t; E)$. The dashed curve is determined by 7 and 13 TeV TCM parameters in Table 1.](image)

D. Energy evolution of spectrum hard component

Figure 15 (left) shows TCM hard-component ratios for three energy pairs (curves), where $H(p_t; E) \approx \alpha(E)\rho_0^2(E)\bar{H}_0(p_t; E)$ relates hard-component yield data to unit-normal model functions. The curves are determined by the parameters in Table 1. The 13 vs 7 TeV comparison (points) reveals little about the energy evolution of $H(p_t; E)$, but larger energy intervals demonstrate that the hard-component width near its mode broadens significantly, and the high-$p_t$ power-law tail falls much less rapidly at higher collision energies. Those trends are quantitatively consistent with the measured $p$-$p$ collision-energy dependence of underlying jet energy spectra.

![Figure 15: Left: Hard-component ratios for three energy pairs (curves), where $H(p_t; E) \approx \alpha(E)\rho_0^2(E)\bar{H}_0(p_t; E)$ relates hard-component yield data to unit-normal model functions. The curves are determined by the parameters in Table 1. Right: A survey of spectrum hard components over the currently accessible energy range from threshold of dijet production (10 GeV) to LHC top energy (13 TeV). The curves are determined by parameters in Table 1 except for the 200 GeV fine solid curves determined also by the $\sigma_{p_t}$ and $q$ trends in Fig. 14 (left). The points are from Refs. 2 (200 GeV) and 10 (13 TeV).](image)
data from four energies (13, 7 and 0.2 TeV above and 0.9 TeV in App. B). The 17.2 GeV extrapolation described below indicates no significant jet contribution to yields and spectra at that energy (dashed curve in Fig. 12—left) and explains why no excess $p_t$ fluctuations were observed at that energy [28-29]. However, evidence for SPS jets is visible in 17 GeV azimuth correlations as a more sensitive detection method [30]. The 200 GeV summary includes parametric variation of $\tilde{H}_0(y_t; q, \sigma_{y_t})$ for seven multiplicity classes (thin solid curves) as described in Sec. IV A. Corresponding data (solid points) represent NSD $p$-$p$ collision energy over three orders of magnitude.

E. Spectrum TCM parameter summary

Table I summarizes NSD $p$-$p$ TCM parameters for a broad range of energies. The entries are grouped as soft-component parameters ($T, n$), hard-component parameters ($y_t, \sigma_{y_t}, q$), hard-soft relation parameter $\alpha$ and soft density $\bar{\rho}_s$. Slope parameter $T = 145$ MeV is held fixed for all cases consistent with observations. Its value is determined solely by a lowest-$y_t$ interval where the hard component is negligible. The interpolated Lévy exponent $n$ values are derived from Fig. 12 (right) (open circles). Interpolated hard-component $y$ values are derived from Fig. 13 (left) (open circles), $\bar{\rho}_s$ values are derived from the universal trend in Fig. 2 (dotted curve) inferred from correlation and yield data. All 0.9 and 2.76 TeV values are predicted via interpolations. All remaining (unstarred) entries are obtained from fits to data.

Greek geV in Ref. I and for 13 TeV in the present study, are denoted by solid triangles in Fig. 16 (left). A model for $\alpha$ energy dependence can be determined by the following argument. $\alpha$ is related to jet systematics via $\bar{\rho}_h$ as

$$\bar{\rho}_{h, NSD} = \epsilon(\Delta y)f_{NSD}2\bar{n}_{ch,j}$$

for a given collision energy, where $2\bar{n}_{ch,j}$ is the mean hadron fragment multiplicity per dijet averaged over a jet spectrum for that energy [31] and $f_{NSD} = (1/\sigma_{NSD})d\sigma_{jet}/d\eta$ [3]. The energy trends for those quantities, inferred from reconstructed-jet data, and $\bar{\rho}_h$ from App. C can be used to predict the energy trend for $\alpha(\sqrt{s})$. Certain defined kinematic quantities are useful: $y_{max} = \ln(2E_{jet}/m_j)$ is a logarithmic representation of jet energy, and $y_t = \ln(\sqrt{s}/m_\pi)$ similarly represents the $p$-$p$ collision energy. $\Delta y_2 = \ln(\sqrt{s}/10$ GeV $)$ represents an observed cutoff of dijet production near $\sqrt{s} = 10$ GeV, and $\Delta y_{max} = \ln(\sqrt{s}/6$ GeV $)$ responds to an inferred infrared cutoff of jet spectra near $E_{jet} = 3$ GeV. Reference [5] reports $d\sigma_{jet}/d\eta \approx 0.026\Delta y^2_{jet}\Delta y_{max}$ and $\sigma_{NSD} \approx 0.83(2 + \Delta y^2_2)$. From Fig. 22 $\bar{\rho}_h \approx 0.81\Delta y_2$. Dijet acceptance factor $\epsilon \approx 0.6$ is an estimate for $\Delta y = 1.5 - 2$ IB. Combining various elements the $\alpha(\sqrt{s})$ trend is

$$\alpha(\sqrt{s}) \approx \epsilon(\Delta y)d\sigma_{jet}2\bar{n}_{ch,j}\frac{\sigma_{NSD}}{\bar{\rho}_h^2}$$

(15)

It remains to determine the systematic variation of dijet fragment multiplicity $2\bar{n}_{ch,j}$ with $p$-$p$ collision energy.

Figure 16 (right) shows the energy trend for factor $2\bar{n}_{ch,j}(\sqrt{s})$ in Eq. (15) estimated in two ways. The first estimate is based on published FFs and jet spectra.
each collision energy a parametrization of the jet spectrum for that energy from Ref. [5] is used to obtain the weighted mean $2\bar{n}_{ch,j}(\sqrt{s})$ of dijet fragment yields $2n_{ch,j}(E_{jet})$ for $p$-$p$ collisions from Fig. 6 of Ref. [3]. The weighted means $2\bar{n}_{ch,j}(\sqrt{s})$ for six collision energies (points) are parametrized by $2\bar{n}_{ch,j}(\sqrt{s}) \approx 3(1 + \Delta y_{max}/10)$ (dashed). When inserted into Eq. (15) that expression produces the dashed curve in the left panel which describes substantially from the $\alpha(\sqrt{s})$ trend inferred from spectrum analysis (solid triangles).

An alternative estimate is the simple proportionality $2\bar{n}_{ch,j}(\sqrt{s}) \approx 0.7\Delta y_{max}$ (dash-dotted) in the right panel. When inserted into Eq. (15) that result produces the dash-dotted curve in the left panel that describes well the $\alpha(\sqrt{s})$ trend inferred from spectrum analysis. The same expression is used to generate the solid curve in Fig. 22. This comparison establishes an absolute relation between jet fragments within reconstructed jets and jet fragments manifesting as spectrum hard components. The comparison suggests that FFs from $p$-$\bar{p}$ collisions, already substantially modified (sharply reduced at lower fragment momentum) as compared to $e^+e^-$ collisions [3], may still overestimate total fragment yields at lower $p$-$p$ collision energies by 50-100%. Those issues may be related to factor-2 disagreements between NLO pQCD theory predictions and measured $p_t$ spectra [32].

VII. SYSTEMATIC UNCERTAINTIES

The main purpose of this study is to extend the $p$-$p$ $p_t$ spectrum TCM established at the RHIC to the highest available collision energies. Previous studies of TCM $n_{ch}$ dependence have been based on isolated $p_t$ spectra [1, 2] whereas spectrum ratios are presently available at LHC energies as in Ref. [10]. Spectrum ratios do not clearly distinguish two dominant hadron production mechanisms and are thus difficult to parametrize accurately or to interpret. However, within the TCM framework spectrum ratios can be processed to isolate soft and hard components as demonstrated in this study. In this section the accuracy of the extended TCM is evaluated.

A. 200 GeV spectrum TCM and spectrum ratios

Systematic uncertainties for 200 GeV $p$-$p$ $p_t$ spectra as in Fig. 1 are described in Sec. VIII of Ref. [1]. The most uncertain aspect of the TCM is the shape of the low-$p_t$ part of hard component $H(p_t)$ (below the mode) that results from subtraction of inferred model $S_0(p_t)$, and the uncertainty is greatest for the lowest multiplicity class. While adjusting the amplitude of $S_0(p_t)$ by a few percent may alter the structure below the mode the result does not match the shape variation with $n_{ch}$ observed there.

Given the imposed $n_{ch}$ bins Table I uncertainties for $\hat{\rho}_0$ and $\hat{\rho}_s$ depend on the estimates for $\xi = 0.66 \pm 0.02$ and $\alpha = 0.006 \pm 0.001$. The value of $\xi$ is estimated from the integrals of $\hat{S}_0(p_t)$ and $\hat{S}_p(p_t)$ (dotted curves in Fig. 1 – left) that describe the data. The value of $\alpha$ is derived from an iterative process described in Refs. [11, 2] where it was established empirically that $\hat{p}_h = \hat{\rho}_s^2$ with $\alpha \approx 0.006$ and $\hat{\rho}_h$ is the integral of spectrum hard component $H(p_t)$. The variations in $\hat{H}_0(p_t)$ introduced in Sec. IV A do not change that relation since small changes to the high-$p_t$ tail do not significantly affect the integral. Figure 5 (left) indicates that the resulting spectrum scaling is self-consistent to a few percent.

A new aspect of this study is the $n_{ch}$ dependence of the TCM hard component. The 200 GeV parameter values in Fig. 1 (left) provide an indication of the precision. The points represent best-fits-by-eye, and the values were not adjusted after the fit procedure. Small deviations from the simple $\hat{\rho}_s$ trends in Eq. (15) suggest parameter precision at the few-percent level. Fig. 1 (right) appears to confirm Fig. 3 (left) as to the quality of the TCM description over a 100-fold variation in dijet production rate and ten-fold variation in soft-hadron density.

The 200 GeV spectrum ratios are in principle precise since $p_t$-dependent tracking inefficiencies cancel and $\hat{\rho}_s$ is well-defined. The relation of absolute spectra to the TCM with varying hard component is suggested by Fig. 3 (left) with r.m.s. deviation 3% for $n = 2-6$ spectra. However, the apparent precision can be misleading as discussed in the next subsection and Sec. IV A. Spectrum ratios alone representing only a fraction of the available information must be supplemented by isolated spectra as in Sec. IV C. And comparisons between models and data must demonstrate the degree of statistical precision. Qualitatively, the $n_{ch}$ dependence of spectrum ratios and the precision of inferred $T(p_t; n'_{ch}, \sqrt{s})$ for several collision systems do confirm a peaked spectrum hard component with mode near 1 GeV/c in agreement with Ref. [1].

B. Spectrum precision and statistical significance

It is useful to specify the r.m.s. Poisson errors for spectrum $\hat{\rho}_0(y_t; n'_{ch})$ on $y_t$ (assuming uncorrected spectra)

$$\delta \hat{\rho}_0(y_t; n'_{ch}) = \frac{\sqrt{\hat{\rho}_0(y_t; n'_{ch})}}{\sqrt{y_t dy_t \Delta \eta N_{ext}(n'_{ch})}}$$

where $N_{ext}(n'_{ch})$ is the number of events for event class $n'_{ch}$. If spectra are normalized by soft charge density $\hat{\rho}_s$, the errors for $\hat{\rho}_0(y_t; n'_{ch})/\hat{\rho}_s$ are

$$\delta \hat{\rho}_0(y_t; n'_{ch})/\hat{\rho}_s = \frac{\sqrt{\hat{\rho}_0(y_t; n'_{ch})/\hat{\rho}_s}}{\sqrt{y_t dy_t \Delta \eta \hat{\rho}_s N_{ext}(n'_{ch})}}$$

The form $\Delta \rho/\sqrt{\rho_{ref}}$ equivalent to Pearson’s normalized covariance [33] has been introduced previously as a per-particle measure of two-particle angular correlations [22]. In Eq. (16) the equivalent quantity for SP spectra, introduced with similar structure to assess the statistical significance of spectrum structure, is based on the Poisson error expression in Eq. (10) and measures fit residuals.
relative to statistical errors as in Fig. 20 (b). Comparing spectra in the form of a ratio $\rho_{\text{dat}}(p_t)/\rho_{\text{ref}}(p_t) \sim \Delta \rho/\sqrt{\rho_{\text{ref}}}$ + 1 as in Fig. 3 (left) introduces an additional factor $1/\sqrt{\rho_{\text{ref}}}$ compared to $\Delta \rho/\sqrt{\rho_{\text{ref}}}$ that strongly suppresses apparent residuals at lower $p_t$ given typical spectrum variation with $p_t$ over 6-8 orders of magnitude.

In Ref. 11 Figs. 2 (left) and 6 show residuals scaled by statistical errors in the form $\Delta \rho/\sqrt{\rho_{\text{ref}}}$, with “data” referring to normalized spectra as in Eq. 17 above and $\Delta \eta = 1$. The assumed prefactor $\sqrt{y_t N_{\text{evt}}(n'_{ch})}$ in those figures then omits factors $dy_t p_t n'_{ch}$. The main effect of the additional factors would be to reduce the scaled residuals for the lower event classes while increasing those for the higher, but the $p_t$ or $y_t$ structure is comparable.

In Fig. 18 (right) the $n'_{ch}$ dependence of $H_0(y_t, n'_{ch})$ at higher $p_t$ appears substantial whereas no such structure is evident in Fig. 6 of Ref. 11 describing the same collision system. There are two reasons: (a) The event number for Ref. 11 was 3M distributed across eleven multiplicity classes whereas for Ref. 2 the event number was 6M distributed across seven multiplicity classes, and the acceptance for the former was $\Delta \eta = 1$ whereas for the latter it was $\Delta \eta = 2$, the combination producing a 3-fold difference in the spectrum r.m.s. statistical error. (b) Figure 6 of Ref. 11 is in the form $\Delta \rho/\sqrt{\rho_{\text{ref}}}$ whereas Fig. 18 (right) of the present study is in the form $\Delta \rho/\sqrt{\rho_{\text{ref}}} + 1$. The extra factor $1/\sqrt{\rho_{\text{ref}}}$ in the latter (varying over several orders of magnitude) strongly suppresses deviations at lower $p_t$ and exaggerates deviations at higher $p_t$. Spectrum comparisons in the form $\Delta \rho/\sqrt{\rho_{\text{ref}}}$ may indicate a certain level of precision but provide no indication of statistical significance. Highly significant deviations at smaller $p_t$ or $y_t$ may be strongly suppressed leading to misconceptions. Spectrum comparisons in the form $\Delta \rho/\sqrt{\rho_{\text{ref}}}$ relate data-model differences to statistical uncertainties presented uniformly over the full $p_t$ acceptance. The differences in two descriptions of the same data can be dramatic as demonstrated in Sec. I V.

C. 13 TeV $n'_{ch}$ binning and $n'_{ch}$ energy dependence

In Refs. 11, 2 accurate independent determination of $\bar{\rho}$ for each multiplicity class facilitated the 200 GeV analysis. In order to provide equivalent information for analysis of Ref. 10 data information on multiplicity distributions from Ref. 34 was employed, specifically the NBD parameters in Table 11. The NBD fits to MD data are accurate to a few percent up to $n'_{ch}/\Delta \eta = 20$ which includes almost all events. Based on multiplicity bins defined in Ref. 10 (denoted here by A, B and C) the bin means were determined from the NBD parameters in Table 11. Figure 22 shows various charge-density measurements vs collision energy and can be used to estimate systematic uncertainties. The solid dots are mean values for several energies obtained from the NBD parameters in Table 11. They are described by the TCM trend (solid curve) within 5% and are consistent with results from Table 12 of Ref. 34 with 4% uncertainties. The open squares summarize data for three event types (INEL, NSD, INEL > 0) from Table 7 of Ref. 34. The combined uncertainty is less than 3% for each point. The summary suggests that the NBD parametrizations in Table 11 provide mean values with accuracy better than 5%.

Based on the bins defined in Ref. 10 and indicated in Fig. 21 (right) the corrected bin means as $\bar{\rho}_t = n'_{ch}/\Delta \eta$ are 3, 9 and 15 and the ensemble mean is 6.2. Bin-mean ratios to the ensemble mean are 0.48, 1.45 and 2.4 for bins A, B and C respectively and should be independent of efficiency or acceptance. Assuming $\xi \approx 0.6$ for Ref. 10 data the corresponding $N_{\text{acc}}^{n'_{ch}}$ values within $\Delta \eta = 1.6$ should be approximately 3, 8.6 and 14.5 with ensemble mean $\langle N_{\text{acc}}^{n'_{ch}} \rangle \approx 6$ (compared to 6.73 from Ref. 10). The bin means themselves are not reported in Ref. 10.

Returning to the energy trends in Fig. 22 the NBD averages (solid dots) appear to favor the TCM trend (solid curve) whereas direct density measurements (open squares) seem to favor the “power law” trend (dashed curve). While the two data sets are consistent within their systematic uncertainties those estimates may not reflect point-to-point uncertainties that could be significantly smaller. MB dijet production continues to scale as $\bar{\rho}_t/\bar{s}\approx \log(\sqrt{s}/10 \text{ GeV})$ as low as 62 GeV [12] and presumably still depends on low-$x$ gluon participants at that point represented by $\bar{s}_x$. Differential spectrum analysis at lower energies might help resolve apparent conflicts.

D. 13 TeV spectrum ratios

Figure 8 (left) shows three spectrum ratios from Ref. 10. The ratio data should in principle be precise due to cancellation of instrumental effects but do present certain difficulties. In the simpler form of Eq. 2 (fourth line) quantity $X(p_t) \rightarrow \rho_{\text{act}}/\rho_{\text{ref}}$ at higher $p_t$ for bin $x$, and that ratio should be accurately known to a few percent. But in Fig. 3 (right) the approximate ratios are 0.25, 0.85 and 1.75 for bins A, B and C respectively whereas 0.48, 1.45 and 2.4 are expected based on results from the previous subsection. That the bin B ratio is asymptotically less than 1 suggests that inverses of described spectrum ratios might actually be plotted. As demonstrated in Fig. 8 (right) the same information should emerge from a TCM analysis in either case.

In Eq. 6 (fourth line) the asymptotic limit at lower $p_t$ is determined by the same ratio $\rho_{\text{act}}/\rho_{\text{ref}}$ as at higher $p_t$, thus providing a consistency check on the data. To obtain consistency the ratio for bin A was rescaled by factor 0.97 and that for bin C was rescaled by factor 1.03, both adjustments within the systematic uncertainties ($\pm 5\%$) of the spectrum data in Fig. 3 of Ref. 10. Bin B data were not processed further because deviations from unity (the physically relevant structure) are not significant.

Those issues notwithstanding, the TCM description of bin A and C ratio data is quite accurate as indicated by Fig. 11 (right): fit residuals well within $\pm 5\%$ up to 16
GeV/c and consistent with individual data uncertainties. The TCM descriptions of other quantities [e.g. \( R(p_t) \), \( X(p_t), Y(p_t) \)] are of similar quality. The 13 TeV parameters extracted for the varying \( H_0(p_t; n'_{ch}) \) model as shown in Fig. 3 (left) seem consistent within a few percent.

### E. TCM parameter energy systematics

Table I summarizes the TCM over a large energy interval. The model accuracy can be assessed both by the simplicity of parameter variations and by comparisons with data that did not contribute to parameter inference.

The simplicity of TCM parameter variations with collision energy is demonstrated by Figs. 12 (right), 13 (left), 16 (left) and 22. Except for Lévy exponent \( n \) the TCM parameters vary approximately linearly with \( \log(s/s_0) \).

In Fig. 12 (right) the \( 1/n \) uncertainties are typically the size of the solid points (17.2 GeV 5%, 200 GeV 2%, 13 TeV 3%). In Fig. 13 (left) \( 1/q \) uncertainties are again the size of the points (about 2%). Hard-component mode \( \hat{y}_t \) is approximately constant at \( \hat{y}_t \approx 2.6 \) with at most 4\% variation across all energies. One should distinguish between 200 GeV fixed-\( H_0(p_t) \) parameters consistent with the \( n = 3 \) multiplicity class (see Fig. 4 – left) and the 200 GeV NSD values in Table II. The fixed reference for 13 TeV is based on the table values for that energy. The values in Table II apply only to spectrum data below 10 GeV/c. For any collision energy the effective power-law parameter \( q \) is expected to increase at higher \( p_t \) consistent with the energy trend of underlying jet spectra [5].

The TCM can also be evaluated as a predictive model by comparison with data not contributing to its definition, as in Fig. 13 where a 0.9 TeV spectrum and spectrum ratios are addressed. All TCM parameters for that energy are the result of TCM predictions with no adjustment. The data descriptions in Fig. 13 are of similar quality to those for 200 GeV and 13 TeV albeit with reduced statistics.

Figure 20 provides more information on spectrum \( n_{ch} \) dependence at 0.9 TeV in the form of power-law Eq. (B1) fit parameters for spectrum data [open points in panels (c) and (d)]. The solid points are obtained from power-law fits to the 0.9 TeV TCM as defined by Table IV. They show good agreement except for the lowest \( n_{ch} \) values. Figure 21 (b) suggests that the power-law model describes spectrum data adequately since the fit residuals are consistent with statistical uncertainties, but the conclusion is misleading. The small number of collision events translates to large statistical uncertainties – the 0.9 TeV data cannot test models effectively.

The 200 GeV study in Ref. 11 was based on 3M events distributed over eleven multiplicity classes within \( \Delta \eta = 1 \) whereas the study in Ref. 2 and the present analysis are based on 6M events distributed over seven multiplicity classes within \( \Delta \eta = 2 \). The two lowest multiplicity classes in Ref. 2 include more than 2M events each as in Table I. The difference relative to Ref. 11 is a nearly 3-fold decrease in r.m.s. statistical errors that reveals new details of hard-component evolution with \( n_{ch} \) (Sec. IV).

In contrast, the 0.9 TeV analysis in Ref [21] is based on less than 0.3M events distributed over more than twenty multiplicity classes instead of seven. The lowest multiplicity classes included only 40K events each (the equivalent for Ref. 10 that introduced the power-law fit model was 10K events). The r.m.s. statistical error is then more than 7-fold larger than for the present study, and fine details of TCM component evolution cannot be resolved.

### VIII. DISCUSSION

Reference 11 mentions “rich features” exhibited by the \( n_{ch} \) dependence of 13 TeV \( p-p \) spectra based on spectrum ratio data represented by Fig. 2 (left) of the present study. There is indeed substantial new information conveyed by the \( n_{ch} \) systematics of \( p_t \) spectra but a TCM context is required to fully access that information.

#### A. Monte-Carlo comparisons with data

Reference 10 includes comparisons of spectrum data with several MC models. Given the increasing availability of high-statistics data, systematic deviations of many tens of statistical error bars may appear between MCs and data suggesting that such models should be rejected. Large deviations persist despite efforts to tune model parameters to accommodate data. Complex MC structure and the many parameters make interpretation problematic. Reference 11 concludes that spectrum data are in “fair agreement” with MCs “but not in all details.”

In contrast, the TCM provides an accurate description of most data to the limits of statistical precision. Separate components are individually comparable with QCD theory (as in Ref. 3), and the few parameters have simple physical interpretations. Systematic variation of model parameters over a broad range of charge multiplicity and collision energy is limited, smooth and controlled by the \( \log(s/s_0) \) trends expected for QCD. TCM analysis of MC output as in Ref. 11 and the present study could allow isolation of MC soft and hard components for independent examination. And MCs offer the possibility to switch off and on specific collision mechanisms to investigate correspondence with TCM observables.

Based on TCM results from the present study MC deviations from data can be interpreted physically. In Fig. 5 of Ref. 10 [Fig. 5 (left) of this study] MCs diverge more quickly than data from the common intercept at unity and then saturate at larger \( p_t \) whereas the data do not. The shape of ratio \( R(p_t) \) is controlled by the hard/soft ratio \( T(p_t) = H_0(p_t)/S_0(p_t) \). Figure 5 of Ref. 10 then suggests that the MC hard/soft ratio rises more quickly than data in that \( p_t \) interval. The corresponding MC \( T_0(p_t) \) in Fig. 10 (left) confirms that the MC hard component is shifted to lower \( p_t \) and is substantially larger...
in amplitude than implied by spectrum data, a result established for PYTHIA in Fig. 9 of Ref. [1].

Saturation of MC spectrum ratios at larger \( p_t \) appears consistent with a fixed hard-component model independent of \( n_{ch} \), whereas spectrum data reveal strong bias of the underlying jet spectrum with increasing event multiplicity. Similar MC-data conflicts have been encountered with angular correlations [12] and ensemble-mean-\( p_t \) data [27] [55].

**B. A universal TCM for \( p-p \) hadron production**

The TCM provides a unifying description of two distinct hadron production mechanisms in high-energy nuclear collisions and is required by a large body of data, of which the present study and Ref. [10] provide examples. The two elements of the TCM, both directly related to the QCD structure of projectile nucleons (i.e. parton distribution functions), represent projectile dissociation and large-angle parton scattering followed by dijet formation near midrapidity. The TCM as invoked in this study applies to yields and spectra near midrapidity. As reported in Ref. [2], accurate description of 2D angular correlations also requires a third (nonjet quadrupole) component.

The few TCM parameters have simple and interconnected energy and multiplicity dependences permitting accurate parameter prediction by interpolation and extrapolation, as in Figs. 4 (left), 6 (left), 12 (right), 13 (left) and 16 (left) as summarized in Table [11] Energy dependence relates_t \( \log(s/s_0) \) with \( \sqrt{s_0} = O(10 \text{ GeV}) \) representing a kinematic limit on dijet production from low-\( x \) gluons. Hard and soft components are directly connected by the empirical relation \( \rho_s \approx \alpha(\sqrt{s})\rho_t^2 \), and there is quantitative correspondence between spectrum hard components and eventwise-reconstructed jets [3] [4].

The TCM parameters and their dependences are \( \alpha(\sqrt{s}), \rho_s(\sqrt{s}), [T, n(\sqrt{s})], \{y_t, \sigma_{y_t}(n'_{ch}, \sqrt{s}), q(n'_{ch}, \sqrt{s})\} \), where the square brackets enclose soft and hard model-function parameters respectively and \( T \) does not vary significantly over a broad range of collision systems. Lévy exponent \( \nu \) describes the “power-law” tail of soft component \( S_0(p_t) \) and in the form \( 1/n(\sqrt{s}) \) increases with energy as \( \log(s/s_0) \). The newly-observed soft-component energy dependence may provide insight on the internal structure of projectile nucleons prior to collision. Two hard-component parameters in the form \( 1/q(\sqrt{s}) \) and \( \sigma_{y_t}(\sqrt{s}) \) increase linearly with \( \log(s/s_0) \) as in Figs. 13 (left) and 16 (left), just as expected from the QCD \( \sqrt{s} \) systematics of an underlying jet energy spectrum and buttressing a jet interpretation for the hard component. The hard-component mode \( y_t \approx 2.6 (\bar{p}_t \approx 1 \text{ GeV/c}) \) is approximately independent of collision conditions and may reflect a universal “infrared cutoff” of jet energy spectra near 3 GeV [3].

Another observation newly derived from high-statistics 200 GeV \( p-p \) data is multiplicity bias of the hard-component shape. Hard-component parameters \( q(n_{ch}) \) and \( \sigma_{y_t}(n'_{ch}) \) vary as in Fig. 4 (left) to determine hard-component evolution with \( n_{ch} \) above the hard-component mode. 13 TeV spectrum ratios suggest similar variations although a more limited \( n_{ch}' \) interval is spanned. Below the hard component mode \( \sigma_{y_t}(n'_{ch}) \) varies as in Fig. 6 (left) in a manner (anti)correlated with \( \sigma_{y_t}(n'_{ch}) \) above the mode. There is apparently no corresponding mechanism incorporated in current MCs, as in Fig. 8 (left) for instance.

The TCM hard component has been extended in the present study to the largest available collision-energy range, from ISR to LHC energies. The extent of correlation of the inferred spectrum hard component with measured properties of eventwise-reconstructed jets over the same interval can be assessed by comparing Fig. 15 (right) of this study with Figs. 5 and 9 of Ref. [5] which presents a universal curve describing measured jet (scattered parton) energy spectra for \( p-p \) collision over the same energy interval. Complementary to that quantitative correspondence the TCM soft component, representing the majority of produced hadrons in any system, appears to be essentially independent of the collision system. That property seems to conflict with models based on significant particle rescattering and thermalization that depend on \( A-A \) centrality and collision energy.

**IX. SUMMARY**

A two-component (soft + hard) model (TCM) of hadron production near midrapidity from high-energy \( p-p \) collisions was derived previously from the charge-multiplicity \( n_{ch} \) dependence of \( p_t \) spectra from 200 GeV \( p-p \) collisions at the relativistic heavy ion collider (RHIC). Based on comparisons with QCD theory and other forms of data the two components have been interpreted to represent longitudinal projectile-nucleon dissociation and minimum-bias (MB) transverse dijet production.

The hard component of the TCM (a peaked distribution on transverse rapidity \( y_t \) with exponential tail) was previously assumed to be approximately independent of \( n_{ch} \). Reanalysis of high-statistics 200 GeV spectrum data in the present study reveals that the entire hard-component shape varies significantly and smoothly with \( n_{ch} \). The parameter variations are described by simple functions. The hard-component model with mode near \( p_t = 1 \text{ GeV/c} \) is observed to broaden above the mode and narrow below the mode with increasing \( n_{ch} \), suggesting that imposition of a condition on event multiplicity biases the underlying jet spectrum: requiring larger \( n_{ch} \) biases to a harder jet spectrum with greater fragment yield.

Spectrum \( n_{ch} \) dependence at large hadron collider (LHC) energies was presented recently in the form of spectrum ratios that retain only partial information about spectrum structure. To extrapolate the TCM to LHC energies new analysis techniques are required. The LHC ratio method is applied to high-statistics 200 GeV spectrum data for which spectrum structure is well un-
understood. A method is devised to derive TCM soft and hard components individually from spectrum ratios. The same methods are then applied to LHC spectrum ratios to derive separate soft and hard components for those spectra. The present analysis reveals similar $n_{ch}$ evolution of the hard-component shape at 13 GeV. Lower-statistics data at 0.9 TeV are consistent with that trend. The $n_{ch}$ dependence of popular Monte Carlo models presented in the same ratio format is consistent with a fixed hard component and differs substantially from data (a larger jet-fragment contribution peaking at lower $p_t$).

Given TCM results at 200 GeV and 13 TeV the general energy dependence of soft and hard TCM spectrum components is determined accurately from 17 GeV to 13 TeV, an energy interval where low-$x$ gluons play a major role in hadron production at midrapidity. The newly-observed energy evolution of the soft component, modeled by a Lévy distribution on transverse mass $m_T$, suggests that the growth of soft-component $p_T$ fluctuations (measured by Lévy exponent $n$ in the form $1/n$) may result from Gribov diffusion. The energy dependence of hard-component parameters is controlled by QCD parameter $\log(s/s_0)$ as expected for jet-related structure. Energy trends observed at the RHIC extrapolate to LHC energies in a manner consistent with trends for event-wise-reconstructed jet spectra.

Although TCM components can be recovered from spectrum ratios using the techniques developed in this study, ratio methods are problematic for several reasons: (a) Spectrum ratios alone can only define a hard/soft ratio of TCM components: At least one complete spectrum must be introduced to achieve separation. (b) Without a TCM context spectrum ratios cannot be interpreted physically because they confuse at least two hadron production mechanisms. (c) Spectrum ratios tend to exaggerate structure at higher $p_T$ and suppress important structure at lower $p_T$ where most jet fragments appear. (d) In conventional data/model ratio comparisons significant model discrepancies at lower $p_T$ are suppressed. Direct comparison of data-model differences to statistical errors over the largest possible $p_T$ interval is imperative.

Determination of an accurate TCM for isolated spectra (rather than ratios) over a broad range of event multiplicities and collision energies as in the present study establishes an accurate and efficient representation of a large volume of spectrum data. The TCM provides explanatory power, quantitative links to data manifestations from a variety of collision systems and direct links to QCD theory not provided by any other data model.

Appendix A: 200 GeV $p_T$ spectra vs $n_{ch}$

To establish a context for spectrum-ratio results from Ref. [10] at LHC energies the same ratio method is applied to 200 GeV $p_T$ spectrum data that have a previously-established TCM and physical interpretations [11-12]. The hard-component model $\hat{H}_0(p_T)$ is assumed to be fixed for this treatment as in the cited references. In Sec. IV that constraint is relaxed and a revised $\hat{H}_0(p_T; n'_{ch})$ varying with uncorrected $n'_{ch}$ to accommodate data is inferred.

1. Spectrum ratios vs $n_{ch}$ and model ratio $T(p_T)$

The 13 TeV spectrum-ratio data from Ref. [10] in Fig. 2 can be better interpreted by applying a similar ratio analysis to the 200 GeV data from Ref. [1] in Fig. 1 where the spectrum structure is well understood. The spectrum-ratio data may be used to estimate TCM model ratio

$$T_0(p_T) = \frac{\hat{H}_0(p_T)}{S_0(p_T)}$$

assuming fixed hard-component model $\hat{H}_0(p_T)$, but the individual model components are not accessible from spectrum ratios alone. The intermediate quantities $X(p_T)$ and $Y(p_T)$ below facilitate inference of TCM data ratio $T(p_T)$ from measured spectrum ratios.

Because 200 GeV spectra are normalized by corrected soft-component density $\hat{\rho}_s$ as in Fig. 1 (left) and Refs. [1-2] the relevant spectrum data ratio is the first line of

$$X(p_T; n'_{ch1}, n'_{ch2}) = \left( \frac{\hat{\rho}_{s2}}{\hat{\rho}_{s1}} \right) \frac{\hat{\rho}_0'(y_T; n'_{ch1})}{\hat{\rho}_0'(y_T; n'_{ch2})}$$

$$X_0(p_T; n'_{ch1}, n'_{ch2}) = \left( \frac{\hat{\rho}_{s2}}{\hat{\rho}_{s1}} \right) \frac{\hat{\rho}_{s1}S_0(p_T) + \hat{\rho}_{h1}\hat{H}_0(p_T)}{\hat{\rho}_{s2}S_0(p_T) + \hat{\rho}_{h2}\hat{H}_0(p_T)}$$

$$= \frac{1 + \alpha \hat{\rho}_{s1}T_0(p_T)}{1 + \alpha \hat{\rho}_{s2}T_0(p_T)} \rightarrow \frac{\hat{\rho}_{s1}}{\hat{\rho}_{s2}} \quad \text{for larger } p_T.$$

The corresponding TCM for spectrum ratios based on a fixed hard-component model is the second line of Eq. [A2] assuming inefficiencies cancel in data ratios.

Figure 17 (left) shows spectrum ratios for $n_1 = 2$ over $n_2 = 6$ (open circles) and its inverse (solid points). Both cases are included to illustrate that apparently very different trends plotted in this way may be equivalent. Results for other combinations are similar. The soft-component densities are $\hat{\rho}_s = 3.38$ and 12.6 respectively from Table I (NSD is $\hat{\rho}_s \approx 2.5$). The TCM in Eq. [A2] (third line) is represented by the solid and dashed curves. The density ratios (fourth line) are represented by upper and lower dotted lines as limiting cases of TCM $X_0(p_T; n_1, n_2)$. The discrepancy between ratio data and the TCM at larger $p_T$ is addressed in Sec. IV A.

Figure 17 (right) shows transformed data (points) as the first line of

$$Y(p_T; n'_{ch1}, n'_{ch2}) = \frac{1}{\hat{\rho}_{s1}/\hat{\rho}_{s2} - 1} [X(p_T; n'_{ch1}, n'_{ch2}) - 1]$$

$$Y_0(p_T; n'_{ch1}, n'_{ch2}) = \frac{\alpha \hat{\rho}_{s2}T_0(p_T)}{1 + \alpha \hat{\rho}_{s2}T_0(p_T)} = \frac{H_2(p_T)}{S_2(p_T) + H_2(p_T)}.$$
The curves $Y_0$ going asymptotically to 1 represent a fixed $H_0(y_t)$ model and intercept 0.5 (dotted line) where $\alpha \rho T_0(p_t) = 1 \times (x = 1$ or 2), shifting left or right depending on multiplicity class $n$ in the denominator of $X(p_t)$. Data deviations relative to the TCM at higher $p_t$ indicate that a fixed hard-component model does not describe ratio data adequately, but with minor adjustment the TCM becomes an accurate representation as in Sec. IV.

Figure 18 (left) shows data (points) in the form

$$\alpha \rho T(p_t) = \frac{Y(p_t)}{1 - Y(p_t)} \quad (A4)$$

scaled to 200 GeV NSD $p$-$p$ collisions ($\rho_s \approx 2.5$). Model $\alpha \rho T_0(p_t)$ (curves) intercepts unity near $p_t = 3$ GeV/c ($y_t \approx 3.75$) corresponding to the hard-soft crossover in Fig. 1 (left). Data with $Y(p_t) > 1$ are undefined. The open and solid points coincide in this format. The superposed solid and dashed curves represent a common TCM $T_0(p_t)$ with fixed hard component corresponding to the bold dashed curves in Fig. 1 to the dash-dotted and dotted lines in Fig. 4 (left) and approximately to convolution of a measured 200 GeV MB jet spectrum with measured $p$-$p$ fragmentation functions [3]. As noted, the strong discrepancies between data and TCM ratios indicate that the assumption of a fixed hard-component model should be revisited. However, ratio data alone do not provide sufficient information to reformulate the TCM.

Figure 18 (right) shows ratios of spectrum data from Fig. 1 (left) to the corresponding TCM expression in Eq. (11) with fixed hard component. The hatched band indicates $\pm 5\%$ deviations. Above $y_t = 3$ significant systematic variation (10% increase per multiplicity class at 4 GeV/c) suggests the requirement for a decreasing power-law exponent $q$ with increasing $n_{ch}$, as might be expected if demand for larger event multiplicities biases the jet spectrum. Although the ratio deviations from unity at lower $y_t$ are smaller in absolute magnitude they are statistically more significant as discussed in Sec. VII B. Hard-component $n_{ch}$ dependence (model parameter variations) for 200 GeV $p$-$p$ collisions is established in Sec. IV via direct comparisons between isolated spectra and the TCM.

Appendix B: 0.9 TeV spectrum TCM

The TCM energy parametrization summarized in Table I and multiplicity dependences determined in Sec. IV can be tested by comparison with additional data from an intermediate energy. Reference 21 presents a $p_t$ spectrum from 0.3M 0.9 TeV INEL $p$-$p$ collisions (Fig. 3) and corresponding multiplicity-dependent ratios (Fig. 6, lower), the ratios as in Fig. 8 (left) of the present study.

1. Predicting a 0.9 TeV spectrum TCM

Figure 19 (left) shows spectrum data (points) from Fig. 3 of Ref. 21. The solid curve is the corresponding TCM defined by the parameters in Table I with no adjustments. Also shown are predicted soft $S(p_t)$ (dotted) and hard $H(p_t)$ (dashed) TCM spectrum components for 0.9 TeV NSD collisions and the hard component for 200 GeV NSD collisions (dash-dotted) for comparison. The plot format is the same as for Fig. 11 (left).

Figure 19 (right) shows spectrum ratio $R(p_t; n_{ch})$ data (points) from Fig. 6 (lower panel) of Ref. 21. The reported bin-mean $n_{ch}$ ($n_{acc}$) values for three multiplicity bins are 3, 7 and 17. The data for central bin B convey no significant information (the spectra in ratio are too similar) and are omitted. The common reference for three ratios is the INEL spectrum. Based on the bin-B ratio data
in Ref. [21] $n_{acc} \approx 6$ is the effective INEL value. Given the asymptotic ratio limits at larger $p_t$ (dotted lines) the effective bin means should then be $n_{acc} \approx 0.35 \times 6 = 2.1$ (A), $\approx 1.15 \times 6 \approx 7$ (B) and $\approx 2 \times 6 = 12$ (C).

The dashed curves (approaching the dotted lines) represent Eq. (6) with TCM parameters for 0.9 TeV and fixed $H_0(p_t)$ from Table II unaltered. The data description is generally good with deviations from the TCM at larger $p_t$ similar to those noted in Figs. 17 (left) and 8 (left). The limited event number precludes detailed analysis of hard-component variation with $n_{ch}$ via spectrum ratios. The solid curves result from $q = 4.7$ and 4.3 and $\sigma_n = 0.525$ and 0.535 for bins A and C respectively. The variations are consistent with interpolation of the parameter trends in Fig. 4 (left) between 200 GeV and 13 TeV.

This comparison confirms that the overall spectrum TCM described in the present study serves as an accurate representation of $p_t$ spectrum data over large $p_t$, energy and multiplicity ranges, with isolation of two components representing distinct hadron production mechanisms that may be compared independently with relevant theory.

2. Model fits with a power-law spectrum model

Table 2 of Ref. [21] does provide significant information on spectrum multiplicity dependence in the form of model fits to spectra for various multiplicity classes. The model function [Eq. (1) of Ref. [21]] is described as a “modified Hagedorn function” (i.e. power-law model)

$$ \rho(p; n'_{ch}) = \frac{p_t}{m_t} \frac{A(p_0, b)}{(1 + p_t/p_0)b}. $$

(B1)

Factor $p_t/m_t$ makes no difference to the fits described below and is omitted for simplicity. Parameter $p_0$ can be expressed as $p_0 = bT$ with $b$ ($\rightarrow n$) and $T$ obtained independently for comparison with Eq. (2). An indirect comparison can be made between 0.9 TeV data and TCM predictions via those fit results: The same fit model is applied to the 0.9 TeV TCM defined by Table II and the parameter trends are compared with those from Ref. [21].

Note that Eq. (B1) defined on $p_t$ is significantly different at lower $p_t$ from Lévy distribution Eq. (2) defined on $m_t - m_h$, and neither function can in isolation meaningfully represent the $n_{ch}$ dependence of an intact $p_t$ spectrum because they cannot describe the spectrum hard component as predicted by measured jet properties [3, 5, 9]. Also note that Fig. 20 compares Eq. (B1) to the full spectrum TCM, not to Eq. (2) alone.

Figure 20 (a) shows TCM spectra with fixed hard component based on 0.9 TeV NSD parameters from Table II (solid) for 20 multiplicity bins as defined in Ref. [21]. For each corrected $n_{ch}$ value the corresponding $p_t$ was obtained using $\alpha = 0.01$ from Fig. 16 (left). The power-law model Eq. (B1) with parameters $(A, T, b)$ (dashed) was fitted to the TCM spectra.

Figure 20 (b) shows fit residuals as spectrum differences relative to $p_t$-bin statistical errors [1] in the form

$$ \frac{\Delta \rho_0}{\sqrt{\rho_{0,ref}}} \to \frac{\Delta n'_{ch}(p_t)}{n'_{ch}(p_t)} $$

(B2)

assuming Poisson bin errors, and $n'_{ch}(p_t)$ is the accepted charge multiplicity in a $p_t$ bin summed over the event ensemble. Equation (B2) measures bin-wise information conveyed by a data spectrum relative to a reference (fit model, Monte Carlo). Panel (b) shows fit residuals in the form relevant to $\chi^2$ minimization. The power-law model seems to describe the 0.9 TeV TCM adequately (no excess beyond statistics shown by the hatched band) but the event number is only 0.3M. For the spectrum data in Ref. [2] the hatched band would be reduced 7-fold and the same residuals then become very significant (see Secs. VIII B and VIII E). The apparently-oscillating residuals from fits with the power-law model function have been described as “log-periodic oscillations” in Ref. [22].

Figure 20 (c) and (d) show fit results from Table 2 of Ref. [21] for two model parameters (open circles) representing $b$ and $T \equiv p_0/b$. The solid points represent fits of Eq. (B1) applied to the TCM as described above. In panel (c) the upper hatched band represents fixed $\hat{S}_0(m_c)$ Lévy exponent $n \approx 9.8$ and the lower hatched band represents fixed $H_0(y_k)$ parameter $q \approx 4.5$ (with $q + 2 = 6.5$ as the relevant exponent for $p_t$). As event multiplicity increases from left to right Eq. (B1) attempts to accommodate in turn first the soft component alone then the hard component. In panel (d) the hatched band represents fixed $\hat{S}_0(m_c)$ parameter $T = 145$ MeV. For larger multiplicities the fitted parameter drops away from the
multiplicity classes (odd-number classes are plotted) as defined in Ref. [21]. (b) Fitted slope parameter $T = p_{t0}/b$ from fits to the TCM (solid) vs those from fits to data (open).

TCM value as the power-law model attempts to accommodate the increasing hard-component amplitude. The good agreement between most open and solid points indicates that the TCM is a satisfactory representation of 0.9 TeV spectrum data for all but the smallest multiplicities.

The data from Ref. [21] (open circles) rise substantially above fits to the TCM with fixed hard component (solid points) for the lowest multiplicity classes. Such differences should arise if the TCM does not model low-$p_t$ excursions of the hard component for smaller event multiplicities, as in Fig. 1 (right). That similar excursions may appear over a range of collision energies is indicated by Fig. 4 (left) of Ref. [27] where ensemble-mean-$p_t$ hard-component trends for all collision energies are reduced for the lowest multiplicity classes.

Appendix C: p-p Charge Multiplicities

In this appendix negative-binomial-distribution (NBD) models of multiplicity distributions (MDs) for several energies are summarized and energy dependence of charge yields relevant to the TCM is inferred. This material is presented to provide context for the analysis of LHC spectrum ratios in Secs. V and VI.

The LHC spectrum study in Ref. [10] refers to imposed charge-multiplicity conditions not fully specified. To better provide a comparison with RHIC data information on LHC charge multiplicities from Ref. [34] can be used. In particular there are issues of consistency between direct $dN_{ch}/d\eta$ density measurements and multiplicity-distribution (MD) mean values $\bar{n}_{ch}$ that should coincide. Charge densities on $\eta$ are reported in Table 7 of Ref. [34] for three event classes and several energies. In this study results for NSD events and $|\eta| < 0.5$ are emphasized as a common reference. Results for other conditions are scaled accordingly. Table 12 of Ref. [34] includes mean values $\bar{n}_{ch}$ from parametrizations of data MDs that may be contrasted with the direct density measurements.

1. Multiplicity distributions and binnings

Probability distributions on p-p event multiplicity can be described by a single negative binomial distribution (NBD) at lower collision energies, but at higher energies a double NBD is required by data [33]. The relevant definitions are provided by Eqs. (14) and (16) from Ref. [34]

$$P(n_{ch}; \bar{n}, k) = \frac{\Gamma(n_{ch} + k)}{\Gamma(k)\Gamma(n_{ch} + 1)} \left( \frac{\bar{n}}{\bar{n} + k} \right)^{n_{ch}} \left( \frac{k}{\bar{n} + k} \right)^k$$

The NBD parameters for four energies are taken from Table 10 of Ref. [34] for $|\eta| < 0.5$ ($\Delta\eta = 1$) and reproduced in Table III below. The parameter values for 13 TeV are extrapolated from the lower energies.

| Energy (TeV) | $\lambda$ | $\alpha$ | $n_1$ | $k_1$ | $n_2$ | $k_2$ | $\rho_{00}$ | $\rho_{20}$ |
|-------------|----------|----------|--------|-------|--------|-------|------------|------------|
| 0.9         | 0.94     | 0.55     | 2.4    | 2.6   | 6.0    | 3.3   | 3.8        | 3.61       |
| 2.76        | 0.93     | 0.51     | 2.5    | 2.6   | 8.0    | 3.1   | 4.8        | 4.55       |
| 4           | 0.94     | 0.70     | 3.6    | 1.8   | 12     | 4.1   | 5.6        | 5.35       |
| 8           | 0.93     | 0.57     | 3.1    | 2.0   | 11     | 3.2   | 5.8        | 5.46       |
| 13          | 0.935    | 0.58     | 3.3    | 2.0   | 12     | 3.5   | 6.2        | 5.87       |

The double-NBD parameters for NSD p-p collisions at several energies and $|\eta| < 0.5$, from Table 9 of Ref. [34]. The 13 TeV (starred) entries are extrapolated from lower-energy data. The $\rho_{00}$ entries are the actual means of the NBD models. The $\rho_{20}$ entries are taken from Fig. 22 (dotted).

Figure 21 (left) shows double NBDs from Ref. [33] fitted to NSD data MDs on $n_{ch}$. The fit residuals are at the percent level. Up to $n_{ch}/\Delta\eta$ there is no significant difference between single- and double-NBD fits. The double-NBD mean values are denoted by $\bar{\rho}_{00}$ in Table III. Entries for $\bar{\rho}_{20}$ are inferred from the dotted curve in Fig. 22.

Figure 21 (right) shows the same distributions on a linear scale with the multiplicity bin system defined in Ref. [10]. Referring to the 13 TeV distribution (bold solid
curve) the left solid vertical line marks the 13 TeV distribution mean $\bar{\rho}_{00} = 6.2$. The right solid line is at twice that value, together defining three multiplicity bins A, B and C in Ref. [10]. Those definitions are based on the accepted multiplicity $N_{acc}^\text{ch}$ within $\Delta \eta = 1.6$, but a single common efficiency factor should apply to all values. The vertical dashed lines indicate three bin means located approximately at 3, 9 and 15 compared to ensemble mean $\bar{\rho}_{00} = 6.2$. The ratios to $\bar{\rho}_{00}$ are then approximately 3/6.2 = 0.48 for A, 9/6.2 = 1.45 for B and 15/6.2 = 2.4 for C.

2. Event-multiplicity energy dependence

The energy dependence of $p$-$p$ event multiplicities at collision energies up to 13 TeV can be inferred from several sources, including INEL > 0 data from Ref. [10] and data for several trigger conditions from Ref. [34].

Figure 22 shows data (open squares) for five energies from Table 7 of Ref. [34] scaled to NSD (yields in Table 7 scale as INEL:NSD:inelastic = 0.81:1.00:1.04). The solid square is 13 TeV INEL > 0 datum 6.46 from Ref. [10] scaled down by 1.04 to NSD value 6.21. The solid dots are weighted means of NBD distributions in Fig. 21 from fits to measured MDs in Ref. [22] reported as $\bar{\rho}_{00}$ in Table 11. The solid triangle is an estimate of the NSD value for 200 GeV, and the open circle is an extrapolation to 17.2 GeV. Curves are described in the text.

3. Accepted vs corrected multiplicities and ratios

In App. A1 accurate knowledge of the $\bar{\rho}$ values for different event classes was essential to process spectrum...
ratios and interpret the results. The corresponding information in Ref. [10] seems incomplete. The relation to accepted multiplicities $N_{ch}^{acc}$ should be as follows. The acceptance factor corresponding to a $p_t$ acceptance cut-off near 0.15 GeV/c is $0.80 \pm 0.02$ according to Fig. 2 (right) of Ref. [27]. That factor should change little with collision energy since it is determined mainly by fixed slope parameter $T \approx 145$ MeV. The factor corresponding to mean tracking efficiency is 0.57. Collision-energy dependence should again be small. The overall acceptance factor should then be $0.56 \pm 0.04 \approx 0.6$.

In Ref. [10] the reported fully-corrected charge density for the INEL > 0 event class is $dn_{ch}/d\eta = 6.46$ within $|\eta| < 0.5$ and 6.61 within $|\eta| < 1.0$. For the spectrum study the accepted charge density within $|\eta| < 0.8$ is $n'_{ch}/\Delta\eta = 6.73/1.6 = 4.2$ while the density (possibly) corrected for tracking efficiency but within $p_t > 0.15$ GeV/c is $n''_{ch}/\Delta\eta = 9.41/1.6 = 5.88$ (“from the spectrum in Fig. 3”). Interpolated to $|\eta| < 0.8$ the fully-corrected charge density should be $dn_{ch}/d\eta \approx 6.55$, and the implied overall acceptance factor is then 4.2/6.55 = 0.64, somewhat higher than the expected 0.57. The ratio 5.88/6.55 = 0.90 is also substantially higher than the expected $p_t$-acceptance factor 0.80. However, 4.2/5.88 = 0.71 is consistent with the expected tracking efficiency 0.70. Although not self-consistent the acceptance factors should cancel in the multiplicity ratios 0.48 (A), 1.45 (B) and 2.4 (C) calculated above. However, there seem to be substantial differences between those values and what can be inferred directly from spectrum ratios in Sec. [V].

[1] J. Adams et al. (STAR Collaboration), Phys. Rev. D 74, 032006 (2006).
[2] T. A. Trainor and D. J. Prindle, Phys. Rev. D 93, 014031 (2016).
[3] T. A. Trainor, Phys. Rev. C 80, 044901 (2009).
[4] T. A. Trainor and D. T. Kettler, Phys. Rev. C 83, 034903 (2011).
[5] T. A. Trainor, Phys. Rev. D 89, 094011 (2014).
[6] M. Gyulassy and L. McLerran, Nucl. Phys. A 750, 30 (2005).
[7] E. Aven, Y. Hatta, C. Fnelsburg, J. Y. Ollitrault and T. Ueda, J. Phys. G 38, 124053 (2011).
[8] K. Jiang et al., Phys. Rev. C 91, 024910 (2015).
[9] T. A. Trainor, Int. J. Mod. Phys. E 17, 1499 (2008).
[10] J. Adam et al. (ALICE Collaboration), Phys. Lett. B 753, 319 (2016).
[11] T. A. Trainor, Phys. Rev. D 87, 054005 (2013).
[12] G. Agakishiev et al. (STAR Collaboration), Phys. Rev. C 86, 064902 (2012).
[13] R. J. Porter and T. A. Trainor (STAR Collaboration), J. Phys. Conf. Ser. 27, 98 (2005).
[14] R. J. Porter and T. A. Trainor (STAR Collaboration), PoS C FRNC2006, 004 (2006).
[15] G. Wilk and Z. Wlofczyk, Phys. Rev. Lett. 84, 2770 (2000).
[16] C. Albajar et al. (UA1 Collaboration), Nucl. Phys. B 309, 405 (1988).
[17] G. J. Alner et al. (UA5 Collaboration), Z. Phys. C 32, 153 (1986).
[18] D. Acosta et al. (CDF Collaboration), Phys. Rev. D 68, 012005 (2003).
[19] T. A. Trainor, J. Phys. G 37, 085004 (2010).
[20] T. Sjostrand and M. van Zijl, Phys. Rev. D 36, 2019 (1987).
[21] K. Aamodt et al. (ALICE Collaboration), Phys. Lett. B 693, 53 (2010).
[22] G. Wilk and Z. Wlofczyk, Entropy 17, 384 (2015).
[23] C. Tsallis, J. Stat. Phys. 52, 479 (1988).
[24] C. Y. Wong, G. Wilk, L. J. L. Cirto and C. Tsallis, Phys. Rev. D 91, 114027 (2015).
[25] C. Alt et al. (NA49 Collaboration), Eur. Phys. J. C 45, 343 (2006).
[26] Y. L. Dokshitzer and D. E. Kharzeev, Ann. Rev. Nucl. Part. Sci. 54, 487 (2004).
[27] T. A. Trainor, Phys. Rev. C 90, 024909 (2014).
[28] H. Appelshauser et al. (NA49 Collaboration), Phys. Rev. D 93, 024915 (2016).
[29] G. Agakishiev et al. (CERES/NA45 Collaboration), Phys. Rev. Lett. 92, 032301 (2004).
[30] T. A. Trainor and D. T. Kettler, Phys. Rev. D 74, 034012 (2006).
[31] B. Abelev et al. (ALICE Collaboration), Eur. Phys. J. C 73, no. 12, 2662 (2013).
[32] B. S. Everitt and A. Skrondal, “The Cambridge Dictionary of Statistics,” 4th Ed., (Cambridge University Press, Cambridge, 2010), p. 107.
[33] J. Adank et al. (ALICE Collaboration), arXiv:1509.07541.