Measurement of charged particle multiplicities and densities in pp collisions at $s = 7$ TeV in the forward region

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Measurement of charged particle multiplicities and densities in \( pp \) collisions at \( \sqrt{s} = 7 \) TeV in the forward region

The LHCb Collaboration
CERN, 1211 Geneva 23, Switzerland

Abstract Charged particle multiplicities are studied in proton–proton collisions in the forward region at a centre-of-mass energy of \( \sqrt{s} = 7 \) TeV with data collected by the LHCb detector. The forward spectrometer allows access to a kinematic range of \( 2.0 < \eta < 4.8 \) in pseudorapidity, momenta greater than 2 GeV/c and transverse momenta greater than 0.2 GeV/c. The measurements are performed using events with at least one charged particle in the kinematic acceptance. The results are presented as functions of pseudorapidity and transverse momentum and are compared to predictions from several Monte Carlo event generators.

1 Introduction

The phenomenology of soft quantum chromodynamic (QCD) processes such as light particle production in proton–proton (\( pp \)) collisions cannot be predicted using perturbative calculations, but can be described by models implemented in Monte Carlo event generators. The calculation of the fragmentation and hadronization processes as well as the modelling of the final states arising from the soft component of a collision (underlying event) are treated differently in the various event generators. The phenomenological models contain parameters that need to be tuned depending on the collision energy and colliding particles species. This is typically achieved using soft QCD measurements. The LHCb collaboration reported measurements on energy flow [1], production cross-sections [2,3] and production ratios of various particle species [4] in the forward region, all of which provide information for event generator optimization.

A fundamental input used for the tuning process is the measurement of prompt charged particle multiplicities. In combination with the study of the corresponding momentum spectra and angular distributions, these measurements can be used to gain a better understanding of hadron collisions. An accurate description of the underlying event is vital for understanding backgrounds in beyond the Standard Model searches or precision measurements of the Standard Model parameters. Previous measurements of charged particle multiplicities performed with \( pp \) collisions at the Large Hadron Collider (LHC) were reported by the ATLAS [5,6], CMS [7] and ALICE [8,9] collaborations. All of these measurements were performed in the central pseudorapidity region. The forward region was studied with the LHCb detector, where an inclusive multiplicity measurement without momentum information was performed [10].

In this paper, \( pp \) interactions at a centre-of-mass energy of \( \sqrt{s} = 7 \) TeV that produce at least one prompt charged particle in the pseudorapidity range of \( 2.0 < \eta < 4.8 \), with a momentum of \( p > 2 \) GeV/c and transverse momentum of \( p_T > 0.2 \) GeV/c, are studied. A prompt particle is defined as a particle that either originates directly from the primary vertex or from a decay chain in which the sum of mean lifetimes does not exceed 10 ps. As a consequence, decay products of beauty and charm hadrons are treated as prompt particles. The information from the full tracking system of the LHCb detector is used, which permits the measurement of the momentum dependence of charged particle multiplicities. Multiplicity distributions, \( P(n) \), for prompt charged particles are reported for the total accessible phase space region as well as for \( \eta \) and \( p_T \) ranges. In addition, mean particle densities are presented as functions of transverse momentum, \( dn/dp_T \), and of pseudorapidity, \( dn/d\eta \). The paper is organised as follows. In Sect. 2 a brief description of the LHCb detector and an overview of track reconstruction algorithms are provided. The recorded data set and Monte Carlo simulations are described in Sect. 3, followed by a discussion of the definition of visible event and the data selection in Sect. 4. The analysis method is described in Sect. 5, and systematic uncertainties are given in Sect. 6. The final results are compared to event generator predictions in Sects. 7 and 8, before summarising in Sect. 9.
2 LHCb detector and track reconstruction

The LHCb detector [11] is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed for the study of particles containing $b$ or $c$ quarks. The detector includes a high-precision tracking system consisting of a silicon-strip vertex detector (VELO) surrounding the interaction region, a large-area silicon-strip detector located upstream of a dipole magnet with a bending power of about 4 Tm, and three stations of silicon-strip detectors and straw drift tubes placed downstream. The combined tracking system provides a momentum measurement with relative uncertainty that varies from 0.4% at 2 GeV/$c$ to 0.6% at 100 GeV/$c$, and impact parameter resolution of 20 $\mu$m for tracks with large transverse momentum. The direction of the magnetic field of the spectrometer dipole magnet is reversed regularly. Different types of charged hadrons are distinguished by information from two ring-imaging Cherenkov detectors. Photon, electron and hadron candidates are identified by a calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic calorimeter and a hadronic calorimeter. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers. The trigger consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which applies full event reconstruction.

The reconstruction algorithms provide different track types depending on the sub-detectors considered. Only two types of tracks are used in this analysis. VELO tracks are only reconstructed in the VELO sub-detector and provide no momentum information. Long tracks are reconstructed by extrapolating VELO tracks through the magnetic dipole field and matching them with hits in the downstream tracking stations, providing momentum information. This is the highest-quality track type and is used for most physics analyses. Requiring charged particles to stay within the geometric acceptance of the LHCb detector after deflection by the magnetic field further restricts the accessible phase space to a minimum momentum of around 2 GeV/$c$. The LHCb detector design minimizes the material of the tracking detectors and allows a high track-reconstruction efficiency even for particles with low momenta. However, the limited number of tracking stations results in the presence of misreconstructed (fake) tracks. A reconstructed track is considered as fake if it does not correspond to the trajectory of a genuine charged particle. The fraction of fake long tracks is non-negligible as the extrapolation of a track through the magnetic field is performed over a distance of several metres, resulting in wrong association between VELO tracks and track segments reconstructed downstream. Another source of wrong track assignment arises from duplicate tracks. These track pairs either share a certain number of hits or consist of different track segments originating from a single particle.

3 Data set and simulation

The measurements are performed using a minimum-bias data sample of $pp$ collisions at a centre-of-mass energy of $\sqrt{s} = 7$ TeV collected during 2010. In this low-luminosity running period, the average number of interactions in the detector acceptance per recorded bunch crossing was less than 0.1. The contribution from bunch crossings with more than one collision (pile-up events) is determined to be less than 4% and is considered as a correction in the analysis. The data consists of 3 million events recorded in equal proportion for both magnetic field polarities. The low luminosity and interaction rate of the proton beams allowed the LHCb detector to be operated with a simplified trigger scheme. For the minimum-bias data set of this analysis, the hardware stage of the trigger system accepted all events, which were then reconstructed by the higher-level software trigger. Events with at least one reconstructed track segment in the VELO were selected.

Fully simulated minimum-bias $pp$ collisions are generated using the PYTHIA 6.4 event generator [12] with a specific LHCb configuration [13] using CTEQ6L [14] parton density functions (PDFs). This implementation, called the LHCb tune, contains contributions from elastic and inelastic processes, where the latter also include single and double diffractive components. Decays of hadrons are performed by EVTGEN [15], in which final-state radiation is generated using PHOTOS [16]. The interaction of the generated particles with the detector and its response are implemented using the GEANT4 toolkit [17,18], as described in Ref. [19]. Processing, reconstruction and selection are identical for simulated events and data. The simulation is used to determine correction factors for the detector acceptance and resolution as well as for quantifying background contributions and reconstruction performance.

The measurements are compared to predictions of two classes of generators, those that have not been optimized using LHC data and those that have. The former includes the Perugia 0 and Perugia NOCR [20] tunes of PYTHIA 6, both of which rely on CTEQ5L [21] PDFs, and the PHOJET event generator [22]. PHOJET describes soft-particle production by relying on the dual-parton model [23], which comprises semi-hard processes modelled by parton scattering and soft processes modelled by pomeron exchange. PYTHIA 8 [24] is available in both classes. An early version of PYTHIA 8 is represented by version 8.145. In more recent versions, the default configuration has been changed to Tune 4C, which is based on LHC measurements in the central rapidity region. Both PYTHIA 8 versions utilize the CTEQ5L PDFs. The
results of the latest available version, PYTHIA 8.180, are used to represent Tune 4C. PYTHIA 8.180, together with recent versions of HERWIG++ [25], represent the class of recent event generators. In contrast to the PYTHIA generator, where hadronisation is described by the Lund string fragmentation, the HERWIG++ generator relies on cluster fragmentation and the preconfinement properties of parton showers. Predictions of two versions of HERWIG++ are chosen, each operated in the minimum-bias configuration, which uses the respective default underlying-event tune. For HERWIG++ version 2.6.3, this corresponds to tune UE-EE-4-MRST (UE-4), while version 2.7.0 [26] relies on tune UE-EE-5-MRST (UE-5). Both tunes were also optimized to reproduce LHC measurements in the central rapidity region and rely on the MRST LO++ [27] PDF set.

4 Event definition and data selection

In analogy with similar approaches adopted in previous measurements [6,9], an event is defined as visible if it contains at least one charged particle in the pseudorapidity range of $\eta < 0$ with $p_T > 0.2\text{GeV}/c$ and $p > 2\text{GeV}/c$. These criteria correspond to the typical kinematic requirements for particles traversing the magnetic field and reaching the downstream tracking stations. In order to compare the data directly to predictions from Monte Carlo generators without having a full detector simulation, the visibility definition is based on the actual presence of real charged particles, regardless of whether they are reconstructed as tracks or not.

The tracks are corrected for detector and reconstruction effects to obtain the distribution of charged particles produced in $pp$ collisions. Only tracks traversing the full tracking system are considered. The kinematic criteria are explicitly applied to all tracks to restrict the measurement to a kinematic range in which reconstruction efficiency is high. The track reconstruction requires a minimum number of detector hits and a successful track fit. To retain high reconstruction efficiency, no additional quality requirement for suppressing the contribution from misreconstructed tracks is applied. To ensure that tracks originate from the primary interaction, it is required that the smallest distance of the extrapolated track to the beam line is less than 2 mm. The position of the beam line is determined independently for each data taking period from events with reconstructed primary vertices. Additionally, a track is required to originate from the luminous region; the distance $z_0$ of the track to the centre of this region has to fulfill $z_0 < 3\sigma_L$, where the width $\sigma_L$ is of the order of 40 mm, determined from a Gaussian fit to the longitudinal position of primary vertices. This restriction also suppresses the contamination from beam-gas background interactions to a negligible amount. The distribution of the $z$-position of tracks at the closest point to the beam line shows that in both high-multiplicity and single-track events, beam-gas interactions are distributed over the entire $z$-range of the VELO, whereas the distribution of tracks originating from $pp$ collisions peaks in the luminous region. There is no explicit requirement for a reconstructed primary vertex in this analysis. Together with the chosen definition of a visible event, this allows the measurement to also be performed for events with only single particles in the acceptance.

5 Analysis

The measured particle multiplicity distributions and mean particle densities are corrected in four steps:

1. reconstructed events are corrected on an event-by-event basis by weighting each track according to a purity factor to account for the contamination from reconstruction artefacts and non-prompt particles;
2. the event sample is further corrected for unobserved events that fulfill the visibility criteria but in which no tracks are reconstructed;
3. in order to obtain measurements for single $pp$ collisions, a correction to remove pile-up events is applied;
4. the effects of various sources of inefficiencies, such as track reconstruction, are addressed.

While correction factors for the multiplicity distributions and mean particle densities are the same, their implementation differs and is discussed in the following.

5.1 Correction for reconstruction artefacts and non-prompt particles

The selected track sample includes three significant categories of impurities: approximately 6.5% are fake tracks, less than 1% are duplicate tracks and about 4.5% are tracks from non-prompt particles. The individual contributions are determined using fully simulated events. Henceforth, all impurity categories are collectively referred to as background tracks.

The probability of reconstructing a fake track, $P_{\text{fake}}$, is dependent on the occupancy of the tracking detectors and on the track parameters. The occupancy dependence is determined as a function of the track multiplicity measured by the VELO and as a function of the number of hits in the downstream tracking stations. This accounts for the increasing probability of reconstructing a fake track depending on the number of hits in each of the tracking devices involved. $P_{\text{fake}}$ also depends on $\eta$ and $p_T$; this is taken into account in an overall four-dimensional parametrisation.

Duplicate tracks are reconstruction artefacts, they have only a weak dependence on tracking-detector occupancy but
exhibit a pronounced kinematic dependence. The probability of reconstructing a duplicate track, $P_{\text{dup}}$, is estimated as a function of $\eta$, $p_{\text{T}}$ and VELO track multiplicity.

The probability that a non-prompt particle is selected, $P_{\text{sec}}$, is also estimated as a function of the same variables as for duplicate tracks. The predominant contribution is due to material interaction, such as photon conversion, and depends on the amount of material traversed in the detector. Low $p_{\text{T}}$ particles are more affected.

For each track, a combined impurity probability, $P_{\text{bkg}}$, is calculated, which is the sum of the three contamination types, $P_{\text{bkg}} = P_{\text{fake}} + P_{\text{dup}} + P_{\text{sec}}$, and depends on the kinematic properties of the track, the occupancy of the tracking detectors and the track multiplicity. When measuring the mean particle densities, it is sufficient to assign a per-track weighting factor of $(1 - P_{\text{bkg}})$ to correct for the impurities mentioned above. However, correcting particle multiplicity distributions in the same way would lead to non-physical fractional event multiplicities. To obtain the background-subtracted multiplicity distributions, the procedure described below is applied. The description only corresponds to the full kinematic range, but the procedure is performed in each of the $\eta$ and $p_{\text{T}}$ sub-ranges separately. The impurity probability, $P_{\text{bkg},i}$, of each track, is summed for all tracks in an event to obtain a total event impurity correction, $\mu_{\text{ev}}$. This corresponds to a mean number of expected background tracks in the event and permits to calculate the probability to reconstruct a certain number of background tracks in each event, assuming Poisson statistics. The number of background tracks $k$ in an event with $n_{\text{ev}}$ observed tracks obeys the probability distribution

$$P_{\text{bkg}}(\mu_{\text{ev}}, k) = \frac{\mu_{\text{ev}}^k}{k!} e^{-\mu_{\text{ev}}}, \quad \text{with} \quad \mu_{\text{ev}} = \sum_{i=1}^{n_{\text{ev}}} P_{\text{bkg},i}. \quad (1)$$

From this relation we derive the probability that an event contains a given number of real prompt particles. Summing the normalized probability distribution of all events we obtain the multiplicity distribution corrected for background tracks.

5.2 Correction for undetected events

Defining a visible event based on the properties of the actual charged particles present in the event rather than on the reconstructed tracks introduces a fraction of spuriously undetected events. These are events that should be visible but contain no reconstructed tracks and thus remain undetected. These unobserved events are most likely to occur when few charged particles are within the kinematic acceptance. The reconstruction of a track can fail due to multiple scattering, material interaction, or inefficiencies of the detector or of the reconstruction algorithms. In order to determine the amount of undetected events that nevertheless fulfil the visibility definition, a data-driven approach is adopted.

The true multiplicity distribution for visible events, $T(n)$, where $n$ is the number of charged particles, starts at $n = 1$. Since some of these events have no reconstructed tracks, they follow a multiplicity distribution $U(n)$ starting from $n = 0$. As an event can only be detected if at least one track is reconstructed, $U(0)$ cannot be determined directly. However, the number of undetected events can be estimated from the observed uncorrected distribution $U(n)$, if the average survival probability, $P_{\text{sur}}$, for a single particle in the kinematic acceptance is known. Assuming that the survival probability, which is determined from simulation, is independent for two or more particles, the observed distribution is approximated in terms of the still unknown actual multiplicity distribution $T$,

$$U(k) = \sum_{n \geq k} \binom{n}{k} P_{\text{sur}}^k (1 - P_{\text{sur}})^{n-k} T(n). \quad (2)$$

This equation is only valid under the assumption that reconstruction artefacts, such as fake tracks, which increase the number of observed tracks with respect to the number of true tracks, can be ignored. Following this approach, an event with a certain number of particles is only reconstructed with the same number of tracks or fewer, but not with more tracks. The uncertainties due to these assumptions are evaluated in simulation and are accounted for as systematic uncertainties. Equation 2 allows $U(0)$ to be estimated from the true distribution $T$. All actual elements $T(k)$ can also be expressed using the corresponding uncorrected measured bin $U(k)$ and correction terms of $T(n)$ at higher values of $n > k$,

$$U(0) \approx \sum_{k=1}^{r} (1 - P_{\text{sur}})^k T(k) \quad \text{with} \quad (3)$$

$$T(k) \approx \frac{U(k)}{P_{\text{sur}}^k} - \sum_{n=k+1}^{k+r} \binom{n}{k} (1 - P_{\text{sur}})^{n-k} T(n). \quad (3)$$

Combining the formulas in Eq. 3 results in a recursive expression for $U(0)$, which can be calculated numerically up to a given order $r$. The procedure is tested in simulation, where the estimated and actual fractions of undetected events agree within an uncertainty of 13%. This is considered as a systematic uncertainty related to the assumptions made in the calculation. The fraction of undetected events obtained for data is 2.3% compared to 3.1% in simulation. The fraction estimated in data is added to the measured multiplicity distributions and is also considered in the event normalisation of the mean particle density measurement.
5.3 Pile-up correction

The average number of interactions per bunch crossing in the selected data taking period is small, resulting in a limited bias from pile-up. The measured particle multiplicity distributions are mainly composed of single \( pp \) collisions and a small fraction of additional second \( pp \) collisions. Therefore events with larger pile-up can be neglected. To obtain the particle multiplicity distribution of single \( pp \) collisions the iterative approach used in Ref. [10] is applied. The procedure typically converges after two iterations when the change of the multiplicity distribution is of the order of the statistical uncertainty. The pile-up correction changes the mean value of the multiplicity distribution by 3.3 %. The measurements of the mean particle densities are normalised to the total number of \( pp \) collisions.

5.4 Efficiency correction and unfolding procedure

The final correction step accounts for limited efficiencies due to detector acceptance \( (\epsilon_{\text{acc}}) \) in the kinematic range of \( 2.0 < \eta < 4.8 \) and track reconstruction \( (\epsilon_{\text{tr}}) \). For particles fulfilling the kinematic requirements, the detector acceptance describes the fraction that reach the end of the downstream tracking stations and are unlikely to interact with material or to be deflected out of the detector by the magnetic field. This fraction and the overall reconstruction efficiency are evaluated independently using simulated events. Correction factors are determined as functions of pseudorapidity and transverse momentum. No multiplicity dependence is observed.

The mean particle densities are corrected by applying a combined correction factor of \( 1/(\epsilon_{\text{acc}}\epsilon_{\text{tr}}) \) to each track in the same way as described in Sect. 5.1. In order to correct the particle multiplicity distributions, an unfolding technique based on a detector response matrix is employed. The response matrix, \( R_{m,n} \), accounts for inefficiencies due to the detector acceptance and track reconstruction. It is constructed from the relation between the distribution of true prompt charged-particles \( T(n) \) and the distribution of measured tracks \( M(m) \), subtracted for background and pile-up,

\[
M(m) = \sum_n R_{m,n} T(n). \tag{4}
\]

The matrix is obtained from simulated events. The simulated number of charged particles per event, \( n \), is compared to the corresponding number of reconstructed and background subtracted tracks, \( m \). Thus each possible value of simulated particle multiplicity is mapped to a distribution of reconstructed tracks. For very high multiplicities, the available number of events from the Monte Carlo sample is not sufficient to populate the entire matrix. The mapping is well described by a Gaussian distribution with mean value \( \bar{m} \) and standard deviation \( \sigma_m \). The distribution of \( \bar{m} \) and \( \sigma_m \) for a true multiplicity bin \( n \) can be parametrized by combinations of polynomial and logarithmic functions. This allows an extrapolation of the matrix up to large values of \( n \) and simultaneously suppresses the effect of statistical fluctuations in the entries of the matrix. For further information the reader is referred to the Appendix, where an example of the detector response matrix is shown in Fig. 8.

To extract the true particle multiplicity distribution \( T(n) \) from the measured distribution \( M(m) \), a procedure based on \( \chi^2 \)-minimisation [28,29] of the measured distribution \( M(m) \) and the folded distribution \( R_{m,n} \bar{T}(n) \) for different hypotheses of the true distribution, \( \bar{T}(n) \), is adopted. The range of variation of \( \bar{T}(n) \) is constrained by parametrising the multiplicity distributions. To avoid introducing model dependencies to the unfolded result, six different models with up to eight floating parameters are used. Five models are based on sums of exponential functions combined with polynomial functions of various order in the exponent and as a multiplier. In addition, a model based on a sum of negative binomial distributions is used. While particle multiplicities in \( \eta \) and \( p_T \) bins can be well described by two negative binomial distributions, this is not sufficient for the multiplicity distribution in the full kinematic range, where this model has not been employed. All the parametrisations used are capable of describing the simulated multiplicity distributions. The floating parameters of the hypothesis \( \bar{T}(n) \) are varied in order to minimise the \( \chi^2 \)-function

\[
\chi^2(\bar{T}) = \sum_m \frac{1}{E(m)^2} \left( M(m) - \sum_n R_{mn} \bar{T}(n) \right)^2, \tag{5}
\]

where \( E(m) \) represents the uncertainty of the measured distribution \( M(m) \). The parametrisation model yielding the best \( \chi^2 \)-value is chosen as the central result, the other models are considered in the systematic uncertainty determination. Both the binned and total event unfolding procedures using simulated data are found to reproduce the generated distributions satisfactorily. The uncertainty of the unfolded distribution is determined through pseudo-experiments. Each pseudo-experiment is generated from the analytical model with the parameters randomly perturbed according to the best fit and the correlation matrix.

As a consistency check, a Bayesian unfolding technique [30] is used. The unfolded distributions of both methods in all kinematic bins are found to be in agreement.

The unfolded distribution for the total event is truncated at a value of 50 particles and the binned distributions at a value of 20 particles. This corresponds to the limit where, even with the extended detector-response matrix, larger particle multiplicities cannot be fully mapped to the range of the
Fig. 1 Charged particle density as a function of $\eta$. The LHCb data are shown as points with statistical error bars (smaller than the marker size) and combined systematic and statistical uncertainties as the grey band. The measurement is compared to several Monte Carlo generator predictions, a Pythia 6 and Phojet, b Pythia 8 and Herwig++. Both plots show predictions of the LHCb tune of Pythia 6, which is used in the analysis.

Fig. 2 Charged particle density as a function of $p_T$. The LHCb data are shown as points with statistical error bars (smaller than the marker size) and combined systematic and statistical uncertainties as the grey band. The measurement is compared to several Monte Carlo generator predictions, a Pythia 6 and Phojet, b Pythia 8 and Herwig++. Both plots show predictions of the LHCb tune of Pythia 6, which is used in the analysis.

The properties of fake tracks are studied in detail by using fully simulated events. The agreement between data and simulation is verified by estimating the fake-track fraction in both samples by probing the matching probability of track segments in the long-track reconstruction algorithm. The results are in good agreement and the differences amount to an overall 2% systematic uncertainty on the applied correction factors.

The systematic uncertainty introduced by differences in the fraction of duplicate tracks in data and simulation is determined by studying the number of track pairs with small opening angles. The observed excess of duplicate tracks in data results in a relative systematic uncertainty on the duplicate-track fraction of 9%. As the total amount of this type of reconstruction artefacts is small, this results in an overall 0.1% systematic uncertainty on the final result.

6 Systematic uncertainties

The precision of the measurements of charged particle multiplicities and mean particle densities are limited by systematic effects. The bin contents of the particle multiplicity distribution for the full event typically have a relative statistical uncertainty in the range of $10^{-4}$ to $10^{-2}$ for low and high multiplicities, respectively. The systematic uncertainties are typically around 1–10%, the largest contribution arising from the uncertainty of the amount of detector material. All individual contributions are discussed below.
To account for differences between the actual track reconstruction efficiency and that estimated from simulation, a global systematic uncertainty of 4% in average is assigned [31,32].

The uncertainty on the detector acceptance can be split in two components: the uncertainty on the knowledge of the detector material and the uncertainty related to the requirement for particles to have trajectories within the acceptance of the downstream tracking stations. The momentum distributions of charged particles in data and in simulation are in good agreement, therefore the second effect is negligible. The remaining uncertainty related to material interaction leads to a relative systematic uncertainty on the correction factors of 3% and is assigned as an individual factor for each track.

A modified response matrix is used to estimate the impact on the multiplicity distributions of systematic uncertainties due to the track reconstruction and detector acceptance. The systematic uncertainties of both efficiencies are combined quadratically and result in a 5% uncertainty on the response matrix. A response matrix with an efficiency decreased by this value is generated. The whole unfolding procedure (Sect. 5.4) is repeated with this matrix and the full difference to the nominal result is assigned as uncertainty.

Model dependencies due to the parametrisations used to unfold the true particle multiplicity distributions are determined by sampling six different parametrisation models for each of the multiplicity distributions. The model corresponding to the minimum $\chi^2$ value of the unfolding fit is taken as the central result, while the maximum difference in each bin between all models and the central result is taken as the systematic uncertainty. This difference is small compared to the uncertainty due to the modified response matrix.

Uncertainties related to the correction for undetected events (Sect. 5.2) are dominated by the 13% systematic uncertainty arising from the assumptions made in the calculation model. In addition, the average survival probability used in this model is affected by uncertainties of the amount of detector material, detector acceptance and track reconstruction efficiency. This sums to a maximum uncertainty of 15% on the number of undetected events. Only bins from one to three tracks are affected, where the variation is dominated by this uncertainty. For the particle densities, the impact is negligible with respect to other uncertainties. For the particle multiplicity distributions it results in a small change of 0.4% of the truncated mean.

Uncertainties related to the pile-up fraction are evaluated to be negligible compared to all other contributions as the total size of the corrections is already small.

The effect of non-zero beam crossing angles is determined to be insignificant, as well as the background induced by beam gas interactions.

Uncertainties introduced by the correction for non-prompt particles depend predominantly on the knowledge of the amount of material within the detector. The agreement with the amount of material modelled in the simulation, on average, is found to be within 10%. In order to estimate the effects of non-prompt particles still passing the track selection, their composition is studied. Around 40% of the wrongly selected particles arise from photon conversion and is related to the uncertainty of the amount of material. Another third of the particles are decay products of $K_S^0$ mesons, whose production cross-section has previously been measured by LHCb [2] to be in good agreement with simulation. Around 20% of the particles originate from decays of $\Lambda$ baryons and hyperons. These are measured to disagree by approximately 40% with the production cross-sections used in the simulation. Combining these contributions results in a 12% systematic uncertainty on the fraction of non-prompt particles.

Fig. 3 Observed charged particle multiplicity distribution in the full kinematic range of the analysis. The error bars represent the statistical uncertainty, the error band shows the combined statistical and systematic uncertainties. The data are compared to several Monte Carlo predictions, a Pythia 6 and PHOJET, b Pythia 8 and Herwig++. Both plots show predictions of the LHCb tune of Pythia 6, which is used in the analysis.
charged particle densities

The fully corrected measurement of mean particle densities in the kinematic region of $p > 2 \text{ GeV}/c$, $p_T > 0.2 \text{ GeV}/c$ and $2.0 < \eta < 4.8$ is presented as a function of pseudorapidity in Fig. 1 and as a function of transverse momentum in Fig. 2; the corresponding numbers are presented in the Appendix. The data points show a characteristic drop towards larger pseudorapidities but also a falling edge for $\eta < 3$, which is caused by the minimum momentum requirement in this analysis. This is qualitatively described by all considered Monte Carlo event generators and their tunes.

The first group of generators that are compared to our measurements are different tunes of PYTHIA 6 and PHOJET and are shown in Figs. 1a and 2a. The default configuration of PYTHIA 6.426 underestimates the amount of charged particles from roughly 20% at large $\eta$ up to 50% at small $\eta$. The descending slopes towards small and large pseudorapidities are also insufficiently modelled. The Perugia NOCR tune shows a slight improvement in shape and in the amount of charged particles; Perugia 0 predicts an even smaller mean particle density over the whole kinematic range. Predictions of the PHOJET generator are similar to the tunes of PYTHIA 6. In this group of predictions, the LHCb tune of PYTHIA 6 provides the best agreement with the data but still underestimates the charged-particle production rate by 10–40%. This behaviour is also observed in the $p_T$ dependence, where all configurations underestimate the number of charged particles. The aforementioned generator predictions were optimized without input of LHC measurements.

Predictions from the more recent generators PYTHIA 8 and HERWIG++ are shown in Figs. 1b and 2b. PYTHIA 8.145 with default parameters was released without tuning to LHC measurements and is not better than the LHCb tune of PYTHIA 6. In contrast, PYTHIA 8.180, which was optimized on LHC data, describes the measurements significantly better than the previous version. The predictions of HERWIG++ are also in reasonably good agreement with data, although the charged-particle production rate is underestimated at small pseudorapidities. The HERWIG++ generator version 2.7.0, which uses tune UE-5, overestimates the number of prompt charged particles in the low $p_T$ range but underestimates it at larger transverse momenta. The predictions of HERWIG++ in version 2.6.3, which relies on tune UE-4, show a more complete...
Fig. 5  Observed charged particle multiplicity distribution in different \( \eta \) bins. Error bars represent the statistical uncertainty, the error bands show the combined statistical and systematic uncertainties. The data are compared to Monte Carlo predictions, a–c Pythia 6 and Phojet, d–f Pythia 8 and Herwig++. All plots show predictions of the LHCb tune of Pythia 6, which is used in the analysis.

8 Multiplicity distributions

The charged particle multiplicity distribution in the full kinematic range of the analysis is shown in Fig. 3, compared to the predictions from the event generators. The corresponding mean value, \( \mu \), and the root-mean-square deviation, \( \sigma \), of the distribution, truncated in the range from 1 to 50 particles, is measured to be \( \mu = 11.304 \pm 0.008 \pm 0.091 \) and \( \sigma = 9.496 \pm 0.006 \pm 0.021 \), where the uncertainties are statistical and systematic, respectively. Using the full range gives consistent results with the value obtained from the particle densities. All generators that do not use LHC data input underestimate the multiplicity distributions. In this comparison, the Phojet generator predicts the smallest probabilities...
Fig. 6 Observed charged particle multiplicity distribution in different $p_T$ bins. Error bars represent the statistical uncertainty, the error bands show the combined statistical and systematic uncertainties. The data are compared to Monte Carlo predictions, a, b PYTHIA 6 and PHOJET, c, d PYTHIA 8 and HERWIG++. All plots show predictions of the LHCb tune of PYTHIA 6, which is used in the analysis.

To observe a large multiplicity event, being in disagreement with the measurement. This can be understood since PHOJET mostly contains soft scattering events. All PYTHIA 6 tunes underestimate the charged particle production cross-section significantly. The prediction from the LHCb tune is closest to the data, but the mean value of the distribution is still about 15% too small. Calculations from more recent generators are in better agreement with the measurement. While PYTHIA 8.145 gives the same insufficient description of the data as its predecessor, the prediction of version 8.180 using Tune 4C shows a reasonable agreement. The HERWIG++ event generator using the underlying event tune UE-4 shows good agreement with the measurement and reproduces the data better than the more recent UE-5 tune.

Charged particle multiplicity distributions for bins in pseudorapidity are displayed in Figs. 4 and 5. The comparison with the predictions from Monte Carlo generators shows the same general features as discussed for the integrated distribution. The predictions of PHOJET and PYTHIA 6 all underestimate the particle multiplicity. The difference in particle production is most prominent at small $\eta$, where the minimum $p$ requirement in this analysis significantly reduces the amount of particles. Even though the LHCb tune is in better agreement with the data, the difference remains large. Recent generator predictions match the data better. Both PYTHIA 8 and HERWIG++ show good agreement with data at larger pseudorapidity, only the range from $2 < \eta < 3$ being still underestimated.

Charged particle multiplicities for bins of transverse momentum are shown in Figs. 6 and 7. The LHCb tune describes the data better than the other tunes. It is interesting to note that at large transverse momenta, where the discrepancies are most prominent, PYTHIA 6.426 in the default configuration matches the shape of the distribution. PYTHIA 8 in the recent configuration shows a reasonably good agreement to the measurement in the mid- and high-$p_T$ range, where also the HERWIG++ generator describes the data. Predictions using the UE-4 tune are closer to the measurement than using the UE-5 tune. Towards larger $p_T$, HERWIG++ predictions underestimate the amount of particles while the PYTHIA 8 prediction is slightly better. PYTHIA 8 underestimates the data towards lower $p_T$, while HERWIG++ overestimates it.

The mean value and the root-mean-square deviation for the multiplicity distributions in $\eta$ and $p_T$ bins are tabulated in the Appendix.
9 Summary

The charged particle multiplicities and the mean particle densities are measured in inclusive $pp$ interactions at a centre-of-mass energy of $\sqrt{s} = 7$ TeV with the LHCb detector. The measurement is performed in the kinematic range $p > 2\text{ GeV}/c$, $p_T > 0.2\text{ GeV}/c$ and $2.0 < \eta < 4.8$, in which at least one charged particle per event is required. By using the full spectrometer information, it is possible to extend the previous LHCb results [10] to include momentum dependent measurements. The comparison of data with predictions from several Monte Carlo event generators shows that predictions from recent generators, tuned to LHC measurements in the central rapidity region, are in better agreement than predictions from older generators. While the phenomenology in some kinematic regions is well described by recent PYTHIA and HERWIG++ simulations, the data in the higher $p_T$ and small $\eta$ ranges of the probed kinematic region are still underestimated. None of the event generators considered are able to describe the entire range of measurements.
Acknowledgments We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff at the LHCb institutes. We acknowledge support from CERN and from the national agencies: CAPES, CNPq, FAPERJ and FINEP (Brazil); NSFC (China); CNRS/IN2P3 and Region Auvergne (France); BMBF, DFG, HGF and MPG (Germany); SFI (Ireland); INFN (Italy); FOM and NWO (The Netherlands); SCSR (Poland); MEN/IFA (Romania); MinES, Rosatom, RFBR and NRC “Kurchatov Institute” (Russia); MinECo, XuntaGal and GENCAT (Spain); SNSF and SER (Switzerland); NAS Ukraine (Ukraine); STFC (United Kingdom); NSF (USA). We also acknowledge the support received from the ERC under FP7. The Tier1 computing centres are supported by IN2P3 (France), KIT and BMBF (Germany), INFN (Italy), NWO and SURF (The Netherlands), PIC (Spain), GridPP (United Kingdom). We are indebted to the communities behind the multiple open source software packages we depend on. We are also thankful for the computing resources and the access to software R&D tools provided by Yandex LLC (Russia).

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

Funded by SCOAP3 / License Version CC BY 4.0.

Appendix

See Fig. 8, Tables 1, 2, 3, 4

![LHCb simulation](image)

**Fig. 8** Example of the parametrized detector response matrix in the full kinematic range. The matrix is obtained from fully simulated events showing the relation between the true charged particle multiplicity and the reconstructed and background subtracted track multiplicity.

| Table 1 Charged particle density as a function of pseudorapidity. The first quoted uncertainty is statistical and the second systematic |
|---------------------------------|------------------|
| Pseudorapidity range            | \( dn/d\eta \)  |
| 2.0 \( \leq \eta < 2.2 \)       | 3.600 ± 0.048 ± 0.463 |
| 2.2 \( \leq \eta < 2.4 \)       | 4.032 ± 0.050 ± 0.460 |
| 2.4 \( \leq \eta < 2.6 \)       | 4.428 ± 0.055 ± 0.367 |
| 2.6 \( \leq \eta < 2.8 \)       | 4.754 ± 0.056 ± 0.277 |
| 2.8 \( \leq \eta < 3.0 \)       | 4.943 ± 0.057 ± 0.285 |
| 3.0 \( \leq \eta < 3.2 \)       | 4.977 ± 0.055 ± 0.267 |
| 3.2 \( \leq \eta < 3.4 \)       | 4.734 ± 0.052 ± 0.213 |
| 3.4 \( \leq \eta < 3.6 \)       | 4.500 ± 0.050 ± 0.207 |
| 3.6 \( \leq \eta < 3.8 \)       | 4.267 ± 0.049 ± 0.200 |
| 3.8 \( \leq \eta < 4.0 \)       | 4.026 ± 0.047 ± 0.194 |
| 4.0 \( \leq \eta < 4.2 \)       | 3.845 ± 0.046 ± 0.186 |
| 4.2 \( \leq \eta < 4.4 \)       | 3.613 ± 0.047 ± 0.263 |
| 4.4 \( \leq \eta < 4.6 \)       | 3.358 ± 0.043 ± 0.179 |
| 4.6 \( \leq \eta < 4.8 \)       | 3.281 ± 0.045 ± 0.174 |

| Table 2 Charged particle density as a function of transverse momentum. The first quoted uncertainty is statistical and the second systematic |
|---------------------------------|------------------|
| Transverse momentum range [ GeV/c ] | \( dn/dp_T [0.1 GeV/c]^{-1} \) |
| 0.20 \( \leq p_T < 0.30 \)       | 1.908 ± 0.024 ± 0.116 |
| 0.30 \( \leq p_T < 0.40 \)       | 1.866 ± 0.026 ± 0.099 |
| 0.40 \( \leq p_T < 0.50 \)       | 1.678 ± 0.022 ± 0.093 |
| 0.50 \( \leq p_T < 0.60 \)       | 1.347 ± 0.009 ± 0.092 |
| 0.60 \( \leq p_T < 0.70 \)       | 1.082 ± 0.007 ± 0.091 |
| 0.70 \( \leq p_T < 0.80 \)       | 0.817 ± 0.006 ± 0.064 |
| 0.80 \( \leq p_T < 0.90 \)       | 0.617 ± 0.006 ± 0.042 |
| 0.90 \( \leq p_T < 1.00 \)       | 0.481 ± 0.005 ± 0.044 |
| 1.00 \( \leq p_T < 1.10 \)       | 0.366 ± 0.005 ± 0.019 |
| 1.10 \( \leq p_T < 1.20 \)       | 0.290 ± 0.004 ± 0.015 |
| 1.20 \( \leq p_T < 1.30 \)       | 0.228 ± 0.004 ± 0.012 |
| 1.30 \( \leq p_T < 1.40 \)       | 0.180 ± 0.004 ± 0.009 |
| 1.40 \( \leq p_T < 1.50 \)       | 0.144 ± 0.003 ± 0.007 |
| 1.50 \( \leq p_T < 1.60 \)       | 0.113 ± 0.002 ± 0.007 |
| 1.60 \( \leq p_T < 1.70 \)       | 0.092 ± 0.002 ± 0.006 |
| 1.70 \( \leq p_T < 1.80 \)       | 0.075 ± 0.001 ± 0.005 |
| 1.80 \( \leq p_T < 1.90 \)       | 0.061 ± 0.001 ± 0.004 |
| 1.90 \( \leq p_T < 2.00 \)       | 0.053 ± 0.001 ± 0.003 |
Table 3 Truncated mean value and root-mean-square deviation for charged particle multiplicities in different $\eta$-bins. The range is from 0 to 20 particles. The first quoted uncertainty is statistical and the second is systematic.

<table>
<thead>
<tr>
<th>Pseudorapidity range</th>
<th>Mean value</th>
<th>Root-mean-square</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2.0 \leq \eta &lt; 2.5$</td>
<td>$2.010 \pm 0.002 \pm 0.118$</td>
<td>$2.460 \pm 0.002 \pm 0.115$</td>
</tr>
<tr>
<td>$2.5 \leq \eta &lt; 3.0$</td>
<td>$2.424 \pm 0.002 \pm 0.097$</td>
<td>$2.736 \pm 0.002 \pm 0.094$</td>
</tr>
<tr>
<td>$3.0 \leq \eta &lt; 3.5$</td>
<td>$2.409 \pm 0.002 \pm 0.100$</td>
<td>$2.668 \pm 0.002 \pm 0.113$</td>
</tr>
<tr>
<td>$3.5 \leq \eta &lt; 4.0$</td>
<td>$2.121 \pm 0.002 \pm 0.087$</td>
<td>$2.396 \pm 0.001 \pm 0.117$</td>
</tr>
<tr>
<td>$4.0 \leq \eta &lt; 4.5$</td>
<td>$1.852 \pm 0.002 \pm 0.069$</td>
<td>$2.093 \pm 0.001 \pm 0.073$</td>
</tr>
</tbody>
</table>

Table 4 Truncated mean value and root-mean-square deviation for charged particle multiplicities in different $p_T$-bins. The range is from 0 to 20 particles. The first quoted uncertainty is statistical and the second is systematic.

<table>
<thead>
<tr>
<th>Transverse momentum range [GeV/c]</th>
<th>Mean value</th>
<th>Root-mean-square</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0.2 \leq p_T &lt; 0.3$</td>
<td>$1.928 \pm 0.002 \pm 0.073$</td>
<td>$2.083 \pm 0.001 \pm 0.067$</td>
</tr>
<tr>
<td>$0.3 \leq p_T &lt; 0.4$</td>
<td>$1.865 \pm 0.002 \pm 0.065$</td>
<td>$1.971 \pm 0.001 \pm 0.050$</td>
</tr>
<tr>
<td>$0.4 \leq p_T &lt; 0.6$</td>
<td>$2.988 \pm 0.002 \pm 0.098$</td>
<td>$2.855 \pm 0.002 \pm 0.069$</td>
</tr>
<tr>
<td>$0.6 \leq p_T &lt; 1.0$</td>
<td>$2.881 \pm 0.003 \pm 0.103$</td>
<td>$3.029 \pm 0.002 \pm 0.090$</td>
</tr>
<tr>
<td>$1.0 \leq p_T &lt; 2.0$</td>
<td>$1.580 \pm 0.002 \pm 0.096$</td>
<td>$2.195 \pm 0.001 \pm 0.093$</td>
</tr>
</tbody>
</table>

References

7. CMS collaboration, V. Khachatryan et al., Charged particle multiplicities in $pp$ interactions at $\sqrt{s} = 0.9, 2.36$, and $7$ TeV. JHEP 01, 079 (2011). arXiv:1011.5531
11. LHCb collaboration, A.A. Alves Jr. et al., The LHCb detector at the LHC. JINST 3, S08005 (2008)
13. I. Belyaev et al., Handling of the generation of primary events in GAUSS, the LHCb simulation framework. in Nuclear Science Symposium Conference Record (NSS/MIC) (IEEE 2010), p. 1155
32. LHCb collaboration, R. Aaij et al., Measurement of $\sigma(pp \to b\bar{b}X)$ at $\sqrt{s} = 7$ TeV in the forward region. Phys. Lett. B 694, 209 (2010). arXiv:1009.2731