

Particle Identification in Jets and High-Multiplicity pp Events with the ALICE TPC

Benjamin Andreas Heß

— 2015 —

Particle Identification in Jets and High-Multiplicity pp Events with the ALICE TPC

Dissertation
der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Dipl.-Phys. Benjamin Andreas Heß
aus Hünfeld

Tübingen
— 2015 —

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation: 09.10.2015

Dekan:

Prof. Dr. Wolfgang Rosenstiel

1. Berichterstatter:

Prof. Dr. Hans Rudolf Schmidt

2. Berichterstatter:

Prof. Dr. Werner Vogelsang

3. Berichterstatter:

Prof. Dr. Wolfgang Kühn

“I do not know how the Third World War will be fought, but I can tell you what they will use in the Fourth — rocks!”

Albert Einstein, 1949 [1]

Abstract

The spectra of identified particles in a collision experiment comprise crucial information about the underlying physical processes. The ALICE experiment has powerful Particle IDentification (PID) capabilities, which are unique at the Large Hadron Collider (LHC). In this thesis, a statistical PID method based on the specific energy loss dE/dx in the ALICE Time Projection Chamber (TPC) is developed: the TPC Multi-Template Fit (MTF). The MTF allows for the extraction of identified charged particle spectra in a wide momentum range, which extends from about $150\text{ MeV}/c$ to above $20\text{ GeV}/c$. The TPC PID requires a detailed modelling of the TPC dE/dx response for momenta above $2 - 3\text{ GeV}/c$. A framework is developed that allows for the determination of the model parameters and for evaluating the PID information of charged particles.

With the MTF, the transverse momentum (p_T) spectra of charged pions, kaons and protons at mid-rapidity ($|\eta| < 0.9$) are measured for pp collisions at $\sqrt{s} = 7\text{ TeV}$. It is studied how these spectra and the corresponding K/π and p/π ratios evolve as a function of the event multiplicity. The study is motivated by the recent observation that high-multiplicity pp collisions exhibit intriguing similarities to p–Pb and Pb–Pb collisions. For instance, long-range correlations were observed in all three systems and the K/π and p/π ratios in p–Pb and Pb–Pb collisions showed a similar evolution with multiplicity. It is found that the latter ratios in pp collisions evolve as a function of multiplicity in a very similar way as in the larger systems. While a hydrodynamical description can explain these effects and is justified in Pb–Pb collisions, its application to smaller systems seems questionable. The new results can help to distinguish between the variety of models that has been suggested to explain the observations.

These “soft” phenomena are just one aspect of the physics at particle colliders. Another are “hard” probes like jets. Jets are objects that are defined for both experimental measurement and theoretical calculation, thus, allowing for directly comparing the two. The calculation of hadronic production cross-sections usually involves Fragmentation Functions (FFs), that need to be determined from experiment and currently exhibit large systematic uncertainties. The MTF is therefore also used to measure the charged hadron composition in charged jets from pp collisions. The spectra have been obtained as a function of the transverse momentum of the constituent, p_T^{track} , and as a function of the reduced momentum $z^{\text{charged}} \equiv p_T^{\text{track}}/p_T^{\text{jet,charged}}$. An increase of the strangeness production with z^{charged} and the suppression of leading baryons ($z^{\text{charged}} \rightarrow 1$) is visible in the K/π and p/π ratios. The results can be used to further constrain the FFs and to distinguish between theoretical models. A comparison to various PYTHIA Perugia tunes shows that all considered tunes roughly reproduce the observed trends, but none of them provides a perfect description of the data.

Zusammenfassung

In Kollisionsexperimenten beinhalten die Spektren identifizierter Teilchen wichtige Informationen über die zugrundeliegenden physikalischen Prozesse. Das ALICE Experiment bietet leistungsstarke Möglichkeiten zur Teilchenidentifikation (PID), die am Large Hadron Collider (LHC) einzigartig sind. In dieser Arbeit wird eine Methode zur statistischen PID entwickelt, die auf dem spezifischen Energieverlust dE/dx in der Spurendriftkammer (TPC) von ALICE beruht: der TPC Multi-Template Fit (MTF). Der MTF ermöglicht die Messung der Spektren von identifizierten, geladenen Teilchen über einen weiten Impulsbereich, welcher sich von ca. 150 MeV/c bis oberhalb von 20 GeV/c erstreckt. Die Teilchenidentifikation in der TPC benötigt eine detaillierte Modellierung des TPC- dE/dx -Ansprechverhaltens für Impulse oberhalb von 2 – 3 GeV/c. Es wird ein Framework entwickelt, welches sowohl die Bestimmung der Modellierungsparameter, als auch die Auswertung der PID-Informationen von geladenen Teilchen ermöglicht.

Der MTF wird zur Messung der Transversalimpulsspektren (p_T) geladener Pionen, Kaonen und Protonen im zentralen Rapiditysbereich ($|\eta| < 0.9$) in pp-Kollisionen mit $\sqrt{s} = 7$ TeV verwendet. Es wird untersucht wie sich diese Spektren und die entsprechenden K/π - und p/π -Verhältnisse als Funktion der Multiplizität des Events verhalten. Motiviert wird diese Studie durch die kürzliche Beobachtung, dass pp-Kollisionen mit hoher Multiplizität verblüffende Ähnlichkeiten mit p–Pb- und Pb–Pb-Kollisionen aufweisen. Beispielsweise wurden langreichweitige Korrelationen in allen drei Systemen beobachtet und die K/π - und p/π -Verhältnisse in p–Pb- und Pb–Pb-Kollisionen zeigten ein ähnliches Verhalten mit sich ändernder Multiplizität. Es wird beobachtet, dass sich die letzteren Verhältnisse in pp-Kollisionen in einer sehr ähnlichen Weise als Funktion der Multiplizität verhalten wie in größeren Systemen. In Pb–Pb-Kollisionen kann eine hydrodynamische Beschreibung diese Effekte erklären und ist gerechtfertigt, wohingegen eine Anwendung auf kleinere Systeme fragwürdig erscheint. Die neuen Ergebnisse können dabei helfen zwischen der Vielzahl an Modellen zu unterscheiden, die zur Erklärung der Beobachtungen vorgeschlagen wurden.

Diese „weichen“ Phänomene sind nur ein Aspekt der Physik an Teilchenbeschleunigern. Ein anderer sind „harte“ Sonden wie etwa Jets. Jets sind Objekte, die sowohl für experimentelle Messungen, als auch für theoretische Berechnungen definiert sind. Sie ermöglichen daher einen direkten Vergleich dieser. Die Berechnung hadronischer Produktionswirkungsquerschnitte geht üblicherweise mit der Verwendung von Fragmentationsfunktionen (FFs) einher, die experimentell bestimmt werden müssen und derzeit große systematische Unsicherheiten aufweisen. Der MTF wird daher auch dazu verwendet, um die Zusammensetzung geladener Jets aus geladenen Hadronen in pp-Kollisionen zu messen. Die Spektren werden sowohl als Funktion des Transversalimpulses der Bestandteile, p_T^{track} , als auch als Funktion des reduzierten Impulses $z^{\text{charged}} \equiv p_T^{\text{track}}/p_T^{\text{jet,charged}}$ bestimmt. Anhand der K/π - und p/π -Verhältnisse zeigt sich eine erhöhte Strangeness-Produktion mit

zunehmendem z^{charged} und eine Unterdrückung führender Baryonen ($z^{\text{charged}} \rightarrow 1$). Die Resultate können dazu verwendet werden, um die FFs weiter einzuschränken und um zwischen theoretischen Modellen zu unterscheiden. Ein Vergleich mit verschiedenen PYTHIA Perugia Tunes ergibt, dass alle untersuchten Tunes annäherungsweise die beobachteten Trends beschreiben. Jedoch bietet keines der Tunes eine perfekte Beschreibung der Daten.

Contents

1. Introduction	1
1.1. The Standard Model	1
1.1.1. The Running Couplings	4
1.2. The Phases of QCD	5
1.3. Parton Fragmentation and Jets	6
1.3.1. Medium Modifications	9
1.4. Collective Phenomena	11
1.5. Outline of This Thesis	15
2. Experimental Access	17
2.1. The Large Hadron Collider	17
2.1.1. Experiments at the LHC	17
2.2. ALICE at the LHC	18
2.2.1. Detector Setup	18
2.2.2. Software for Simulation, Reconstruction and Analysis . . .	24
2.3. Jet Reconstruction	26
2.3.1. Jet Reconstruction in ALICE	28
3. Particle Identification with the ALICE TPC	31
3.1. Overview	31
3.2. Energy Loss of Charged Particles in Matter	33
3.3. Technical Aspects	34
3.4. Bethe-Bloch Fit	35
3.4.1. Inclusion of V0 Particles	38
3.4.2. V0 Purity	40
3.4.3. Low-Momentum Correction	43
3.4.4. Spline Accuracy	44
3.4.5. Systematic Uncertainty of the Splines	45
3.5. Eta-Maps	48
3.5.1. Systematic Uncertainty of the Eta Correction	51
3.6. Resolution Maps	54
3.6.1. Systematic Uncertainty of the Resolution Maps	56
3.7. Occupancy Effects	59
3.8. Final Remarks	59
4. The TPC Multi-Template Fit	61
4.1. Analysis Outline	61

Contents

4.2.	Template Generation	61
4.2.1.	Asymmetry of the dE/dx Response	63
4.2.2.	Special Templates for High dE/dx	67
4.2.3.	Handling of Light Nuclei	69
4.3.	Fitting Procedure	71
4.3.1.	Implementation of a Binned Maximum Likelihood Fit	72
4.3.2.	Simultaneous Fitting of Individual Delta' (species)	73
4.3.3.	Regularisation	75
4.3.4.	Handling of Leptons	79
4.4.	MC Blind Test	80
5.	Monte Carlo Corrections	85
5.1.	MC Corrections for the Jet Analysis	85
5.1.1.	Efficiency, Acceptance and Transverse Momentum Resolution	86
5.1.2.	Secondary Contamination	87
5.1.3.	Muon Contamination	90
5.1.4.	Total Correction Factors	91
5.2.	MC Corrections for the Inclusive Analysis	92
5.2.1.	GEANT-FLUKA Correction	92
5.2.2.	Efficiency for Inclusive Particle Production	92
5.2.3.	Secondary and Muon Contamination	93
5.2.4.	Total Correction Factors	94
5.2.5.	MC Correction Factors for the Multiplicity Dependent Analysis	94
5.3.	Systematic Uncertainty of the MC Correction Factors	101
5.3.1.	Tracking Efficiency and Momentum Resolution	101
5.3.2.	Shape Dependence	104
5.3.3.	Secondary Particle Contamination	105
5.3.4.	GEANT-FLUKA Correction	115
5.3.5.	Strangeness Scaling	116
5.3.6.	Muon Contamination	116
5.3.7.	Multiplicity Dependence	116
5.3.8.	Error Propagation to the To-Pion Ratios	116
5.3.9.	Total Uncertainties	117
5.4.	MC Closure Check	122
5.5.	Limitations of the MC Correction Factors	124
6.	Identified Particle Production in pp Collisions in ALICE	127
6.1.	Event and Track Selection	127
6.2.	Discussion of Uncorrected Results for Jets	129
6.2.1.	Typical Uncorrected Results	129
6.2.2.	Fit Quality Checks	130

6.2.3. Comparison of Uncorrected Results to the TPC Coherent Fit and MC	133
6.3. Estimation of Systematic Uncertainties	137
6.3.1. Splines	138
6.3.2. Eta Correction	138
6.3.3. Resolution Map	139
6.3.4. Shape of the Detector Response	140
6.3.5. PID Weighting	140
6.3.6. Pile-up Rejection	143
6.3.7. Special Treatment of Errors for the Multiplicity-to-Minimum Bias Ratios	144
6.3.8. Total Systematic Errors	144
6.3.9. Systematic Uncertainty Verification via Eta-Differential Jet Analysis	145
6.3.10. Robustness Against Cut Variations	147
6.4. Inclusive Minimum Bias Particle Production	155
6.4.1. Comparison of Results to Other Analyses	159
6.5. Inclusive Particle Production in Multiplicity Bins	161
6.5.1. Multiplicity Estimators	161
6.5.2. Quality Checks	162
6.5.3. Results	164
6.5.4. Charge Dependence	167
6.6. Particle Production in Jets	171
6.6.1. Results	173
6.6.2. Comparison to MC Models	177
6.6.3. Charge Differential Analysis	178
6.7. Discussion	184
6.7.1. Jets	184
6.7.2. pp Versus Multiplicity	185
6.8. Conclusion About Analysis Method and Outlook	190
7. Summary	193
A. Remarks About the Ratios of two Results	195
B. Run Lists for Analysis	197
B.1. List of Monte Carlo Runs	197
C. Track Selection and Cuts	199
C.1. V0 Candidate Selection	199
C.2. ITS-TPC Standard Track Cuts 2010	199
C.3. Hybrid Tracks	200
C.3.1. Default Hybrid Tracks	200
C.3.2. Modified Hybrid Tracks	200

Contents

C.4. Geometrical Cut	201
C.5. Global Tracks and Cut Variations	201
D. MC Blind Test Results	203
E. Systematic Uncertainties of MC Corrections	209
E.1. Secondary Correction	209
E.2. Relative Systematic Uncertainties from Various Sources	215
F. Fit Quality Plots	221
F.1. Multiplicity Analysis	221
F.2. Jet Analysis	227
G. Plots for the Eta-differential Jet Analysis	229
H. Cross-Checks	233
H.1. TPC Coherent Fit	233
H.2. Combined PID and Lund	237
I. Robustness of the Analysis Against Cut Variations	239
J. Charge Differential Analysis Results	247
J.1. Multiplicity-Binned Analysis	247
J.2. Jet Analysis	249
K. Comparison of Corrected Particle Yields in Jets with MC	251
Bibliography	263

List of Figures

1.1.	Running of the strong coupling constant	5
1.2.	QCD phase diagram	6
1.3.	Illustration of semi-inclusive identified jet production	7
1.4.	Charged particle medium modification factors in p–Pb and Pb–Pb collisions	11
1.5.	CMS measurement of two-particle correlations in different collision systems	12
1.6.	Centrality evolution of K/π and p/π in p–Pb and Pb–Pb collisions .	14
2.1.	Schematic view of the LHC with the four large experiments	18
2.2.	Layout of the ALICE detector system	19
2.3.	Overview of the particle identification capabilities of the ALICE detector system	20
2.4.	Schematic view of the TPC.	21
2.5.	Illustration of the TPC working principle	22
2.6.	Event reconstruction flow	25
2.7.	Illustration of infrared and collinear (un-)safety	26
2.8.	Illustration of different jet reconstruction algorithms	28
3.1.	Specific energy loss dE/dx in the ALICE TPC	32
3.2.	Bethe-Bloch fit	37
3.3.	Comparison of primary electrons and V0 electrons with different cuts	39
3.4.	Armenteros-Podolanski plot of V0 mothers	40
3.5.	Invariant mass distributions of V0 mothers	41
3.6.	V0 purity via TPC dE/dx distributions	42
3.7.	Residuals of the Bethe-Bloch fit	44
3.8.	Spline accuracy	45
3.9.	Comparison of different Bethe-Bloch parametrisations	46
3.10.	Fit residuals of different Bethe-Bloch parametrisations	47
3.11.	Influence of the polar angle on the effective track length per pad row	48
3.12.	η dependence of dE/dx for different species at different momenta .	49
3.13.	η -map	51
3.14.	η -map quality	52
3.15.	Resolution fits	55
3.16.	Resolution map	57
3.17.	Accuracy of the resolution map	58

List of Figures

3.18. Limitations of statistical PID with the TPC	60
4.1. Template generation with and without pre-PID	62
4.2. Correlation between transverse momentum and TPC momentum	63
4.3. Asymmetric dE/dx response in data	64
4.4. Asymmetric dE/dx response in MC	67
4.5. Demonstration of special dE/dx templates	69
4.6. Demonstration of the special treatment of light nuclei	70
4.7. Template fits to data for different Δ'_{species}	74
4.8. Comparison of kaon fractions with and without simultaneous fitting	75
4.9. Effect of the regularisation demonstrated in MC	77
4.10. Comparison of different dE/dx crossings in MC	78
4.11. Effect of regularisation in data	78
4.12. Comparison of splines extracted with and without MC information	81
4.13. MC blind test for pions fitted versus p_T (jets)	82
4.14. MC blind test for kaons fitted versus p_T (jets)	83
4.15. MC blind test for protons fitted versus p_T (jets)	84
5.1. Efficiency correction factors (jets)	87
5.2. Yield ratio between MC and CMS measurement for different strange particles	89
5.3. Primary fractions and strangeness scaling (jets)	89
5.4. Correction factors for the muon contamination	90
5.5. Combined MC correction factors for the yields and to-pion ratios in jets	91
5.6. GEANT-FLUKA correction factors	93
5.7. Efficiency correction factors (inclusive)	93
5.8. Primary fractions (inclusive)	94
5.9. Combined MC correction factors for the inclusive yields and to-pion ratios	95
5.10. Efficiency correction factors for different bins of V0M multiplicity and MB	97
5.11. Efficiency correction factors for different bins of reference multiplicity and MB	98
5.12. Primary fractions for different bins of V0M multiplicity and MB	99
5.13. Primary fractions for different bins of reference multiplicity and MB	100
5.14. Single track efficiencies for positive particles	102
5.15. Tracking efficiency systematics estimated via the fast simulation	103
5.16. Modified fragmentation distributions	105
5.17. Ratio loose/tight DCA without kaon correction	108
5.18. Ratio loose/tight DCA after kaon correction	109
5.19. Double ratio for protons in jets with $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$	110
5.20. Double ratio for protons in jets with different p_T^{jet}	111
5.21. Double ratio for pions in jets with different p_T^{jet}	112

5.22.	Summary plot for the systematic uncertainty estimation of the secondary correction for protons in jets as a function of p_T	113
5.23.	Summary plot for the systematic uncertainty estimation of the secondary correction for pions in jets as a function of p_T	114
5.24.	Relative systematic uncertainties of the MC correction factors for the inclusive yields and to-pion ratios	119
5.25.	Relative systematic uncertainties of the MC correction factors for $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$	120
5.26.	Relative systematic uncertainties of the MB MC correction factors in the multiplicity analysis	121
5.27.	MC closure check of the corrections	123
6.1.	Uncorrected particle fractions in jets	130
6.2.	Template fits to data in different p_T^{track} intervals for jets	131
6.3.	Comparison of the template fits for minimum bias data and jet constituents	132
6.4.	Comparison of uncorrected pion fractions as a function of p_T to the TPC Coherent Fit and MC	134
6.5.	Comparison of uncorrected kaon fractions as a function of p_T to the TPC Coherent Fit and MC	135
6.6.	Comparison of uncorrected proton fractions as a function of p_T to the TPC Coherent Fit and MC	136
6.7.	Systematic error of the uncorrected fractions from uncertainties of the splines	138
6.8.	Systematic error of the uncorrected fractions from uncertainties of the η correction	139
6.9.	Systematic error of the uncorrected fractions from uncertainties of the resolution	140
6.10.	Systematic error of the uncorrected fractions from uncertainties of the shape of the dE/dx response	140
6.11.	Systematic error of the uncorrected fractions from uncertainties of the PID weighting	141
6.12.	Shape uncertainty from pre-PID weighting at lowest p_T	142
6.13.	Total systematic error of the uncorrected fractions	145
6.14.	Comparison of uncorrected particle fractions in η slices for $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ in data	146
6.15.	Single track efficiencies for the standard cut and the TPC geometrical cut	148
6.16.	Comparison of splines for default cut and TPC geometrical cut	149
6.17.	Comparison of corrected particle yields for the TPC geometrical cut and the ncl cut for $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ in data	150
6.18.	Cut variations for pions in jets	152
6.19.	Cut variations for kaons in jets	153
6.20.	Cut variations for protons in jets	154

List of Figures

6.21.	Impact of vertex cuts on the inclusive charged primary p_T spectra in MC	156
6.22.	Comparison of the normalisation factors for rapidity and pseudo-rapidity	158
6.23.	Comparison of MB yields to other analyses for negatively charged particles	160
6.24.	Statistics for the multiplicity-binned analysis	162
6.25.	Multiplicity dependence of the mean dE/dx for V0 pions	163
6.26.	Template fits to data for different reference multiplicity bins for $p_T = 4 - 4.5 \text{ GeV}/c$	164
6.27.	π and K spectra for different multiplicity bins	166
6.28.	p spectra for different multiplicity bins	167
6.29.	K/π ratios versus multiplicity	168
6.30.	p/π ratios versus multiplicity	169
6.31.	Comparison of positively and negatively charged particles for different bins of V0M multiplicity	170
6.32.	Reconstructed jet spectrum	171
6.33.	Comparison of the corrected charged $\pi/K/p$ yields for different p_T^{jet} in data	175
6.34.	Comparison of the corrected to-pion ratios for different p_T^{jet} in data as a function of p_T^{track}	176
6.35.	Comparison of the corrected to-pion ratios for different p_T^{jet} in data as a function of z	176
6.36.	Comparison of corrected pion yields in jets to different MC predictions as a function of p_T	178
6.37.	Comparison of corrected kaon yields in jets to different MC predictions as a function of p_T	179
6.38.	Comparison of corrected proton yields in jets to different MC predictions as a function of p_T	180
6.39.	Comparison of corrected K/π ratios in jets to different MC predictions as a function of p_T	181
6.40.	Comparison of corrected p/π ratios in jets to different MC predictions as a function of p_T	182
6.41.	Comparison of positively and negatively charged particles in jets	183
6.42.	ALICE measurement of charged particle yields in charged jets	185
6.43.	Correlation between p_T^{jet} and multiplicity	187
6.44.	Comparison of to-pion ratios for jets and different multiplicity estimators	189
D.1.	MC blind test for fitting versus p_T (inclusive)	204
D.2.	MC blind test for fitting versus z	205
D.3.	MC blind test for fitting versus ξ	206
D.4.	MC blind test for electrons (jets)	207

E.1.	Double ratio for pions in jets with $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$ without applying the kaon correction	210
E.2.	Double ratio for the inclusive case	211
E.3.	Summary plot for the systematic uncertainty estimation of the secondary correction in the inclusive case	212
E.4.	Summary plot for the systematic uncertainty estimation of the secondary correction for protons in jets as a function of z	213
E.5.	Summary plot for the systematic uncertainty estimation of the secondary correction for pions in jets as a function of z	214
E.6.	Relative systematic uncertainties of the MC correction factors for the pion yields	216
E.7.	Relative systematic uncertainties of the MC correction factors for the kaon yields	217
E.8.	Relative systematic uncertainties of the MC correction factors for the proton yields	218
E.9.	Relative systematic uncertainties of the MC correction factors for the to-pion ratios of kaons	219
E.10.	Relative systematic uncertainties of the MC correction factors for the to-pion ratios of protons	220
F.1.	Template fits to data for different reference multiplicity bins for $p_T = 0.8 - 0.9 \text{ GeV}/c$	222
F.2.	Template fits to data for different reference multiplicity bins for $p_T = 8 - 10 \text{ GeV}/c$	223
F.3.	Template fits to data for different bins of V0M percentile for $p_T = 0.8 - 0.9 \text{ GeV}/c$	224
F.4.	Template fits to data for different bins of V0M percentile for $p_T = 4 - 4.5 \text{ GeV}/c$	225
F.5.	Template fits to data for different bins of V0M percentile for $p_T = 8 - 10 \text{ GeV}/c$	226
F.6.	Template fits to data in different p_T^{track} intervals for jets	227
F.7.	Template fits to data in different z intervals for jets	228
G.1.	Comparison of uncorrected particle fractions in η slices for the inclusive case in data	230
G.2.	Comparison of uncorrected particle fractions in η slices for $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$ in data	231
G.3.	Comparison of uncorrected particle fractions in η slices for $p_T^{\text{jet}} = 15 - 20 \text{ GeV}/c$ in data	232
H.1.	Comparison of uncorrected pion fractions as a function of z to the TPC Coherent Fit	234
H.2.	Comparison of uncorrected kaon fractions as a function of z to the TPC Coherent Fit	235

List of Figures

H.3.	Comparison of uncorrected proton fractions as a function of z to the TPC Coherent Fit	236
H.4.	Comparison of MB yields to other analyses for positively charged particles	238
I.1.	Comparison of corrected pion yields for the TPC geometrical cut and the ncl cut	240
I.2.	Comparison of corrected kaon yields for the TPC geometrical cut and the ncl cut	241
I.3.	Comparison of corrected proton yields for the TPC geometrical cut and the ncl cut	242
I.4.	Cut variations for pions in jets as a function of z	243
I.5.	Cut variations for kaons in jets as a function of z	244
I.6.	Cut variations for protons in jets as a function of z	245
I.7.	Cut variations in the inclusive case	246
J.1.	Comparison of positively and negatively charged particles for different bins of reference multiplicity	248
J.2.	Comparison of positively and negatively charged particles in jets with $p_T^{\text{jet}} = 15 - 20 \text{ GeV}/c$	249
K.1.	Comparison of corrected pion yields in jets to different MC predictions as a function of z	252
K.2.	Comparison of corrected kaon yields in jets to different MC predictions as a function of z	253
K.3.	Comparison of corrected proton yields in jets to different MC predictions as a function of z	254
K.4.	Comparison of corrected pion yields in jets to different MC predictions as a function of p_T (logarithmic scale)	255
K.5.	Comparison of corrected kaon yields in jets to different MC predictions as a function of p_T (logarithmic scale)	256
K.6.	Comparison of corrected proton yields in jets to different MC predictions as a function of p_T (logarithmic scale)	257
K.7.	Comparison of corrected pion yields in jets to different MC predictions as a function of z (logarithmic scale)	258
K.8.	Comparison of corrected kaon yields in jets to different MC predictions as a function of z (logarithmic scale)	259
K.9.	Comparison of corrected proton yields in jets to different MC predictions as a function of z (logarithmic scale)	260
K.10.	Comparison of corrected K/π ratios in jets to different MC predictions as a function of z	261
K.11.	Comparison of corrected p/π ratios in jets to different MC predictions as a function of z	262

List of Tables

1.1.	Fundamental fermions of the Standard Model	2
1.2.	Interactions of the Standard Model	2
2.1.	Pad sizes of the ALICE TPC	23
5.1.	Relative systematic uncertainties of the secondary contamination . .	115
5.2.	Overview of the systematic uncertainty estimation of the MC correction factors.	118
5.3.	p_T and z cut-offs due to limitations of the MC correction factors . .	125
6.1.	Statistics of various analysis data samples	128
6.2.	Track cuts used in the individual parts of the analysis	129
6.3.	Variations for systematic uncertainty estimation	137
6.4.	Comparison of the normalised p_T spectra ratios with the N_{evt} ratios for the vertex cut	156
6.5.	Comparison of the normalised p_T spectra ratios with the N_{evt} ratios for the Z_{vtx} cut	156

1. Introduction

Mankind has always had the ambition to understand nature. Over the centuries, which saw the progressive invention of ingenious technologies, scientists around the world contributed to a description of the universe, they lived in. The past two centuries witnessed a revolutionary series of particle physics experiments, which allowed to gain deep insight into the matter the universe is built of. The discovery of the fundamental constituents of matter, like electrons and quarks, and their interactions between each other culminated in the formulation of a theory which is able to explain most of the observations made until today: the Standard Model of particle physics.

Despite the great efforts of former generations, many mysteries are still unresolved. The Standard Model is not complete because it relies on parameters, such as particle masses, that are not predicted by this theory, but need to be determined by experiment. Among other issues, the Standard Model can neither explain the indications for a huge amount of cold dark matter, nor does it include gravitation at all. The long yearned-for *Theory of Everything* still awaits its discovery.

A detailed knowledge of the microcosm is essential to better understand the macrocosm. An open question for the latter is the evolution of the early universe, around $10\text{ }\mu\text{s}$ after the Big Bang. Within this period, the temperature was above a critical value of about $T_{\text{crit}} \approx 160\text{ MeV}$ [2] and the matter is expected to have existed in a state called Quark-Gluon Plasma (QGP). Understanding this phase of matter and its transition to normal hadronic matter is crucial. On the one hand, it furthers our knowledge about the evolution of the early universe. On the other hand, it allows to verify Standard Model predictions.

The Large Hadron Collider (LHC) has ushered in a new era of high-energy physics experiments, that may help to resolve some of the remaining mysteries. This thesis is intended to add one small piece to the resolution of the big puzzle.

1.1. The Standard Model

In the Standard Model of particle physics, matter consists of a set of fundamental fermions with spin $1/2$ — leptons and quarks — and their anti-particles. The fermions can be arranged in three generations or families as shown in Tab. 1.1. They are subject to the electromagnetic, weak and strong interactions. The Standard Model implements these forces as an exchange of gauge bosons with spin 1,

1. Introduction

see Tab. 1.2. These gauge bosons couple to a specific charge, which, e.g., is for strong interactions the colour charge. Leptons do not carry colour charge and, thus, do not participate in strong interactions. Gravity as the fourth fundamental interaction is not part of the Standard Model and there is no accepted theory on quantum gravitation yet.

Table 1.1.: Fundamental fermions of the Standard Model. Anti-particles are not shown. The charge q is given in units of the electron charge e . The mass values are approximate values based on [3]¹. At the time the Standard Model was constructed, the neutrinos were assumed to be massless.

Generation	Lepton	q/e	$m \cdot c^2$	Quark	q/e	$m_{\text{bare}} \cdot c^2$
First	ν_e	0	$< 225 \text{ eV}$	u	$2/3$	$\sim 2.3 \text{ MeV}$
	e^-	-1	511 keV	d	$-1/3$	$\sim 4.8 \text{ MeV}$
Second	ν_μ	0	$< 0.19 \text{ MeV}$	c	$2/3$	$\sim 1.3 \text{ GeV}$
	μ^-	-1	106 MeV	s	$-1/3$	$\sim 95 \text{ MeV}$
Third	ν_τ	0	$< 18.2 \text{ MeV}$	t	$2/3$	$\sim 173 \text{ GeV}$
	τ^-	-1	1.78 GeV	b	$-1/3$	$\sim 4.2 \text{ GeV}$

Table 1.2.: Interactions of the Standard Model. The interactions between particles are described by the exchange of gauge bosons that couple to their specific charge. In the Standard Model, the electromagnetic and weak interactions are unified and described by the electroweak theory.

Interaction	Gauge boson(s)	Coupling to
Electromagnetic	γ	Electric charge
Weak	Z^0, W^\pm	$\left\{ \begin{array}{l} \text{Weak charge} \\ W^\pm \text{ also to electric charge} \end{array} \right.$
Strong	8 gluons	Colour charge

The Standard Model is based on *Quantum Field Theories* (QFTs) [4, 5]. The most prominent QFT is *Quantum ElectroDynamics* (QED), which describes electromagnetic interactions via the exchange of photons. QED has proven to reproduce (or predict) the values of high-precision experiments. An intriguing example is the calculation of the anomalous magnetic moment of the electron, which is found to be in relative agreement with experiment at the level of $7.7 \cdot 10^{-10}$ [6]. Mathematically, QED is an abelian gauge theory with the symmetry group $U(1)$.

¹The limits for the neutrino masses are stated with 95% CL for ν_e and ν_τ and with 90% CL for ν_μ (note that smaller mass limits exist for $\bar{\nu}_e$).

and is found to be renormalisable. The success of QED triggered huge efforts of finding a Grand Unified Theory (GUT) which describes all forces — electromagnetic, weak and strong — and merges them into a single force at high energies around some GUT scale. Big progress was achieved with the formulation of the electroweak theory by Sheldon Glashow, Abdus Salam and Steven Weinberg, which unifies electromagnetic and weak interactions in a gauge theory with the symmetry group $SU(2) \times U(1)$. Finally, it was found that the strong interaction can be described by *Quantum ChromoDynamics* (QCD) [7], a non-abelian gauge theory with symmetry group $SU(3)$. In summary, the interactions in the Standard model are governed by an $SU(3) \times SU(2) \times U(1)$ symmetry.

However, there was still an open issue. All bosons (and fermions) are massless in order to maintain gauge invariance, which contradicts the experimental findings, in particular, the W^\pm and Z^0 bosons were found to be massive. Gauge invariance is crucial for a theory to be renormalisable. This introduces the need of a mechanism to endow the bosons with mass without explicitly breaking gauge invariance.

A solution was already proposed in 1964 by Peter Higgs [8, 9], François Englert and Robert Brout [10], as well as Gerald Guralnik, Carl Hagen and Tom Kibble [11]. Their idea was to introduce a new field, today usually called Higgs field. The Higgs field, which is a $SU(2)$ doublet of complex scalar fields (4 real components in total), leads to spontaneous breaking of the local $SU(2)$ symmetry. Three out of the four degrees of freedom of the Higgs field describe massless Goldstone bosons, which are absorbed as an additional spin degree of freedom of the W^\pm and Z^0 bosons, thus, making them massive. Fermions can also acquire mass via their Yukawa coupling to the Higgs field. The fourth degree of freedom of the Higgs field manifests itself in a neutral scalar particle, the Higgs boson H^0 , which is an excitation of the Higgs field.

It was not before 2012 that the discovery of the Higgs boson at the LHC was announced by the ATLAS [12] and CMS collaborations [13]. A great challenge was that theory did not predict the Higgs mass, but only constraints from direct searches or indirectly from cross-sections containing the Higgs in loop processes existed. In the end, the mass of the Higgs boson was found to be around $125\text{ GeV}/c$ [3] and its properties are consistent with Standard Model predictions, which strongly corroborates the hypothesis of mass generation via the Higgs mechanism. The discovery completes the set of fundamental particles in the Standard Model and explains why some of these particles are massive.

At the time the Standard Model was constructed, neutrinos were assumed to be massless. However, the observation of neutrino oscillation (see e.g. [14]) suggests that neutrinos are massive. If they are produced (and detected) via the weak interaction, they are in a flavour eigenstate, which is a linear combination of mass eigenstates. The change of the quantum mechanical phases of the mass eigenstates during the propagation through space depends on the mass of each state. As a result, the mixture of mass eigenstates and, consequently, the flavour

1. Introduction

eigenstate change during the travel. Hence, the observation of neutrino oscillations suggests that the mass eigenstates have non-vanishing and different masses. There are extensions to the Standard Model to implement massive neutrinos (see e.g. [15]).

Despite all efforts, the LHC experiments have not found significant deviations from Standard Model predictions yet. In particular, no new particles like super-partners or further Higgs bosons predicted by some theories beyond the Standard Model have been discovered up to now.

1.1.1. The Running Couplings

In accordance with Heisenberg’s uncertainty principle, the vacuum is filled with short-lived “virtual” pairs of particles and anti-particles. In QFT, these virtual particles can screen or anti-screen the bosons mediating the forces, which modifies the coupling constant of the corresponding force. The strength of the effect depends on the momentum transfer Q^2 . Hence, the coupling constants are not constant, but “running” with Q^2 . For the electromagnetic force, the coupling strength is given by the fine-structure constant $\alpha \approx 1/137$ at low Q^2 and increases only slightly with Q^2 . The smallness of the coupling justifies the use of perturbation theory to calculate cross-sections expanded in powers of α .

The situation is fundamentally different for the strong coupling α_s in QCD. In that case, the mediators of the force, the gluons, carry colour charge themselves, to which they couple. To leading order, this yields [3]:

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2n_f) \cdot \ln(Q^2/\Lambda_{\text{QCD}}^2)}, \quad (1.1)$$

where n_f is the number of light quark flavours² and Λ_{QCD} is the scale at which the coupling diverges. Λ_{QCD} is the only free intrinsic parameter of QCD and experimentally $\Lambda_{\text{QCD}} \sim 200$ MeV is found. The momentum transfer dependence of α_s is shown in Fig. 1.1 together with a summary of measurements. Obviously, the trend in QCD is opposite to that in QED, since the coupling decreases with Q^2 . At distances larger than $\hbar c/\Lambda_{\text{QCD}} \approx 1$ fm, which corresponds to small Q^2 , the coupling becomes very strong and this could mean that it requires an infinite amount of energy to pull apart two coloured objects. Although it is not analytically proven, this could explain the experimental fact that fractionally charged objects (as quarks) have not been observed [16]. This observation is called *confinement*. Instead, quarks form colourless particles, namely baryons (qqq) and mesons ($q\bar{q}$).

The large value $\alpha_s \sim 1$ prohibits the application of perturbative QCD in the low energy regime. However, this is the region in which hadronisation and collective phenomena take place and is also related to the matter in neutron stars. So,

²The mass of these light quarks is much smaller than some renormalisation scale μ_R . The equation holds for an effective theory, in which all heavier quarks decouple from the theory.

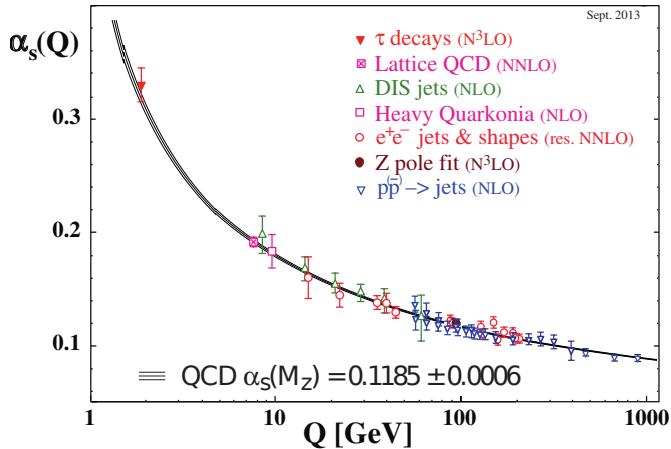


Figure 1.1.: Running of the strong coupling constant α_s as a function of the momentum transfer Q . Figure from [3].

it is important to study this region. To gain theoretical access to the low energy regime, effective theories like chiral perturbation theory [17] may be used. Another approach is lattice QCD [18], which performs numerical QCD calculations on discrete space-time points utilising powerful computing resources.

Going to infinite momentum transfers or, equivalently, to infinitely small distances, the strong coupling tends to zero. This behaviour is called *asymptotic freedom*. In particular, it justifies the applicability of perturbative QCD in the high energy regime.

1.2. The Phases of QCD

The asymptotic freedom of QCD has striking implications. QCD calculations predict that for matter at sufficiently high energy densities and/or temperatures, the confinement is repealed [19, 20]. The partons, i.e. quarks and gluons, become the relevant degrees of freedom in that case and can move freely. Such a state is called *Quark-Gluon Plasma* (QGP) and is believed to have existed in the early universe just a few microseconds after the Big Bang [2].

In general, the running of the strong coupling and, as a consequence, the phase of matter depend on the temperature T and the baryon chemical potential μ_B , which is related to the net baryon density. This results in different phases as a function of T and μ_B . A contemporary view of such a QCD phase diagram is shown in Fig. 1.2. Nuclear matter is in the hadron gas phase, whereas the dense core of neutron stars might reach into a QGP or a colour superconductor phase.

A number of current and future experiments tries to map the phase boundary between hadron gas and QGP and to locate the critical point — if it exists —, at which the first order phase transition changes to a smooth crossover. The experiments study different regions of μ_B via various projectiles and collision

1. Introduction

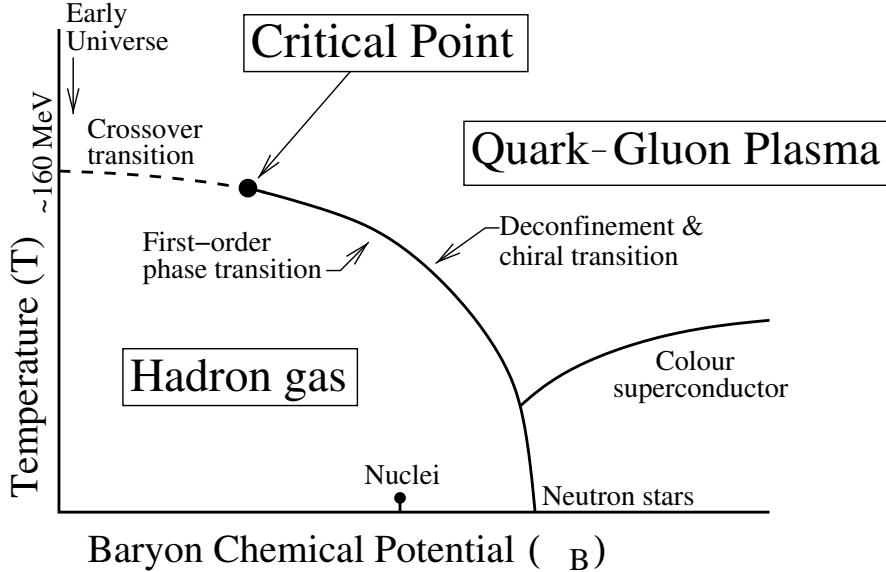


Figure 1.2.: QCD phase diagram with the temperature on the ordinate and the baryon chemical potential, which is directly connected to the baryon density, on the abscissa. Figure from [21].

energies. Among such experiments are e.g. (with increasing reach in μ_B) ALICE, ATLAS and CMS at the LHC (France/Switzerland), PHENIX and STAR at RHIC (USA) and CBM at FAIR (Germany).

1.3. Parton Fragmentation and Jets

The QCD confinement has the implication that each parton from hard-scattering processes at some point needs to hadronise into colour-neutral objects. The high-momentum parton from a hard-scattering process can produce a parton shower via gluon emission and gluons producing quark-anti-quark pairs. Once the emerging partons reach some soft scale, which is around λ_{QCD} , the hadronisation (or fragmentation) starts. To date, the theoretical description of the fragmentation process is still an open challenge and is treated via non-perturbative functions that need to be determined from experiment: the so-called *Fragmentation Functions* (FFs). The ultimate goal, namely to theoretically describe hard-scattering processes and calculate their cross-sections, crucially depends on the knowledge of the FFs.

An unpleasant complication is that theoretically the FFs deal with partons, which cannot be measured directly in experiment. This introduces the need of an object that approximately recovers the 4-momentum of the parton of interest. Jets are such objects. In a simple picture, a jet is a spray of particles that is collimated inside a cone of certain radius. Considering a parton a that originates from a hard-scattering, a parton shower may evolve with subsequent hadronisation. If

the 4-momenta of all emerging hadrons within a cone of sufficiently large radius are summed up, the 4-momentum of parton a is recovered. The important aspect is that the jet object can be constructed in both experiment and theoretical model and, thus, provides a link between the two. Note, however, that the jet recovers the original parton momentum only in the ideal case. Experimentally, detector inefficiencies may cause that some hadrons of the jet are not detected. Also, particles from other processes could be emitted into the jet cone and assigned to the jet. Or, vice versa, particles may be emitted out of the cone. These effects need to be taken into account carefully and introduce uncertainties.

The FFs depend on the hadron species h the parton fragments into. Hence, the study of different processes with $h = \pi^\pm, K^\pm, \dots$ determines different aspects of the hadronisation process. Measurements of identified FFs provide the best constraints, whereas charged particle FFs, i.e. without identification, only contain information about the (weighted) sum of identified FFs. An instructive example for identified FFs is the semi-inclusive identified jet production process illustrated in Fig. 1.3. A parton a is hard-scattered off a parton b , which is denoted by $d\hat{\sigma}$. The parton a carries the momentum $p_a = x_a \cdot P_a$, which is the (longitudinal) fraction x_a of the proton momentum P_a , and similarly for parton b . The hard-scattering results in a jet containing the parton c . Any other emerging partons are denoted by X . There can be gluon radiation (denoted by X'), such that the momentum of parton c , p_c , is smaller than that of the original jet, p_{jet} . Eventually, hadronisation takes place and produces besides some other hadrons (denoted by X'') the hadron h carrying the collinear momentum fraction $z_c = p_h/p_c$.

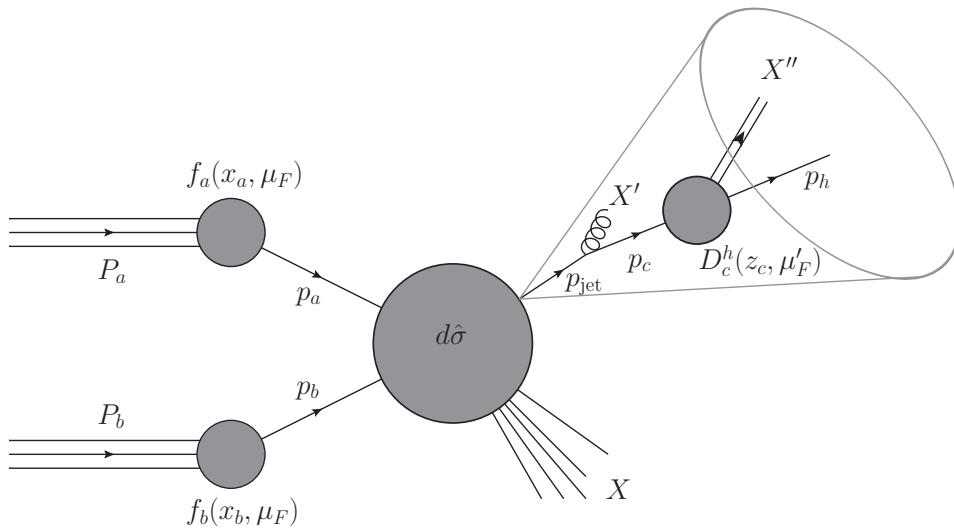


Figure 1.3.: Illustration of the semi-inclusive identified jet production. The partons a and b of the incoming protons experience a hard-scattering. The emerging parton c fragments into the hadron h , which is part of a jet.

1. Introduction

The calculation of the cross-section for this process uses the factorisation theorem (cf. [22] for a detailed discussion). It states that large-momentum-transfer reactions may be factorised into short-distance and long-distance parts. The former involve only large energy scales from the large-momentum transfer and can be calculated via perturbative QCD. The latter are considered to be universal, i.e. they are identical for any inelastic reaction. They cannot be calculated perturbatively, but need to be measured. A factorisation scale needs to be introduced that separates short from long-distance effects (also see [23]). Using the factorisation theorem, the semi-inclusive cross-section for producing an (identified) hadron h inside a jet (cf. Fig. 1.3) schematically reads:

$$\begin{aligned} d\sigma_{pp \rightarrow \text{jet}(h)X} &\propto \sum_{a,b,c} \iiint dx_a dx_b dz_c f_a(x_a, \mu_F) f_b(x_b, \mu_F) d\hat{\sigma}_{ab \rightarrow \text{jet}(c)X} D_c^h(z_c, \mu'_F), \\ d\hat{\sigma}_{ab \rightarrow \text{jet}(c)X} &= d\hat{\sigma}_{ab \rightarrow \text{jet}(c)X}(p_a = x_a \cdot P_a, p_b = x_b \cdot P_b, p_c = \frac{p_h}{z_c}, \mu_F, \mu_R, \mu'_F). \end{aligned} \quad (1.2)$$

The factorisation scales μ_F and μ'_F are introduced to absorb initial and final-state collinear singularities into the *Parton Distribution Functions* (PDFs) $f_{a,b}(x_{a,b}, \mu_F)$ and the *Fragmentation Function* (FF) $D_c^h(z_c, \mu'_F)$. $d\hat{\sigma}_{ab \rightarrow \text{jet}(c)X}$ is the (perturbatively calculable) short-distance scattering cross-section for the partonic subprocess $ab \rightarrow \text{jet}(c)X$. Amongst others, it depends on the momenta of the involved partons and the scale μ_R used for the renormalisation of the running strong coupling constant α_s . In the simple parton model, the PDF $f_a(x_a, \mu_F)$ gives the probability density to find a parton a with momentum fraction x_a inside the hadron (proton, in this case) at given energy scale μ_F . Similarly, the FF $D_c^h(z_c, \mu'_F)$ is the probability density for a parton c to produce a hadron h which carries the fraction z_c of the parton momentum p_c , at energy scale μ'_F . The energy scale dependence of both, PDFs and FFs, is described by the DGLAP equations³ [24, 25, 26].

The DGLAP equations take into account that the parton splitting processes $q \rightarrow q+g$, $g \rightarrow g+g$ and $g \rightarrow q+\bar{q}$, described by splitting functions as a function of the momentum fraction z , alter PDFs and FFs when changing the energy scale. The splitting functions have a perturbative expansion, which breaks down for $z \rightarrow 0$ due to divergent logarithmic terms. Consequently, the perturbative QCD calculation could not access the region of low momenta or low z . Fortunately, it turns out to be possible to carry out a resummation procedure giving finite results. The procedure is called *Modified Leading Logarithmic Approximation* (MLLA) [27, 28] and expands the validity range of the calculation to the low z (or p_T) regime.

Thanks to the universality of the FFs, results from different collision systems can be combined. A clean environment is provided in e^+e^- annihilations, since there are no initial-state hadron remnants. This enables a precise measurement

³Named after its discoverers Yuri Dokshitzer, Wladimir Naumowitsch Gribow, Lew Nikolaevitsch Lipatow, Guido Altarelli and Giorgio Parisi.

of quark FFs D_q^h . At leading order, there are only quarks produced. That is why hadron-hadron collisions are by far more sensitive to gluon FFs D_g^h than e^+e^- annihilations, but to the price of providing a more complicated environment. In that sense the measurements of FFs in hadron-hadron collisions are complementary to those in e^+e^- annihilations. On the one hand, rather recent measurements of identified charged hadron production from e^+e^- annihilations have been reported by the BaBar [29] and Belle [30] collaborations. The corresponding production in quark and gluon jets has already been measured earlier by the DELPHI collaboration [31]. On the other hand, the CDF collaboration measured identified neutral jet fragments in $p\bar{p}$ collisions [32]. Despite these efforts, the uncertainties of the FFs still give a sizeable contribution to the overall uncertainty of model predictions. The LHC experiments allow to access an unprecedented kinematic regime, which will help to further the knowledge about FFs and improve their accuracy — an overview discussion is presented in [33]. ATLAS, CMS and ALICE can all measure unidentified charged hadron yields. But the capabilities of ALICE to distinguish between different hadron types are unique at the LHC. This allows for crucial tests of the current understanding of FFs and the underlying perturbative QCD framework.

Recently, the charged jet cross-section and properties, including unidentified charged particle spectra, were reported by the ALICE collaboration [34] — in that publication, further references to previous jet measurements at the Tevatron and recent results from the LHC can be found.

1.3.1. Medium Modifications

Besides the vacuum FFs in pp collisions, the FFs from nucleus-nucleus collisions can be measured. In such collisions, a medium can be created, which could modify the FFs as long as the fragmentation happens inside the medium. The medium-modified FFs, denoted as nFFs, could differ significantly from the vacuum FFs due to medium-induced final-state soft exchanges taking place after the hard-scattering process. For example, such processes may be induced gluon radiation or elastic scattering. Hence, the nFFs contain information about the medium properties.

A possible theoretical approach for the nFF $D_{c/A}^h$ is to relate it to the vacuum FF D_c^h via a convolution [35]:

$$D_{c/A}^h(z, Q_0^2) \equiv \int_z^1 \frac{dy}{y} W_c^h(y, A, Q_0^2) D_c^h\left(\frac{z}{y}, Q_0^2\right). \quad (1.3)$$

The convolution is carried out at some initial scale Q_0 for colliding projectiles with atomic mass number A . The weight functions W_c^h parametrise all nuclear modifications and must retain the information of the vacuum FFs for $A = 1$. Once $D_{c/A}^h$ and D_c^h have been measured, the weight functions can be determined and the medium properties derived from them. However, whether the factorisation of hard and soft processes still holds in a nuclear environment, needs to be tested.

1. Introduction

The main features of FFs are usually determined from e^+e^- collisions providing very precise data. Evidently, this information source is not available for nFFs. But measurements by the HERMES collaboration on deuterium, helium, neon, krypton and neon targets [36] and the measurement of single-inclusive identified hadron yields in d–Au collisions at mid-rapidity by e.g. STAR [37] show evidence for medium-induced effects. The LHC experiments and, in particular, the ALICE experiment can provide a valuable contribution to the measurement of identified nFFs — see [33] for an overview; several references to measurements at the LHC for the study of medium modifications can be found in [34].

A quantitative way to study the medium modification is to consider the *nuclear modification factor* R_{AA} . It is defined as the ratio of the particle yield in nucleus-nucleus collisions and the expectation from scaling nucleon-nucleon collisions (also see [38]):

$$R_{AA}(p_T) \equiv \frac{d^2N_{AA}(p_T)/dp_Tdy}{\langle N_{coll} \rangle \cdot d^2N_{pp}(p_T)/dp_Tdy}. \quad (1.4)$$

N_{coll} is the number of binary collisions obtained from Glauber Monte Carlo simulations [39]. In Fig. 1.4, ALICE measurements of the nuclear modification factors in Pb–Pb collisions are shown. A strong suppression ($R_{PbPb} < 1$) is observed in central Pb–Pb collisions. The suppression of high- p_T particles is referred to as *jet quenching*. The applicability of N_{coll} scaling is justified by measurements showing that particles that do not interact strongly (like photons) have $R_{AA} = 1$. However, this measurement alone does not allow to judge whether the observed energy loss is a final or an initial-state effect. To disentangle these two, the nuclear modification factors have been measured in p–Pb collisions as well. As shown in the figure, initial-state effects are important at low and intermediate p_T . At high p_T , the R_{pPb} is consistent with unity. Therefore, the observed suppression has to be a final-state effect.

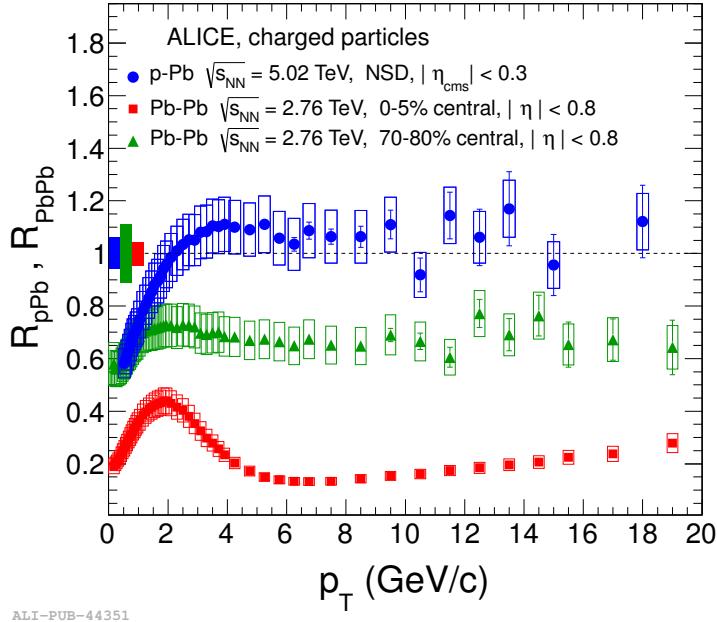


Figure 1.4.: ALICE measurement of charged particle medium modification factors in p–Pb and Pb–Pb collisions. In addition, central (red squares) and peripheral (green triangles) Pb–Pb collisions are compared. Figure from [40].

1.4. Collective Phenomena

The paradigm for the description of hard probes like jets is to take pp results (vacuum case) and scale with the number of binary collisions N_{coll} to compare with larger collision systems (see previous section and [38]). For soft physics, the situation is opposite. Soft effects increase with the number of participants N_{part} . This makes large systems like Pb–Pb a good baseline, since such effects are enhanced. Furthermore, it was studied in detail how the initial geometry in Pb–Pb collisions influences the soft physics like hydrodynamical flow. By now, hydrodynamics is well established for Pb–Pb collisions. The question is what happens to the soft physics in smaller systems.

A prominent example is the “ridge” observed in two-particle correlations as shown in Fig. 1.5a for Pb–Pb collisions measured by CMS [41]. The ridge is visible as a pronounced correlation along $\Delta\eta$ at the near side ($\Delta\phi$ around 0). This figure contains many interesting physical aspects. The peak at $(\Delta\phi \sim 0, \Delta\eta \sim 0)$ is due to local correlations like from particles within a single jet. The ridge at the away-side ($\Delta\phi \sim \pi$) is primarily associated with global correlations like a common event plane in Pb–Pb collisions. If the medium created in such collisions behaves like a liquid, the common event plane gives rise to collective motion independent of η . The ridge at the near side ($\Delta\phi \sim 0$) can also be attributed to such collective motion. However, it is cleaner at large $\Delta\eta$ than the away-side ridge, since the

1. Introduction

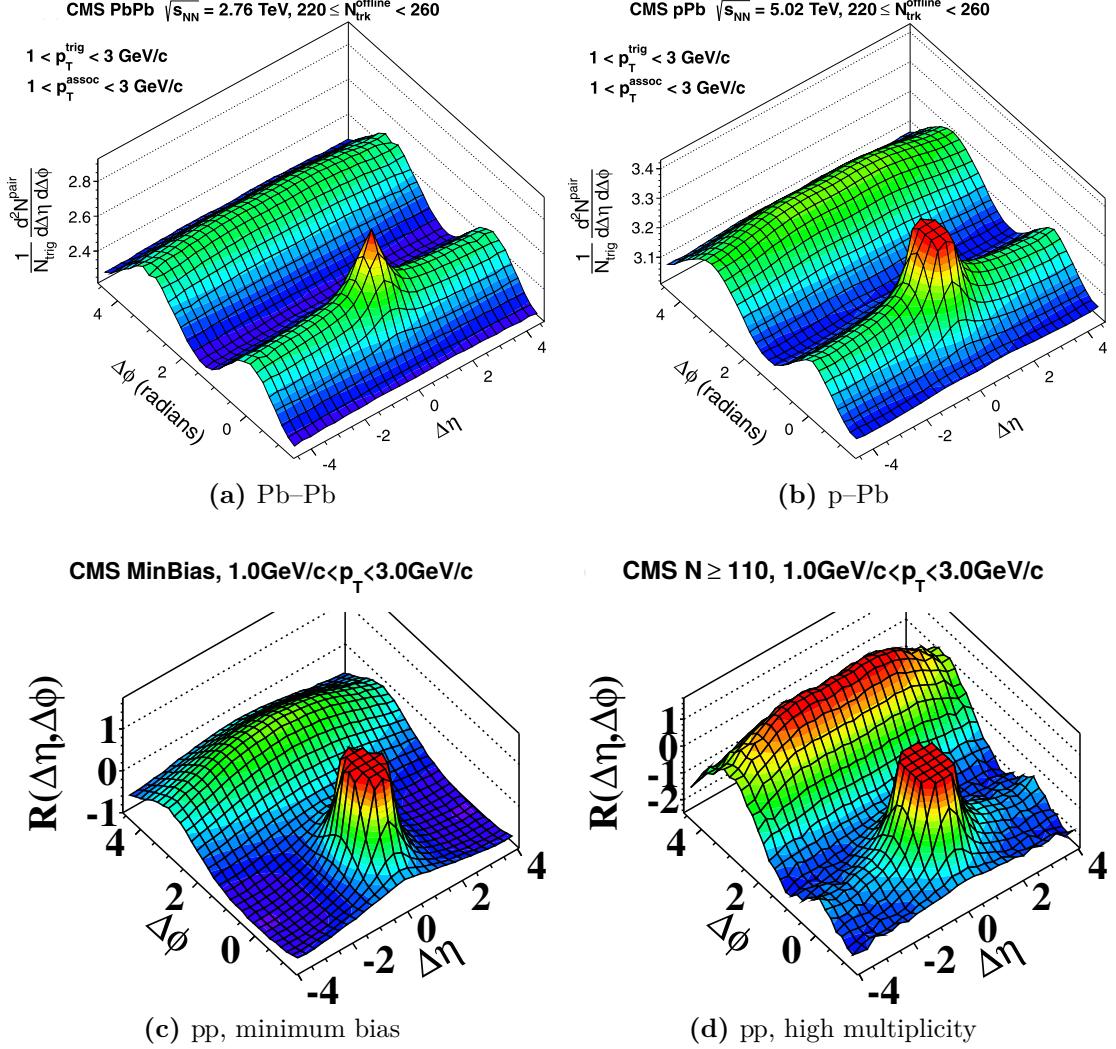


Figure 1.5.: CMS measurement of two-particle correlations in different collision systems at intermediate transverse momentum. The near-side ridge is clearly visible in **a)** Pb–Pb and **b)** p–Pb collisions (figures from [41]). In pp collisions, it is not observed for **c)** minimum bias, but seems to appear for **d)** high multiplicity (figures from [42]).

latter has additional contributions from the correlation of particles in back-to-back jets (the longitudinal jet momentum determining the η is asymmetric in dijets).

The collective motion can be described by hydrodynamics, which is considered to be applicable to macroscopic systems. It appears therefore reasonable that the contribution from collective motion disappears in small systems. To get a rough feeling about the system size in the different collision systems, two-pion femtoscopy may be used to determine the radii of the particle-emitting region

at freeze-out. According to measurements by ALICE, the product of the three femtoscopy radii can become as large as 300 fm^3 in Pb–Pb [43] and are typically a few fm^3 in both p–Pb [44] and pp [45] collisions. Hence, it seems unlikely that hydrodynamic evolutions play a role in the small pp and p–Pb systems. Indeed, the ridge is not present in minimum bias pp collisions measured by CMS [42] as shown in Fig. 1.5c, which seems to confirm this expectation. Surprisingly, very similar two-particle correlations as in Pb–Pb are observed in p–Pb collisions, see Fig. 1.5b. Even in high-multiplicity pp collisions a small near-side ridge is visible, see Fig. 1.5d. How can such small systems exhibit long-range (large $\Delta\eta$) correlations?

This clearly demonstrates the contrast to hard probes mentioned in the beginning. The soft physics phenomena are well understood in Pb–Pb collisions, but their observation in small systems like pp needs further study. A variety of theoretical models has been proposed (see [46] for a brief review) to explain the observation. Eventually, the experiment as the highest judge has to decide which of these models gives the best description of nature.

Another interesting observation is the multiplicity evolution of the to-pion ratios for kaons (K/π) and protons (p/π). Also here, some observed trends can be attributed to collective phenomena and it is instructive to study how these effects change depending on the collision system. In Fig. 1.6, the ALICE measurement [47] of the K/π and p/π ratios of central and peripheral Pb–Pb collisions are compared to those in p–Pb. Some similar trends are observed in p–Pb and Pb–Pb collisions. In both systems, K/π has only a mild centrality dependence. In addition, the observed crossing point of p/π of different centralities is at roughly the same position (1.5 and 1.6 GeV/c for Pb–Pb and p–Pb collisions, respectively). For both K/π and p/π the multiplicity dependence is larger in the bigger system.

The p/π enhancement with multiplicity at intermediate p_T in Pb–Pb collisions can be explained by hydrodynamical flow [48] and parton coalescence or recombination models [49, 50, 51, 52] — where due to coalescence low- p_T baryons gain more transverse momentum from flow than mesons and are pushed to higher p_T . It should be noted that this enhancement is caused by the (azimuthally symmetric) radial flow, whereas the ridge structures in Fig. 1.5a are due to asymmetric azimuthal flow. In Pb–Pb collisions, the K/π ratio increases with multiplicity above $p_T \sim 1 \text{ GeV}/c$. This observation can be explained by a reduced canonical suppression of the strangeness production in larger freeze-out volumes [53] or by an enhanced strangeness production in a quark-gluon plasma [54]. Both effects should be irrelevant for the small pp system and also for p–Pb.

Extrapolating the trends from Pb–Pb to p–Pb to pp collisions leads to the naive expectation that K/π becomes basically independent of multiplicity and that p/π still has a crossing point at the same position, but the multiplicity dependence may be further reduced. ALICE is perfectly suited to carry out the corresponding pp measurement for different multiplicities and to check whether this expectation is realised in nature.

1. Introduction

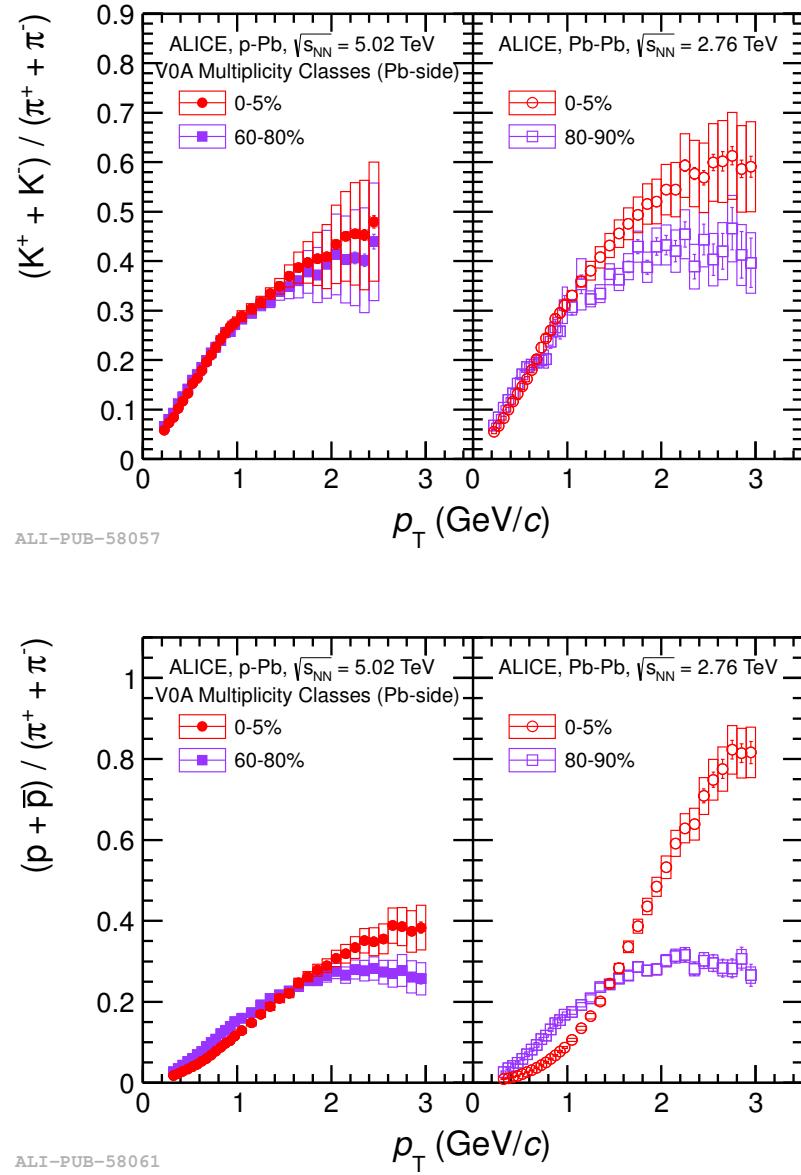


Figure 1.6.: Centrality evolution of K/π (top) and p/π (bottom) in $p\text{-Pb}$ and Pb-Pb collisions measured by ALICE. Some similar trends are observed in $p\text{-Pb}$ and Pb-Pb as discussed in the text. Figures from [47].

1.5. Outline of This Thesis

This thesis is structured as followed. The experimental setup and general data processing are presented in Chapter 2. Chapter 3 discusses the particle identification with the ALICE Time Projection Chamber (TPC). It further describes the modelling of the TPC dE/dx response, which is utilised in Chapter 4 to generate dE/dx templates with the expected dE/dx response for each species. These templates are used to extract the particle composition in both the inclusive case (all particles, no jet reconstruction, different event multiplicities) and in charged jets.

With Monte Carlo (MC) simulations the results are corrected for detector effects like efficiency and resolution, secondary particle contamination, etc. The correction procedure is presented in Chapter 5. Finally, Chapter 6 discusses the inclusive particle production in different multiplicity bins from two complementary multiplicity estimators at first. Secondly, the particle composition in jets is discussed and compared to MC models and to the inclusive case.

Important remarks about the way results are compared to each other in this thesis can be found in Appendix A. It is common practice in the ALICE collaboration to discuss analysis procedures and results in (usually ALICE-internal) analysis notes, which contain all technical details and performed cross-checks. This is also the case for the analysis described in this thesis.

2. Experimental Access

In the last century, the development and improvement of particle accelerators enabled a detailed study of the physics of particle interactions. They are a mean to collide particles in a laboratory with controlled (high) energy and at high rates.

2.1. The Large Hadron Collider

The Large Hadron Collider (LHC) is currently the world's highest energy particle accelerator that has been designed to collide either protons or heavy ions. It has been installed at the European Organisation for Nuclear Research (CERN), near Geneva. The LHC has a circumference of about 27 km and is located about 100 m below the surface. The nominal beam energy is 7 TeV for protons and 2.76 TeV/nucleon for lead ions [55, 56]. The design luminosity is $\mathfrak{L}_{pp} = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ in proton-proton (pp) and $\mathfrak{L}_{\text{PbPb}} = 10^{27} \text{ cm}^{-2}\text{s}^{-1}$ in lead-lead (Pb–Pb) collisions, respectively. It was considered later to operate the LHC also in proton-lead (p–Pb) mode [57, 58].

During LHC run 1 (2009–2013) the LHC delivered pp, Pb–Pb and p–Pb collisions with nucleon-nucleon centre-of-mass energies of $\sqrt{s_{\text{NN}}} = 7, 2.76$ and 5.02 TeV , respectively [59]. There have also been periods of pp operation with $\sqrt{s_{\text{NN}}} = 0.9$ and 2.76 TeV . In 2012, the energy in pp was increased to $\sqrt{s_{\text{NN}}} = 8 \text{ TeV}$.

2.1.1. Experiments at the LHC

The LHC beams are brought to collision at four interaction points, where the four large LHC experiments have been installed¹, see Fig. 2.1. Two of them, ATLAS and CMS, use general purpose detectors, whereas LHCb is specialised on B physics and the ALICE detector was specifically designed for the needs of heavy-ion collisions.

The *ATLAS* (A Toroidal LHC ApparatuS) experiment [64] has been optimised for the search of the Higgs boson, dark matter candidates and physics beyond the Standard Model, like extra dimensions or supersymmetric particles.

The *CMS* (Compact Muon Solenoid) experiment [65] pursues the same scientific goals as the ATLAS experiment. However, other technical solutions and,

¹In addition, there are 3 smaller experiments at the LHC, namely TOTEM [60], LHCf [61] and MoEDAL [62].

2. Experimental Access

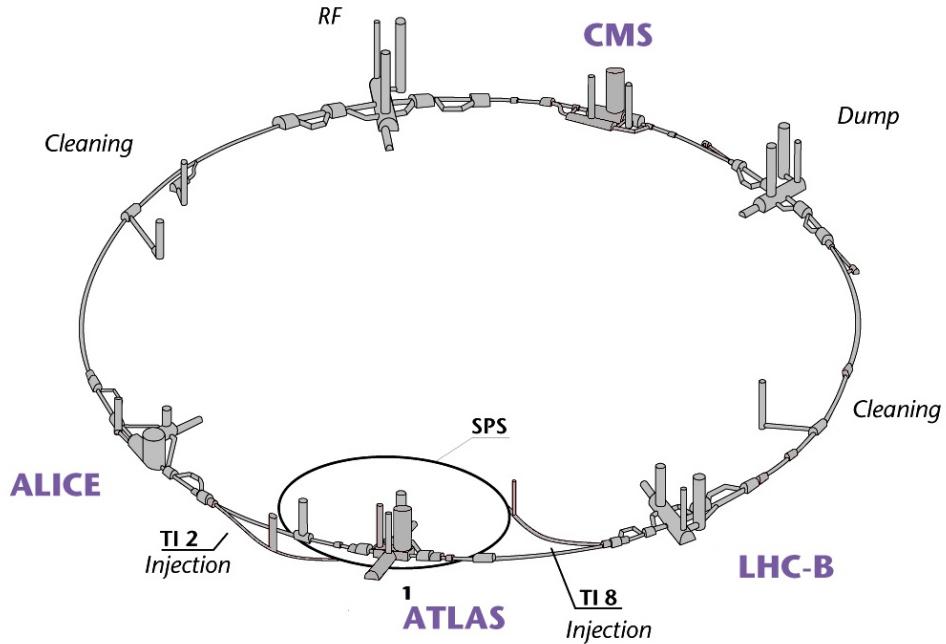


Figure 2.1.: Schematic view of the Large Hadron Collider and the four large experiments: ATLAS, ALICE, CMS and LHCb. The figure also shows the Super Proton Synchrotron (SPS), which lies at the end of a chain of accelerators and injects particles into the Large Hadron Collider (figure adapted from [63]).

in particular, a different design of the magnet system are used for this purpose. This allows for mutual verification of the results acquired by both experiments.

ALICE (A Large Ion Collider Experiment) [66] is specialised on heavy-ion collisions. The experiment studies the Quark-Gluon Plasma (QGP) that is created in Pb–Pb collisions and investigates the QCD confinement. ALICE has unique particle identification capabilities at the LHC and can, due to a lower magnetic field than ATLAS and CMS, measure particles with low transverse momenta. This enables complementary measurements in all collision systems.

The *LHCb* (Large Hadron Collider beauty) experiment [67] searches for the origin of the observed matter-anti-matter asymmetry in the universe by studying the CP-violation in B-meson decays.

2.2. ALICE at the LHC

2.2.1. Detector Setup

The ALICE detector has been optimised for heavy-ion collisions in order to study the properties of the QGP. The Particle IDentification (PID) capabilities of ALICE are unique at the LHC. Hence, the detector is also used to study proton-proton and proton-lead collisions.

The experiment has a total weight of 10,000 t and its overall dimensions are $16 \times 16 \times 26$ m³ [66]. It can be divided into two parts: the central barrel, which focusses on measurements at mid-rapidity, and the forward muon spectrometer at pseudo-rapidity $-4.0 \leq \eta \leq -2.5$. A schematic view of the detector and its constituents is shown in Fig. 2.2. Combining the PID information of all detectors, ALICE is capable of identifying particles in a wide momentum range, see Fig. 2.3.

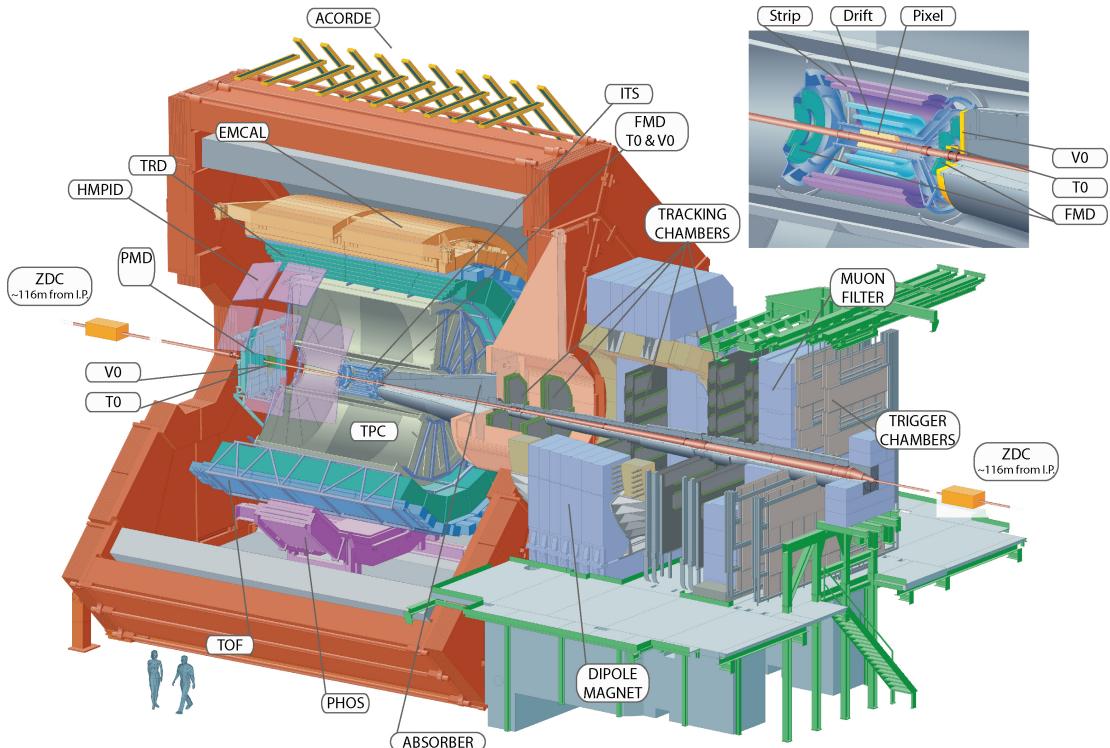


Figure 2.2.: Layout of the ALICE detector system in LHC run 1. The detector can be divided into 2 parts: the central barrel and the muon arm in forward direction. The central barrel is embedded in the L3 solenoid and consists of several subdetectors. The inset shows the ITS.

The individual detectors will be described shortly in the following. A comprehensive description of all detectors with references to the corresponding design reports is given in [66].

The Central Barrel

The central barrel is located inside the L3 solenoid, which provides a magnetic field of up to 0.5 T. The detector has an onion shell structure. Moving from the collision vertex radially to the outside, the produced particles traverse the following detectors: the Inner Tracking System (ITS), the Time Projection Chamber (TPC), the Transition Radiation Detector (TRD) and the Time Of Flight (TOF)

2. Experimental Access

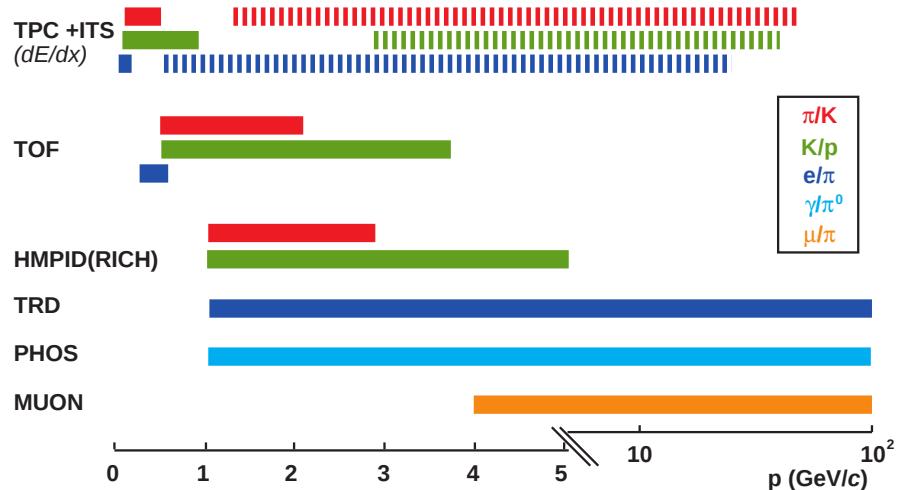


Figure 2.3.: Overview of the particle identification capabilities of the ALICE detector system. The values of this plot are based on [68]. The broken bars indicate momentum regions, in which only PID on statistical basis can be performed by the corresponding detectors. Taking all detectors together, ALICE allows to identify particles in a wide momentum range. Figure from [69].

detector. Further detectors at large radii are the High Momentum Particle IDentification (HMPID) detector, the PHOton Spectrometer (PHOS) and the ElectroMagnetic Calorimeter (EMCal). All detectors, except for HMPID, PHOS and EMCal, which have a smaller overall acceptance, cover the full azimuthal angle.

The global central barrel tracking [70] is based on the information from ITS and TPC. The bending of charged particle trajectories in the magnetic field allows for measuring the momentum. The central barrel's PID information is mainly obtained from the dE/dx measurement in the TPC and in the ITS, the TOF data and the measurement of dE/dx and transition radiation in the TRD.

The Inner Tracking System

The main purpose of the Inner Tracking System (ITS) [71] is the localisation of the primary vertex with a resolution of better than $100\text{ }\mu\text{m}$ and to reconstruct the secondary vertices of short-lived particles. It is also capable of tracking and identifying particles with momenta below $100\text{ MeV}/c$, which will not reach the TPC.

The ITS consists of 6 layers: The two innermost layers are Silicon Pixel Detectors (SPD), followed by two layers of Silicon Drift Detectors (SDD) and completed by two outermost layers of Silicon Strip Detectors (SSD). The SDD and SSD information includes up to four samples of the particle's specific energy loss dE/dx , which can be used for PID.

The Time Projection Chamber

To date, the ALICE Time Projection Chamber (TPC) [72, 66] is the largest TPC ever built. It is the main tracking device of the central barrel and has been designed to reconstruct up to 20,000 trajectories of charged particles per event, which corresponds to a charged particle multiplicity density of $dN_{\text{ch}}/d\eta = 8000$. Another crucial purpose is the identification of particles via their specific energy loss dE/dx . The TPC PID is the main topic of this thesis. This section gives an overview of the working principle. The PID procedure will be discussed in detail in Chapter 3.

The TPC has a cylindrical shape with a length of 5 m. Its inner and outer radii are about 80 cm and 250 cm, respectively. It is filled with a mixture of noble gas and quencher. For the 2010 data taking relevant for this thesis, the gas proportions were 90 Ne / 10 CO₂ / 5 N₂. The TPC volume is divided into two parts by the central electrode (cf. Fig. 2.4), which is set to a negative voltage of -100 kV . The cathode planes close to the end-plates on each side are kept at 0 kV, see Fig. 2.5. The strength of the resulting drift field is 400 V/cm. The homogeneity of the electric field is achieved by a field cage with equipotential strips, which are connected by a voltage divider.

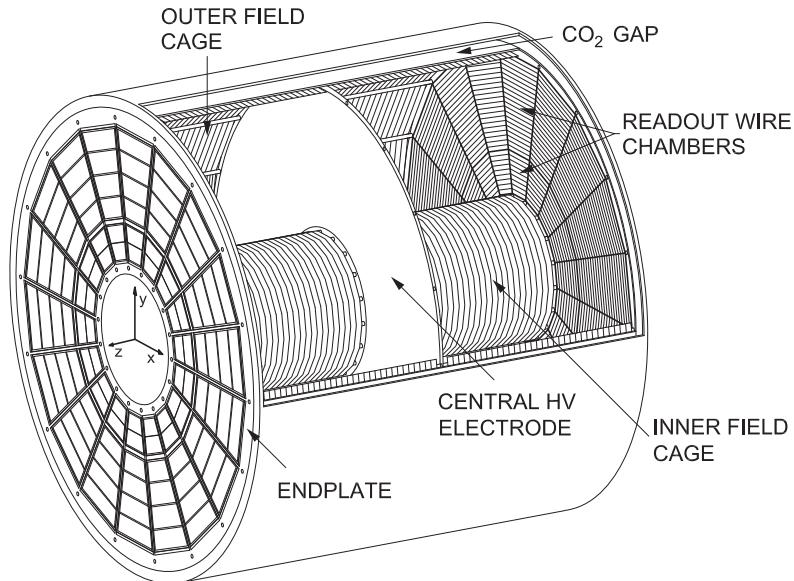


Figure 2.4.: Schematic view of the TPC. Figure from [73].

A charged particle traversing the TPC volume ionises the gas atoms along its trajectory, see Fig. 2.5. The acceleration of the ionisation electrons by the drift field is interrupted by collisions with the gas atoms. Macroscopically, this leads to a constant drift velocity of about 2.7 cm/ μs and a maximum drift time around 90 μs . The electron clouds also diffuse during their drift in the gas. Due to their larger mass, the ion drift velocity is lower by three orders of magnitude. This

2. Experimental Access

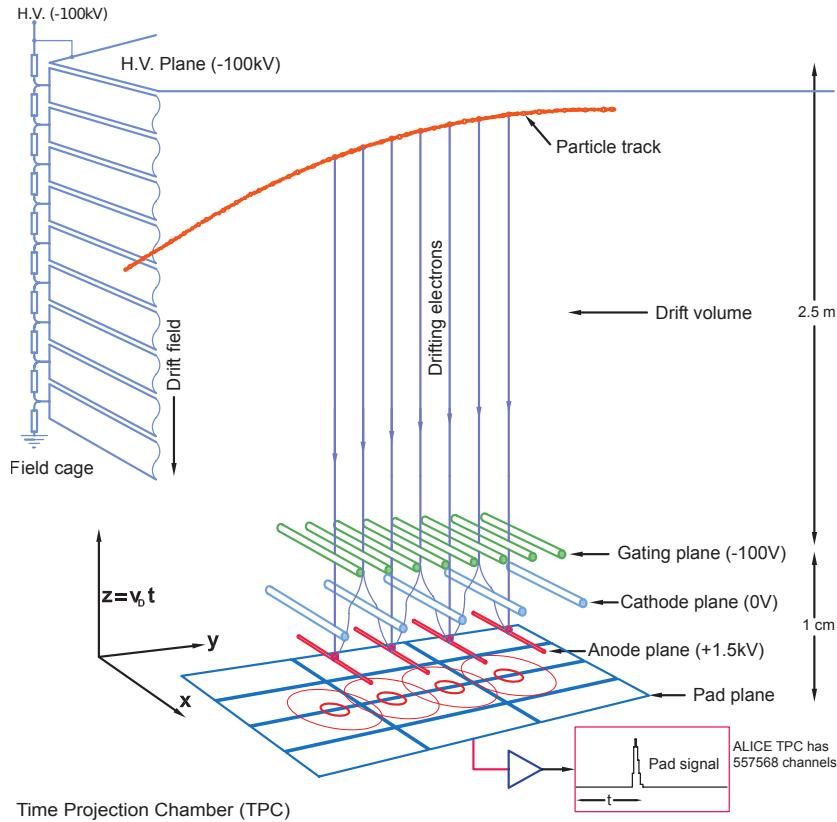


Figure 2.5.: Illustration of the TPC working principle. The field lines are shown for the case that the gating grid is “open”. Figure from [74].

may lead to a substantial amount of space charge in the drift volume distorting the field lines.

The drift volume ends in a Multi-Wire Proportional Chamber (MWPC), the amplification region, as illustrated in Fig. 2.5. It consists of the gating, cathode and anode wire planes and the pad plane. A high positive voltage of order 1.5 kV is set on the anode wire grid. In the vicinity of an anode wire, the electric field strength increases proportionally to the inverse distance. This leads to an amplification of the drifting electrons: they are accelerated towards the anode wire and create secondary ionisation electrons. Those also get accelerated, in turn ionising further gas atoms, such that an avalanche starts. The amplification voltage is chosen such that the induced signal in the pad plane is proportional to the number of (primary) drift electrons and, thus, proportional to the energy loss of the track.

Photons are produced in the avalanche process as well. If their mean free path length in the gas is too large, they can trigger a second full avalanche that is spatially separated from the original one. This can lead to instabilities of the MWPC, if on average every avalanche creates at least one further avalanche. As a solution, quench gases — CO₂ and N₂ are used in case of the ALICE TPC —

are added to the gas mixture. They exhibit large photoabsorption cross-sections over a broad wavelength region. This reduces the mean free path length of the photons, such that they only contribute to the original avalanche.

The readout plane of the MWPC is segmented into Inner (IROC) and Outer ReadOut Chambers (OROC). The occupancy decreases with increasing distance to the beam pipe. Accordingly, the pad sizes increase with the radius as listed in Tab. 2.1. In total, there are three different pad sizes.

Table 2.1.: Pad sizes of the ALICE TPC. There are three different pad types.

Pad type	Number of rows	size (mm ²)
Short (IROC)	63	4 × 7.5
Medium (OROC)	64	6 × 10
Long (OROC)	32	6 × 15

The readout chambers are normally closed by the gating grid. It is only opened by a trigger for the duration of one drift time. Due to this gating, only electrons of the desired event can enter the amplification region and the large amount of ions produced there is prevented from reaching the drift volume, which avoids the accumulation of space charge.

The Transition Radiation Detector

The main purpose of the Transition Radiation Detector (TRD) [75] is the e/π -separation at momenta in excess of $1\text{ GeV}/c$. Due to its position at a larger radius than the TPC, the TRD improves the resolution of the momentum measurement. In addition, the TRD is a fast tracker allowing for a level 1 trigger based on tracks with PID information.

The TRD consists of radiators and drift volumes with MWPCs, filled with a gas mixture of 85% Xe and 15% CO₂. It is arranged in 18 supermodules. If a charged particle with $\gamma \gtrsim 1000$ traverses the radiator, transition radiation may be emitted. Due to their small mass, only electrons have such high γ values up to high momenta, which allows for discriminating them from other particle species.

The Time Of Flight Detector

The Time Of Flight (TOF) detector [76] provides time of flight measurements with a resolution of about 80 ps [70]. This renders PID possible for momenta up to $2.5\text{ GeV}/c$ for pions and kaons and up to $4\text{ GeV}/c$ for protons. The modules of the detector are Multi-gap Resistive Plate Chambers (MRPC) and arranged in 18 supermodules.

2. Experimental Access

The Other Central Barrel Detectors

Other detectors in the central barrel are the High Momentum Particle IDentification (HMPID) detector, the PHOton Spectrometer (PHOS) and the Electro-Magnetic Calorimeter (EMCal).

The HMPID detector is a Ring Imaging CHerenkov (RICH) detector, that is used for hadron identification up to momenta of about $5 \text{ GeV}/c$ [77].

The PHOS has been designed for the measurement of low- p_T direct photons and for the study of jet quenching via high- p_T π^0 and γ -jet correlations [78]. The PHOS also contributes to the trigger, e.g. it provides a trigger on high- p_T photons. It is an electromagnetic calorimeter made of lead-tungstate (PbWO_4) crystals.

Finally, the EMCal [79] improves the jet quenching measurements and the jet energy resolution. It is also capable of measuring high- p_T photons, neutral hadrons and electrons. In addition, it allows for a fast trigger on high-energy jets. The EMCal has been designed as a layered Pb-scintillator sampling calorimeter.

The Muon Spectrometer

The muon spectrometer consists of absorber plates, tracking and trigger chambers and has its own (warm) dipole magnet with a field integral along the beam axis of 3 Tm [80]. The spectrometer is used to measure the spectrum of the ψ and Υ families, as well as that of the ϕ meson via their $\mu^+\mu^-$ decay channels. Besides, the production of open charm and beauty is studied. The muon arm also provides a trigger on high- p_T (di-)muons.

The Trigger and Forward Detectors

There is a set of forward detectors which are located close to the beam pipe: the Zero Degree Calorimeter (ZDC), the Photon Multiplicity Detector (PMD), the Forward Multiplicity Detector (FMD), the V0 detector and the T0 detector. These detectors allow for fast global event characterisation or for triggering [81].

The V0 detector consists of two plastic scintillator discs, V0A and V0C, mounted 340 cm and 90 cm away from the interaction point, respectively, on opposite sides. V0A and V0C cover the forward rapidity ranges $2.8 < \eta < 5.1$ and $-3.7 < \eta < -1.7$, respectively.

Finally, the ALICE COsmic Ray DEtector (ACORDE) is mounted on top of the L3 magnet and can be used to trigger on cosmic rays [82]. It consists of an array of plastic scintillators.

2.2.2. Software for Simulation, Reconstruction and Analysis

The data recorded by ALICE is processed within the software framework AliRoot [83], which is based on ROOT [84, 85, 86]. It comprises means to process data

and to carry out Monte Carlo simulations. In the latter case, primary collisions and particle production can be simulated via a variety of event generators, which includes e.g. PYTHIA [87] and HIJING [88]. The generated particles are then transported through the detector via GEANT3 [89], GEANT4 [90] or Fluka [91].

The recorded raw data is not used directly for physics analyses because it would mean to repeat many CPU-intensive processes for each analysis. Instead, the raw data is reconstructed and calibrated. The event reconstruction in the central barrel [70], which is schematically shown in Fig. 2.6, starts with the conversion of the raw data of individual detectors into clusters. After a preliminary vertex position has been determined from the SPD, the TPC track finding is performed starting with seed clusters at the outer TPC radius. TPC clusters at smaller radii are attached to the tracks via a Kalman filter technique [92]. If present, ITS hits are matched to the TPC tracks. Afterwards, the tracks are propagated outwards to TRD, TOF, EMCAL, PHOS and HMPID. A final inward refit allows for obtaining the ultimate precision. Subsequently, secondary vertices from photon conversions and decays of strange hadrons (K_S^0 , Λ) are reconstructed, followed by the search for cascades (Ξ^\pm , Ω^\pm). The former are denoted as V0 candidates.

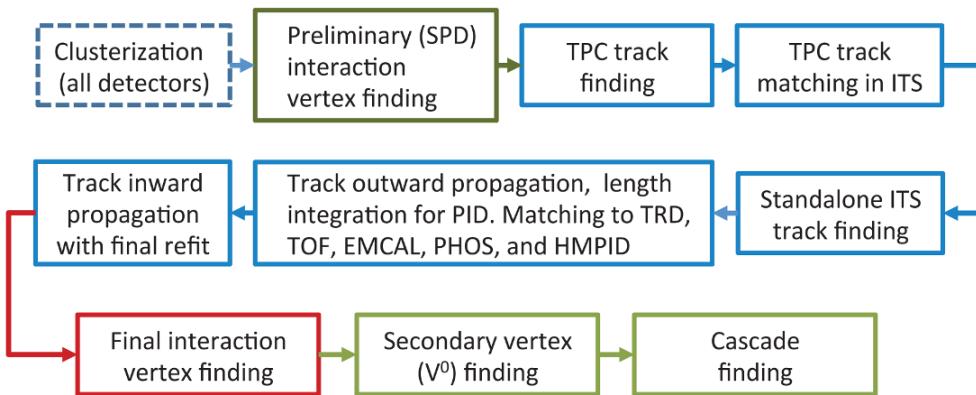


Figure 2.6.: Event reconstruction flow for the central barrel. Figure from [70].

The result of the reconstruction is stored as *Event Summary Data* (ESD) and includes for each event information about primary and secondary vertices, the tracks with (calibrated) PID results from various detectors, etc. The data can be compressed further to *Analysis Object Data* (AOD). The AOD is created from ESD in a filtering process, which selects certain classes of events and stores only tracks that pass a specific set of cuts. Furthermore, AOD may contain advanced objects like jets reconstructed with different algorithms. Hence, running analyses on AOD reduces the I/O overhead and has the benefit of quickly selecting objects like tracks with a filter mask instead of applying all selection criteria. On the other hand, using ESD is more flexible with respect to changing cuts or e.g. the jet finding parameters. Analyses can be run on both inputs, ESD and AOD.

2.3. Jet Reconstruction

Intuitively, one thinks of a jet as a spray of particles that is collimated inside a cone of certain radius. Practically, jets are not well-defined objects. A definition is required that meets experimental and theoretical needs and is valid for all levels, i.e. parton level on the theory side and detector level on the experimental side. However, it proves difficult to define jet finding algorithms fulfilling the list of requirements — see [93] for further discussion. For instance, the result of the jet finding must not change, if soft particles are emitted or collinear radiation takes place. The former demand is called *infrared safety*, the latter *collinear safety*. Algorithms not fulfilling these requirements may lead to divergent cross-sections at higher order perturbative calculations.

The issue of InfraRed and Collinear (IRC) (un-)safety can be demonstrated with a naive seeded cone-type algorithm. This algorithm takes seeds — e.g. the tracks with the highest momentum —, uses the 4-momentum of the seed as a trial direction for the cone and clusters all particles into a jet with some radius R (in the $\eta\phi$ plane) around this trial direction. The 4-momenta of all particles in the jet are then summed up and used as a new trial direction. The procedure is iterated until the trial direction of the cone does not change any longer, i.e. a stable cone has been found. The particles associated to this jet are removed from the further procedure and the algorithm continues with the next jet.

At first glance, such an algorithm may appear to be robust. However, it is neither infrared, nor collinear safe as illustrated in Fig. 2.7. If there are two tracks with rather high momentum and with distance larger than the radius parameter R used for the jet finding, the algorithm yields two jets. If, however, an arbitrarily soft particle is added in the middle of these tracks, such that the distance from each track is smaller than R , this can change the jet finding result: the soft particle may be used as a seed resulting in only one found jet. Hence, such an algorithm is infrared unsafe.

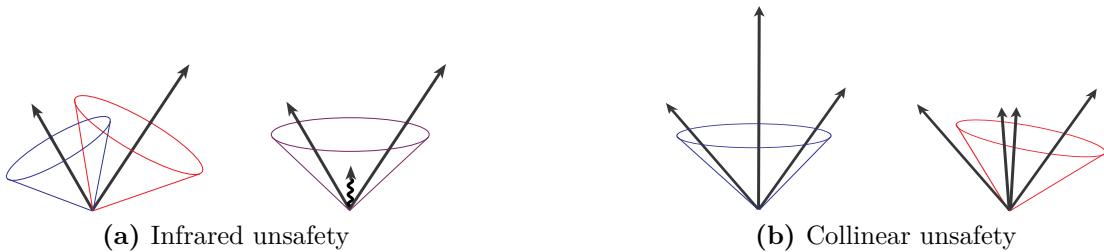


Figure 2.7.: Illustration of an infrared and collinear unsafe jet finder. The result of the jet finding is changed if **a)** an arbitrarily soft particle (curly line) is added (infrared unsafe) or **b)** the leading track is split into two collinear tracks with similar momenta (collinear unsafe), cf. discussion in the text. Figure from [93].

To mitigate the problem in the above scenario, the algorithm can be implemented to take the seed with the highest momentum first. But this will lead to collinear unsafety as can be demonstrated with the scenario in Fig. 2.7b. There are three tracks, the one in the middle being less than R away from the other two and having the largest momentum. The distance between the outer tracks shall be slightly below $2R$. Furthermore, the track on the left shall have the smallest momentum. In this case, the algorithm will choose the track in the middle as seed and find a single jet containing all tracks. If the track in the middle is split into two almost collinear tracks with similar momenta (this may happen due to detector deficiencies or also due to QCD splitting processes), then the track on the right has the highest momentum and becomes the first seed. The resulting jet will not contain the track on the left, which becomes a second jet for its own.

Over the years, there has been great progress in improving jet finding algorithms. Usually, they can be applied to both hits in calorimeter cells and charged tracks. An example for an IRC safe cone-type algorithm is *SISCone* [94], which is a seedless algorithm with much improved performance with respect to processing time. Besides the cone-type algorithms, cluster-type algorithms are widely used at present. The latter cluster objects with respect to their distance d defined as:

$$\begin{aligned} d_{iB} &\equiv (p_{T,i})^{2p}, \\ d_{ij} &\equiv \min \left\{ (p_{T,i})^{2p}, (p_{T,j})^{2p} \right\} \frac{\Delta R_{ij}^2}{R^2}, \\ \Delta R_{ij}^2 &\equiv (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2. \end{aligned} \quad (2.1)$$

This distance measure combines the (transverse) momentum ($p_{T,i}$) with the geometrical (ΔR_{ij}) scales. d_{iB} is the distance of object i to the beam, d_{ij} is the distance between objects i and j , R as a distance parameter and the value of p specifies the algorithm. It turns out that only the sign of p matters, but not the absolute value (except for zero). The k_T algorithm² [95, 96] uses $p = 1$, $p = 0$ is utilised by the *Cambridge-Aachen algorithm* [97] and the *anti- k_T algorithm* [98] is defined with $p = -1$. In any case, the procedure is as follows. All objects are put into a list. The distances between two objects and between object and beam are calculated for every object in the list. If d_{iB} is the smallest distance, object i is considered as a jet and removed from the object list. If d_{ij} is smallest, the objects i and j are merged to a new object (in case of tracks, usually the 4-momenta are added) in the list and the old daughter objects are removed. This procedure is iterated until the object list is empty.

At present, the anti- k_T algorithm is commonly used for jet finding in high-energy collisions. To estimate the underlying soft event in heavy-ion collisions, the k_T algorithm is a typical choice. In the following, the focus will be on these two types.

²Sometimes the transverse momentum is denoted by k_T instead of p_T , which is the origin of the algorithm's name.

2. Experimental Access

The behaviour of the k_T and anti- k_T algorithms is illustrated in Fig. 2.8³. A parton level event was clustered with these algorithms and about 10^4 random soft ghost particles have been added to the event. Since these algorithms are infrared (and collinear) safe, the ghost particles do not affect the jet finding process. But they allow to determine the actual areas of the reconstructed jets, which is indicated by the coloured cells. Note that also in data ghosts can be inserted as tracer particles to investigate effects on the jet reconstruction. Similarly, a Monte Carlo pp event can be embedded into a measured Pb–Pb events to study the impact of the underlying event on the jet reconstruction.

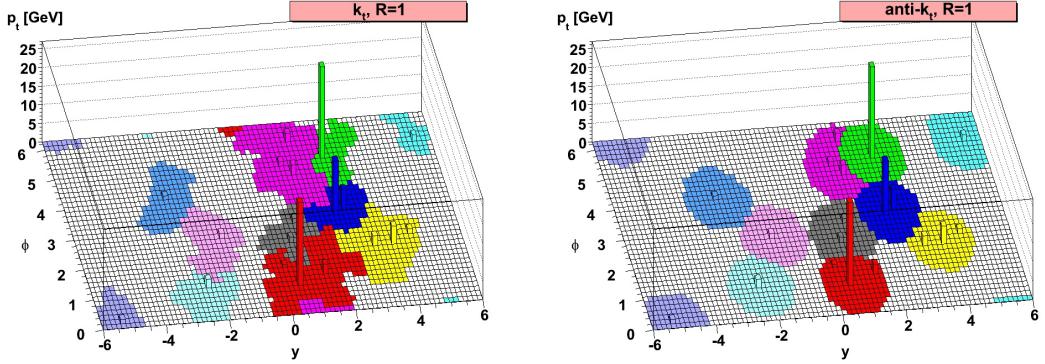


Figure 2.8.: The k_T (left) and anti- k_T (right) jet reconstruction algorithms are illustrated for a sample parton level event (generated with HERWIG [99]). By adding many random soft ghost particles, the jet areas can be determined (coloured cells). Figure from [98].

In contrast to the k_T algorithm, the anti- k_T algorithm yields circular hard jets, which is beneficial for the analysis of e.g. area-related properties of jets. Furthermore, circular jets somehow meet the naive expectation of the jet shape. Note that in case of two overlapping jets, the anti- k_T algorithm leads to a circular shape only for the harder one and clips part of the area of the soft one. This is clearly visible in the figure for the circular green area neighbouring the magenta crescent.

2.3.1. Jet Reconstruction in ALICE

In ALICE, there are two ways to reconstruct jets. The first one deals with *charged jets*, which are based on (charged) tracks reconstructed with the TPC and the ITS. In this case, full azimuthal acceptance is provided for tracks with $|\eta| < 0.9$. Yet, these do not contain any information about neutral particles, i.e. the jet

³The figure shows the rapidity y instead of the pseudo-rapidity η on one axis. Correspondingly, η has to be replaced by y in Eq. 2.1.

energy is unknown — the reconstructed p_T^{jet} can only serve as a lower limit for the full jet energy.

The second way to reconstruct jets is to combine the energy deposition in the EMCAL with the charged tracks from TPC and ITS. This provides access to *full jets* including information about the neutral energy fraction. However, this method although has caveats compared to the charged jets. Firstly, the full jet acceptance is limited by that of the EMCAL to $|\eta| < 0.7$ and $\Delta\phi = 107^\circ$, which decreases statistics considerably. Secondly, a subtraction method needs to be applied to the jet energy to avoid double-counting of the contribution of charged tracks that also deposit energy in the calorimeter. This requires a careful matching between charged tracks and calorimeter hits.

In LHC run 2, ALICE will record EMCAL-triggered jets, which should provide sufficient statistics for full jet studies. Furthermore, the PHOS detector was expanded by a Dijet Calorimeter (DCal) [100] during the long shutdown 1 in 2013 and 2014: 3 DCal supermodules were placed adjacently to 3 PHOS supermodules on either side, such that they are opposite in azimuth to the EMCAL. The resulting acceptance is $|\eta| < 0.7$ and $\Delta\phi = 60^\circ$. This detector configuration increases the acceptance for back-to-back hadron-jet and γ -jet measurements and allows to study dijets with good acceptance for a jet finding with resolution parameter $R \leq 0.4$.

3. Particle Identification with the ALICE TPC

A charged particle traversing a gas deposits energy in it via ionisation. The energy deposition per unit path length dE/dx in a gas is described by the Bethe-Bloch function [101]. The ALICE TPC [73] measures this energy deposition for particle identification and reconstructs the track. Its dE/dx resolution of about 5% for pp collisions allows to discriminate between different particle species from momenta as low as a few hundred MeV/c up to above 20 GeV/c.

3.1. Overview

For a given gas, the Bethe-Bloch function describing the specific energy loss dE/dx of a charged particle only depends on $\beta\gamma = p/(mc)$, where p is the particle momentum and m its mass. If the Bethe-Bloch curves are plotted as a function of momentum, the different particle masses of the individual species lead to a scale factor along the abscissa. Due to detector effects and the track reconstruction¹ there are momentum dependent corrections to the pure Bethe-Bloch function.

The dE/dx of the tracks from pp collisions at $\sqrt{s} = 7$ TeV is plotted as a function of the track momentum at the TPC inner wall in Fig. 3.1. The black lines are parametrisations of the Bethe-Bloch curves for the individual species. Whereas at low momenta ($p \lesssim 0.5$ GeV/c) the hadrons are well separated and can be identified via the TPC dE/dx track-by-track, the separation becomes small and approximately constant as a function of momentum in the so-called relativistic rise ($p \gtrsim 5$ GeV/c). At even higher momenta, the pions start to reach the plateau region with constant dE/dx and the separation decreases. At high momentum — and in particular above 2 GeV/c —, a track-by-track identification is not possible, since the dE/dx distributions of different species overlap at a given momentum. Yet, the identified yields can still be determined on a statistical basis, e.g. using the Multi-Template Fit (MTF) described in Chapter 4.

The MTF requires a precise description of the TPC dE/dx response. If this response is known, templates with the expected dE/dx distributions can be generated for each species. Finally, the sum of the templates can be fitted to the

¹At low momenta, protons and kaons have a large dE/dx . Thus, their energy decreases during their passage through the gas, which increases the dE/dx even more. However, the reconstruction algorithm yields only one value for the momentum and there is also only a single value for the dE/dx of the whole track.

3. Particle Identification with the ALICE TPC

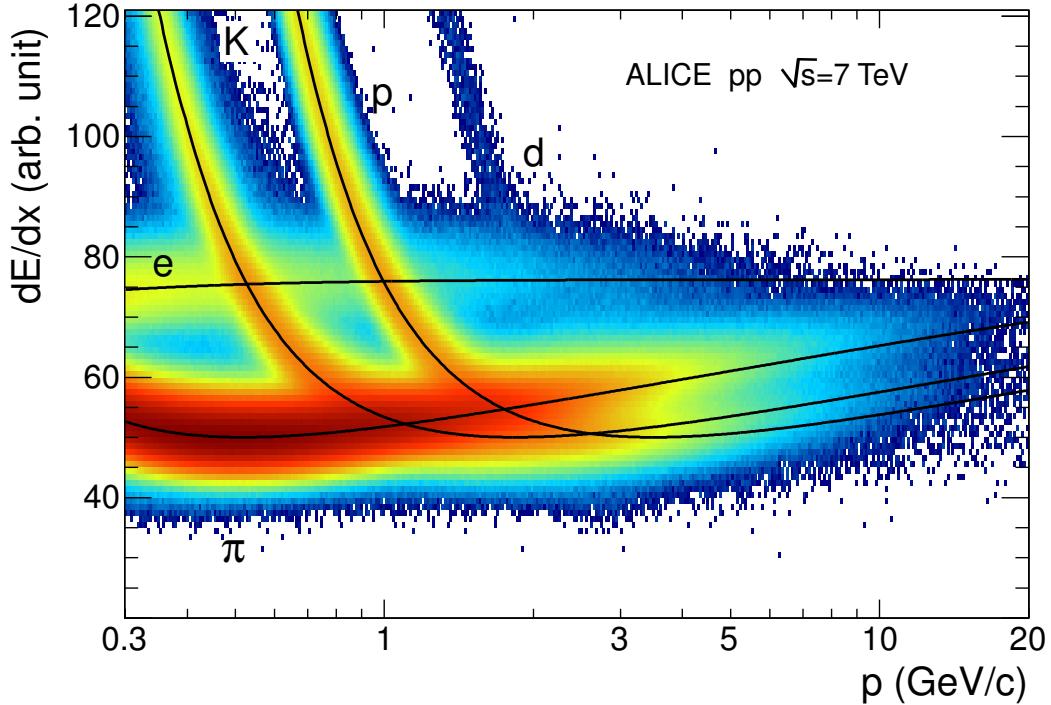


Figure 3.1.: Specific energy loss dE/dx of tracks in the ALICE TPC. The black lines are fits with a Bethe-Bloch parametrisation to the individual particle species.

measured dE/dx distribution in each momentum slice. The amplitudes of the individual templates are then the yields of the corresponding particle species.

The TPC dE/dx response for the MTF is modelled from the following inputs:

- mean dE/dx from the Bethe-Bloch fit (including low- p corrections),
- dependence of dE/dx on η (which corresponds to the polar angle),
- dE/dx resolution and
- dE/dx dependence on multiplicity/occupancy (only relevant for p–Pb and Pb–Pb collisions).

These items will be discussed in the following sections.

In this work, only charged particles will be considered. In particular, neutral particles are *not* taken into account for the jet reconstruction. The final outcome of the analysis will be the fully corrected to-pion ratios and particle yields as a function of $p_{T\text{track}}$ and the momentum fraction z , where it is understood that these quantities are related to charged particles and jets. The results will be confronted with various Monte Carlo (MC) models.

Similarly, the fully corrected to-pion ratios and particle yields will be studied in inclusive events in different bins of multiplicity. The evolution of these quantities with multiplicity will be compared to p–Pb and Pb–Pb collisions.

3.2. Energy Loss of Charged Particles in Matter

If a charged particle propagates through matter, it deposits energy via ionisation. The average energy loss per unit path length of a heavy particle traversing material was calculated for the first time by Bethe [101]. It reads:

$$\langle dE/dx \rangle = \frac{4\pi Ne^4}{m_e c^2} \frac{z^2}{\beta^2} \left(\ln \left(\frac{2m_e c^2}{I} \beta^2 \gamma^2 \right) - \beta^2 - \frac{\delta(\beta)}{2} \right). \quad (3.1)$$

In this equation, N is the electron number density of the traversed matter, e the elementary charge, m_e the electron rest mass, z the charge of the propagating particle, β its velocity in terms of the speed of light c , γ the corresponding Lorentz factor and I the mean excitation energy of the atom. The values of I were calculated by Bloch using the Thomas-Fermi theory of the atom. Eq. 3.1 is called *Bethe-Bloch formula*. With increasing $\beta\gamma$, $\langle dE/dx \rangle$ first decreases with $1/\beta^2$ until it reaches the *minimum ionising region*, which is typically at $\beta\gamma \approx 3.6$, and finally rises again due to the logarithmic term. The latter region is called the *relativistic rise*. The origin of this effect is twofold. Firstly, the maximum transferable energy from particle to atom increases with $\beta\gamma$, which also causes a higher mean energy loss. Secondly, the Lorentz contraction increases the electromagnetic field in the transverse direction leading to a higher cross-section for excitation and ionisation.

The $\delta(\beta)$ term was introduced later by Fermi [102] to take into account the *density effect*: the surrounding atoms are polarisable and shield the field of the propagating particle. The exact behaviour of $\delta(\beta)$ depends on the material, but in general it is a linear function of $\ln \gamma$ for $\beta \rightarrow 1$ [103]. It cancels partly, but not completely the relativistic rise.

The Eq. 3.1 is only valid for particles heavier than electrons, since the latter have different kinematic limits and are identical to their collision partners. However, even for heavier particles the formula cannot be used to calculate track ionisation in drift chambers because the formula includes all high-energy transfers that are kinematically possible, but do not contribute to a track. This happens if the transferred energy is larger than a certain threshold: the electron is knocked out of the gas atom (δ -electron) and forms a second track. If the range of the δ -electron in the gas is sufficiently high, it can be recognised as a new track and its ionisation is no longer attributed to the original track. The actual energy threshold above which this happens depends on the gas and the detector electronics. As a consequence, the energy loss for tracks in drift chambers is restricted by an effective cut-off E_{\max} , such that the correspondingly modified Bethe-Bloch formula reads [103]:

$$\langle dE/dx \rangle_{\text{restricted}} = \frac{4\pi Ne^4}{m_e c^2} \frac{z^2}{\beta^2} \left(\ln \left(\frac{\sqrt{2m_e c^2 E_{\max}} \beta \gamma}{I} \right) - \frac{\beta^2}{2} - \frac{\delta(\beta)}{2} \right). \quad (3.2)$$

3. Particle Identification with the ALICE TPC

The common cut-off E_{\max} replaces the different kinematic limits of diverse species and extends the validity of this formula also to electrons. It can be shown (see [104] and references therein) that the restricted energy loss exhibits a complete cancellation of the relativistic rise: for $\beta \rightarrow 1$, the energy loss approaches a constant value, which is called *Fermi plateau*. Since β can be expressed in terms of $\beta\gamma$ via

$$\gamma = \frac{1}{\sqrt{1 - \beta^2}} \Leftrightarrow \beta = \frac{\beta\gamma}{\sqrt{1 + (\beta\gamma)^2}}, \quad (3.3)$$

the restricted energy loss depends only on the charge and the $\beta\gamma$ of the propagating particle for a given detector system.

It is noteworthy that the separation between Fermi plateau and the minimum ionising region strongly depends on the used detector material. In gas detectors, the separation is usually sufficient to distinguish between particles in these regions, whereas it is too small for this purpose in solid state detectors like silicon.

3.3. Technical Aspects

Experimentally, a track in a drift chamber is reconstructed from a set of ionisation clusters. In case of the ALICE TPC, there are up to 159 such clusters corresponding to the number of pad rows in the readout plane. In each pad row, some charge ΔQ deposited by the track via ionisation is detected. Doing some calibration, the deposited energy ΔE in each pad row can be calculated from ΔQ . Taking into account further corrections due to e.g. the inclination angle, the effective track length Δx over the pad row can be determined [105]. The result is a single measurement of the energy loss per unit path length $\left(\frac{\Delta E}{\Delta x}\right)_i$ for cluster i . The probability for the deposited energy follows a Landau distribution. Accordingly, $\left(\frac{\Delta E}{\Delta x}\right)_i$ fluctuates from pad row to pad row. To make use of all available information and to assign a single dE/dx value to the whole track, the results of all n clusters of the track need to be combined in a reasonable way. The most straight forward estimator for the track dE/dx would just be the average of the n energy loss values. In fact, this is a bad estimator because the underlying ionisation distribution, the Landau distribution, has neither a finite mean, nor a finite variance (see [104] for further discussion). This leads to large fluctuations from track to track for such an estimator. A reasonable choice of a good estimator is the *truncated mean* $\langle S \rangle_\alpha$. It is characterised by a parameter $\alpha \in]0, 1]$ and defined as the average of the αn lowest values among the n values $\left(\frac{\Delta E}{\Delta x}\right)_i$:

$$\langle S \rangle_\alpha \equiv \frac{1}{m} \sum_{i=1}^m \left(\frac{\Delta E}{\Delta x} \right)_i, \quad (3.4)$$

where m is the integer closest to αn and $\left(\frac{\Delta E}{\Delta x}\right)_i \leq \left(\frac{\Delta E}{\Delta x}\right)_{i+1} \forall i \in [1, n - 1]$. The value of α is chosen to maximise the separation power², which yields $\alpha = 0.6$ [105]. It turns out that the dE/dx distribution determined with the truncated mean method with such a value for α exhibits at least approximately a Gaussian shape. The additional tail of the distribution can be parametrised as will be discussed in Section 4.2.1.

Due to these beneficial properties, the truncated mean $\langle S \rangle_{\alpha=0.6}$ is used as the ALICE TPC dE/dx signal. Hereafter, the TPC dE/dx always refers to this truncated mean definition.

3.4. Bethe-Bloch Fit

The mean TPC dE/dx as a function of $\beta\gamma$, $\mu(\beta\gamma)$, can be described by a parametrisation of the Bethe-Bloch function in Eq. 3.2. In this work, the so-called *Lund* model [106] with parameters a_0, \dots, a_4 is used for μ_0 and an additional saturation term with parameter $a_5 > 0$ is introduced (*saturated Lund* model):

$$\mu(\beta\gamma; a_0, \dots, a_5) = \exp\left(-\frac{a_5}{\mu_0(\beta\gamma; a_0, \dots, a_4)}\right) \mu_0(\beta\gamma; a_0, \dots, a_4), \quad (3.5)$$

$$\begin{aligned} \mu_0(\beta\gamma; a_0, \dots, a_4) &= \frac{a_0}{\beta^{2a_3}} + \frac{a_1}{a_2} \ln \frac{(1 + \beta\gamma)^{a_2}}{1 + d(a_0, a_1, a_2, a_4) \cdot (1 + \beta\gamma)^{a_2}}, \\ d(a_0, a_1, a_2, a_4) &= \exp\left(\frac{a_2(a_0 - a_4)}{a_1}\right), \end{aligned} \quad (3.6)$$

where Eq. 3.3 may be used to express β in terms of $\beta\gamma$. The saturation weight with its additional parameter adds an extra degree of freedom to better describe the data. Mainly the low $\beta\gamma$ region, where the dE/dx falls steeply with $\beta\gamma$, is affected by the saturation weight.

To extract the parameters of μ in Eq. 3.5, data points over a broad $\beta\gamma$ range are required. Since the recorded data statistics decreases exponentially with rising momentum, it is necessary to select different particle species in certain momentum ranges where these species can be clearly identified. To identify the particles, the following techniques are applied:

- TPC: If the TPC dE/dx separation is large, a broad inclusion cut (typically 4σ) can be applied with respect to the (roughly) estimated mean dE/dx .
- TOF (+ TPC): After a selection of tracks via TOF (3σ inclusion cut), an additional TPC cut as in the TPC only case is applied.

²In this case, the separation power is defined as the dE/dx distance between minimum ionising and the plateau divided by the average dE/dx resolution in these two regions.

3. Particle Identification with the ALICE TPC

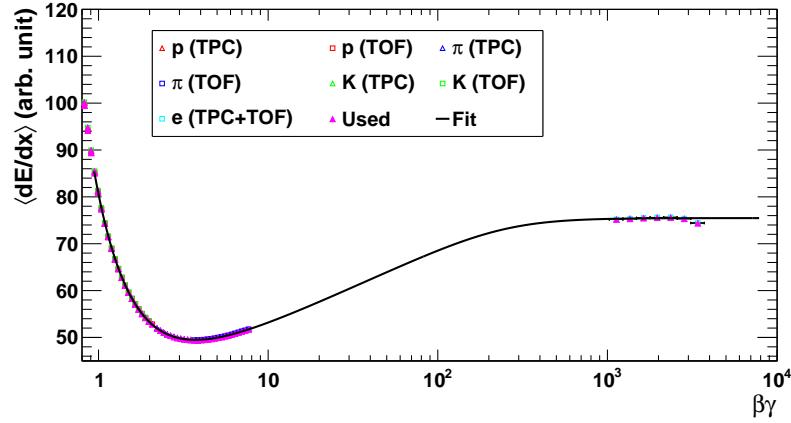
- V0: Tracks are selected solely based on topology (γ conversions and weak decays of K_S^0 , Λ and $\bar{\Lambda}$).
- V0 + TOF: V0 tracks (based on topology) are selected and an additional TOF selection (3- σ inclusion cut) is imposed.
- MC ID: In case of MC, tracks can be selected solely via their MC ID.

More details of each technique will be discussed in the following. TPC only can be used for protons up to approximately $0.6 \text{ GeV}/c$, for kaons up to $0.3 \text{ GeV}/c$ and for pions up to $0.5 \text{ GeV}/c$. Above these thresholds, TOF + TPC is used. For electrons, only momenta above $2 \text{ GeV}/c$ and only TOF + TPC are used for a first iteration without V0's. Note that the actual thresholds slightly depend on the data set and are tuned individually.

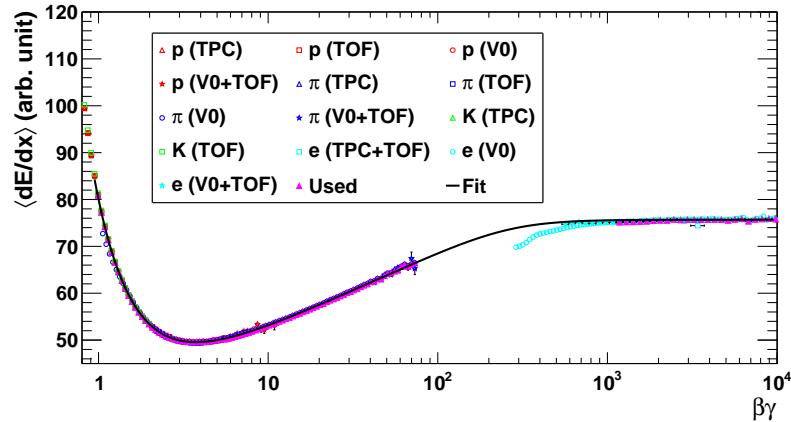
Having selected pure samples the dE/dx distribution in each momentum slice is fitted with a Gaussian. The mean values as a function of the momentum can be converted to a function of $\beta\gamma$ using the corresponding particle mass. Note that the TPC dE/dx response is not a pure Gaussian, but has a tail towards larger dE/dx as will be discussed in Section 4.2.1. Thus, the Gauss fit is restricted to the region around the maximum. This range is spanned by the first and the last bins in the dE/dx histogram having at least a content of 10% of the value of the maximum. The maximum is defined here as the average bin content of the bin with the largest content and its next neighbours. This averaging is done to smooth possible fluctuations.

For the Bethe-Bloch fit, momentum slices with less than 10 entries for the considered species or with relative error larger than 2% are excluded. The result for the fit without V0's in pp collisions at $\sqrt{s} = 7 \text{ TeV}$ is shown in Fig. 3.2a. The data points from samples with different identification techniques are close to each other, such that typically no difference is visible for the chosen scale of the plot. The magenta triangles indicate the data points that are used for the fit. As can be seen, there are no non-V0 data points in the $\beta\gamma$ range 10 to 1000. Consequently, the fit without V0 information is not constrained in this region. In addition, there might be small biases due to impurities even after TOF selection, which is not acceptable for a high precision analysis as the MTF. Nevertheless, the result already gives a good description of the mean dE/dx up to the minimum ionising region around $\beta\gamma \approx 3.5$ and reasonably well describes the plateau ($\beta\gamma \gtrsim 1000$). This is visible from the fit (black line) to the data points. Note that at this stage an accuracy of order several per cent is acceptable, as will be discussed below.

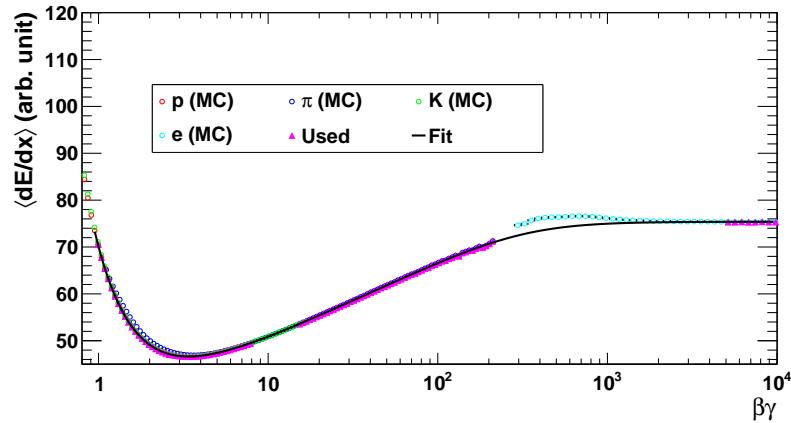
For comparison, the Bethe-Bloch fit for MC is shown in Fig. 3.2c. Since this parametrisation is used officially for TPC PID and no bias should be introduced for other analyses, the species have been selected purely on basis of the MC identity. Evidently, the shape is different from what is observed in data. In particular, this holds true for the dE/dx separation of minimum ionising and plateau.



(a) Bethe-Bloch fit without V0 information



(b) Bethe-Bloch fit with V0 information



(c) Bethe-Bloch fit for MC

Figure 3.2.: Bethe-Bloch fit (*saturated Lund* model) **a)** without and **b)** with usage of V0's. For comparison, **c)** shows the Bethe-Bloch fit for MC. The magenta triangles indicate which data points are used for the fit. The V0 data points included in b) significantly extend the $\beta\gamma$ range with respect to a) (compare magenta triangles). This improves the accuracy for $\beta\gamma > 10$. The impact is not visible in this figure.

3. Particle Identification with the ALICE TPC

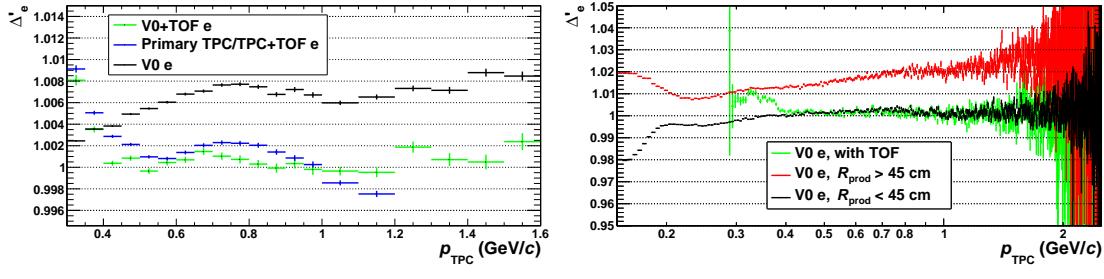
Also note that detector effects bias the dE/dx of low-momentum particles. This causes the discrepancy between fit and electron data points for $\beta\gamma = O(100)$, which is why these data points are not used for the fit. The correction of these effects will be discussed in Section 3.4.3.

3.4.1. Inclusion of V0 Particles

Using V0's (tracks selected via topology) considerably extends the $\beta\gamma$ range in which a pure sample can be selected, see Fig. 3.2b. The comparison of Fig. 3.2a and Fig. 3.2b shows that both relativistic rise ($\beta\gamma$ region between minimum ionising and plateau) and plateau are much better constrained if the V0's are included to the fit. However, the topology of V0 tracks is different from that of primaries, which, together with the η dependence of the TPC dE/dx , can change the mean dE/dx . Therefore, the parameters for μ are extracted without V0 information first. This result is used to create a correction map for the η dependence as will be described in Section 3.5. The η -map is created from the proton sample and depends on the dE/dx . To cover the full dE/dx range (with protons), the η -map creation requires a good mean dE/dx parametrisation only up to the minimum ionising region. Hence, the result from this first fit iteration can be used reliably. As the η dependence is only slowly changing with dE/dx , an accuracy for μ from the first iteration of $O(\%)$ is sufficient. In the next iteration, the parameters are extracted for the η -corrected data including the information from V0's. This procedure (extracting the mean dE/dx for the current η correction, using it to get a new η correction) is iterated until the mean dE/dx of all species does not change by more than 0.1%. Typically, this convergence is already reached after one additional iteration.

For the Bethe-Bloch fit with V0 information, the kaons are not used at all, since there are no V0's for kaons and the purity after TOF selection is rather limited. For protons, only V0 selection is used, whereas for electrons, V0 + TOF gives a more reliable result. The reason is that the TOF requirement for electrons ensures that the V0 topology is not too far away from that of a primary track³ and, thus, the dE/dx is the same. It turns out that the dE/dx of V0 electrons from photon conversions is different from that of V0 + TOF electrons (and primary TOF electrons) if the conversion radius is too large. This observation is demonstrated in Fig. 3.3. The vertical axis shows the dE/dx normalised to some common reference, which is close to the expected electron dE/dx . As can be seen from Fig. 3.3a, the mean dE/dx of primary electrons (blue) and V0 electrons with TOF selection (green) agrees well, whereas that of V0 electrons without additional TOF selection (black) is off by up to 1%. The reason for this discrepancy can be read off from Fig. 3.3b, which splits the V0 only sample into two sets. The red points are for production radii $R_{\text{prod}} > 45 \text{ cm}$ and the black points for

³Technically, the TOF requirement for the considered data set demands an ITS refit, such that only tracks with (matched) ITS hits are accepted.



(a) Primary compared to V0 electrons with and without TOF selection

(b) V0 electrons with different cuts

Figure 3.3.: Comparison of the mean dE/dx normalised to some common reference of primary and V0 electrons. In a), the pure V0 selection without TOF (black) deviates from that with TOF (green), where the latter is in agreement with primary electrons (blue). The reason of the deviation is illustrated in b), which splits the pure V0 electron sample into two sets with different cuts on the production radius R_{prod} .

$R_{\text{prod}} < 45$ cm. Obviously, the mean dE/dx of V0 electrons with smaller production radii agrees well with that of V0 + TOF electrons, which in turn agrees with the one of primary electrons. However, the red data points are biased by more than 1%. The discrepancy between green and black data points at momenta below 0.4 GeV/c is because the TOF efficiency rapidly drops and depends on η , such that the results become unreliable.

In general, V0 electrons have a different topology than primary electrons. Since the η correction has already been applied, the mean dE/dx of primary electrons and V0 electrons with small production radii becomes similar — with possible deviations due to residual η dependence. Also, there might be some rather small bias for primary electrons due to contamination even after the TOF selection. However, if the V0 production radius is large, i.e. $R_{\text{prod}} \gtrsim 45$ cm, the topology is very different from that of primaries. All the corrections that are applied during the reconstruction are optimised for primaries. Therefore, the dE/dx is only reliable if the topology is not too different from primaries. In case of V0 electrons, this is only true if $R_{\text{prod}} \lesssim 45$ cm. It has been checked that for V0 pions and V0 protons the mean dE/dx agrees with that of primaries (selected via TPC and TPC + TOF) usually better than 1% in regions where a very clean primary selection is possible. For pions, this is around $p \approx 0.5$ GeV/c and for protons, this is true for $p \lesssim 2$ GeV/c.

As a result, there is no need for an additional TOF cut on V0 pions and V0 protons. However, for V0 electrons, a cut on the production radius is required, namely it is demanded to lie between 3 and 45 cm. To be on the safe side, the Bethe-Bloch fit takes V0 + TOF electrons at high momenta, which is a stronger criterion than just cutting on the production radius.

3.4.2. V0 Purity

The extraction of the mean dE/dx relies on clean samples from the V0 selection. In this section, the V0 selection will be discussed in more detail. Moreover, the V0 purity will be investigated to estimate possible dE/dx biases.

To ensure a high purity of the V0 candidates, cuts on the invariant mass and on the *Armenteros-Podolanski* variables [107] are applied. Details about these cuts can be found in Appendix C.1. The Fig. 3.4 shows the plot in the *Armenteros-Podolanski* variables of the V0's of which the daughter particles are used for the TPC PID response extraction. Such a plot illustrates the yields as a function of q_T and α . q_T is defined as the transverse momentum of the positively charged particle with respect to the mother particle. α is the asymmetry of the longitudinal daughter momenta with respect to the mother particle. In formulas, this translates to:

$$q_T \equiv p_T^+, \quad (3.7)$$

$$\alpha \equiv \frac{p_L^+ - p_L^-}{p_L^+ + p_L^-}. \quad (3.8)$$

p_L^+ (p_L^-) is the longitudinal momentum of the positively (negatively) charged daughter particle with respect to the momentum of the mother. Similarly, p_T^+ (p_T^-) are the corresponding transverse momenta with respect to the mother momentum.

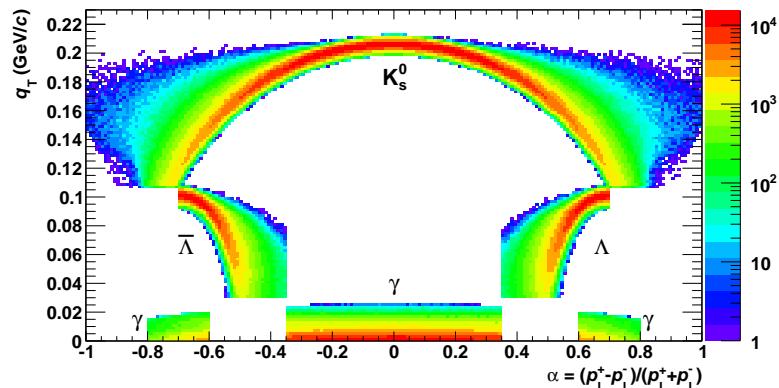


Figure 3.4.: Armenteros-Podolanski plot of V0 mothers of particles used to extract the TPC PID response. The cuts have already been applied (see text). The contributions from K_S^0 , Λ , $\bar{\Lambda}$ and γ are indicated.

The contributions from K_S^0 , Λ , $\bar{\Lambda}$ and γ are explicitly indicated in the figure. The cuts on invariant mass, q_T and α have already been applied. This removes the left (right) branch of $\bar{\Lambda}$ (Λ), which overlaps with the one of K_S^0 . Similarly, the regions in which the γ contribution overlaps with that of Λ or $\bar{\Lambda}$ are removed around $q_T = 0$.

The choice of variables allows for a pretty clean selection of V0 particles. γ 's have zero rest mass, hence their q_T is close to zero. Without cuts, they would be distributed symmetrically about $\alpha = 0$ because the decay particles have equal masses. The same is true for K_S^0 candidates. In contrast, Λ decays to proton and pion, which have very different masses. This shifts the symmetry axis away from $\alpha = 0$ because the proton on average takes a larger part of the momentum (similarly for $\bar{\Lambda}$).

Besides the cuts on the Armenteros-Podolanski variables, a cut on the invariant mass of the decay products is applied. The invariant mass distributions after all cuts are shown in Fig. 3.5. These figures suggest a rather high purity, but do not allow for a quantitative statement. Hence, it is instructive to investigate the dE/dx distributions of the daughter particles of all V0 mothers fulfilling the selection criteria. The resulting dE/dx distributions for V0 electrons, V0 pions and V0 protons are shown in Fig. 3.6.

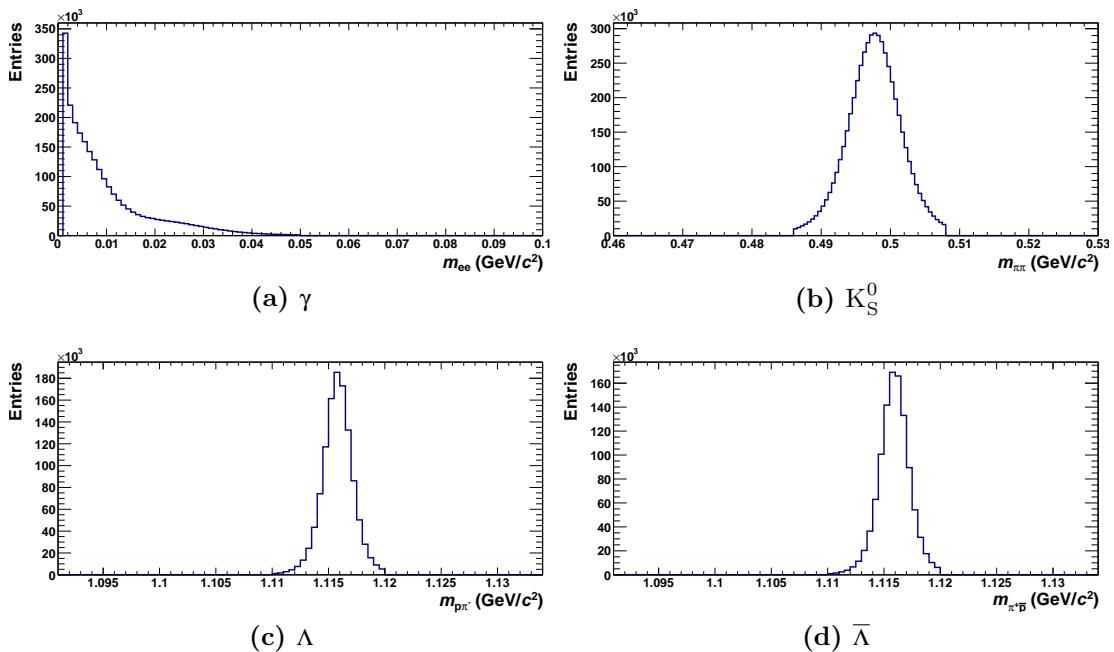


Figure 3.5.: Invariant mass distributions of V0 mothers of particles used to extract the TPC PID response. The cuts have already been applied.

The V0 electrons (Fig. 3.6a) have negligible contamination from kaons and protons. The pion contamination is well visible as the lower dE/dx branch, but the peak height of this branch is a factor 100 lower than that of the electrons. Indeed, comparing the integrals in momentum bins with clean separation, the contamination is typically 1% or less. Due to such a small contamination and the clean separation above $p = 0.3 \text{ GeV}/c$, there is no dE/dx bias expected due to impurities of the sample. Below momenta of about $0.2 \text{ GeV}/c$, the pion and electron distributions largely overlap, such that a bias cannot be excluded, but should be

3. Particle Identification with the ALICE TPC

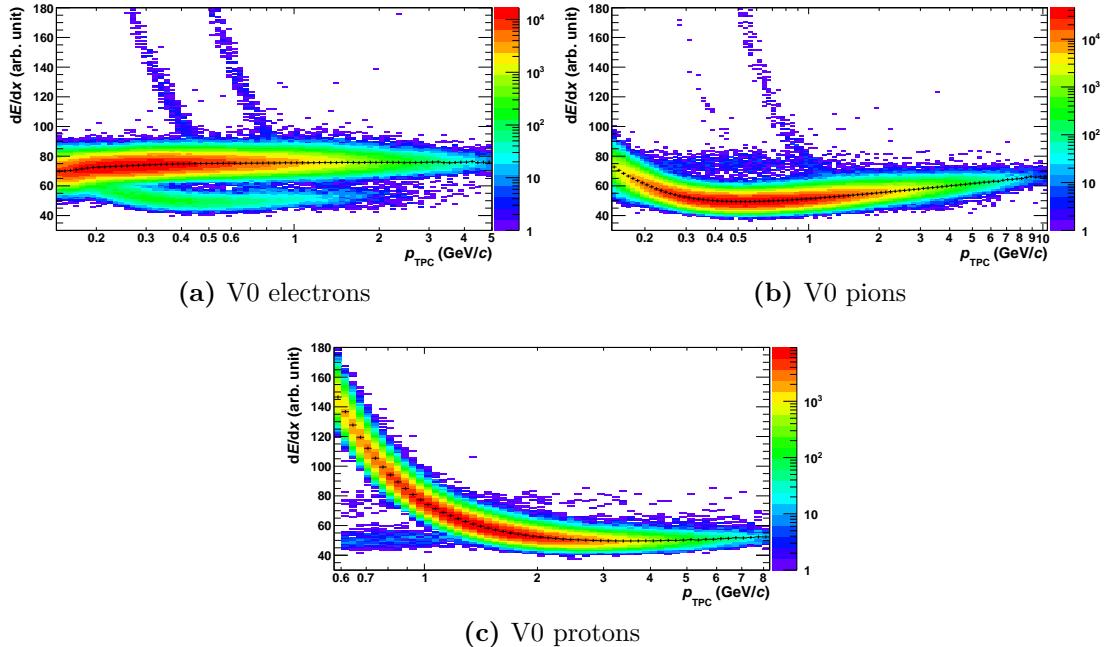


Figure 3.6.: The V0 purity is demonstrated via the dE/dx distributions of the V0 daughters. The black data points indicate the mean of Gauss fits around the peaks.

extremely small with respect to the small pion fraction in the selected sample. The estimated systematic uncertainties, as will be discussed in Section 3.4.5, will take this into account.

For the V0 pions (Fig. 3.6b), the contamination from other species is negligible in the whole momentum range (e.g. around $p = 0.9 \text{ GeV}/c$ there are less than 100 visible counts from non-pion bands compared to about 260,000 counts from pions, which amounts to a contamination of less than 0.1%). In particular, the proton and kaon distributions for $p > 2 \text{ GeV}/c$ seem to be removed completely. As a result, there is no dE/dx bias from contamination expected for the V0 pion sample.

The main contribution to the impurities of the V0 proton sample (Fig. 3.6c) stems from pions and is typically around one per mille (e.g. about 50 pion counts out of more than 63,000 in total around $p = 0.9 \text{ GeV}/c$). Hence, the dE/dx bias due to contamination is expected to be negligible for the V0 proton sample.

In summary, there is no significant dE/dx bias expected from impurities of the V0 samples. The only problematic case are the V0 electrons with momenta below $0.2 \text{ GeV}/c$, which will be taken into account for the systematic uncertainties of the mean dE/dx parametrisation.

3.4.3. Low-Momentum Correction

The Bethe-Bloch curves are solely a function of $\beta\gamma$. However, the TPC dE/dx response is found to have some momentum dependence, especially at momenta of order $0.1 \text{ GeV}/c$. This is illustrated in Fig. 3.7, which shows the relative deviation of data and fit for the individual species. The effect is $20 - 30\%$ for kaons and protons, and below 10% for pions and electrons. The main reason is that for low momenta, the energy loss significantly changes the momentum, in particular for kaons and protons, which have large dE/dx . Thus, the measured dE/dx is averaged over different momenta. In addition, angular effects are affecting the dE/dx . At low transverse momentum, which is implied by low total momentum, the radius of curvature in the magnetic field gets small, which leads to rather extreme azimuthal inclination angles for a given pad row. This changes the effective track length over the pad row modifying the dE/dx .

It turns out that it is required to correct low-momentum effects up to about $1 \text{ GeV}/c$. Since the Bethe-Bloch fit is a function of $\beta\gamma$, this translates into different $\beta\gamma$ ranges for the individual species. The corrections are directly extracted from the residuals in Fig. 3.7 as the blue curves. The open black points give a rough estimate of the quality of the correction. They have been obtained by applying the correction factor (blue curve) at the bin centre of each data point to the whole bin. Since the magnitude of the correction can change significantly within one bin, the open black points do *not* give an estimate of the systematic error, but allow to judge whether the functional form is adequate.

The high-purity V0 samples allow to extract the correction factor also for pions and electrons down to $p = 0.15 \text{ GeV}/c$. Note that the electron correction is extracted from the V0 sample without additional TOF selection, since the TOF information is not available for $p_T < 0.3 \text{ GeV}/c$. Instead, the cut on the radius of the production vertex becomes effective.

Next, the result of the Bethe-Bloch fit is multiplied with the low-momentum correction factors for each species giving the mean dE/dx , μ , in bins of momentum. Via the particle mass, the momentum bins can be converted to $\beta\gamma$ bins.

Finally, splines are used to get a continuous function μ of $\beta\gamma$ for each species (i.e. each having a dedicated low-momentum correction). In the following, these functions will simply be called ‘splines’.

The low-momentum correction is used for data and for MC as well. In both cases, the correction factors are only significant in regions with dE/dx well above minimum ionising. Thus, a quality cut on the number of PID clusters is expected to have minor to no influence on the correction factors. Since muons are close to pions in terms of dE/dx for $p > 150 \text{ MeV}/c$, it is not possible to extract a dedicated low-momentum correction for them. As an approximation, the pion splines are taken, but the muon mass is used to translate momentum to $\beta\gamma$.

In the following sections, it is convenient to look at the dE/dx relative to the expectation, defined as:

3. Particle Identification with the ALICE TPC

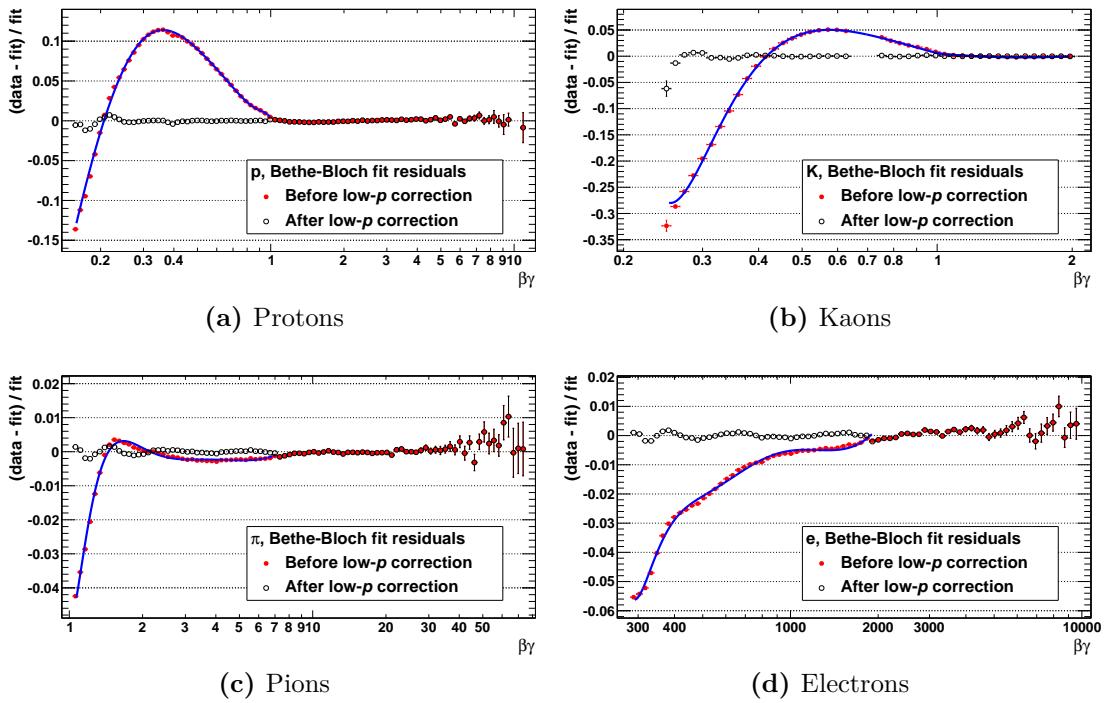


Figure 3.7.: The residuals of the Bethe-Bloch fit for data are shown as red full points. The blue lines are parametrisations of the residuals used to correct for low-momentum effects. The black open points are the difference of the blue lines and the red points.

$$\Delta'_{\text{species}} \equiv \frac{dE/dx}{dE/dx_{\text{expected}}} \cdot \quad (3.9)$$

Note that the expected dE/dx is not just the value of the splines, but includes corrections like for the polar angle as discussed in Section 3.5.

3.4.4. Spline Accuracy

The accuracy of the splines is illustrated in Fig. 3.8, which shows the Δ' of various species for both primary and V0 particles. The horizontal red lines indicate the $\pm 0.2\%$ region around unity. As can be seen, the accuracy of the splines is about $\pm 0.2\%$. The primary kaons (green) and protons (black) start to deviate for $p \gtrsim 1$ and $2.5 \text{ GeV}/c$, respectively. This is the region, where the TOF cannot provide a clean selection anymore and contamination biases the fitted mean dE/dx . There are a few outliers beyond $\pm 0.2\%$, but they have large statistical uncertainties and in such cases the fit becomes unreliable. This concerns V0 electrons (magenta) at high momenta and V0 protons (blue) at very low and high momenta. In summary, the red lines are an envelope for most of the data points, especially they fully include the typical fluctuations.

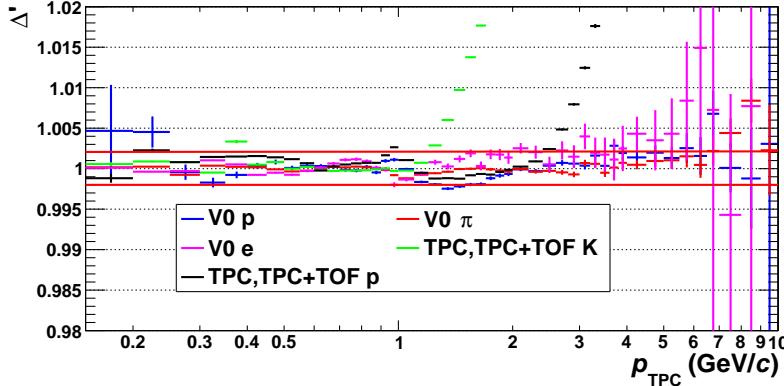


Figure 3.8.: Accuracy of the splines in data. The Δ' on the vertical axis uses as reference the type of the corresponding species. Note that for the data points from primary particles (green and black), contamination shifts the Δ' away from unity above some momentum threshold. The horizontal red lines show the $\pm 0.2\%$ region around unity, which largely covers the most extreme deviations of all species.

3.4.5. Systematic Uncertainty of the Splines

The systematic uncertainty of the splines is rather tricky to estimate, since different error sources contribute. The splines have an intrinsic accuracy of 0.2% as discussed in Section 3.4.4. This accuracy is estimated from clean samples. At high momentum, only V0's are left as a clean sample, but the statistics is too low to draw conclusions from that. Therefore, different models for the specific energy loss are compared: *Saturated Lund* (cf. Eq. 3.5), *Lund* (Eq. 3.5 with $a_5 = 0$) and *ALEPH*. The latter is a slight modification of the parametrisation proposed by Blum and Rolandi [104] for the ALEPH TPC [108] and defined as:

$$\mu(\beta\gamma; a_0, \dots, a_5) = \frac{a_0}{\beta^{a_3}} \left(a_1 - a_2 - a_5 \beta^{a_3} - \ln \left(1 + \frac{1}{(\beta\gamma)^{a_4} e^{a_2}} \right) \right). \quad (3.10)$$

Note that with respect to the original parametrisation the parameter a_2 was redefined for better numerical stability and the parameter a_5 was added (it equals unity in the original formula) to have an additional degree of freedom to adjust the separation between minimum ionising and plateau region. In the following, *ALEPH* refers to Eq. 3.10 with $a_5 = 1$, whereas a_5 floats for *Modified ALEPH*.

In Fig. 3.9, different Bethe-Bloch parametrisations are compared. The low-momentum correction is applied to each parametrisation individually. Except for a small outlier in case of the *Lund* model, all three models are consistent within 2% for $p < 10 \text{ GeV}/c$. The outlier is in the region where the low-momentum correction is switched on. More important is the different relativistic rise of the dE/dx for the *Lund* and the *ALEPH* models: the pions deviate up to about 1.3% around $p = 40 \text{ GeV}/c$. The same holds true for kaons and protons at higher momenta, when they reach a similar $\beta\gamma$.

3. Particle Identification with the ALICE TPC

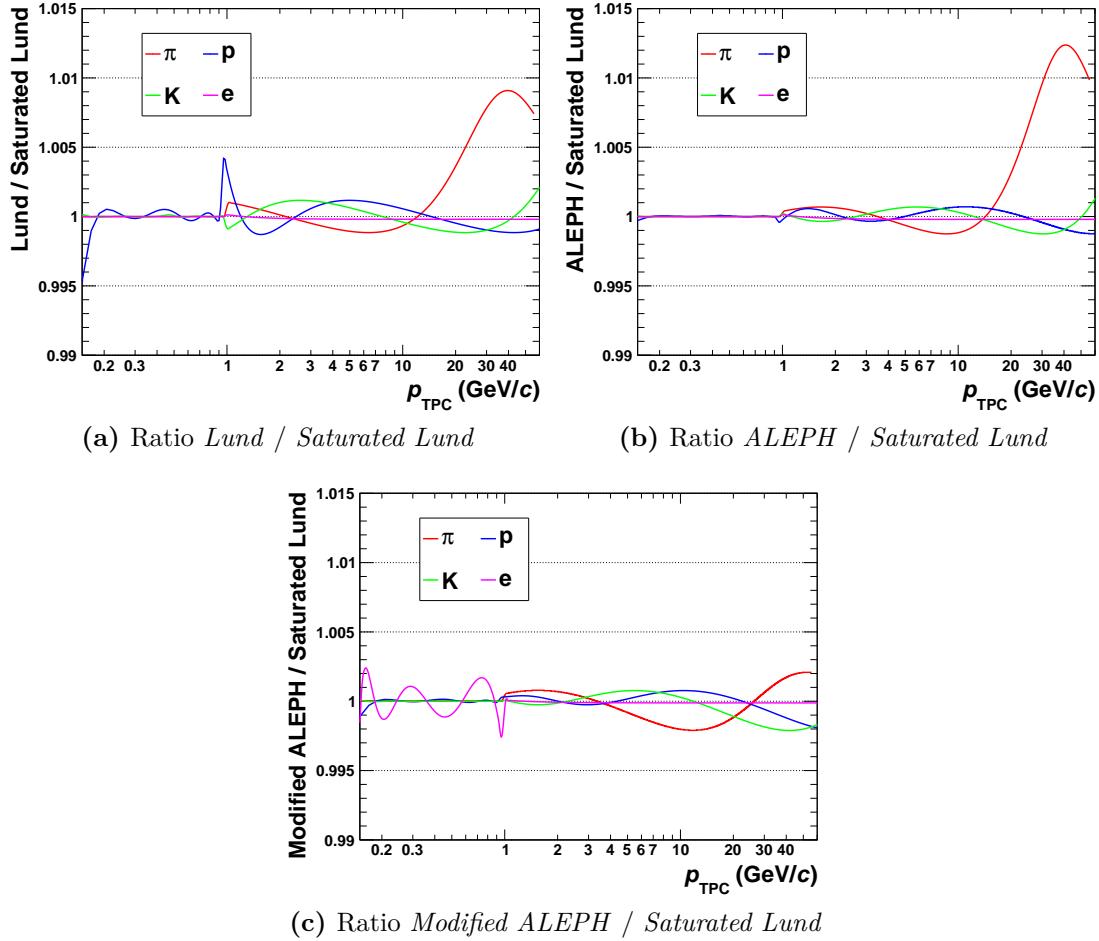


Figure 3.9.: Comparison of **a)** *Lund*, **b)** *ALEPH* and **c)** *Modified ALEPH* Bethe-Bloch parametrisations to the *Saturated Lund* model. Note that the low-momentum corrections are already included for all parametrisations.

Another systematic error source is the “gap” region with $\beta\gamma = \mathcal{O}(100)$, where no data points are present to constrain the Bethe-Bloch shape, see Figs. 3.2b and 3.10a. The V0 pions and V0 electrons approach this gap from both sides. For the latter, low-momentum systematic effects contribute, such that only the data points above $\beta\gamma = 1000$ can be safely used. For the former, the limited statistics causes large statistical uncertainties and fluctuations above $\beta\gamma = 40$. All fit residuals from V0 pions and V0 electrons are within -0.6% and $+1.3\%$ in Fig. 3.10. The lower boundary is dictated by the electrons, but is also a conservative estimate for the pions, which still touch this value with their error bars.

To take into account all these contributions (accuracy of the splines, difference of the models, uncertainty due to the gap), the systematic uncertainty of the splines is estimated as $\pm 0.2\%$ at low and as $+1.3\%$ and -0.6% at high $\beta\gamma$.

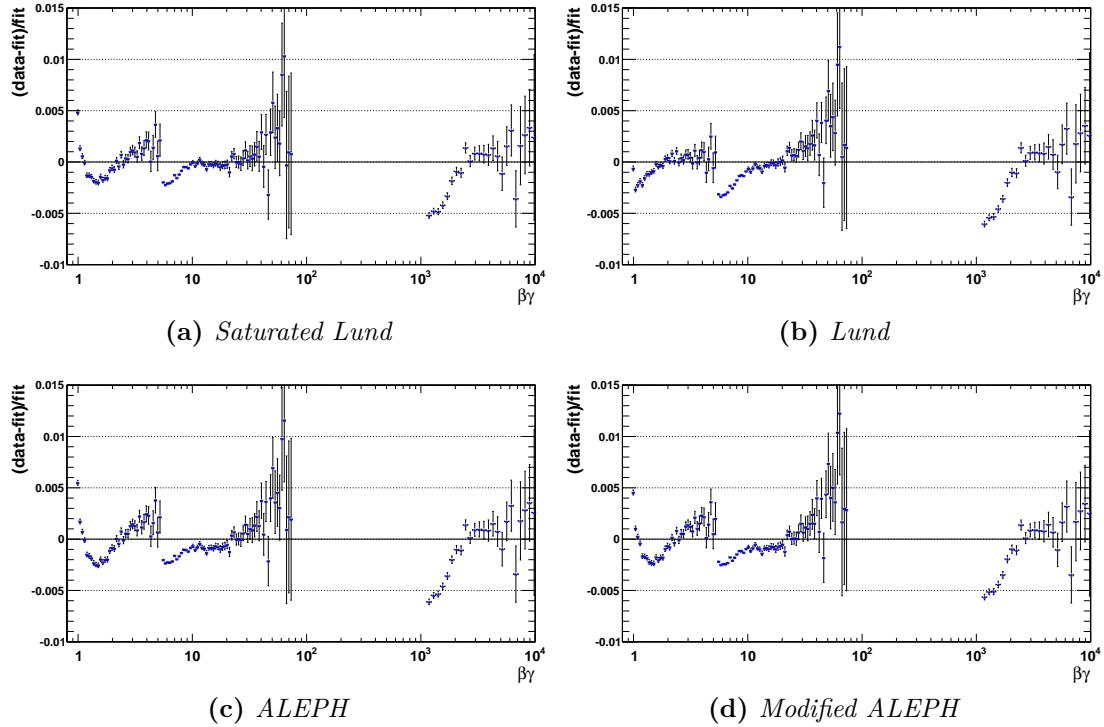


Figure 3.10.: Fit residuals of a) *Saturated Lund*, b) *Lund*, c) *ALEPH* and d) *Modified ALEPH*. The residuals are for the pure fit, without any low-momentum correction.

The uncertainty is cranked up smoothly around $\beta\gamma = 50$, such that roughly the systematic uncertainty for pions grows from 0.2% at $p = 5.6 \text{ GeV}/c$ to 1.3% at $8.5 \text{ GeV}/c$.

Note that also the electrons will pick up the larger error even at low momentum. This is intended because it is rather difficult to get clean electron samples, since the dE/dx distributions of the by far more abundant other species are almost in the whole momentum range touching the electron distribution. This can introduce some bias to the electron splines. On the one hand, using V0's to extract the electron splines strongly reduces such biases (cf. Section 3.4.2). On the other hand, there may be additional uncertainties from momentum dependencies of the mean dE/dx for low-momentum electrons. Both effects should be on the per mille scale, such that assigning the rather large systematic uncertainty due to the gap also to the electrons should be a conservative estimate. Anyhow, since the electron yield fraction is small, even this conservative estimate will have only small impact on the systematic uncertainty of the hadron fractions.

For the mean value of the splines, the *Saturated Lund* model is chosen. According to Fig. 3.10, which shows the residuals of the pure fit without low-momentum correction, *Saturated Lund* describes better the pions in the $\beta\gamma = O(10)$ region — although the difference is $O(1\%)$ — and this might hint to a better description at higher $\beta\gamma$.

3.5. η -Maps

If a particle traverses the TPC at some polar angle θ with respect to the beam line, the electrons created from the gas ionisation diffuse during their drift towards the cathode plane. This changes the induced charge on the single readout pads, which in turn has an impact on the dE/dx resolution and the mean dE/dx . Furthermore, the closer the angle of incidence of the track on the pad plane is to 90 degrees, the larger is the fraction of path length projected onto a single pad row, see Fig. 3.11. Therefore, the deposited charge per pad row is larger for tracks in forward direction than for those at mid-rapidity. This causes the TPC dE/dx of a track to depend on the pseudo-rapidity

$$\eta \equiv 0.5 \ln \frac{p + p_z}{p - p_z} = -\ln \left(\tan \frac{\theta_{\text{global}}}{2} \right), \quad (3.11)$$

where p is the total momentum of the track and p_z the momentum along the beam line (z -direction). More precisely, the dE/dx depends on the tangent of the local polar angle, $\tan(\theta)$, measured at the entrance of the TPC, whereas the pseudo-rapidity is linked to the global polar angle θ_{global} (with explicit subscript). In the local case, $\theta = 0$ is perpendicular to the beam line, whereas in the global case, it is along the beam line. The local θ is the relevant observable for PID calibration, since the global angle can be different due to multiple scattering.

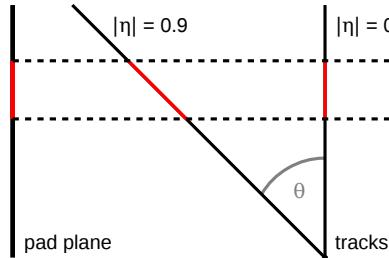


Figure 3.11.: Influence of the local polar angle θ on the effective track length per pad.

A single pad row detects charges from tracks in the region indicated by the dashed lines. The effective track length over this pad row (red) is smaller for a track at mid-rapidity ($|\eta| = 0$) than for a track in forward direction ($|\eta| = 0.9$).

It is observed that the shape of Δ'_{species} as a function of η in given momentum bins is roughly the same for all species, if the momentum bins are chosen such that the mean dE/dx are similar. This is shown in Fig. 3.12, in which the relative change of the dE/dx as a function of η for individual species at different momenta, but similar dE/dx is compared. For each species, primary candidates are selected via TPC+TOF. The TOF requirement leads to almost zero efficiency for tracks with low (total) momentum at large η . This biases the values of bins with large

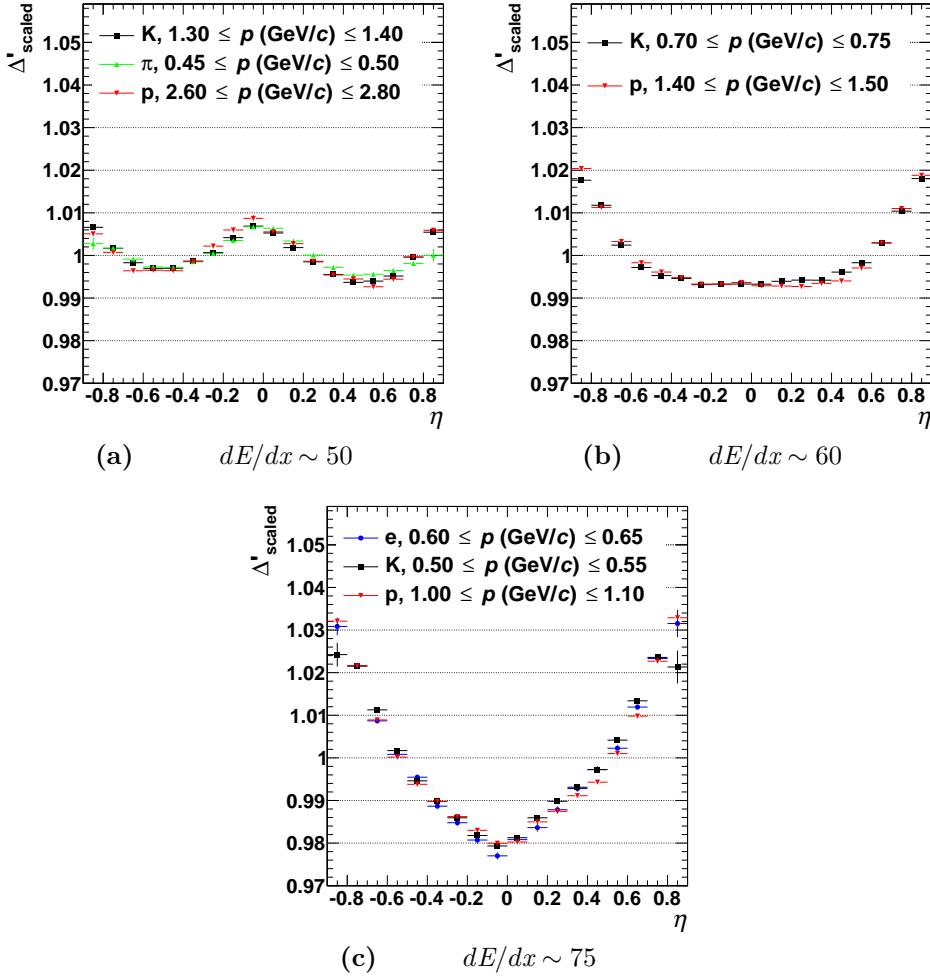


Figure 3.12.: η dependence of dE/dx for different species at different momenta, but similar mean dE/dx . The data points are for primary candidates selected via TPC+TOF. The ordinate shows the corresponding Δ'_{species} scaled such that the integral equals that of a flat distribution at unity.

$|\eta|$ to the values of the neighbouring bins with smaller $|\eta|$ for the data sets with momenta around $0.5 \text{ GeV}/c$ and below.

As can be seen, the η shape varies significantly with dE/dx , but looks similar for the individual species. This means that the shape in good approximation only depends on dE/dx , but not on the momentum or the species. Hence, the η dependence can be extracted from a single species as a function of dE/dx and then be applied to all other species. This results in a map with Δ'_{species} as a function of $\tan(\theta)$ and inverse dE/dx . For the latter dimension, the inverse is taken to have a finer binning at low dE/dx . At high dE/dx , the separation is large for all species and a precise modelling is less important. Additionally, only kaons and protons (and light nuclei like deuterons) have a dE/dx beyond 100

3. Particle Identification with the ALICE TPC

for $p > 0.15 \text{ GeV}/c$. But these species have small detection efficiency (below 1% for the lowest momenta, since they do not reach (sufficiently far in) the TPC), when their dE/dx becomes large, which does anyway not allow to gather sufficient statistics for a fine binning in this region.

The η -map⁴ needs to cover the whole dE/dx range and it must be extracted from a species for which a clean identification is possible. This is true for protons from lowest momenta up to the minimum ionising region around $3 \text{ GeV}/c$. The proton identification is done with different methods for 3 momentum regions:

- TPC (only): Up to $p \approx 0.6 \text{ GeV}/c$, the dE/dx of protons is clearly separated from other species (cf. Fig. 3.1). They are identified solely with the TPC using a momentum dependent dE/dx cut within the gap between the proton and kaon bands.
- TOF (+ TPC): After a selection of tracks via TOF (3σ inclusion cut), an additional TPC cut is applied as in the TPC only case. The resulting sample is typically used up to $p = 2.6 \text{ GeV}/c$.
- V0: Tracks are selected solely based on topology. This sample is used for higher momenta, where the TOF selection may start to suffer from contamination.
- MC ID: In case of MC, tracks can be selected solely via their MC ID.

The following procedure is carried out after the splines with the expected dE/dx have been extracted in the first iteration of the Bethe-Bloch fit, i.e. without using V0's. The Δ'_p distribution in each momentum and $\tan(\theta)$ bin is fitted with a Gaussian to extract the mean value. As for the splines, the fit is restricted around the maximum of the distribution (cf. Section 3.4). Note that the low-momentum correction is already included for the evaluation of Δ'_p — in this step, it is only needed for protons, for which a pure selection at low-momentum without V0's is possible. This procedure yields a map in the space $(p, \tan(\theta))$. Finally, the splines are used to change the first dimension from p to $1/(dE/dx^{\text{expected}})$. For convenience, the superscript “expected” will be dropped in the following for that dimension.

The η -map is shown in Fig. 3.13a. As can be seen, the η dependence is up to 10% around $dE/dx = 100$ and changes its shape as a function of dE/dx . Note that the binning of the η -map depends on the available statistics: the higher the statistics, the finer the binning can be.

The maps have a smooth shape and should not depend too much on the chosen binning. Hence, a refined map with much finer binning is created from the raw map via interpolation, which is shown in Fig. 3.13b. This is done via

⁴More precisely, it should be called “polar-angle”-map, but the name was kept for historical reasons.

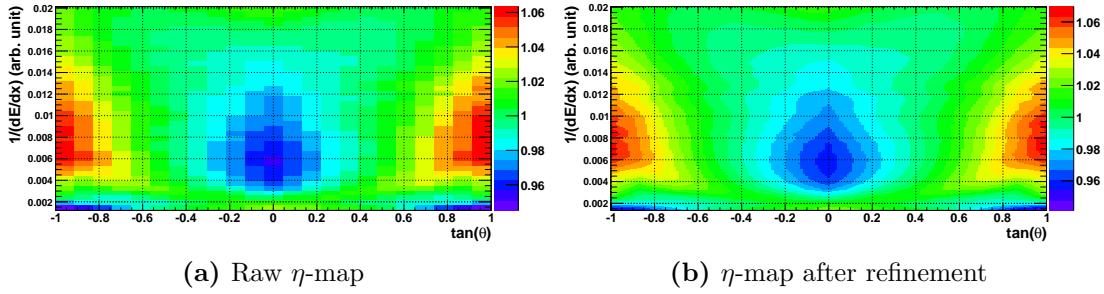


Figure 3.13.: η -map obtained from protons. **a)** shows the raw result from the fits. In **b)**, the refined map using interpolation is shown. The colour code shows Δ'_p .

bilinear interpolation based on the four nearest bin centres, which is provided by ROOT [85]. Instead of using the bin contents of the refined map, it is also possible to interpolate from the raw map every time the η correction is called, which happens several times for each track. It is a trade-off between CPU time and accuracy. For the binning of the refined map, the difference between bin content and interpolated value (from the raw map) is $< 0.1\%$, which is about one order of magnitude better than the estimated uncertainties of the η correction (see below). Therefore, the refined maps are used instead of the interpolation from the raw map.

The splines determine $\Delta' = 1$ in the map, i.e. they “anchor” the map. In order to have as little as possible change of the splines due to the η correction, the map is anchored to the η -averaged splines, that are extracted from all tracks with $|\eta| < 0.9$. The map is used to correct the dE/dx of all particles and extract new splines from it. This procedure is iterated until convergence as already discussed in Section 3.4.1. The anchoring ensures that the change of the splines is minimal in subsequent iterations.

3.5.1. Systematic Uncertainty of the η Correction

In Fig. 3.14, the quality of the η correction is shown. The η correction removes the η dependence of the mean dE/dx at momenta above $2 \text{ GeV}/c$ as desired. Obviously, the residual dependence is similar for pions (red) and protons (blue) at $p = 3$ and $2 \text{ GeV}/c$, respectively. A glance back to Fig. 3.1 shows that the mean dE/dx values are similar for both species at these different momenta. This justifies again that the use of a single η -map for all species at not too low momenta is a good approximation. For very small momenta below $0.3 \text{ GeV}/c$ or so, an additional momentum dependence spoils this approximation. For protons, the residual η dependence at $p \approx 0.25 \text{ GeV}/c$ (green) is still within a few per mille because the η -map was created from a proton sample. Other species, however, suffer from the momentum dependence. This is visible for the pions with $p \approx 0.3 \text{ GeV}/c$ (black), where deviations of almost 2% from the expectation are observed.

3. Particle Identification with the ALICE TPC

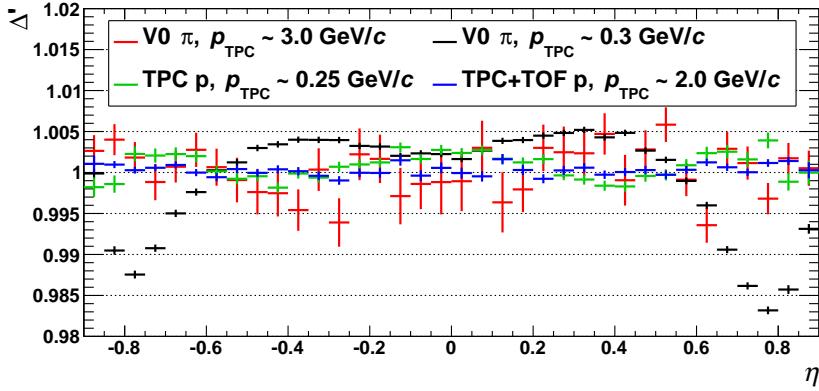


Figure 3.14.: η -map quality for different species at various momenta. The vertical axis is the dE/dx divided by the expected dE/dx for the corresponding species.

Looking also at other species and the whole momentum range, it can be concluded that typically the residual η dependence, $\delta\widetilde{\Delta}'$, is better than $\pm 0.5\%$ for $p > 0.45 \text{ GeV}/c$. Below this threshold it is typically within 3%. In the following, a fixed momentum is considered, such that the variable $\delta\widetilde{\Delta}'$ takes a constant value that gives an average estimate of the residual η dependence for many tracks and all η and dE/dx .

This variable now needs to be related to a change of the correction factor $corr_{\text{para}} = corr_{\text{para}}(\langle dE/dx \rangle, \eta)$ from the map to estimate the systematic uncertainty of the η correction. For better legibility, the dependence of the variables will be dropped in the following equations. The actual correction $c_{\text{para}} = c_{\text{para}}(\langle dE/dx \rangle, \eta)$ is the deviation of $corr_{\text{para}}$ from unity:

$$corr_{\text{para}} \equiv 1 + c_{\text{para}}. \quad (3.12)$$

Choosing an arbitrary species s , the splines parametrise the mean dE/dx of this species averaged over η , denoted by $dE/dx_{\text{splines}} = dE/dx_{\text{splines}}(s)$. Together with the correction factor, this allows to calculate the expected dE/dx as:

$$dE/dx_{\text{para}} = dE/dx_{\text{splines}} corr_{\text{para}}, \quad (3.13)$$

where $dE/dx_{\text{para}} = dE/dx_{\text{para}}(\eta, s)$ for arbitrary η and species s are implicitly understood. Note that the (implicit) first parameter of $corr_{\text{para}}$ becomes $\langle dE/dx \rangle = dE/dx_{\text{splines}}$.

The dE/dx value from the parametrisation can be compared to the observed value dE/dx_{obs} . The anchoring of the maps (region with $corr_{\text{para}} = 1$) is supposed to be unchanged by systematic η variations, since the anchoring is related to the η -averaged dE/dx of the splines, which is an independent error source. For a fixed momentum bin, the splines anyway would give only an overall offset if looking at the η shape as in Fig. 3.14. Evidently, no such offset is visible, such that the above estimate of $\delta\widetilde{\Delta}'$ is not biased by inaccuracies of the splines. Consequently,

it can be assumed for the following consideration that dE/dx_{splines} gives a perfect description of the η -averaged dE/dx . Similarly to Eq. 3.13, the observed dE/dx value can then be decomposed as:

$$dE/dx_{\text{obs}} = dE/dx_{\text{splines}} \text{corr}_{\text{obs}}, \quad (3.14)$$

where corr_{obs} is the ideal η correction factor that reproduces the observation.

It is assumed that the deviation from unity of the ratio of parametrised and observed dE/dx scales with the size of the actual correction c_{para} :

$$\frac{dE/dx_{\text{obs}}}{dE/dx_{\text{para}}} \equiv 1 + c_{\text{para}} \delta\Delta', \quad (3.15)$$

where, by assumption, $\delta\Delta' = \delta\Delta'(dE/dx_{\text{splines}})$ does not depend on η , but is allowed to change with dE/dx . In other words, each dE/dx row of the η -map has a certain shape in η and $\delta\Delta'$ can take different values for each shape.

The ratio in Eq. 3.15 is nothing else than an average Δ' for many tracks (cf. Eq. 3.9). For a fixed dE/dx_{splines} — which corresponds to one histogram in Fig. 3.14 —, a single row of the η -map is considered and $\delta\Delta'$ is a constant. It is important to note that $\delta\Delta'$ does not equal $\widetilde{\delta\Delta'}$ mentioned above. Rather, $\widetilde{\delta\Delta'}$ was defined to reflect the most extreme η residuals, which by assumption occur for the largest correction c_{para} . To relate these two variables, it is convenient to consider the extreme values of the η correction per row of the η -map:

$$\begin{aligned} |c_{\text{extreme}}(dE/dx_{\text{splines}})| &\equiv |\text{corr}_{\text{extreme}}(dE/dx_{\text{splines}}) - 1| \\ &\equiv \max_{\eta} \{|c(dE/dx_{\text{splines}}, \eta)|\}, \end{aligned} \quad (3.16)$$

i.e. $|c_{\text{extreme}}(dE/dx_{\text{splines}})|$ is the largest deviation from unity of all correction factors in the full η range for a given dE/dx row of the η -map. By construction, this implies:

$$\widetilde{\delta\Delta'} = c_{\text{extreme}}(dE/dx_{\text{splines}}) \delta\Delta'(dE/dx_{\text{splines}}), \quad \forall dE/dx_{\text{splines}}. \quad (3.17)$$

This equation holds true for all dE/dx rows. By reasonable assumptions, the equation reduces the unhandy two-dimensional residual η dependence in the $(dE/dx, \eta)$ space to a single value.

Having collected all required pieces, it is now time to put them together. From Eqs. 3.13 and 3.14, it follows:

$$\frac{dE/dx_{\text{obs}}}{dE/dx_{\text{para}}} = \frac{dE/dx_{\text{splines}} \text{corr}_{\text{obs}}}{dE/dx_{\text{splines}} \text{corr}_{\text{para}}} = \frac{\text{corr}_{\text{obs}}}{\text{corr}_{\text{para}}}, \quad (3.18)$$

which leads via Eq. 3.15 to:

$$\begin{aligned} \text{corr}_{\text{obs}} &= \text{corr}_{\text{para}} (1 + c_{\text{para}} \delta\Delta') = \text{corr}_{\text{para}} (1 + \delta\Delta' (\text{corr}_{\text{para}} - 1)) \\ &= \text{corr}_{\text{para}} \left(1 + \widetilde{\delta\Delta'} \frac{\text{corr}_{\text{para}} - 1}{\text{corr}_{\text{extreme}} - 1} \right), \end{aligned} \quad (3.19)$$

3. Particle Identification with the ALICE TPC

where Eqs. 3.12 and 3.17 were used in the second and last step, respectively.

Finally, the systematic variation of the η correction is chosen such that the observed correction factor from Eq. 3.19 is recovered for a given $\delta\widetilde{\Delta}'$:

$$\begin{aligned} \text{corr}_{\text{sys}}(dE/dx_{\text{splines}}, \eta) &\equiv \text{corr}_{\text{para}}(dE/dx_{\text{splines}}, \eta) \\ &\cdot \left(1 + \delta\widetilde{\Delta}' \frac{\text{corr}_{\text{para}}(dE/dx_{\text{splines}}, \eta) - 1}{\text{corr}_{\text{extreme}}(dE/dx_{\text{splines}}) - 1} \right), \end{aligned} \quad (3.20)$$

where all dependencies have been written down explicitly.

By definition, $|\frac{\text{corr}_{\text{para}} - 1}{\text{corr}_{\text{extreme}} - 1}| \leq 1$ holds true. As desired, there is no change of the spline value, i.e. $\text{corr}_{\text{sys}} = \text{corr}_{\text{para}}$ for $\text{corr}_{\text{para}} = 1$, and the systematic variation scales with the deviation of the correction factor from unity. Positive values of $\delta\widetilde{\Delta}'$ increase the η dependence, while negative values decrease it correspondingly.

In the end, Eq. 3.20 allows for propagating a known systematic uncertainty $\delta\widetilde{\Delta}'$ to the η correction factor.

3.6. Resolution Maps

As for the η -maps, the resolution maps are created from clean proton samples identified via TPC, TPC + TOF and V0. While small contamination hardly influences the mean of a Gaussian fit to the dE/dx distribution, it can seriously bias the width of the Gaussian. Thus, V0 protons are already used for momenta larger than $2.0 \text{ GeV}/c$. In order to get a proper mean dE/dx and to get rid of binning effects in η , the resolution maps are created from data that is already corrected with respect to η .

Fig. 3.15 shows an example of how the resolution is extracted in a single bin in momentum and polar angle. Firstly, the sample is put into bins with different number of PID clusters⁵, N_{cl} . To have enough statistics, the binning in momentum and angle is reduced accordingly compared to the η -map. Secondly, the Δ' distribution of each bin is fitted with a Gaussian. Again, to take into account the asymmetric dE/dx shape, the fit is restricted around the maximum (cf. Section 3.4). The means of the fit are shown in the left panel, the widths in the middle panel. Since the fit gives the width of Δ' and not of dE/dx , the middle panel already shows the *relative* width. However, there might be a systematic bias if the mean Δ' is not at unity. This would e.g. be the case if the η correction would not have been applied before. As can be seen, the mean is rather close to unity after correction, but there are fluctuations. To get the “real” relative width,

⁵This is the number of detected PID clusters above threshold, including 1-pad clusters. Note that for the tracking, also clusters close to the sector boundaries of the TPC are used. Such clusters are excluded for the PID, since they cannot be calibrated as precisely as clusters away from the boundaries.

the width from the fit is divided by the mean⁶. The result is shown in the right panel. Finally, these data points are fitted as a function of the number of clusters (blue line), where the relative resolution is parametrised as:

$$\sigma_{\text{rel}}(N_{\text{cl}}) = \sqrt{p_0^2 + \frac{p_1^2}{N_{\text{cl}}}}. \quad (3.21)$$

p_1 is the resolution per single PID cluster and the contribution from this term is assumed to decrease statistically with $1/\sqrt{N_{\text{cl}}}$. The term with p_0 can be considered as a systematic contribution to the resolution, e.g. there could be pressure changes in the TPC that cause a finite resolution even for an infinite number of PID clusters.

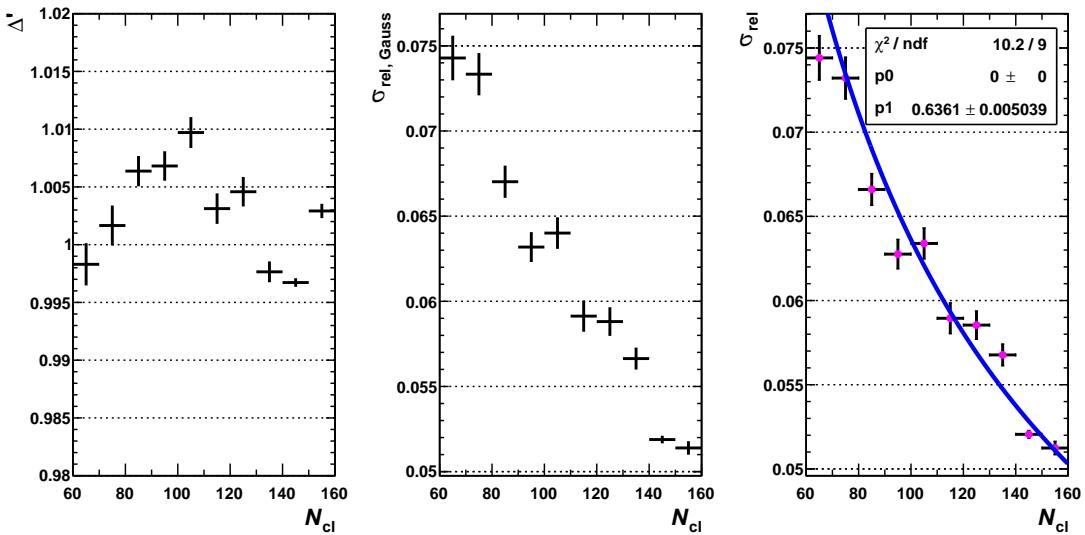


Figure 3.15.: Determination of the resolution as a function of the number of clusters. The left panel shows the mean Δ' of the fitted Gaussian and the middle panel its width. In the right panel, the ratio of middle and left panels are shown, together with a fit of these data points (blue line). The data points are from protons with $p \approx 1.5 \text{ GeV}/c$ and $\tan(\theta) \approx 0.15$.

The actual number of pad rows translates to $N_{\text{cl}} \leq 159$. To get a reliable dE/dx from the truncated mean, N_{cl} should be at least 60. It turns out that this rather small fit range causes strong correlations of p_0 and p_1 , such that from one bin of the resolution map to the other the values can be swapped. In addition, the additional degree of freedom seems not to be necessary as can be seen e.g. from Fig. 3.15. Therefore, p_0 has been fixed to zero.

⁶The width in the middle panel is relative with respect to the expected dE/dx . By dividing these data points by the mean Δ' , the expected dE/dx drops out and the width is relative with respect to the measured dE/dx . In particular, there is no dependence on the splines in this fit anymore.

3. Particle Identification with the ALICE TPC

For the fit with Eq. 3.21, the Gaussian width of bins with very low statistics is typically biased to higher values. To remove such bins, all data points with more than 20% relative error for the Gaussian width are excluded from the fit.

The relative resolution per cluster, p_1 , is put into a resolution map as a function of inverse (spline) dE/dx and $\tan(\theta)$. The raw results from the fits are shown in Fig. 3.16a. Some further processing is necessary to get a reliable map. First of all, there are “white” bins, where the fitting failed — sometimes there is too little statistics in a bin; especially if the sample is changed to TOF or V0 the statistics goes down from one bin to the other. The value of these bins is set to the median of the neighbouring bins, excluding bins where the fit failed. The same procedure is applied to outliers, which deviate from the median of their neighbouring bins by more than some value (typically 20%). To get rid of fluctuations between the bins, the map is smoothed, see Fig. 3.16b. Finally, interpolation is used to get a finer binning as shown in Fig. 3.16c.

Note that protons (and also kaons) have almost zero detection efficiency for momenta below $0.2 - 0.3 \text{ GeV}/c$ (cf. Fig. 5.7). As a result, there is low statistics at very high dE/dx (the very bottom of the figures) that does not allow for a reliable fit of the width of the Gaussian. The systematic uncertainty is estimated from Fig. 3.16a for y -bins below 0.004 (i.e. for $dE/dx > 250$): there is a jump of the relative resolution from about 0.45 to 0.65. Therefore, a systematic error of 50% is assigned for $dE/dx > 250$. This region only affects light nuclei and also protons and kaons at $p \lesssim 0.4$ and $0.2 \text{ GeV}/c$, respectively. Fortunately, the separation is large in these cases, so that the large uncertainty poses only little harm. Also note that the map is not fully symmetric. This is already the case for the η -maps, but not visible from the plots, since the asymmetry is typically less than 1%.

Finally, the resolution is calculated from σ_{rel} as:

$$\sigma^{\text{species}}(N_{cl}) \equiv \sigma_{\text{rel}}(N_{cl}) dE/dx_{\text{expected}}^{\text{species}}, \quad (3.22)$$

where implicit dependencies on track parameters, like momentum and η , are not indicated explicitly.

3.6.1. Systematic Uncertainty of the Resolution Maps

It is pretty difficult to estimate the systematic uncertainty of the resolution (maps). The resolution depends on the number of PID clusters (N_{cl}), η , dE/dx and in case of p–Pb and Pb–Pb collisions additionally on the multiplicity. Note that the resolution depends on the spline dE/dx in two ways. Firstly, there is a direct dependence via Eq. 3.22. Secondly, the splines determine the coordinate in one dimension of the (relative) resolution map, see Fig. 3.16c. Fortunately, both effects turn out to be negligible for the estimated accuracy of the splines in comparison to the total estimated uncertainty of the resolution map.

There is by far too little statistics to check the resolution in each of the above mentioned dimensions. Instead, the accuracy is estimated from looking at the

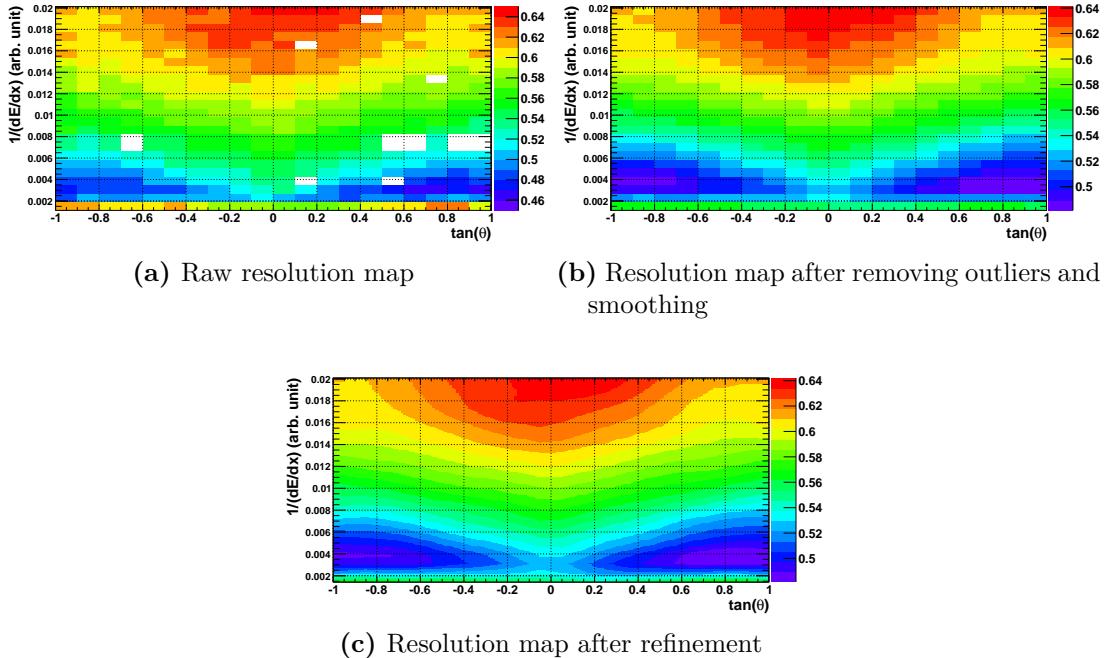


Figure 3.16.: Resolution map obtained from protons. **a)** shows the raw result from the fits and **b)** shows the same map after removing outliers and smoothing. In **c)**, the refined map using interpolation is presented. The colour code shows the relative resolution per cluster, where bins in which the fit failed are shown in white.

widths of the $n_{\sigma,\text{TPC}}$ distributions from different species as a function of momentum, where the number of sigmas for the TPC, $n_{\sigma,\text{TPC}}$, is defined as:

$$n_{\sigma, \text{TPC}}^{\text{species}} \equiv \frac{dE/dx - dE/dx_{\text{expected}}^{\text{species}}}{\sigma_{\text{expected}}^{\text{species}}}. \quad (3.23)$$

All other dimensions are just integrated. The different species at various momenta cover a broad dE/dx region and estimate possible residual momentum dependencies, if two species with same dE/dx , but varying momenta are compared.

Such a comparison is shown in Fig. 3.17. The $n_{\sigma,\text{TPC}}$ distributions have been fitted with a Gaussian in each bin of p_{TPC} . To take into account the shape asymmetry of the response, which biases the fitted width to larger values, the tails are removed by restricting the fit to $\pm 1.5\sigma$ around $n_{\sigma,\text{TPC}} = 0$. It was checked that the difference to results obtained with fits restricted to $\pm 3\sigma$ is small (per mille level), except that for the broader $n_{\sigma,\text{TPC}}$ range, contamination by other species sets in earlier, i.e. the width is biased to larger values already at lower momentum. This restriction can be interpreted as a TPC cut.

The data points compare primary pions, kaons and protons selected by a 3σ TOF inclusion cut and V0 electrons. The reason to take V0 electrons and not primaries is related to the fact that even at low momenta there is large

3. Particle Identification with the ALICE TPC

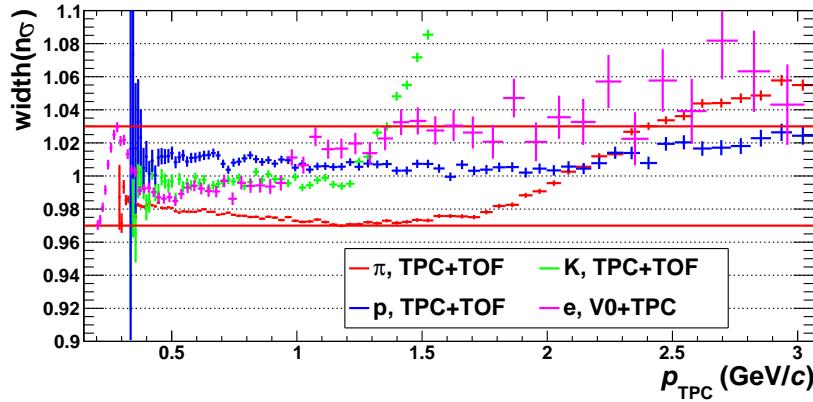


Figure 3.17.: Accuracy of the resolution map in data. The fitted Gaussian width of the $n_{\sigma,\text{TPC}}$ distribution is shown for primary particles selected by TOF (except for the magenta data points, which are from V0's). For every data sample, a TPC cut was applied in the sense that the Gauss fit has been restricted to $\pm 1.5\sigma$ around the expected dE/dx . Note that at some momentum contamination shifts the width to larger values. The horizontal red lines show the $\pm 3\%$ region around unity, which largely covers the deviations of all species.

contamination of the TOF electron sample and the pion dE/dx tail reaches into the electron distribution. Note that at some momentum contamination shifts the width to larger values for the primary particles, when TOF does not provide a clean selection anymore. In case of contamination, the fitted width contains no reliable information and is ignored. Also note that in case of very limited statistics in a momentum bin, the fit tends to overestimate the width. This is visible for V0 electrons (magenta) above 1.5 GeV/c.

Apart from the contamination regions, all deviations are within $\pm 3\%$ (horizontal red lines). Since other dimensions, like η , are integrated in this plot, the 3% envelope is taken as the overall uncertainty of the resolution parametrisation.

It was pointed out above that the estimated resolution becomes problematic for $dE/dx > 250$ due to the low detection efficiency for the data samples available for the extraction of the parametrisation. As a result, the uncertainty is estimated as $\pm 50\%$ in this region.

3.7. Occupancy Effects

The occupancy in the TPC may influence the dE/dx of the reconstructed tracks. The signal induced on the cathode plane by the slowly moving ions in the amplification region exhibits a rather long tail. In case of two tracks with the same azimuth, but slightly different polar angle, the ionisation electrons of these tracks drift towards the same anode wires and the ions produced there to the same cathode wires. For a certain difference of the track polar angle, the induced signals in the pad plane overlap, such that the second signal gets an additional contribution from the tail of the first. Eventually, this biases the calculated dE/dx of the track. If the signals of multiple tracks overlap, this effect is enhanced correspondingly.

In central Pb–Pb collisions with many thousand (primary and secondary) tracks inside the TPC fiducial volume, the dE/dx bias is several per cent. This must be treated properly in analyses requiring a high dE/dx precision. Fortunately, the occupancy effects on the dE/dx are negligible in pp collisions even at highest multiplicities, as will be demonstrated in Section 6.5.2.

3.8. Final Remarks

The model for the TPC dE/dx response and the extracted parameters from this work have been integrated into the official ALICE PID framework and can be used by other analyses. Indeed, several analyses use this model and carry out their own benchmarks of its quality. As a result, there are plenty of other checks besides those presented in this thesis.

It should also be noted that the TPC dE/dx performance can still be improved for the 2010 data sets. Newer data sets have been reconstructed using so-called *subthreshold clusters* and it turned out that this improved the separation power of the TPC. Subthreshold clusters are virtual clusters that are added to the dE/dx calculation. In case there are two pad rows with a signal above threshold and one row in between that has a signal below the threshold, a cluster with the charge of the threshold is added. The subthreshold clusters can and will be used for the analysis of newer data sets and also for the 2010 data periods, if there is a new reconstruction.

In conclusion, the TPC can identify hadrons track-by-track for low momenta and has a stable statistical PID performance in the relativistic rise for $p \gtrsim 3 \text{ GeV}/c$. Though, statistical PID is not possible up to arbitrarily high momenta, since the Bethe-Bloch curves reach the Fermi plateau at some point. In Fig. 3.18, the mean dE/dx parametrisations extracted from pp data are shown as a function of momentum for the individual species. Around $p = 40 - 50 \text{ GeV}/c$, the separation between electrons and pions starts approaching zero, making a (statistical) identification problematic. It is possible to correct the electron contamination in the pion sample e.g. via Monte Carlo simulations. The model dependence of this

3. Particle Identification with the ALICE TPC

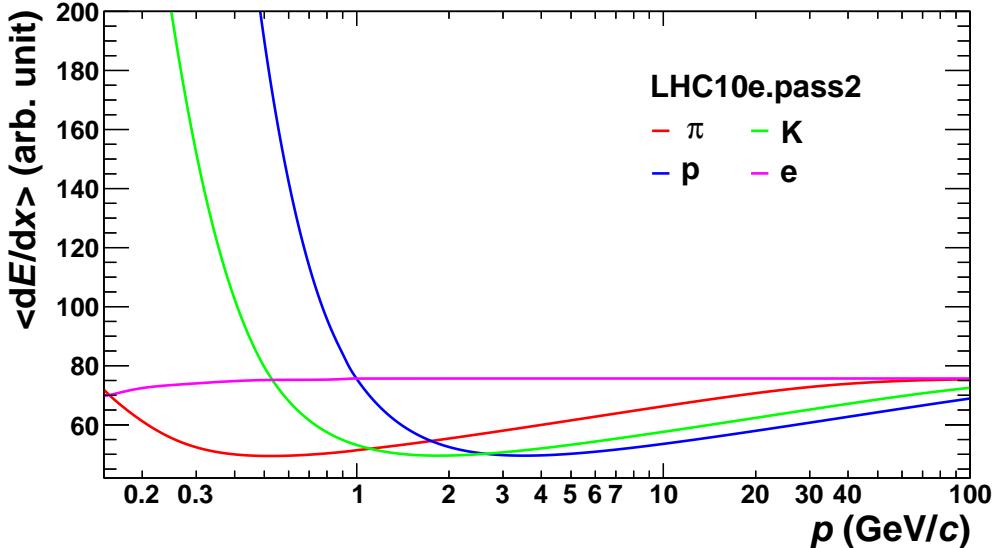


Figure 3.18.: Limitations of statistical PID with the ALICE TPC. The plot shows the mean dE/dx parametrisations extracted for the LHC10e period of the 2010 pp data taking with $\sqrt{s} = 7$ TeV. Around $p = 40 - 50$ GeV/ c , the curves for electrons and pions get close to each other, making a (statistical) identification impossible. The same holds for the (statistical) identification of pions and kaons at $p \sim 100$ GeV/ c .

correction enters then the systematic uncertainty. Since pions are by far more abundant than electrons, the resulting systematic uncertainty from that source should still be acceptable. At $p \sim 100$ GeV/ c , the discrimination between pions and kaons becomes infeasible.

Note that the plot only gives a rough estimate about the upper threshold of the range in which statistical PID is possible. There is some uncertainty of the parametrisations, in particular for the region where the relativistic rise approaches the plateau. Nevertheless, it can be expected that the statistical $\pi/K/p$ identification with the TPC dE/dx can be used up to $p \sim 100$ GeV/ c .

4. The TPC Multi-Template Fit

4.1. Analysis Outline

The TPC Multi-Template Fit (MTF) uses the parametrisations of the TPC PID response to generate dE/dx templates for each species. With these, the particle content of a given sample can be extracted. This thesis investigates as such samples particles in jets and the inclusive particle production in different multiplicity classes. Typically, the particle content will be extracted in bins of the transverse momentum p_T . The MTF fits the templates to the measured dE/dx distribution in each p_T bin, but also correlates the particle fractions of adjacent bins by imposing a regularisation condition. To do so, all p_T bins are fitted in a single global fit with a penalty term for the regularisation. To justify the procedure, it will be run on a MC sample using the MC information only to obtain the MC truth.

4.2. Template Generation

With the detailed knowledge about the TPC dE/dx response templates with the expected response for each species can be generated. The inputs for the template generation are the parameters of the measured tracks and the event itself: track momentum, p_T , $\tan(\theta)$, number of PID clusters and event multiplicity (only relevant for p–Pb and Pb–Pb). The input parameters are different for the individual species, for instance protons at 0.5 GeV/c typically have more PID clusters than pions due to the much higher dE/dx . Therefore, a so-called “pre-PID” is used for the template generation: for each measured track, the probabilities to be of a specific type are derived from the combined PID of ITS, TPC and TOF. The probabilities already take into account priors for the particle abundance, which are the default priors used in ALICE. The priors are defined as the probability to observe a certain particle type for given p_T . The actual values of the priors only matter if there is no clean identification by the detectors. In this case, the dE/dx is anyway similar, such that wrong pre-PID probabilities do not pose any harm.

For each track, 500 dE/dx responses (according to Eq. 4.5, see below) are generated in total and attributed to the individual species weighted with the probabilities from the pre-PID. For instance, if the probability of a particle to be a pion, kaon or proton are 60%, 25% and 15%, respectively, $0.6 \cdot 500 = 300$ responses will contribute to the pion, 125 to the kaon and 75 to the proton

4. The TPC Multi-Template Fit

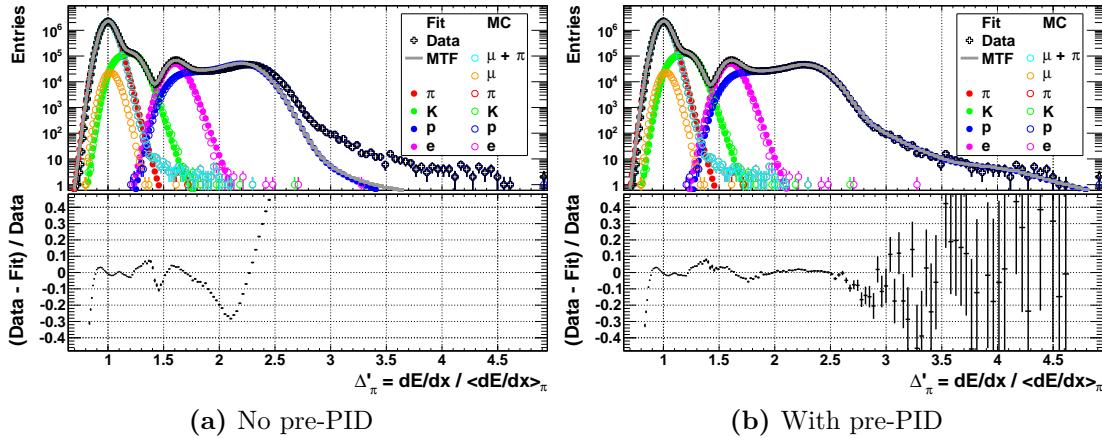


Figure 4.1.: Template generation and fit quality in MC **a)** without and **b)** with using pre-PID. In the upper panels, the dE/dx distribution for $p_T = 0.6 - 0.65 \text{ GeV}/c$ is shown in black together with the templates (full points), the MC true response (open points) and the fit (grey line). The lower panels show the relative deviation of the fit from data.

template. Typical templates using pre-PID are shown in Fig. 4.1b. Note that the total momentum p of the input track is used to derive the expected dE/dx , but the template is then related to the corresponding p_T bin according to the p_T of the input track.

The exact number of generated responses N^{resp} does not matter, but it should not be too small in order to obtain reasonably smooth templates also for bins with low statistics (at high track and/or jet momentum). However, N^{resp} should not become too large, since the template generation with asymmetric shape, which will be discussed in Section 4.2.1, consumes pretty much processing time. It turns out that $N^{\text{resp}} = 500$ is a suitable choice. As a cross-check, N^{resp} has been increased by a factor 10 and the results were found to be unchanged.

The necessity for using pre-PID at low p_T is demonstrated in Fig. 4.1 for the MC case. Not using pre-PID simply means that each track contributes to the templates of all species with equal weight. The agreement between generated templates (full points) and MC true response (open points) is considerably improved if pre-PID is utilised. The impact is largest for protons, since their dE/dx is very different from that of other species for $p_T = 0.6 \text{ GeV}/c$, which is the momentum bin shown in the figure. The improvement of the fit quality can directly be read off from the lower panels.

It is remarkable that the proton templates in Fig. 4.1a exhibit a very different shape than the measured dE/dx distribution. Comparing with Fig. 4.1b, the proton templates are obviously biased towards lower values due to the contribution from non-protons. The origin of this bias is due to the energy loss of the particles between vertex and TPC. At momenta of a few hundred MeV/c, protons

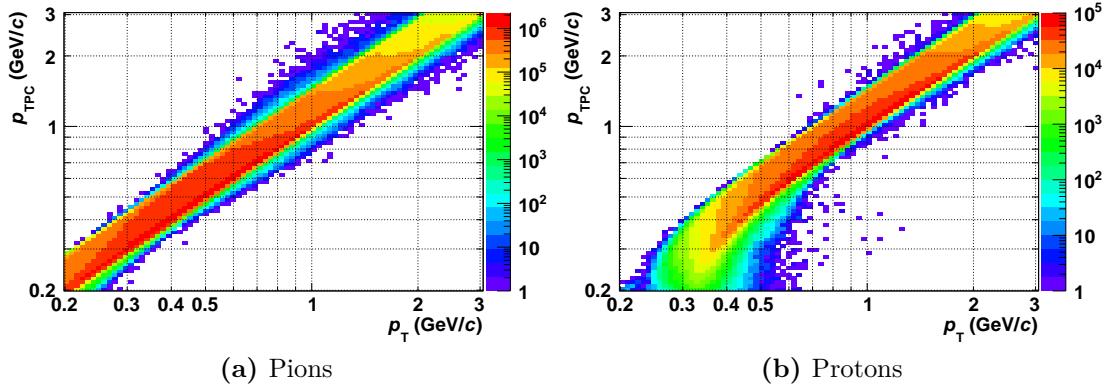


Figure 4.2.: Correlation between transverse momentum, p_T , and the momentum at the TPC inner wall, p_{TPC} , for **a)** pions and **b)** protons in MC. The MC label was used to determine the particle species. The colour code shows the number of accepted tracks within $|\eta| < 0.9$.

have a much higher dE/dx than other species. As a result, their momentum at the TPC inner wall, p_{TPC} , is reduced compared to the momentum at the vertex, the latter being directly related to the transverse momentum, p_T , via the polar angle. The correlation between p_{TPC} and p_T for protons and pions is shown in Fig. 4.2. The pion correlation is diagonal. The width of the distribution is due to the finite momentum resolution and due to the rather broad pseudo-rapidity range, $|\eta| < 0.9$. The peak is slightly shifted downwards in p_{TPC} due to the small energy loss between vertex and TPC. This energy loss is larger for protons at low momenta and increases with decreasing momentum, leading to a bending of the correlation in that region. As a consequence, the p_{TPC} distribution at a fixed low- p_T bin is shifted to lower values for protons compared to pions. Taking now the track parameters of pions to generate the proton templates leads to the usage of too high p_{TPC} values. Due to the steeply falling proton dE/dx as a function of p_{TPC} , this biases the templates to smaller dE/dx values.

In summary, pre-PID is important if species with very different dE/dx contribute. It ensures that only the parameters of proper tracks are used for the template generation for the individual species. For $p_T \gtrsim 2 \text{ GeV}/c$, the dE/dx of hadrons is similar. On the one hand, pre-PID has, thus, only little or no effect. On the other hand, pre-PID can anyway no longer distinguish individual species, such that pre-PID essentially reduces to the priors.

4.2.1. Asymmetry of the dE/dx Response

The energy loss in the TPC follows a Landau distribution. Taking the truncated mean of the energy losses over each pad row largely removes the tail of the distribution, see discussion in Section 3.3. However, the tail is not completely removed. Together with threshold effects in the electronics, this causes a dE/dx response

4. The TPC Multi-Template Fit

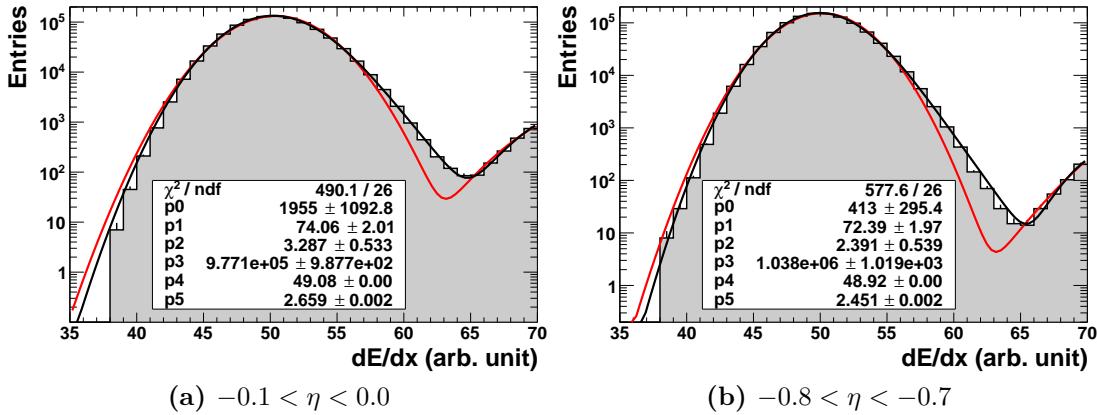


Figure 4.3.: Asymmetry of the dE/dx response around the minimum ionising region in data. The red curve is a double-Gauss fit to the data points, whereas the black curve uses the asymmetric description for the pion response. The asymmetry of the response is visible at **a)** small η as well as at **b)** large eta. The data points are for negative particles with $p_{TPC} = 0.45 - 0.5 \text{ GeV}/c$.

that is pretty much a Gaussian, but with some tail towards higher dE/dx as shown in Fig. 4.3. To exclude effects on the shape from the dependence of the mean dE/dx on $\beta\gamma$ and η , the figure shows the dE/dx distribution of π^- in small η intervals and a very small momentum interval. The latter interval is chosen to be around minimum ionising, such that the mean dE/dx is flat as a function of momentum. Note that standard track cuts have been used, but no η correction was applied. The red curve is a double-Gaussian fit to the dE/dx distribution, which has contributions from pions and electrons in this region. The parameters for the electron Gaussian have been restricted to some range about the values obtained from fitting only the electrons. Obviously, the fit does not describe the data very well, the χ^2/NDF is found to be larger than 200 (the value in the figure is for the asymmetric shape fit). The asymmetry of the dE/dx distribution is obvious for both small and large η .

To get a better description of the data, a new parametrisation for the dE/dx distribution is required. The black curve shows the fit using an asymmetric shape for the pions as will be discussed in the following. It is already visible by eye that the black curve agrees much better with data than the red curve. Note that the plots look qualitatively the same for other η intervals. The ansatz for the new dE/dx parametrisation is to start with a normalised Gaussian distribution G and convolute it with an exponential function e :

$$G(x; \mu, \sigma) \equiv 1/(\sqrt{2\pi}\sigma) \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad (4.1)$$

$$e(x; \sigma, \lambda) \equiv \exp\left(-x\frac{\lambda}{\sigma}\right), \quad (4.2)$$

$$(G \otimes e)(x; \mu, \sigma, \lambda) \equiv \int_0^\infty d\tau G(x - \tau; \mu, \sigma) e(\tau; \sigma, \lambda), \quad (4.3)$$

where $\mu > 0$ is the mean expected dE/dx , $\sigma > 0$ the expected width and $\lambda > 0$ is a parameter for the tail. Note that the convolution is restricted to $[0, \infty[$ and therefore asymmetric because the Landau tails and threshold effects bias the dE/dx only towards larger values.

The strength of the tail is proportional to σ/λ , such that for the relative tail strength

$$s_{\text{rel}}^{\text{tail}} \propto \sigma/\lambda/\mu \equiv \sigma_{\text{rel}}/\lambda \quad (4.4)$$

with $\sigma_{\text{rel}} \equiv \sigma/\mu$ holds true. Eq. 4.4 implies that for fixed λ the tail increases with worse relative resolution. This trend is confirmed by observations similar to those in Fig. 4.3, where the width of the pions (parameter $p5$ in the figures) is larger at mid-rapidity (Fig. 4.3a) than in forward direction (Fig. 4.3b). λ and μ are approximately the same in both cases. Hence, the relative tail strength is smaller in forward direction. Even with such a parametrisation, the tail in forward direction seems to be slightly too large at the very edges. The origin of the remaining deviation is discussed below.

Finally, the convolution is to be normalised to unity, which introduces a normalisation factor σ/λ . After some steps the convolution can be written as:

$$(G \otimes e)(x; \mu, \sigma, \lambda)_{\text{normalised}} = \frac{\lambda}{\sigma} \frac{1}{2} \exp\left(-\frac{\lambda}{\sigma}(x - \mu) + \frac{\lambda^2}{2}\right) \cdot \text{erfc}\left((-x + \mu + \sigma\lambda)/\sqrt{2\sigma}\right). \quad (4.5)$$

In Fig. 4.3, it is demonstrated that using Eq. 4.5 for the shape instead of a pure Gaussian better describes the data. Moving from a Gaussian to the asymmetric shape for pions¹ reduces the χ^2/NDF by a factor 10. By eye, the agreement is much better. There are still some deviations at the edges visible because only a single value for sigma was used in the fit, but the distribution stems from a superposition of almost-Gaussian shapes with different widths. The diverse widths originate from a varying number of PID clusters. For the analysis, this feature will be taken into account by generating the expected distribution for each track with the corresponding width.

¹For electrons, a pure Gaussian fit was kept for simplicity here. Anyway, only the left part of the electron shape is inside the histogram, on which the asymmetric shape has only little impact. Note that the number of degrees of freedom are the same for pure Gaussian and asymmetric shape, since the tail parameter is fixed to what was used for the analysis in pp at $\sqrt{s} = 7$ TeV data. This parameter is the same for all η .

4. The TPC Multi-Template Fit

The tail parameter for each data taking period is extracted from plots similar to Fig. 4.3 in momentum regions where the separation is sufficiently large. Unfortunately, there is quite some correlation between the fit parameters. In addition, there is not enough statistics to fit besides η slices also in bins with different number of clusters. Hence, the fitted distribution is always a superposition of different resolutions to some extent. As a consequence, it is not possible to get an exact value of the tail this way, but a reasonable approximation by demanding a good description in many η bins at the same time. Eventually, the best value for the tail is found by fitting generated distributions with slightly different λ values around the found approximation to data. Typical plots for this procedure are similar to Fig. 4.1b and should be in a region where the templates for the individual species are different. Otherwise, there might be accidentally better agreement, but for particle fractions that are completely off.

To be able to use Eq. 4.5, it is important to realise that the mean and also the width of the convolution are different from the (input) mean and width of the Gaussian. The fit of the tail parameter alone is already pretty complicated as mentioned before. Hence, it is not possible to extract the splines — and also the η - and resolution maps — already by fitting with the asymmetric shape, but a pure Gaussian is used. To reduce the impact of the tail, the Gaussian fits are restricted to the region around the maximum (see discussion in Section 3.4). If the fit results were taken unmodified for the template generation, a pure Gauss fit of the templates with the same restriction would yield a different result. This is because the convolution slightly shifts and distorts the shape even in the restricted region. To get templates with high precision, it is necessary to take this residual effect of the tail into account.

Unfortunately, it is not possible to analytically derive functions that convert the parameters from the restricted Gaussian fit to those of the input Gaussian of the convolution. However, the following relations turn out to hold true:

$$\begin{aligned}\sigma_{\text{ratio}} &\equiv \sigma_{\text{input}} / \sigma_{\text{Gauss,restricted}} = f(\lambda), \\ \mu_{\text{shift}} &\equiv \mu_{\text{input}} - \mu_{\text{Gauss,restricted}} = g(\lambda) \sigma_{\text{input}},\end{aligned}\tag{4.6}$$

with f and g being some unknown functions of λ . Since λ is fixed for each data period, the values of f and g at the corresponding λ can be determined by generating a template for arbitrarily chosen μ_{input} and σ_{input} . The template is fitted with a pure Gaussian in the same restricted region as is used for the splines and maps. Using the relations 4.6 fixes then the function values. Once these values are known, μ_{input} and σ_{input} can be calculated from $\mu_{\text{Gauss,restricted}}$ and $\sigma_{\text{Gauss,restricted}}$, which get their values assigned from the splines and maps.

The improvement of the fit quality with the asymmetric shape is shown in Fig. 4.4 in case of MC. Note that the tail parameter was tuned for data. The asymmetric dE/dx produces templates that are closer to the MC true response and improves the fit quality, best visible in the region $\Delta' = 1.1 - 1.2$. Similar plots exist for data, but there it is hard to judge, since there are no MC true responses

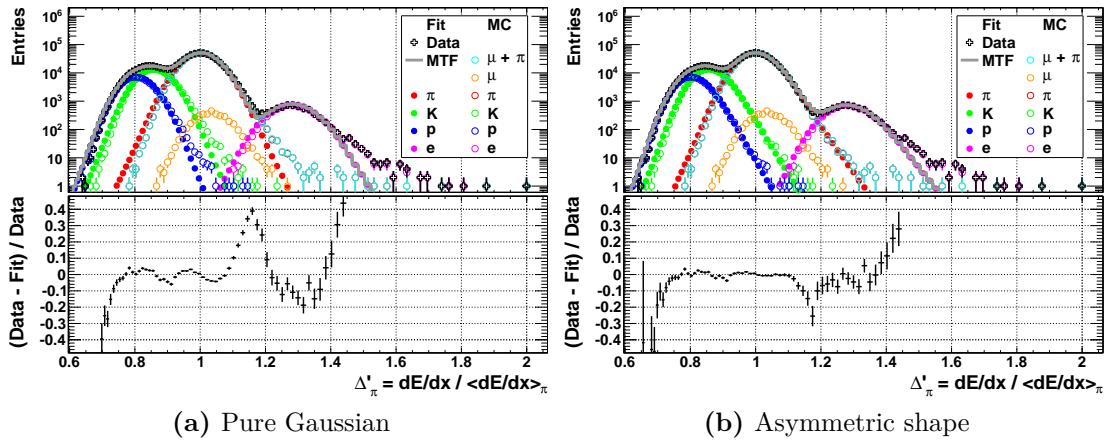


Figure 4.4.: Comparison of **a)** pure Gaussian and **b)** asymmetric response in MC. In the upper panels, the dE/dx distribution for $p_T = 4.0 - 4.5 \text{ GeV}/c$ is shown in black together with the templates (full points), the MC true response (open points) and the fit (grey line). The lower panels show the relative deviation of the fit from data.

that can be compared to. The relative deviation of the fit from data is improved by a similar amount as for the MC example, but it must be kept in mind that the fractions change if the templates are altered. Nevertheless, the observation of the tail in data as shown in Fig. 4.3 and the similar improvement of the fit quality in data and MC give a hint that the asymmetric dE/dx response is the better parametrisation.

In general, the tail can be different in MC and data, so that it cannot just be tuned for MC and then adopted for data. Since it is rather hard to judge which tail gives the “best” result in data, a conservative systematic error estimate is applied for the shape of the dE/dx response. The error is defined as the spread of the results obtained with pure Gaussian response and asymmetric shape.

4.2.2. Special Templates for High dE/dx

For momenta below $0.8 \text{ GeV}/c$ ($0.4 \text{ GeV}/c$), protons (kaons) can be identified track-by-track solely with the TPC dE/dx , since the separation to all other species is large. Though, the steep slope of the dE/dx as a function of momentum for kaons and protons makes the template generation difficult: if the slope of the splines is slightly different from data, the templates will not describe the measured dE/dx accurately. There are even more problems that make the template generation at low momenta difficult. Firstly, the resolution maps for high dE/dx are created with protons, which have tiny efficiency at low momentum (the same holds for kaons). The statistics might be sufficient to fit the mean dE/dx in η slices, but is too small to fit the width of the dE/dx in η and number of PID clusters bins. Consequently, the resolution parametrisation becomes unreliable

4. The TPC Multi-Template Fit

for large dE/dx — the systematic error is estimated to be 50% for $dE/dx > 250$. Secondly, the systematics of the η correction become larger at low momenta, as was discussed in Section 3.5.

Pions and electrons still have templates with reasonable systematics at low momenta because the mean dE/dx is rather flat and they do not suffer from the huge resolution uncertainties present for high dE/dx . Yet, this does not hold for the kaon and proton templates, which suffer from large systematic uncertainties. Therefore, it is reasonable to replace the templates for K and p in low- p_T bins just with the measured dE/dx of tracks that are most probable K or p according to TPC identification. The probabilities are evaluated as the number of sigmas for TPC, $n_{\sigma, \text{TPC}}$ (see Eq. 3.23), and the species with the smallest $|n_{\sigma, \text{TPC}}|$ is chosen. For the given η window of $|\eta| < 0.8 - 0.9$, kaon and proton templates are created with this method up to $p_T = 0.3 \text{ GeV}/c$ and $p_T = 0.45 \text{ GeV}/c$, respectively. On the one hand, a clean identification is possible below these thresholds. On the other hand, the parametrisations create reliable templates above the thresholds.

Note that the systematic error of the pion and proton fractions becomes negligible in the region where the special templates are used. This makes sense, since the separation is large and the identification becomes trivial.

The effect of the special templates is shown in Fig. 4.5. The left plot uses the standard templates. As can be seen, the standard templates deviate from the observed shape to some extent. In the right plot, the special templates are used for kaons and protons. Effectively, the dE/dx distributions of kaons and protons are removed from the fit, since the templates are just the measured dE/dx distributions of these species and their yield is also fixed to that of the special templates. Ultimately, this means that only species without special templates are fitted in the corresponding bin.

Note that there is a tiny overlap between kaons and electrons and it is the last p_T -bin which uses the special templates for kaons. This overlap is even smaller (or does not exist at all) if fixed (total) momentum and small η intervals are considered. The η dependence amounts to roughly 10% in the considered dE/dx region, i.e. both, kaons and electrons, have 10% different mean depending on their η . In the figure, the left tail of the kaons stems from η regions with low dE/dx , whereas the right tail of the electrons is from η regions with high dE/dx . Hence, it can happen that the dE/dx distributions for a broad η interval overlap, while the species are still separated for any given η . This implies that the special templates do not suffer from contamination. In summary, the overlap poses no harm and is anyway negligible.

It turns out that the effect of the special templates is mainly of psychological nature: although the shape of the generated templates deviates from the observed dE/dx distribution, the implementation of the fit as a log likelihood fit — which will be described in detail in Section 4.3.1 — guarantees that the total yield is correct and yields more or less the correct particle fraction. For data, the relative change of the proton fractions when using special instead of standard templates

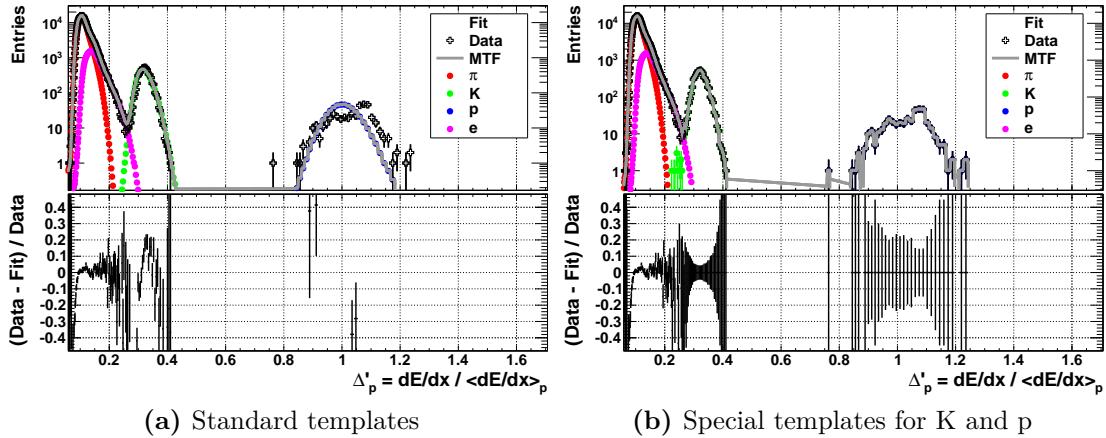


Figure 4.5.: Demonstration of special dE/dx templates: in **a**), the default templates for all species have been used, whereas in **b**), special templates are used for kaons and protons. The plots are for data and show tracks with $p_T^{\text{track}} = 0.25 - 0.30 \text{ GeV}/c$ in jets with $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$.

is smaller than 3% — note that the fractions are at the per mille level in the low- p_T region; once the fractions reach about 1%, the relative difference drops below 1%. For kaons, the relative change is even below 5%. Similar numbers hold true for MC, but a direct comparison to data is not possible, since the dE/dx parametrisations are significantly different. The dE/dx crossing points as well as the thresholds where the distributions start to overlap are different for MC. The thresholds stated above are those for data.

Although the effect of the special templates is small, their results look much more trustworthy than relying solely on the log likelihood fit with templates that have an imperfect shape.

4.2.3. Handling of Light Nuclei

Besides pions, kaons, protons, electrons and muons, there are also light nuclei like deuterons. They only contribute a tiny fraction to the particle yields (per mille level and below). Nevertheless, they contribute to the templates with completely different input parameters. Therefore, there is a special treatment for them in this analysis. Firstly, light nuclei can easily be rejected up to about $1 \text{ GeV}/c$, since they are well separated from all other species in dE/dx (cf. Fig. 3.1). Secondly, if no clean rejection is possible they should contribute to the template of the species which is closest in dE/dx .

To be sure that no other particles than light nuclei are removed, a rather loose rejection is applied:

Reject particle if $\begin{cases} \min\{|n_{\sigma, \text{TPC}}^{\text{p}}|, |n_{\sigma, \text{TPC}}^{\text{e}}|\} > 15, & p \geq 0.3 \text{ GeV}/c \\ \min\{|n_{\sigma, \text{TPC}}^{\text{p}}|, |n_{\sigma, \text{TPC}}^{\text{e}}|\} > 10, & \text{otherwise} \end{cases}, \quad (4.7)$

4. The TPC Multi-Template Fit

i.e. all particles with much higher dE/dx than those of protons and electrons are removed.

As visible from Fig. 3.1, the dE/dx of light nuclei becomes similar to that of other species once it drops below the electron dE/dx . In this case, the contamination of the templates is not dangerous anymore due to similar input parameters. The only remaining region is where the dE/dx is not too far away from that of other species, such that the track cannot be rejected safely by Eq. 4.7, but it is still significantly larger than for the other species. For such cases, the track is kept, but the signal only contributes to the species which is closest in dE/dx . This means that the pre-PID weight of a track is set to 100% for the species with the highest dE/dx (excluding light nuclei, i.e. it is set to either electrons or protons) in the following cases:

- the most probable PID (including light nuclei) is a light nucleus and $dE/dx > dE/dx_{\text{expected}}^e + 5\sigma_{\text{expected}}^e$,
- $dE/dx > dE/dx_{\text{expected}}^{\text{species}} + 5\sigma_{\text{expected}}^{\text{species}}$ for both species = e and p.

Both conditions are rather similar, but are used in order not to depend too much on the splines for the light nuclei, which are not as accurate as those for the other species.

The impact of the treatment of special nuclei is illustrated in Fig. 4.6. Clearly, the peaks from the light nuclei are removed by the cuts and the (logarithmic) plots with the fits look much more convincing. The change of the templates is negligible in this figure.

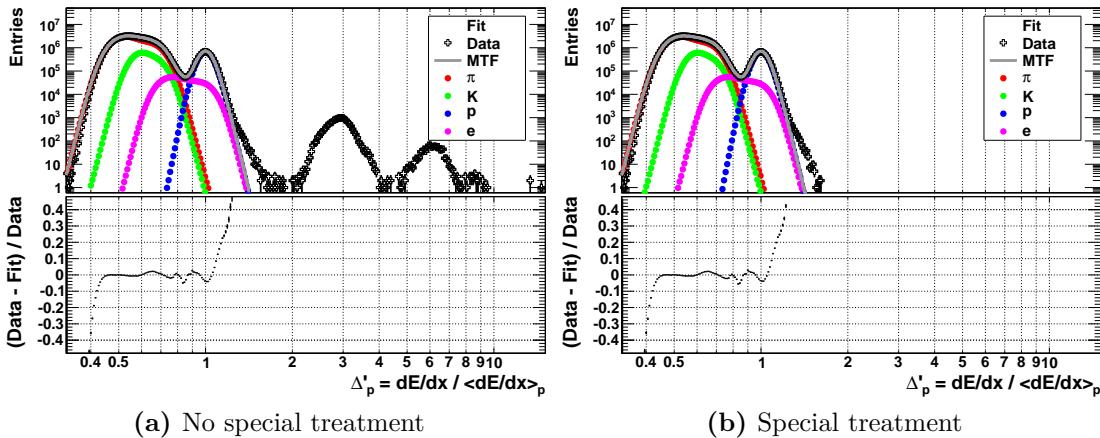


Figure 4.6.: Demonstration of the special treatment of light nuclei: in **a)**, no special treatment is applied to light nuclei, whereas in **b)**, light nuclei are rejected if safely possible and remainders contribute differently to the templates. The plots are for data and show tracks with $p_T^{\text{track}} = 0.7 - 0.8 \text{ GeV}/c$.

To sum up, the whole procedure guarantees that the input parameters of the remaining contribution of light nuclei to the templates is similar to that of the corresponding species and that most of the light nuclei are rejected completely.

4.3. Fitting Procedure

The fitting procedure is carried out individually in each p_T bin. The fit extracts the particle fractions of the species under consideration, which can be converted to particle yields via the total number of observed tracks in that bin. Pions, kaons, protons, electrons and muons are by far the most abundant species. There are also light nuclei (like deuterons and alpha particles). It was described in the last section that they can be largely removed from the sample. Their remaining yield fraction is negligible and can be ignored. The dE/dx resolution of the TPC is about 5% [73], which allows to separate these species from momenta as low as $0.15 \text{ GeV}/c$ to above $20 \text{ GeV}/c$. For higher momenta, the dE/dx of all species starts saturating at the plateau, such that the separation decreases and vanishes at some point (cf. Section 3.8).

One of the strengths of the Multi-Template Fit is the freedom to generate the templates and extract the fractions also in bins of other observables. For the jet analysis in this work, the following observables are interesting:

$$z \equiv z^{\text{charged}} \equiv \frac{p_T^{\text{track}}}{p_{\text{jet,charged}}}, \quad (4.8)$$

$$\xi \equiv \xi^{\text{charged}} \equiv \ln \frac{1}{z^{\text{charged}}} = \ln \frac{p_T^{\text{jet,charged}}}{p_T^{\text{track}}}. \quad (4.9)$$

Note that only charged particles are considered in this work. Especially, the jet reconstruction only uses charged jets, no neutral particles. The index ‘‘charged’’ will be dropped in the following. The principle of the fitting is the same for p_T , z and ξ . This section will focus on the p_T case and mention z and ξ only if there are exceptions.

Basically, the fit can be performed in the two-dimensional (p , dE/dx) space, cf. Fig. 3.1. Since the physical results are invariant in p_T instead of p , the first dimension is changed accordingly. It is also convenient to switch from dE/dx to Δ'_s (see Eq. 3.9) because for given species s , this observable has no distortions of the shape due to the dE/dx slope as a function of momentum, η dependence, etc. — yet, these distortions are still present for any other species. In addition, the dE/dx resolution approximately scales with dE/dx . For Δ' , this dE/dx dependence is divided out. In summary, Δ' turns out to be easier to interpret and is therefore used as the second dimension of the fit.

For a given p_T (bin), the data points are the number of observed tracks in the Δ' bins. These points are fitted with the sum of counts of the generated templates

4. The TPC Multi-Template Fit

for each species in the corresponding bin:

$$v_i(p_T; \vec{A}(p_T), N_{\text{tot}}(p_T)) \equiv \sum_s f_i^s(p_T) A_s(p_T) N_{\text{tot}}(p_T), \quad (4.10)$$

where the sum extends over all species. N_{tot} is the total number of observed counts, f_i^s is the fraction of counts for the template of species s in the i -th Δ' bin and A_s is the particle fraction of species s , summarised in the vector \vec{A} , in which each component corresponds to one species. The species templates in each p_T bin are normalised to the total number of generated responses for this species, such that:

$$\sum_i f_i^s(p_T) = 1, \forall s, \forall p_T, \quad (4.11)$$

where the sum is over all Δ' bins. This means that v_i gives the predicted counts from the model for the i -th Δ' bin. Assuming the number of Δ' bins to be N for a given p_T , the v_i can be considered as the N components of a vector \vec{v} . Likewise, the observed counts $n_i(p_T)$ in the i -th Δ' bin can be written as a vector $\vec{n}(p_T)$.

4.3.1. Implementation of a Binned Maximum Likelihood Fit

The Multi-Template Fit uses the binned maximum likelihood estimation method [109]. The Minuit tool [110] — implemented in the TMinuit package of the ROOT software [85] — is used for maximisation (minimisation) of the (negative) log likelihood and the (statistical) error estimation.

In the following, a single p_T bin will be considered for simplicity, i.e. the dependence on p_T will be dropped. Also, $\vec{v} = \vec{v}(\vec{A}, N_{\text{tot}})$ is understood and i will denote the index of the Δ' bin. Assuming a Poisson probability density function $P(k; \lambda)$ for every Δ' bin, the likelihood function for the fit can be defined as:

$$L_P(\vec{v}; \vec{n}) \equiv \prod_i P(n_i, v_i) = \prod_i \exp(-v_i) \frac{v_i^{n_i}}{n_i!}. \quad (4.12)$$

$P(n_i; v_i)$ gives the probability to observe n_i counts if the expectation is v_i and $L_P(\vec{v}; \vec{n})$ is the likelihood that the observation \vec{n} is described by the model prediction \vec{v} . The maximum likelihood estimate of \vec{v} maximises L_P . It is convenient to consider the likelihood ratio R_λ to get rid of the factorials:

$$R_\lambda(\vec{v}; \vec{n}) \equiv L_P(\vec{v}; \vec{n}) / L_P(\vec{n}; \vec{n}) = \prod_i \exp(-v_i + n_i) \left(\frac{v_i}{n_i} \right)^{n_i}. \quad (4.13)$$

Minimising the negative logarithm of R_λ is equivalent to maximising R_λ because the logarithmic function is continuous and monotonously increasing. This suggests the following definition of the statistical negative log likelihood function²:

$$l_{\text{stat}} \equiv -\ln R_\lambda = -\ln L_P(\vec{v}; \vec{n}) + \ln L_P(\vec{n}; \vec{n}). \quad (4.14)$$

²Note that the factor 2 in the formula in [109] has been absorbed into the error definition of Minuit.

The minimisation process is carried out by TMinuit with respect to \vec{v} , the data \vec{n} is fixed. The minimum of l_{stat} then maximises L_P , since $\ln L_P(\vec{n}; \vec{n})$ is a constant, which does not affect the minimisation/maximisation. It can be shown (see [109] and references therein) that Eq. 4.14 asymptotically obeys a chi-square distribution (up to a factor 2 absorbed into the error definition of Minuit). This allows to better judge the goodness of the fit and motivates its definition a posteriori. Substituting Eq. 4.13, this finally yields:

$$l_{\text{stat}} = \sum_i \left(v_i - n_i + n_i \ln \frac{n_i}{v_i} \right). \quad (4.15)$$

A benefit of using a binned log likelihood fit together with the proper template normalisation (Eq. 4.11) is that it ensures that the fractions sum up to unity within statistical error:

$$\sum_s A_s(p_T) = 1, \forall p_T, \quad (4.16)$$

where the sum extends over all species. For bins with very small statistics, there might be a small deviation from unity of up to a few per cent. To guarantee the exact equality of Eq. 4.16, the fractions of each p_T bin are scaled with $1/\sum_s A_s(p_T)$.

4.3.2. Simultaneous Fitting of Individual Δ'_{species}

As discussed before, it is convenient to use Δ'_{species} as defined in Eq. 3.9 for the fitting. Fig. 4.7 compares fits for different choices of the reference species r of Δ'_r in data around $p_T \approx 3 \text{ GeV}/c$, where the dE/dx of all species is rather similar. As can be seen, the shapes get squeezed or stretched depending on the reference species. Overall, the fit quality looks similar, especially, the trends of the deviations between fit and data are the same. A glance back to Fig. 4.6a shows an extreme case in the sense of very different dE/dx . In that case, the protons, which are the reference species, have still a Gaussian-like shape, whereas the shapes of all other species are completely distorted.

All in all, the Δ' observable allows for an easy interpretation and a close-to-Gaussian shape for one reference species. There is no superior choice of the reference species. Nevertheless, the binning of the Δ' might influence the fitting. To smooth out such effects, a *simultaneous fit* is used: for each p_T bin, the same data, but with the binning for 4 different reference species (π , K , p , e), is fitted simultaneously, i.e. there is only one negative log likelihood. For instance, this means for Fig. 4.7 that all 4 histograms with the data points are fitted in one step with their individual set of templates, but the fit parameters — the particle fractions — are the same for all histograms.

The effect of the simultaneous fitting is depicted in Fig. 4.8 for the kaon fractions in data. The fits using only one reference species (coloured points) have strongly fluctuating fractions in the dE/dx crossing regions. The spread

4. The TPC Multi-Template Fit

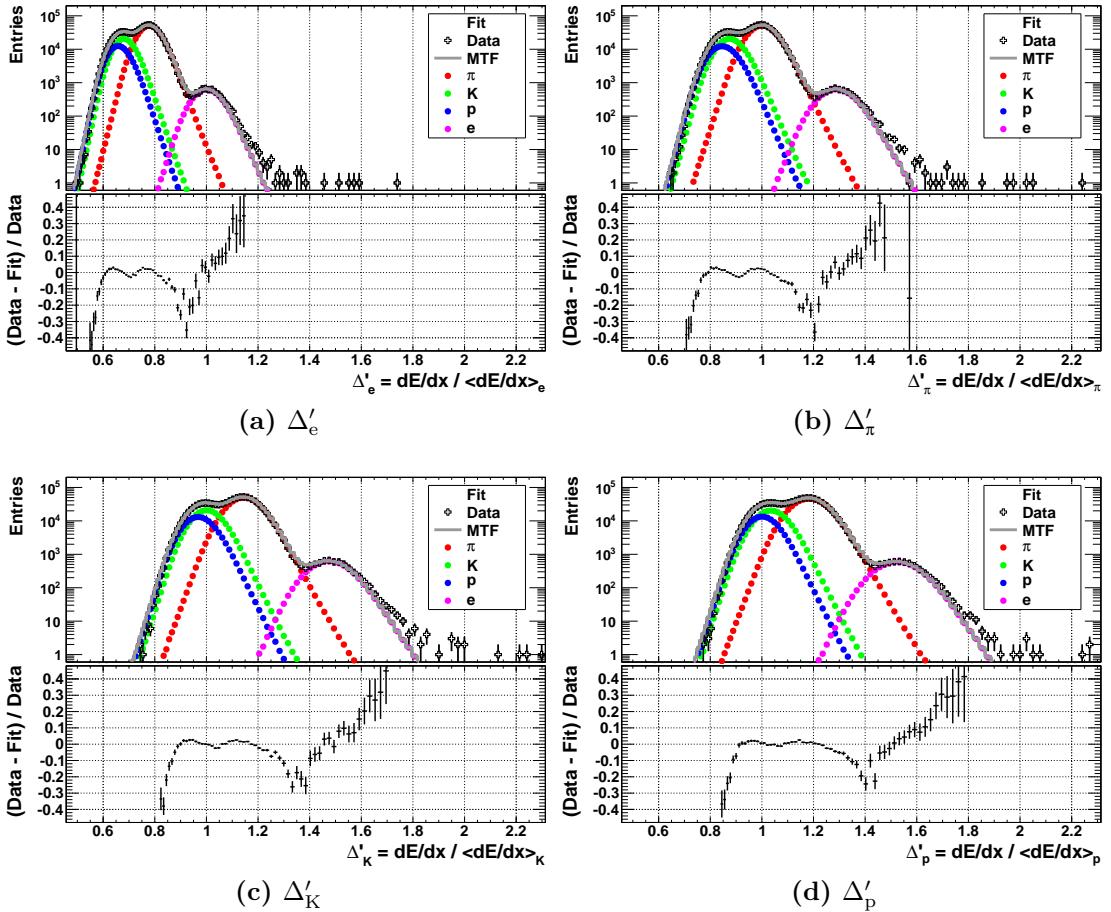


Figure 4.7.: Comparison of template fits to data for different Δ'_{species} for the horizontal axis. In the upper panel of each plot, the dE/dx distribution for $p_T = 3.0 - 3.2 \text{ GeV}/c$ is shown in black together with the templates (full points) and the fit (grey line). The lower panels show the relative deviation of the fit from data.

between the results for the different reference species is much larger than the statistical uncertainty. This is exactly what is expected, since the fit can converge accidentally to a local minimum there and the binning changes this minimum. In addition, the estimated errors have some arbitrariness for the individual fits. Looking at the black data points, which are for the simultaneous fit, it is visible that the fluctuations are reduced. Obviously, the use of different binnings is capable of removing part of the bias in the crossing regions by eliminating local minima. The regularisation that will be discussed in Section 4.3.3 will significantly reduce such fluctuations and biases additionally.

The statistical errors of the fit are provided by TMinuit as the covariance matrix [111]. It is the inverse of the Hessian matrix, which is the second derivative matrix of the parameters, at the point of convergence. The simultaneous fit artificially increases the statistics seen by TMinuit by a factor 4 (the number

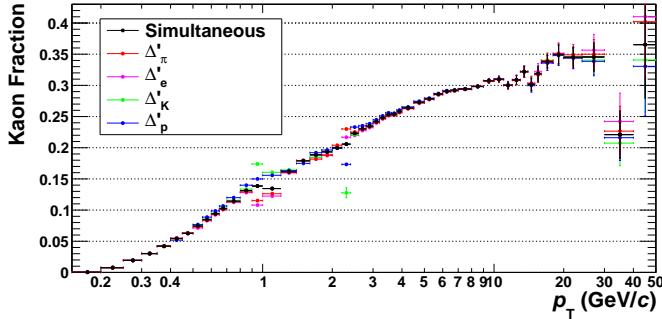


Figure 4.8.: Comparison of kaon fractions with (black points) and without (coloured points) simultaneous fitting for data. The simultaneous fit reduces the fluctuations due to binning effects in the crossing regions. Note that the special templates have been used below $0.3 \text{ GeV}/c$.

of considered reference species), since the information is the same and just the binning is different. To recover the “real” statistical error, the covariance matrix elements are multiplied by 4, which scales the parameter errors by a factor $\sqrt{4} = 2$. It has been checked by comparing the simultaneous fit to that with only one Δ'_{species} that this error scaling works properly — see Fig. 4.8, where this procedure is already applied for the simultaneous fit. In particular, the results (mean and error) agree within statistical uncertainties at high p_T , where all species have a similar dE/dx and binning effects are, thus, expected to be small.

4.3.3. Regularisation

The dE/dx of two species becomes equal in crossing points where their Bethe-Bloch curves cross each other, see Fig. 3.1. Exactly at these points, the separation between those species is zero and the negative log likelihood of the MTF has no unique minimum. In such cases, the yield is shared more or less arbitrarily between the involved species (cf. discussion in Section 4.3.2). Due to the finite momentum bin width, there is some small separation left, which reduces the effect to some extent. In addition, bins in transverse instead of total momentum are considered and the rather broad interval $|\eta| < 0.8 - 0.9$ smears the momenta contributing to a single bin. For the jet analysis, where bins in z or ξ are considered, this smearing is even larger.

Eventually, there are bins close to the dE/dx crossing points which have small, but non-vanishing separation for two species. This leads to fluctuating particle fractions in neighbouring bins, which are not physical. To reduce the fluctuations, a regularisation is introduced that is based on the method described in [112]. The particle fractions are assumed to change smoothly as a function of $\text{Log}(p_T)$ (or, in case of jets, as a function of z and ξ , respectively). The idea is to correlate the particle fractions of neighbouring bins. For each bin, a penalty term is added, which compares the currently fitted fraction with the interpolated

4. The TPC Multi-Template Fit

fraction from the nearest neighbours on both sides. If the difference is larger than the estimated statistical error of this bin, the penalty term becomes large. The likelihood function becomes:

$$l_{\text{tot}} \equiv l_{\text{stat}} + l_{\text{reg}}, \quad (4.17)$$

$$l_{\text{reg}} \equiv \sum_{j,s} N_{\text{refs}} \frac{(A_s(\text{obs}_j) - \tilde{A}_s(\text{obs}_j))^2}{2\delta_s^2(\text{obs}_j)}, \quad (4.18)$$

$$\delta_s(\text{obs}_j) \equiv \sqrt{\frac{A_s(\text{obs}_j) + \tilde{A}_s(\text{obs}_j)}{2}/w_j}, \quad (4.19)$$

with the following definitions:

- The sum for l_{reg} extends over all species s and all obs bins j , where obs can either be p_T , z or ξ .
- $N_{\text{refs}} = 4$ is the number of reference species for the simultaneous fitting. This factor is necessary to get the proper weight of l_{reg} with respect to l_{stat} : l_{stat} contains the contributions from all reference species for each obs bin, but there is only one common fraction A_s in every bin. This means the statistics of l_{stat} increases by a factor N_{refs} , which must also be included in l_{reg} .
- $\tilde{A}_s(\text{obs}_j)$ is the interpolation for A_s at obs_j from the two neighbouring values, $A_s(\text{obs}_{j-1})$ and $A_s(\text{obs}_{j+1})$, which uses *Neville's algorithm* [113].
- The statistical weight w_j of the j -th obs bin is usually set to $N_{\text{tot}}(\text{obs}_j)$ (exceptions are discussed below).
- δ_s is the estimated statistical uncertainty of A_s in the corresponding bin. It is defined as the square root of the mean of the current A_s and the interpolated value, divided by the statistical weight of this bin. With this definition, the regularisation term has a chi-square form. The factor 1/2 in Eq. 4.18 takes into account the error definition of TMinuit, i.e. it ensures that the asymptotic behaviour is the same as for l_{stat} .

The definition of l_{tot} correlates all obs bins, such that the particle fractions in all bins are extracted in one fitting step. Note that only the next neighbours are used for the interpolation \tilde{A}_s , which makes the regularisation to have a local effect.

The impact of the regularisation is shown in Fig. 4.9 for the MC case. Without regularisation, the kaon and proton fractions fluctuate around $p_T = 2 \text{ GeV}/c$. The same is true for the pion and kaon fractions at $p_T \approx 1 \text{ GeV}/c$. It is also visible that the fluctuations above $p_T = 10 \text{ GeV}/c$ are reduced. Obviously, the regularisation improves the agreement between fit (solid points) and MC truth (open points).

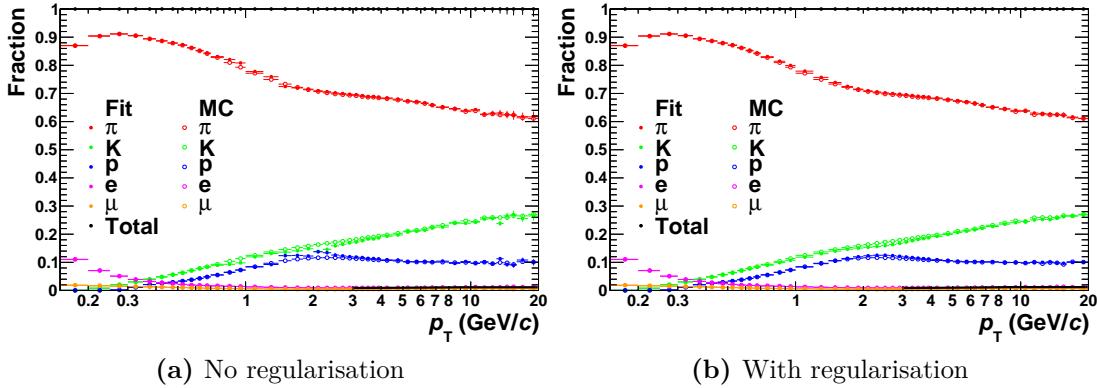


Figure 4.9.: Effect of the regularisation demonstrated for MC. The full points of the fit should be compared with the open points from the MC truth. Note that no regularisation is applied to the MC truth.

In particular, there is no bias introduced by the regularisation. Note that the regularisation reduces the statistical errors, which is best visible at high p_T . The reason is that the statistics of neighbouring bins is correlated.

It is remarkable to see that there is no visible bias in the crossing regions of electrons with kaons and protons at about $p_T = 0.5$ and $1 \text{ GeV}/c$, respectively. This can be understood from Fig. 4.10, where the crossing of electrons and kaons is compared to that of kaons and protons. The former is rather steep, i.e. the dE/dx curves of the two species have very different slopes in the crossing point, whereas the latter is relatively flat. The analysis is done in bins of transverse momentum, not (total) momentum, and the bins have some finite width. Since the Bethe-Bloch curves depend on $\beta\gamma$ and, thus, for a given species on the total momentum, the dE/dx distribution and the templates are smeared in a p_T bin. In case of a steep crossing, this leads to completely different template shapes for the corresponding two species — compare green and magenta points in Fig. 4.10a. In contrast, for a flat crossing, the templates are rather similar — compare green and blue points in Fig. 4.10b. The more different the templates are, the better the fit can distinguish between the species and the smaller the bias in the crossing region becomes.

The regularisation typically removes the bias from the individual fits almost completely for “not too flat” dE/dx crossings. This is the case for pions and kaons around $p_T = 0.9 \text{ GeV}/c$. However, for broad crossings like that of kaons and protons at $p_T = 2 - 3 \text{ GeV}/c$, many bins are biased and the regularisation cannot completely recover the truth. This is by construction, since the regularisation is done locally.

The same conclusions apply to the results for data shown in Fig. 4.11. Note that in this case, the muons are ignored in the fit, i.e. muons are included in the pion fraction, which needs to be corrected at a later stage.

4. The TPC Multi-Template Fit

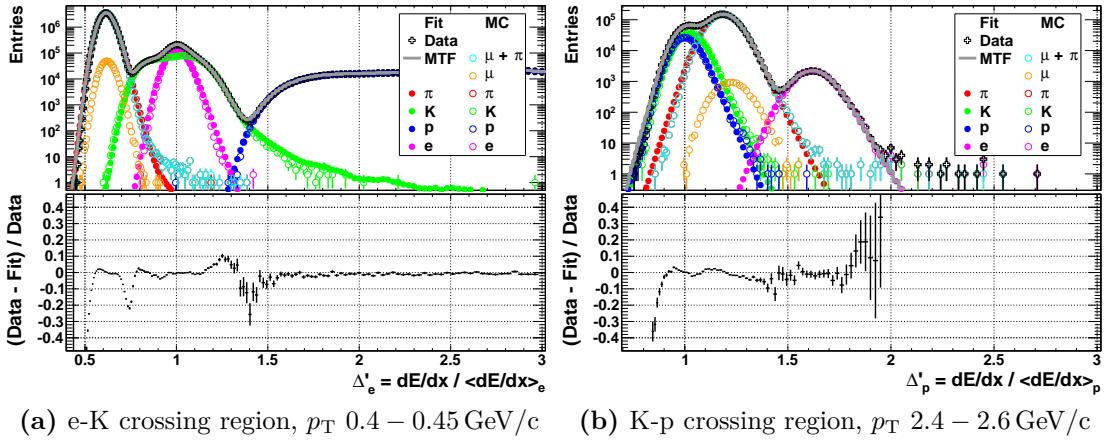


Figure 4.10.: Comparison of templates in the **a)** e-K and **b)** K-p crossing regions for MC. The full points are the generated templates, the open points indicate the dE/dx distribution of the MC truth. The crossing of e and K is steep and the templates are correspondingly very different, whereas the K and p templates look similar for the flat K-p crossing. Note the different Δ'_{species} on the abscissa. Also note that the deviations from zero in the lower panels are restricted to the edges of the pion distribution (a) and of the kaon and proton distributions (b), respectively. Their impact on the integral of the distribution and, thus, the fitted fractions is negligible.

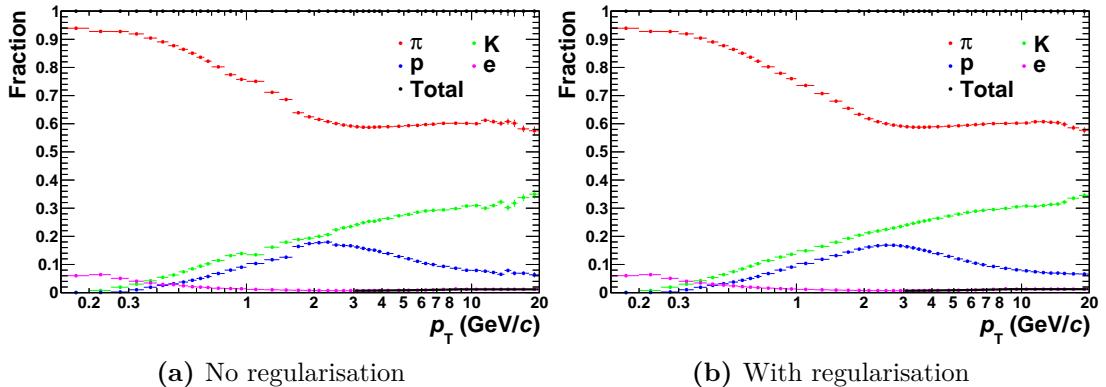


Figure 4.11.: Comparison of raw fractions with and without regularisation as in Fig. 4.9, but for data.

As will be discussed in Chapter 5, the efficiency of kaons and protons drastically changes for $p_T \lesssim 0.3$ GeV/c. This results in a rapid rise of the raw fractions, although the physical fractions are smooth. In such cases, the regularisation would bias the results because it by construction forces a smooth shape. Hence, the regularisation is switched off for kaons and protons for p_T below 0.3 and 0.45 GeV/c, respectively. This is done by setting the corresponding terms in Eq. 4.18 to zero. The regularisation for other species is still used, since there is no such rapid change

of the efficiency. Note that the regularisation can be safely switched off for kaons and protons, since they are well separated from other species, such that there is no danger of misidentification.

In case of fitting versus z (ξ), the threshold below (above) which regularisation is switched off is calculated from the p_T threshold by taking the lower limit of the considered p_T^{jet} range as an effective p_T^{jet} in Eqs. 4.8 and 4.9. This ensures that all contributing p_T bins are in the safe region with stable efficiency.

There is one peculiarity for the fitting versus z . Although the pion efficiency is smooth as a function of p_T , it has a jump in the first two z bins (cf. Fig. 5.1). This can be understood, since changing the efficiency from a function of p_T to one of z basically shifts the curve to the left. Almost all p_T bins with low efficiency end up in the very first z bin. Higher z bins only get contributions from p_T bins with rather similar efficiency. This causes a jump and introduces the necessity to switch off the regularisation for pions in the first two z bins in addition to that for kaons and protons. Since the by far more abundant pions directly influence the other particle fractions, the regularisation is switched off for all species (i.e. also for electrons) in these bins. Note that this is not an issue for ξ , since it has by definition much finer binning for low p_T .

4.3.4. Handling of Leptons

Although electrons are well separated from hadrons starting from $p_T \approx 2 \text{ GeV}/c$ up to rather high momenta like $20 \text{ GeV}/c$, it is not possible to reliably fit their fractions above some momentum threshold. The reason is rapidly decreasing statistics together with a small raw electron fraction of about 1%. Typically, there are too few electron counts in p_T bins beyond $10 \text{ GeV}/c$ for a meaningful fit. According to MC simulations, the raw electron fraction rises approximately linearly between $p_T = 3$ and $10 \text{ GeV}/c$ and then stays constant. Qualitatively, the same behaviour is observed in data if the electron fraction is floating for the fit. To avoid strong, non-physical fluctuations of the fractions, the electron fraction is kept fixed for $p_T > 10 \text{ GeV}/c$. The fraction is fitted in the range $p_T = 3 - 10 \text{ GeV}/c$ with a polynomial of degree 1 and fixed to the value of the fit at the upper edge of the fit range. The uncertainties of the fit are propagated as the statistical error for the fixed electron fraction.

Note that pions and muons cannot be clearly separated via the TPC dE/dx (or via the time of flight) due to the very similar masses. As discussed in Section 3.4.3, it is not possible to extract dedicated splines for muons, such that only approximate splines are available. Besides the systematic uncertainties of these splines, it is not possible to reliably disentangle the muon and pion yields because their templates look almost identical, even at low transverse momenta. The raw muon fraction is typically below 1% for the cuts used in this work. At very low transverse momentum it might go up to several per cent. Fitting both muons and pions would lead to huge fluctuations of their fractions, much larger than the

4. The TPC Multi-Template Fit

actual muon fraction. As a consequence, the fit uses only the pion templates and extracts the “pion plus muon” yield. The muon contamination is corrected at a later stage via MC-based correction factors in each (transverse) momentum bin (cf. Section 5.1.3).

In summary, the MTF has a rather limited precision for the lepton yields. One of the main goals of this thesis is to study the fragmentation of jets. On the one hand, leptons are irrelevant in this respect. On the other hand, any new physics results from lepton spectra would require a very high precision that cannot be achieved with this analysis (and is beyond its scope). Consequently, leptons can be considered as contamination of hadrons in this thesis. To correct for this contamination, the precision for the lepton yields is sufficiently high in this work, since the overall lepton (raw) fraction is typically around 1%.

4.4. MC Blind Test

To prove the fit principle, the procedure is applied to MC. The MC information about the particle type is not used for the fit, i.e. the fit is “blind” as in data. The TPC dE/dx parametrisations created in the frame of this work are the official ones for ALICE. The decision of the collaboration for MC was to use the MC information to extract the parametrisations — which are anyway very different for MC. Therefore, the MC test as a whole is rather a “semi-blind test”, since it uses the MC information to extract the splines and maps. The blind test is carried out on the pp minimum bias MC production (LHC10f6a), see Appendix B.1 for details.

In Fig. 4.12, the splines extracted with the MC information are compared to those extracted as in data, i.e. from V0’s and other clean detector samples (TPC, TPC+TOF) without using any MC information. As can be seen, the splines typically agree within $\pm 2\%$. In the region where the dE/dx relativistic rise approaches the plateau (pions with $p > 20 \text{ GeV}/c$), the deviation is about 1.3%. This is mainly due to the lack of data points in this region, if V0’s are used, whereas there is much higher statistics and, thus, better constraint from (MC identified) pions. This is visible if Figs. 3.2b — this figure is for data, but the gap size is similar for MC — and 3.2c are compared.

The systematic uncertainty of the splines has been discussed in Section 3.4.5. It turns out that the difference of the splines, when using or not using the MC information for the extraction, is covered by the estimated uncertainty. For the MC “semi-blind test” in the following, the splines extracted with MC information will be used.

The raw fractions — i.e. no correction is applied — in jets obtained via this way are compared to the MC true fractions in Figs. 4.13, 4.14 and 4.15 for pions, kaons and protons, respectively. The fractions from the MC truth are obtained by counting the tracks in the corresponding bins and assigning the species according

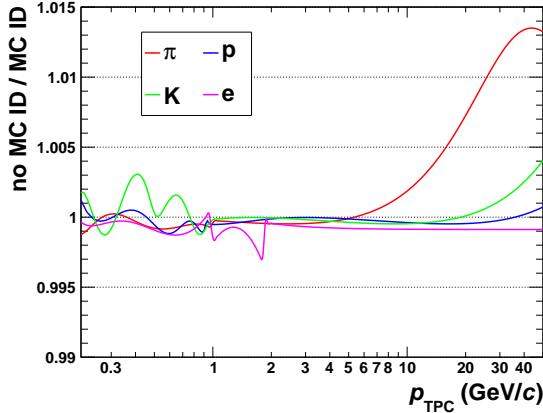


Figure 4.12.: The splines extracted with (*MC ID*) and without (*no MC ID*) MC information are compared. Their ratio is shown as a function of momentum for the individual species. *no MC ID* indicates that the splines were extracted as in data from V0's and other clean detector samples (TPC, TPC+TOF). For *MC ID*, primary particles which are “identified” via the particle level truth have been used.

to the MC label. The figure also shows the systematic errors of the fitted fractions, which are obtained by assuming the same variations of the template shapes as in data — this will be discussed in Section 6.3. Due to the difference of the mean dE/dx shape in MC and data, the thresholds for the special templates would be completely different, since the crossings and regions with clean separation are shifted. So, it makes no sense to use the special templates in MC and they are not used for the semi-blind test.

The MC truth is reproduced by the fit typically within $\pm 2\%$ for pions and $\pm 10\%$ for kaons and protons. Larger deviations in the highest p_T^{jet} bin are mostly covered by the error bars. Similar plots for the inclusive case and for the fit versus z and ξ can be found in Appendix D, which demonstrates that the fitting procedure performs similarly well as a function of p_T , z and ξ . It is also presented there that the fit reproduces the electron fraction of the MC truth within uncertainties.

For an infinitely fine binning in p_T^{jet} , p_T , z and ξ , the information content of a fit versus the latter three would be the same. In particular, z and ξ are very similar, the latter being some kind of zoom into the low- z region. For clarity and in order not to increase the number of plots by about 50%, the results presented in this thesis will focus on the fitting as a function of p_T and z . Nevertheless, it will be mentioned in the following how the ξ case is handled for completeness.

4. The TPC Multi-Template Fit

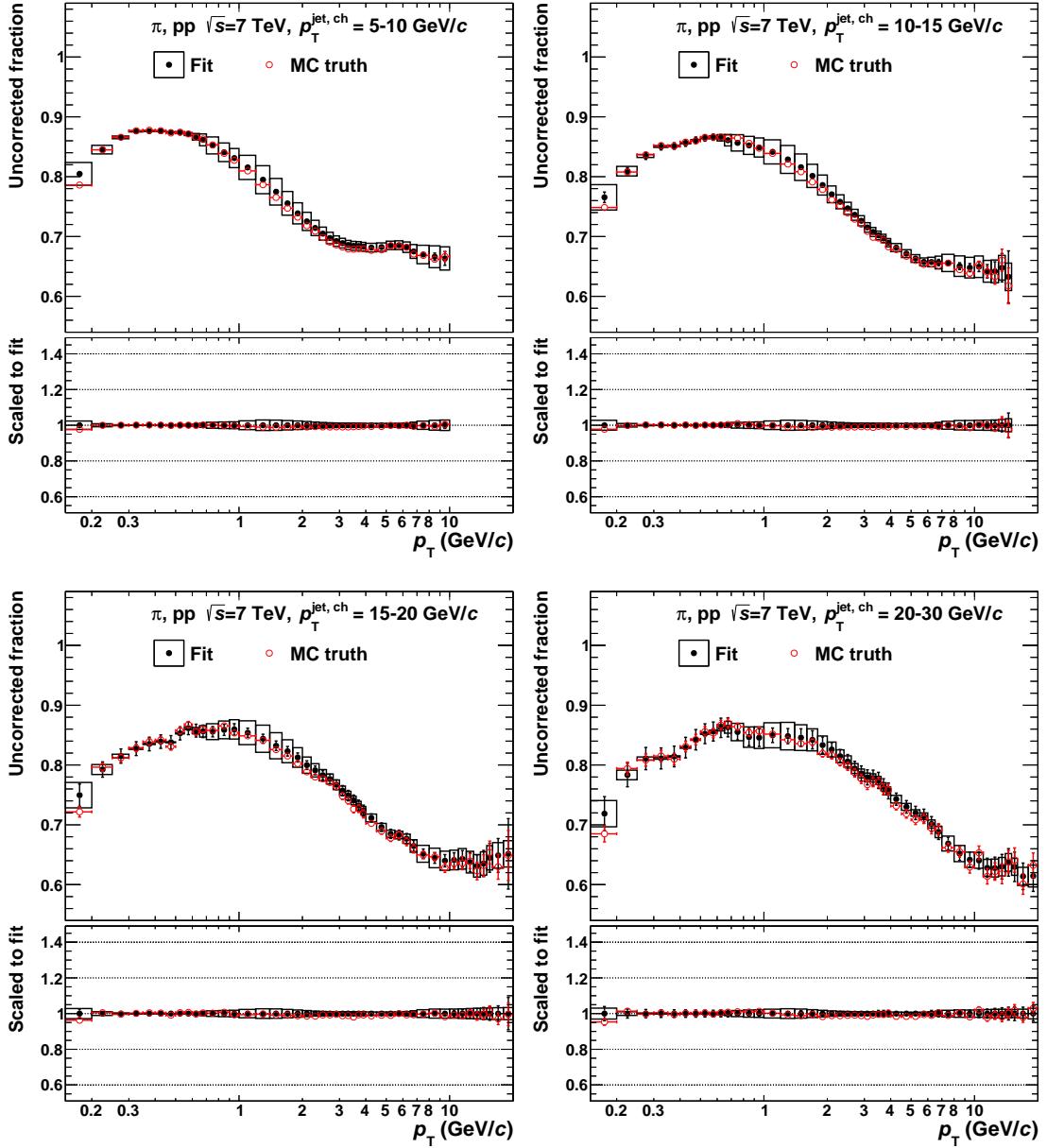


Figure 4.13.: MC blind test for pions fitted versus p_{T} . The raw fractions from fitting the MC output (black) are compared to the MC truth (red). The plots show (from upper left to lower right): $p_{\text{jet}}^{\text{jet}} = 5 - 10, 10 - 15, 15 - 20$ and $20 - 30 \text{ GeV}/c$.

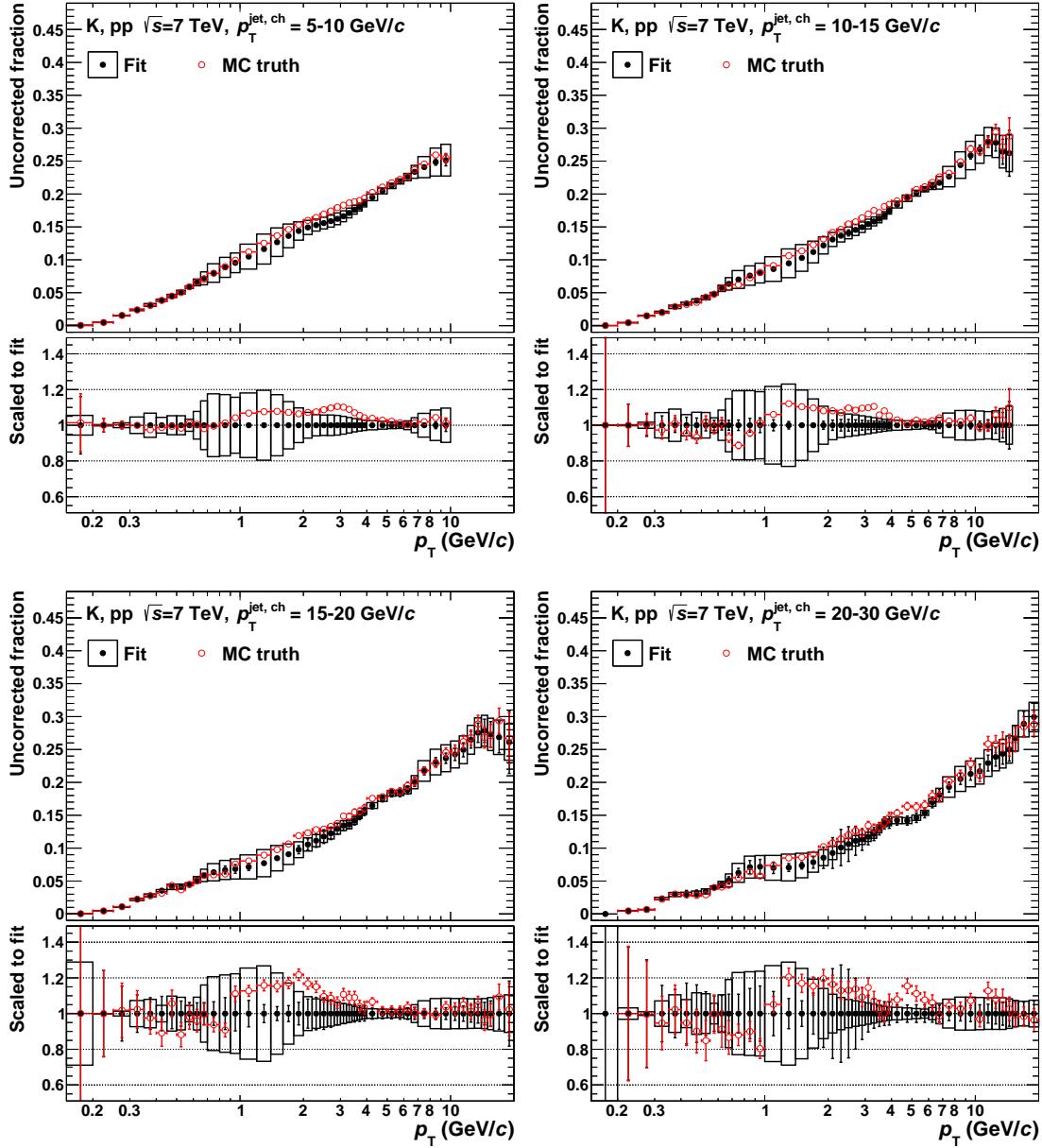


Figure 4.14.: MC blind test for kaons fitted versus p_T . The raw fractions from fitting the MC output (black) are compared to the MC truth (red). The plots show (from upper left to lower right): $p_T^{\text{jet}} = 5 - 10, 10 - 15, 15 - 20$ and $20 - 30 \text{ GeV}/c$.

4. The TPC Multi-Template Fit

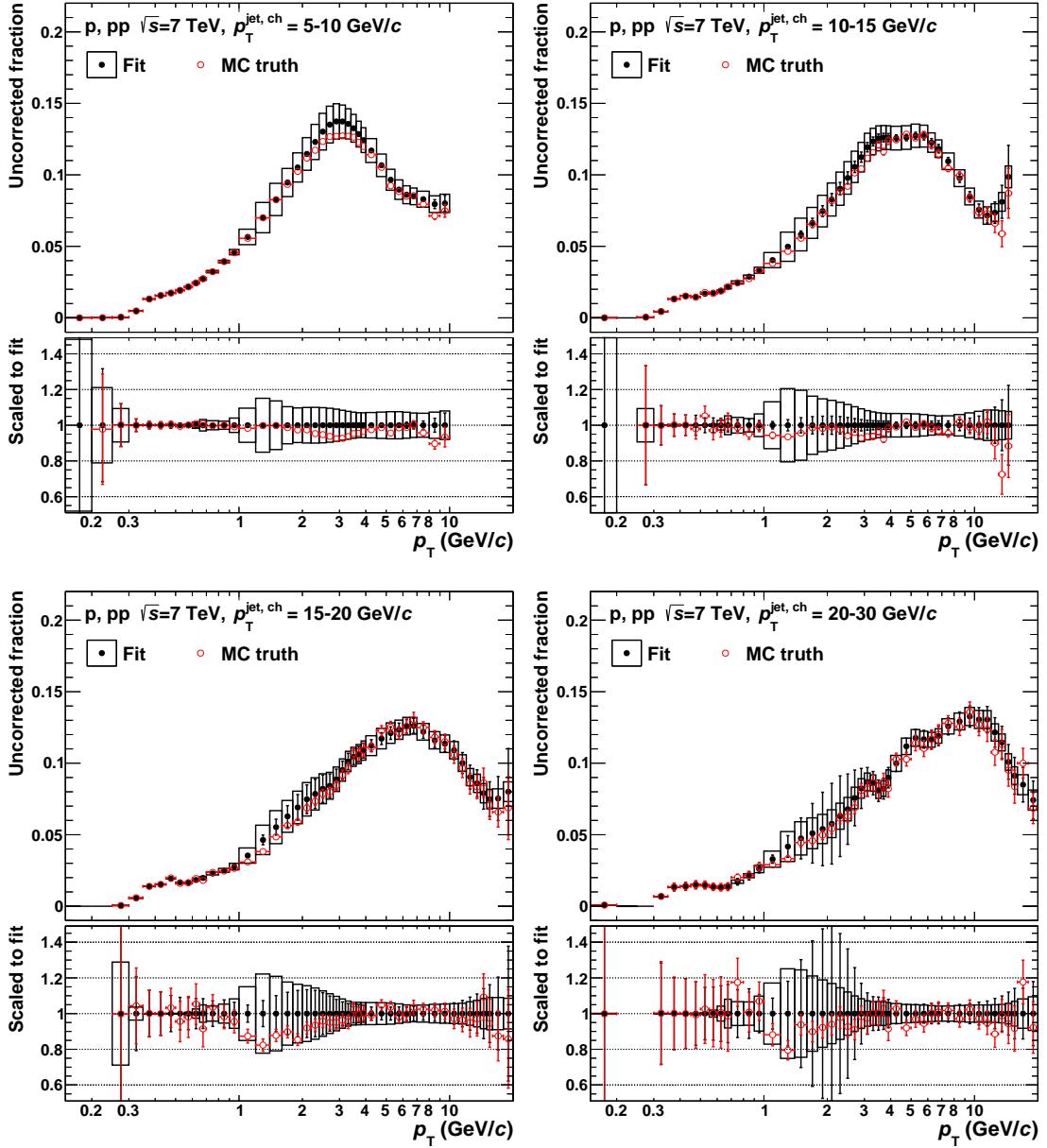


Figure 4.15.: MC blind test for protons fitted versus p_{T} . The raw fractions from fitting the MC output (black) are compared to the MC truth (red). The plots show (from upper left to lower right): $p_{\text{T}}^{\text{jet}} = 5 - 10, 10 - 15, 15 - 20$ and $20 - 30 \text{ GeV}/c$.

5. Monte Carlo Corrections

The charged particle spectra extracted with the Multi-Template Fit (MTF) need to be corrected for detector effects and secondary particle contamination. An MC correction procedure was developed for unidentified charged particles [34], including the machinery to estimate the systematic error. An overview on this method is included in this chapter. For this work, the method has been extended to extract the correction factors and estimate their systematic uncertainties for each particle species and different charges individually. The estimation of the systematic uncertainty of the secondary particle contamination has been changed significantly with respect to the unidentified version. Also, a correction for the muon contamination for the pions have been developed and the so-called “GEANT-FLUKA correction” has been added.

Detector effects like tracking efficiency, acceptance and p_T resolution as well as the contamination by secondary particles are estimated by generating events with PYTHIA [87] (tune Perugia-0 [114]) and transporting the particles through a full detector simulation with GEANT3 [89]. The MC particle ID is used to extract the corrections for each species and each charge separately.

The jet case will be considered first, followed by the inclusive case. The reason is that the inclusive case can be interpreted as a special case of the jets, where each event consists of nothing but one “jet” that is always reconstructed and the bin $p_T^{\text{jet}} \geq 0$ is considered. Many of the formulas presented in the next sections are valid for both inclusive and jets. Hence, they will be defined and discussed for both cases in the next section.

5.1. MC Corrections for the Jet Analysis

In case of jets, the impact of detector effects and secondaries on the track spectra additionally influences the charged jet reconstruction. The p_T smearing due to finite momentum resolution and tracking inefficiencies can change the reconstructed jet momentum. In particular, if high- z fragments get lost due to tracking inefficiency this can lead to bin migration for the reconstructed jet momentum.

To have sufficient statistics for the jet samples, events are generated in 4 parton p_T hard bins with $5 - 11$, $11 - 21$, $21 - 36$ and $36 - 57$ GeV/ c . These bins are merged using the cross-section of each bin as weight. A new MC production, labelled LHC14b6, has been used for this purpose. Every p_T hard bin comprises at least 4 million generated jets. The corresponding run list can be found in

5. Monte Carlo Corrections

Appendix B.1. The jet reconstruction uses the anti- k_T algorithm. The correction factors are extracted for the modified hybrid track cuts (cf. Section 6.1 and Appendix C.3) with an additional cut on at least 60 PID clusters. The details of the cuts do not matter for the following discussion of the technique.

5.1.1. Efficiency, Acceptance and p_T Resolution

To correct for detector effects, the reconstructed particle yields (detector level) are compared to the generated yields (particle level) individually for each species. Since the secondary contamination is corrected in a separate step, only physical primaries are considered for both yields. The jet reconstruction is carried out independently on detector and particle level. In particular, the scaled momentum observables z and ξ are calculated with the p_T^{jet} of the corresponding level. The correction is done bin-by-bin and is two-dimensional: one dimension is the p_T^{jet} , the other is p_T^{track} , z or ξ , respectively. In the following, the variable obs is one of p_T^{track} , z or ξ . The correction factors $corr_s$ are just the ratio of particle and detector level yields, which is basically the inverse of the efficiency. The p_T resolution is corrected by taking the reconstructed obs_{rec} — with index “rec” — for the yield on the detector level and the generated obs_{gen} — with index “gen” — on the particle level:

$$(corr_{s,\varepsilon}(obs))^{-1} \equiv \varepsilon_s(obs) \\ \equiv \begin{cases} \frac{dN_{\text{prim},s}^{\text{rec}}(obs_{\text{rec}})}{dobs_{\text{rec}}} / \frac{dN_{\text{prim},s}^{\text{gen}}(obs_{\text{gen}})}{dobs_{\text{gen}}}, & \text{inclusive} \\ \left(\frac{1}{N_{\text{Jets}}^{\text{rec}}} \frac{dN_{\text{prim},s}^{\text{rec}}(obs_{\text{rec}})}{dobs_{\text{rec}}} \right) / \left(\frac{1}{N_{\text{Jets}}^{\text{gen}}} \frac{dN_{\text{prim},s}^{\text{gen}}(obs_{\text{gen}})}{dobs_{\text{gen}}} \right), & \text{jets} \end{cases}, \quad (5.1)$$

where s is the species and $N_{\text{prim},s}^{\text{rec}}(obs_{\text{rec}})$ and $N_{\text{prim},s}^{\text{gen}}(obs_{\text{gen}})$ are the yields in the corresponding obs bin at the detector and particle level, respectively. In case of jets, the reconstructed spectrum in the numerator is normalised to the number of reconstructed jets, $N_{\text{Jets}}^{\text{rec}}$, and the generated spectrum in the denominator accordingly to the number of generated jets, $N_{\text{Jets}}^{\text{gen}}$. Note that an implicit dependence on p_T^{jet} is understood for all variables.

The p_T resolution smears the efficiency curves. Since the jet reconstruction is carried out independently on generated and reconstructed level, bin migration may occur. For instance, high- z fragments might get lost due to tracking inefficiency, which may change the p_T^{jet} such that it ends up in a lower p_T^{jet} bin (“feed-down”). The bin migration effect is clearly visible for the efficiency of jet particles, as can be seen from Fig. 5.1. At high p_T (and high z), the efficiency rises and goes beyond unity. In that sense, the (inverse) correction factor in Eq. 5.1 is a convolution, namely efficiency \times acceptance \times p_T resolution.

The rising trend at low p_T is a consequence of the decreasing specific energy loss in the material (for hadrons, for electron it is rather constant) and a larger bending radius, which both increase the probability for a sufficiently large track length for

successful reconstruction inside the TPC. Very low- p_T protons get absorbed in the material with a high probability, which reduces the detection efficiency basically to zero. Around $p_T = 0.8 - 1 \text{ GeV}/c$, a slight drop of the efficiency sets in because the bending of the tracks becomes very small and a larger fraction of the track clusters may end up in the dead zones between the TPC sector boundaries. This effect saturates at higher p_T and is over-compensated by the feed-down from higher p_T^{jet} . The kaon efficiency is depleted due to weak decays. The depletion gets smaller for higher p_T due to the boost with respect to the lab frame.

Note that the correction factors (the inverse of the efficiency) for kaons and protons are of order $10 - 100$ in the lowest p_T bins.

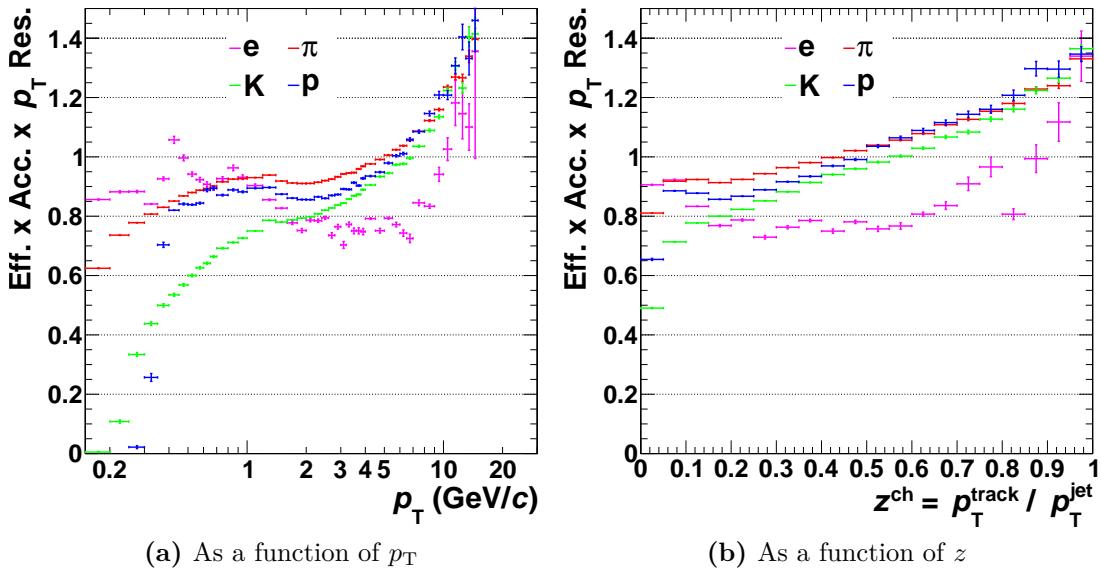


Figure 5.1.: The (inverse) efficiency \times acceptance \times p_T resolution correction factors for particles in jets with $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ for both charges are shown as a function of **a**) p_T and **b**) z . The impact of the p_T resolution is discussed in the text.

5.1.2. Secondary Contamination

Primary particles are selected via a cut on the Distance of Closest Approach (DCA) with respect to the primary vertex. However, there are also secondary particles that can pass this cut. They can be produced by photon conversions, hadronic interactions with detector material and weak decays of charged pions or of strange particles (e.g. Λ , K_S^0). All secondary particles from these sources are considered as contamination of the measurement, since they have either nothing to do with the fragmentation process or the original collision (true for photon conversions and hadronic interactions) or stem from neutral (strange) particles, which are not considered in this analysis of identified charged particles.

5. Monte Carlo Corrections

The impact of secondary particles on the results is twofold: they contribute to the reconstructed p_T^{jet} and to the track spectra. The first effect is already implicitly corrected by the bin-by-bin correction discussed in Section 5.1.1, since on the particle level only physical primaries are used for the jet reconstruction. The secondary contamination of the track spectra is estimated from the fraction of physical primaries in the track sample passing all cuts on the detector level. Again, a bin-by-bin correction factor is used for each species in bins of p_T^{jet} and obs .

A study [115] shows that PYTHIA does not reproduce the observed strangeness production in pp collisions at $\sqrt{s} = 7 \text{ TeV}$. The measured results were obtained with the CMS experiment from non-single-diffractive events in Minimum Bias (MB) collisions for rapidity $|y| < 2$ and the ratio to various MC models can be retrieved from the mcplots web page [116]. To correct this effect, a strangeness scaling is utilised: the spectra of secondary particles originating from strange mothers ($K_S^0, K^-, \Lambda, \Xi^0, \Xi^-$ or associated anti-particles) are scaled by the corresponding factors from the CMS measurement. The yield ratio between MC and CMS data is depicted in Fig. 5.2. The scale factor is the inverse of the corresponding ratio. The measured K_S^0 yield is not corrected for feed-down and the measured Λ yield not corrected for feed-down from Σ . The correction factor is obtained consistently in this work. Note that the scale factors are obtained for MB and it is neither clear whether the strangeness scaling is needed in case of jets, nor whether the scale factors remain the same. Also, the rapidity range of the CMS measurement is larger than for this analysis — though, there is no change expected. Therefore, a large systematic error will be assigned to the strangeness scaling as will be discussed in Section 5.3.

The secondary contamination of each species s is corrected bin-by-bin with the primary fraction correction factor:

$$f_{\text{prim},s}(obs) \equiv \frac{N_{\text{prim},s}^{\text{rec}}(obs)}{N_{\text{prim},s}^{\text{rec}}(obs) + N_{\text{sec},s}^{\text{rec}}(obs)}, \quad (5.2)$$

where $N_{\text{prim},s}^{\text{rec}}(obs)$ and $N_{\text{sec},s}^{\text{rec}}(obs)$ are the number of reconstructed primary and secondary tracks in the considered obs bin, respectively. Note that an implicit dependence on p_T^{jet} is understood for all variables.

The primary fractions for particles in jets with $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ are shown in Fig. 5.3. The strangeness scaling considerably changes the proton primary fraction by several per cent at low p_T . The impact decreases with rising p_T and eventually disappears around $6 \text{ GeV}/c$. There is also some minor impact on the pion primary fraction of up to 1 – 2% at lowest p_T . The primary electrons are dominated by secondary electrons from photon conversion. There are also many secondary protons at low p_T from weak decays and hadronic interactions. In total, the correction factor for protons is a factor 10 different from unity for the lowest p_T bins. If the primary fractions are extracted as a function of z , the steep rise of the protons is shifted to the left and ends up in the very first bin.

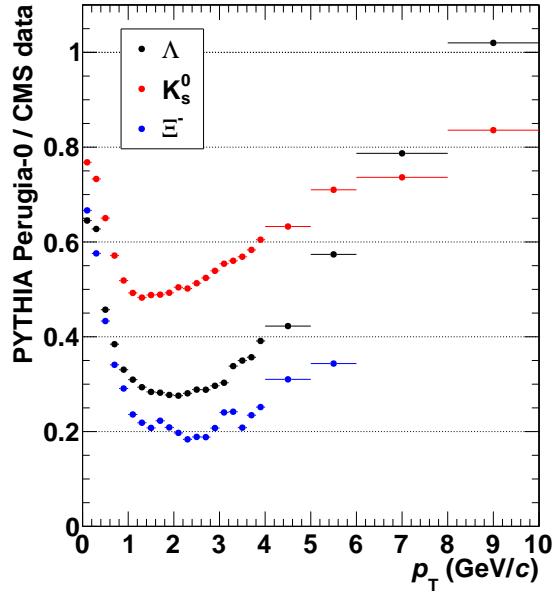


Figure 5.2.: Yield ratio between MC and CMS measurement for different strange particles: Λ , K_s^0 and Ξ^- . The MC yields are for PYTHIA tune Perugia-0, which is the same tune that was used to extract the MC corrections in this work. The inverse of the corresponding ratio is used as a strangeness scale factor for the secondary correction.

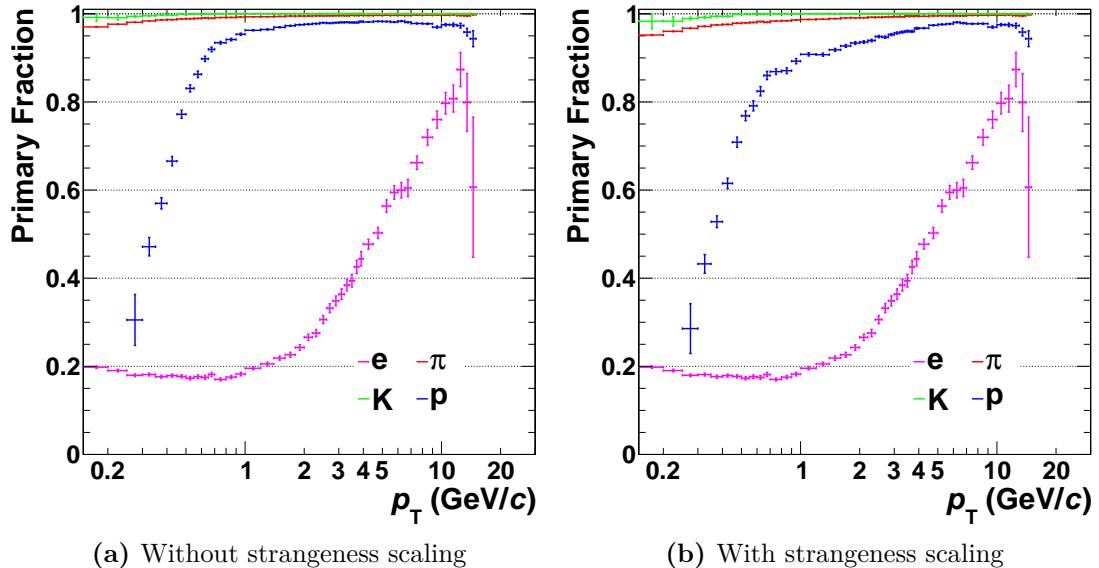


Figure 5.3.: The primary fractions for tracks in jets with $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ are shown as a function of p_T a) without and b) with strangeness scaling. The results are for both charges.

5. Monte Carlo Corrections

5.1.3. Muon Contamination

Since the TPC resolution does not allow for the separation of pions and muons (cf. Section 4.3.4), the MTF extracts the summed raw yield ($\pi_{\text{all}} + \mu_{\text{all}}$)(obs) as a function of obs . The index “all” denotes that both primary and secondary particles contribute. To correct for the muon contamination, the particle fraction of each species s needs to be multiplied by the correction factor $f_{\pi-\mu,s}(obs)$. It is defined as:

$$f_{\pi-\mu,s}(obs) \equiv \begin{cases} g(obs), & s = \pi \\ 1, & \text{otherwise} \end{cases}, \quad (5.3)$$

$$g(obs) \equiv \frac{N_{\text{all},\pi}^{\text{rec}}(obs)}{N_{\text{all},\pi}^{\text{rec}}(obs) + N_{\text{all},\mu}^{\text{rec}}(obs)}, \quad (5.4)$$

which is a symbolical notation meaning that the correction factor is unity for all species other than pion and has a bin-by-bin correction value for pions. Note that the correction factor $g(obs)$ takes into account strangeness scaling and is extracted from the MC true pion ($N_{\text{all},\pi}^{\text{rec}}(obs)$) and muon ($N_{\text{all},\mu}^{\text{rec}}(obs)$) yields of the corresponding bin. Again, an implicit dependence on p_T^{jet} is understood for all variables.

As illustrated in Fig. 5.4, the muon contamination is below 2% at all p_T and smaller than 1.5% in most of the bins. The dominant fraction of the muons stems from weak decays (predominantly of charged pions). The impact of the strangeness scaling on the correction factor has negligible impact. The same observations apply for $obs = z, \xi$.

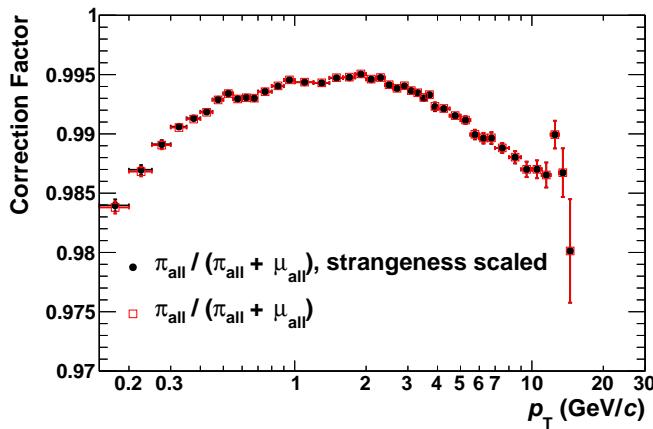


Figure 5.4.: Correction factors for the muon contamination of the pions in jets with $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$. The difference between the correction factors with (black dots) and without (red open boxes) strangeness scaling is negligible.

5.1.4. Total Correction Factors

The combination of all correction factors is shown in Fig. 5.5 for $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$. The dependence on the jet momentum turns out to be weak. Note that the correction factors for kaons and protons (and the corresponding to-pion ratios) become larger than 2 for low p_T^{track} and are not visible in the figures. Due to the steep fall of the efficiency, there can be a huge jump of the total correction factor from one bin to the next.

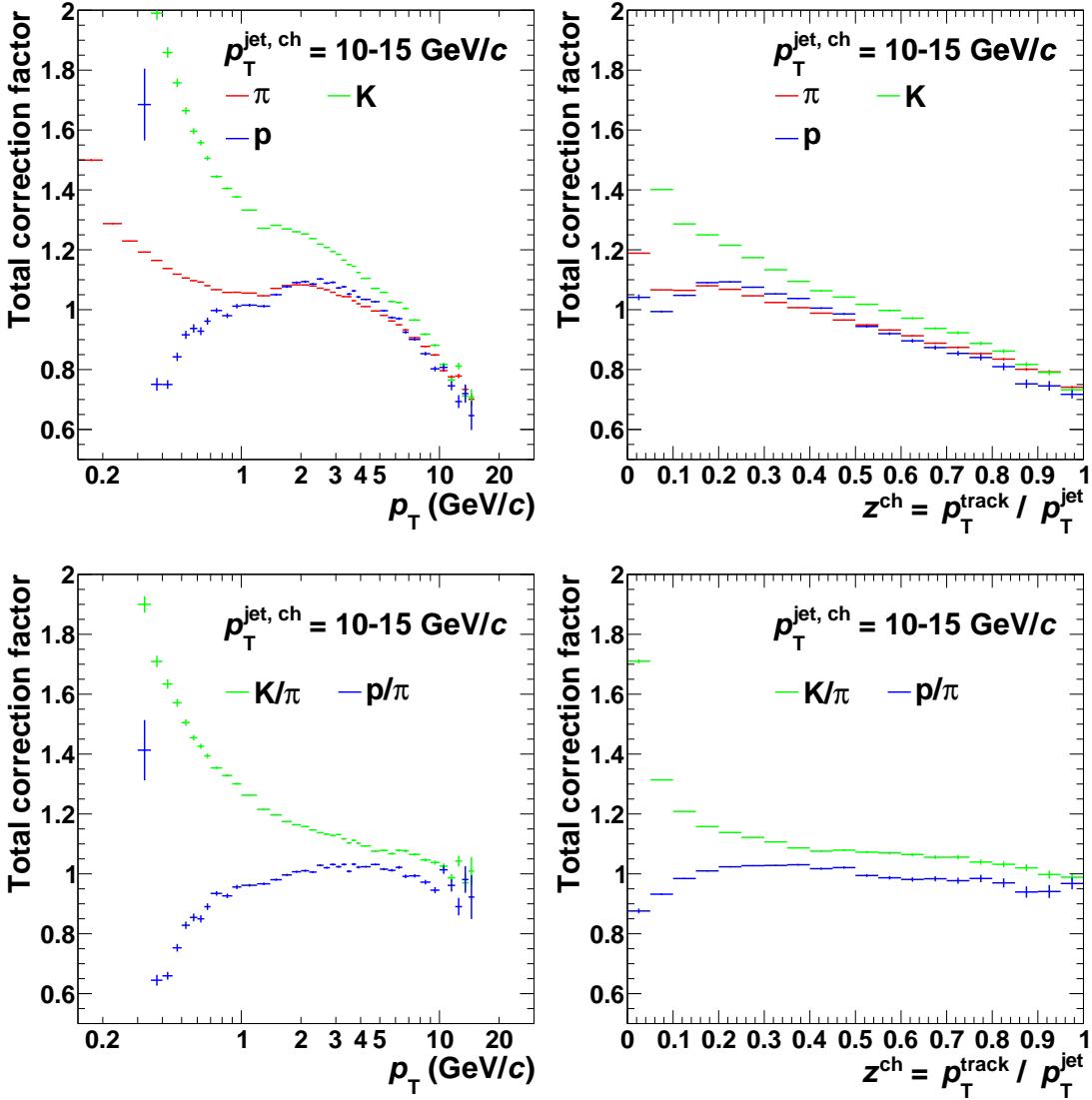


Figure 5.5.: Combined MC correction factors for the yields (top row) and to-pion ratios (bottom row) in jets with $p_T^{\text{jet}, \text{ch}} = 10 - 15 \text{ GeV}/c$ as a function of p_T^{track} (left) and z (right).

5.2. MC Corrections for the Inclusive Analysis

The correction factors for the inclusive case have been extracted from a MC minimum bias production. The run list can be found in Appendix B.1. The same physics selection and event cuts as used for the data analysis are applied to the MC production (cf. Section 6.1), which yields about 103 million events.

There are two set of cuts with corresponding correction factors used for the inclusive analysis. They will be discussed in detail in Section 6.1. Here, only a brief overview is given. The first set is obtained for the modified hybrid track cuts with some additional PID cuts and for tracks with $|\eta| < 0.9$. It will be used as a reference for the jet analysis, which uses the same set of cuts. The second set are the standard cuts for tracks within $|\eta| < 0.8$ and also with some additional PID cuts. The inclusive analysis in multiplicity bins will use this set. The machinery is the same for both cut sets. Since the overall trends of efficiency, etc. are similar, this section will focus on the modified hybrid cuts.

5.2.1. GEANT-FLUKA Correction

It was found [117] that the GEANT3 version used as transport model for the detector simulation overestimates the reaction cross-section in the detector material for anti-protons and negative kaons. The transport model FLUKA [91] gives a more accurate description. Thus, the ratio of the tracking efficiencies from GEANT3 and FLUKA is parametrised [117] and used as a correction factor:

$$R_s^{\text{G-F}}(p_{\text{T}}) \equiv \frac{\epsilon_s^{\text{GEANT3}}(p_{\text{T}})}{\epsilon_s^{\text{FLUKA}}(p_{\text{T}})} = \begin{cases} 1 - 0.13 \cdot \exp(-0.68 p_{\text{T}}), & s = \bar{p} \\ \min\{0.973 + 0.012 p_{\text{T}}, 1\}, & s = K^- \\ 1, & \text{otherwise} \end{cases}, \quad (5.5)$$

where the particle level momentum is used for p_{T} . The correction factors for anti-protons and negative kaons are shown in Fig. 5.6. They are about 2.5% for negative kaons, up to 12% for anti-protons at lowest p_{T} and approach zero for higher p_{T} .

Note that the MC jet production was carried out with a newer GEANT3 version that already uses the correct cross-section. Hence, the GEANT-FLUKA correction is only needed for results obtained from the old minimum bias MC production.

5.2.2. Efficiency for Inclusive Particle Production

The correction factors for the inclusive case are extracted similarly to jets by just taking all accepted particles in the event, without jet reconstruction and without normalising the spectra to the number of jets, see Eq. 5.1. The resulting efficiency is shown in Fig. 5.7 for both positively and negatively charged particles.

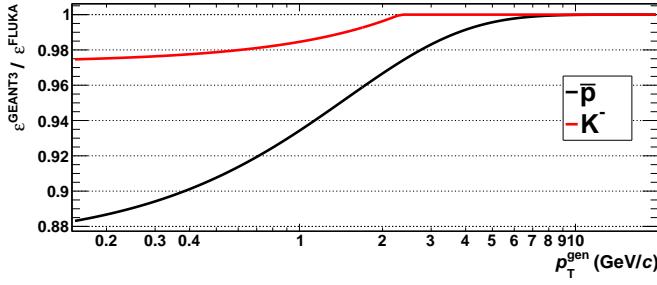


Figure 5.6.: GEANT-FLUKA correction factors for the efficiency of anti-protons (black) and negative kaons (red).

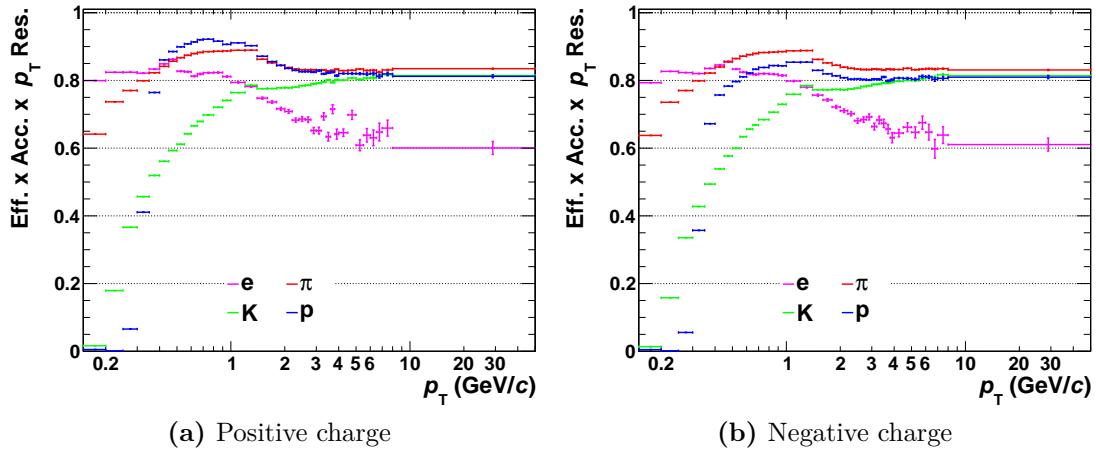


Figure 5.7.: The (inverse) efficiency \times acceptance $\times p_T$ resolution correction factors for inclusive particle production are shown for particles with a) positive and b) negative charge. The GEANT-FLUKA correction is already included.

The GEANT-FLUKA correction is already applied. The charge dependence of the proton and kaon efficiencies is mainly due to absorption in material and isospin effects. The low and intermediate p_T region is similar to the jet case. The dropping trend due to the TPC dead zones around $p_T = 1 \text{ GeV}/c$, that is also visible for the jet case in Fig. 5.1, saturates at about $4 \text{ GeV}/c$.

5.2.3. Secondary and Muon Contamination

The primary fractions (cf. Eq. 5.2) shown in Fig. 5.8 for the inclusive case can be compared to those for particles in jets with $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$, illustrated in Fig. 5.3. The overall trends are similar.

The muon contamination (see Eq. 5.3) for the inclusive case agrees within 1 – 2% with the jet case, which was shown in Fig. 5.4.

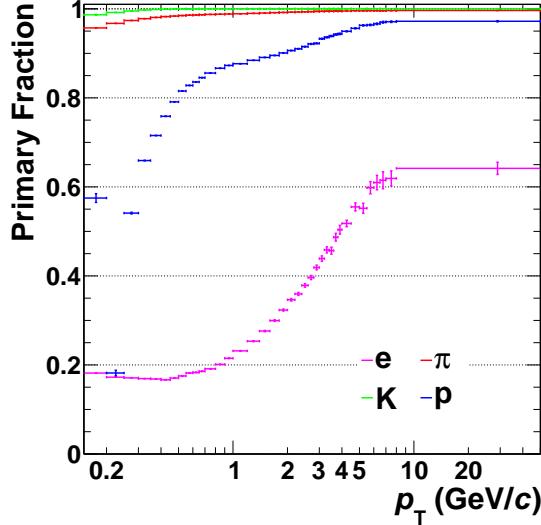


Figure 5.8.: The primary fractions for inclusive spectra are shown for both charges and including strangeness scaling.

5.2.4. Total Correction Factors

The combination of all correction factors is shown in Fig. 5.9. As already observed for the jet case, the correction factors for kaons and protons may become larger than 2 for low p_T^{track} due to the steeply falling efficiency and are not visible. The figure also compares the modified hybrid track cuts to the standard cuts. The latter have larger correction factors due to a smaller efficiency. Note that the small proton primary fraction at low p_T partly balances the low efficiency, the correction factor of which is the inverse. Therefore, the corrections are not necessarily small if the total correction factor is close to unity.

5.2.5. MC Correction Factors for the Multiplicity Dependent Analysis

As discussed at the beginning of the section, the multiplicity dependent analysis uses standard ITS-TPC track cuts. These have a slightly lower efficiency than the modified hybrid track cuts, but a larger primary fraction. The statistics in a given multiplicity bin is rather low, such that there are considerable statistical fluctuations and correspondingly large errors of the correction factors of each multiplicity bin. In this section, it will be shown that the correction factors are largely independent of multiplicity, which justifies to extract them for a single broad multiplicity bin with high statistics and apply them to all bins.

Two different multiplicity estimators will be considered, namely the *Reference Multiplicity* (RM) and the *V0M multiplicity* estimator. They will be described in Section 6.5.1. Here, it is sufficient to know that RM is highly correlated with

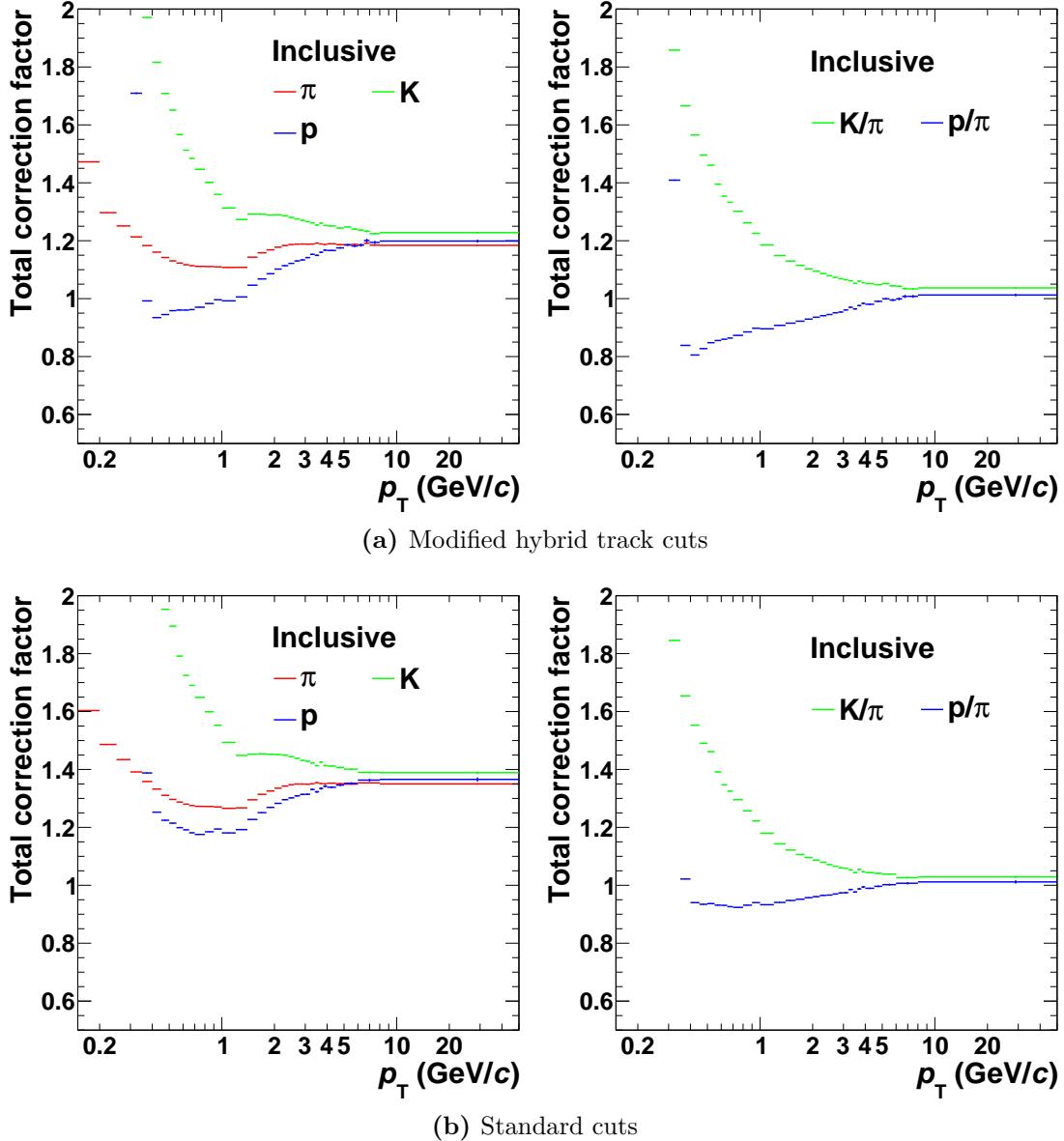


Figure 5.9.: Combined MC correction factors for the inclusive yields (left) and to-pion ratios (right). The values are for **a)** the modified hybrid track cuts (jet analysis) and **b)** the standard cuts (multiplicity analysis), respectively.

the tracks in the analysis region, whereas V0M is rather independent. This may introduce different biases on the MC correction factors.

Note that bins with too low efficiency or too large systematic error of the MC correction are excluded from the analysis because the correction is unreliable in that case. This will be discussed in Section 5.5 and leads to an effective p_T cut-off around 0.35 GeV/c for kaons and protons. Correspondingly, the following comparison will restrict itself to the $p_T > 0.35$ GeV/c region.

5. Monte Carlo Corrections

The bin-by-bin correction factors for efficiency, acceptance and p_T resolution are shown in Fig. 5.10 for MB and various bins of V0M multiplicity. The corresponding factors for RM are shown in Fig. 5.11. Evidently, the correction factors are the same for all classes of V0M multiplicity modulo statistical fluctuations and in particular agree with those from MB (chosen as reference for the ratio in the lower panel of the figure). This justifies to apply the MB correction factors, which have the highest statistics, to all V0M multiplicity bins. However, the situation is different for RM, where an overall 1 – 2% spread between the multiplicity classes is observed. In addition, the 1 – 3 multiplicity bin deviates by up to 6% from other bins. For this multiplicity estimator, the bin $4 - \infty$ (chosen as reference for the ratio in the lower panel of the figure) is used to correct all multiplicity bins. In particular, the chosen range for this bin removes contributions from events with undefined reference multiplicity¹.

In Figs. 5.12 and 5.13, the primary fractions are shown for MB and bins of V0M and RM, respectively. Keeping in mind the p_T cut-off, the primary fractions of kaons and pions of different multiplicity bins agree with better than $\pm 1\%$. However, this observation holds true for protons only for $p_T \gtrsim 0.6 \text{ GeV}/c$. The spread is about $\pm 4 - 8\%$ at $p_T = 0.35 \text{ GeV}/c$. In any case, this affects only the very first few bins of the proton spectra.

In summary, in order not to introduce large statistical fluctuations by using individual correction factors for each multiplicity bin, the correction factors are extracted for the bin $4 - \infty$ and MB for RM and V0M, respectively, and then applied to all other bins.

¹This can happen if an event does not satisfy requirements on the vertex for the multiplicity estimator, which are stronger than those for the analyses.

5.2. MC Corrections for the Inclusive Analysis

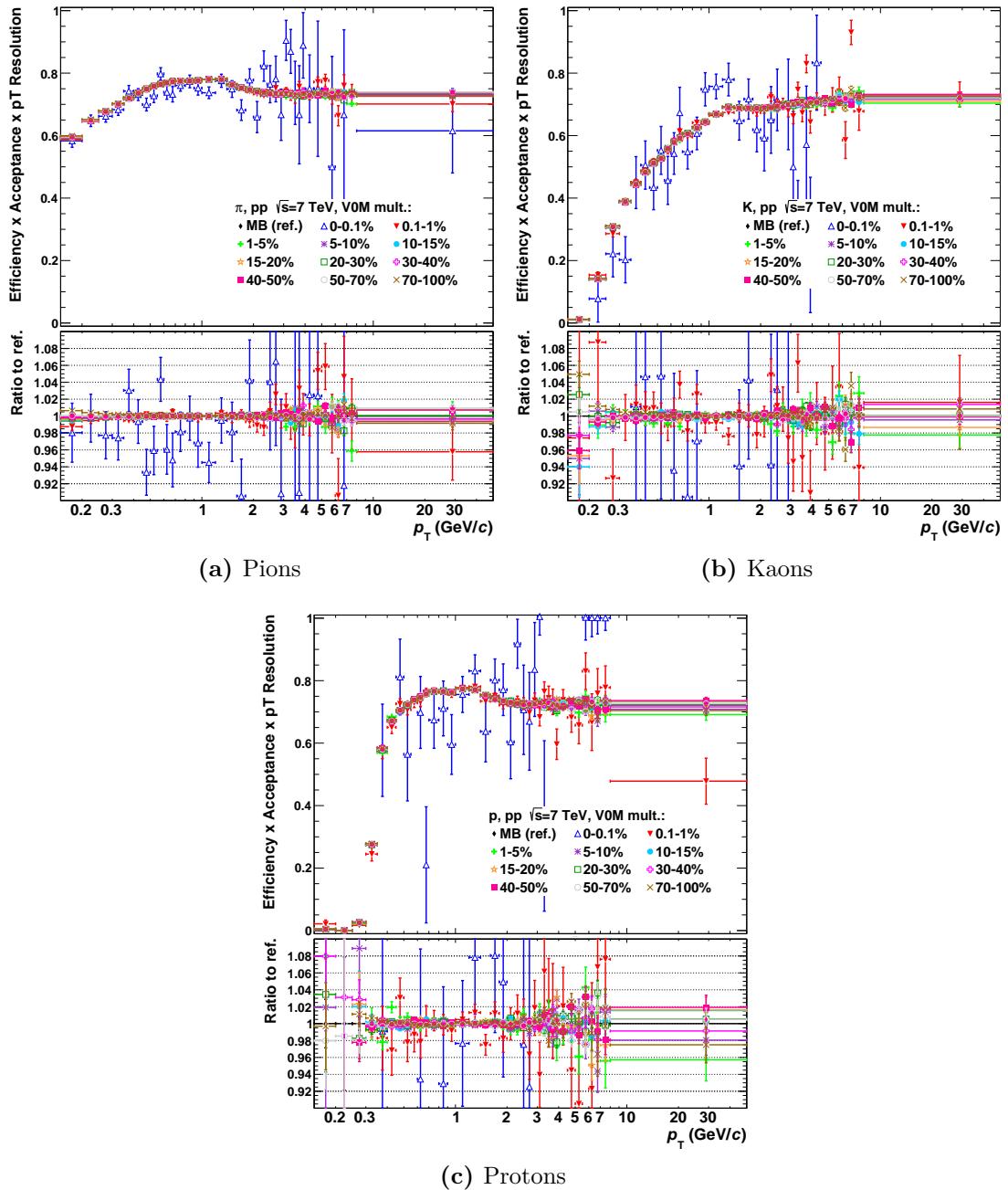


Figure 5.10.: (Inverse) efficiency \times acceptance $\times p_T$ resolution correction factors for different bins of V0M multiplicity and MB.

5. Monte Carlo Corrections

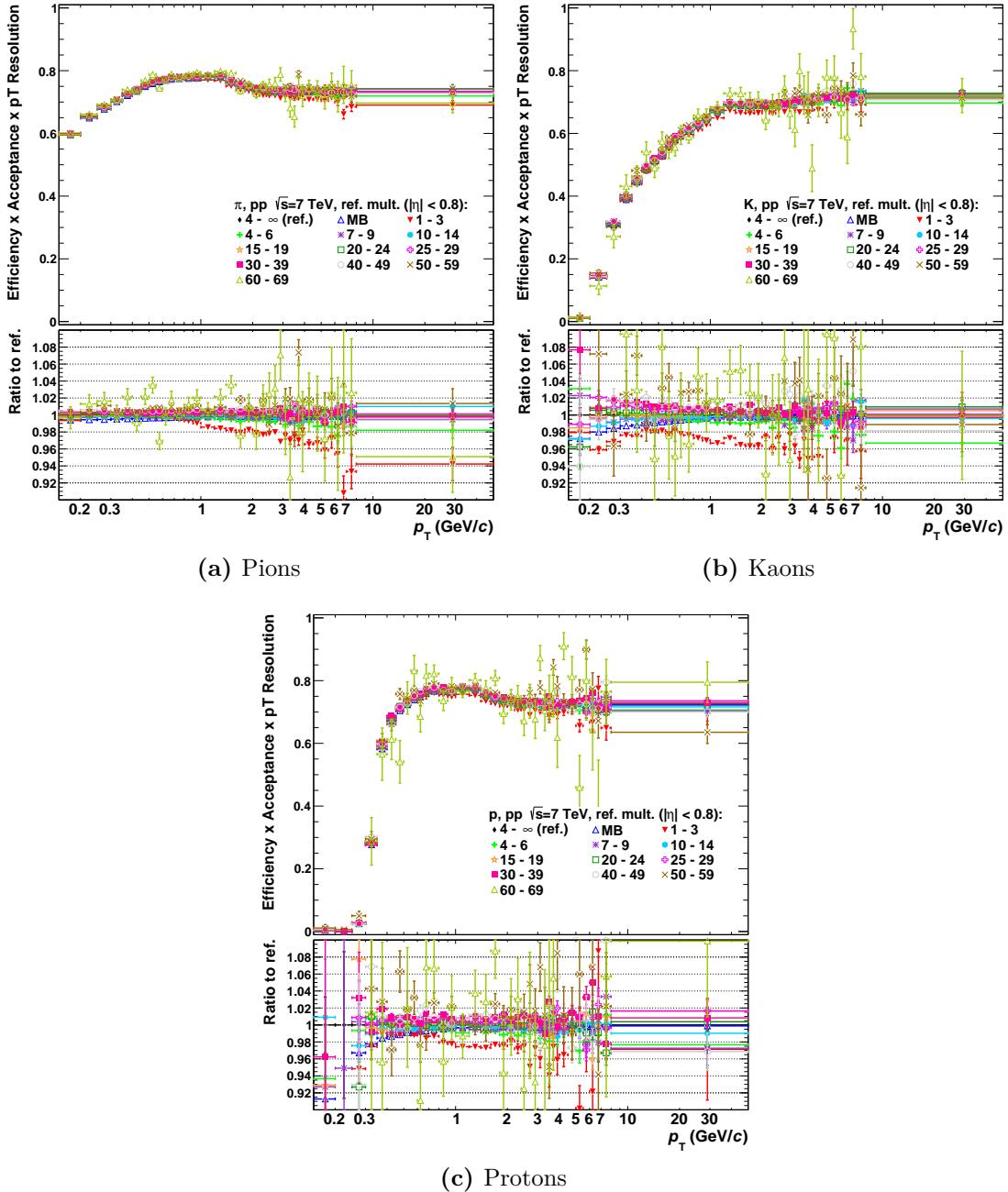


Figure 5.11.: (Inverse) efficiency \times acceptance $\times p_T$ resolution correction factors for different bins of reference multiplicity and MB.

5.2. MC Corrections for the Inclusive Analysis

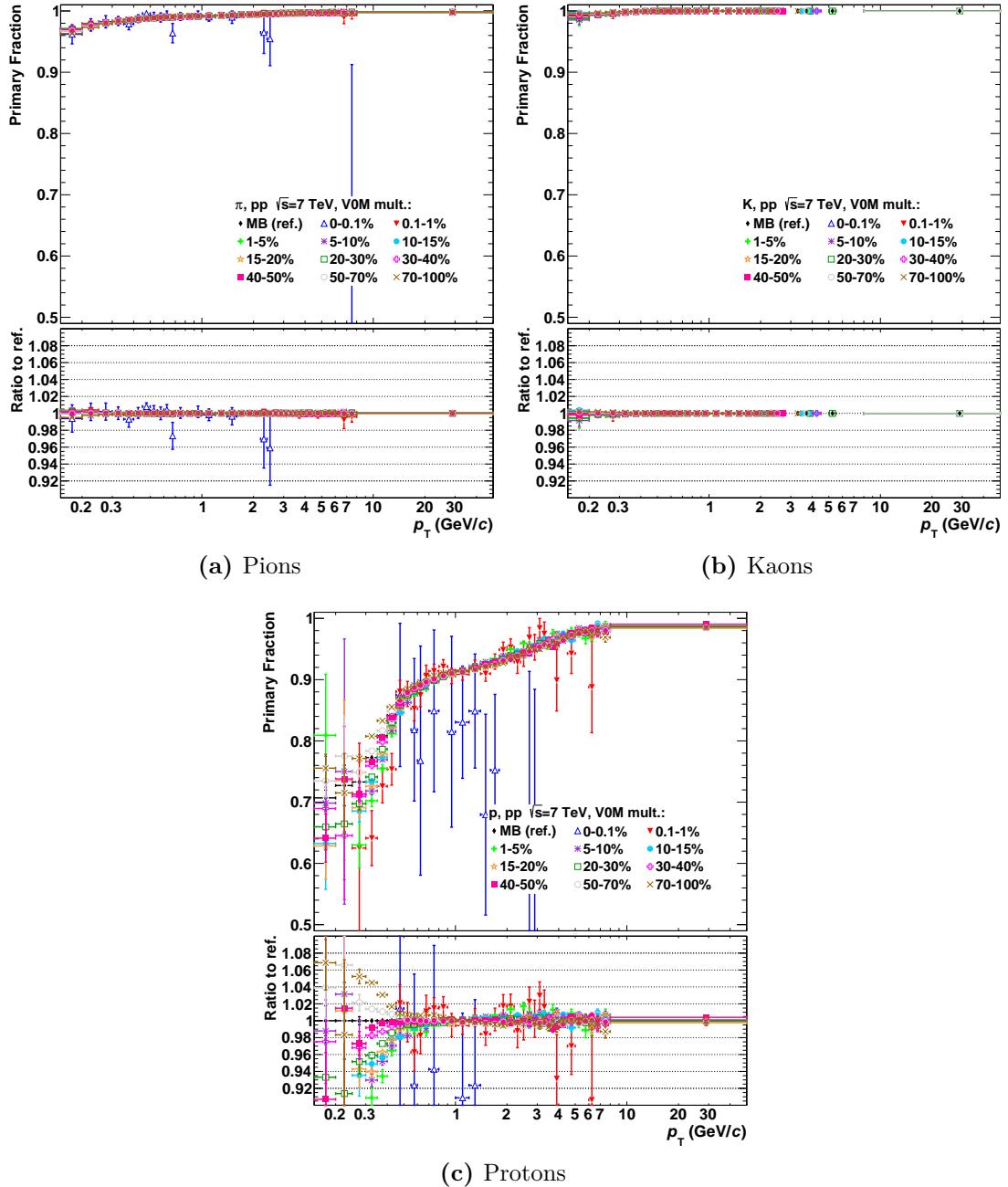


Figure 5.12.: Primary fractions for different bins of V0M multiplicity and MB.

5. Monte Carlo Corrections

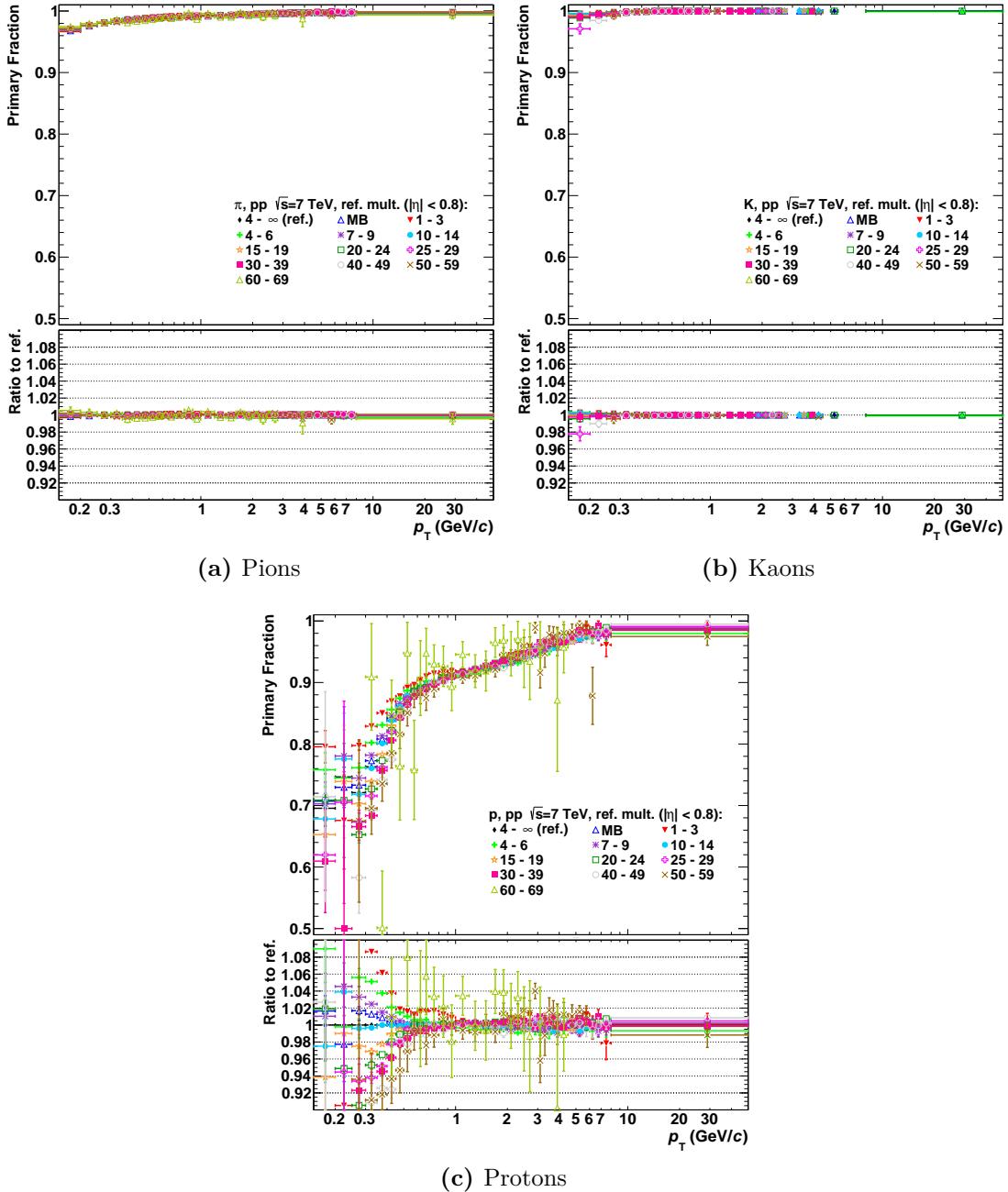


Figure 5.13.: Primary fractions for different bins of reference multiplicity and MB.

5.3. Systematic Uncertainty of the MC Correction Factors

This section discusses the estimation of the systematic uncertainties of the MC bin-by-bin correction factors. The procedure is based on that for charged particles [34], but is carried out for each species (and charge) individually. In addition, the estimation for the uncertainty of the secondary contamination was changed significantly. Note that, in contrast to [34], not only leading, but all jets are analysed for the jet case and there is no subtraction of the underlying event.

For the inclusive analysis, the systematic uncertainties of each source can be estimated by just assigning a reasonable error relative to the correction factor. However, there are cases for jets that are more subtle. The tracking efficiency and momentum resolution have a non-trivial impact on the particle spectra in jets, since they also influence the jet reconstruction. Furthermore, the jet shape depends on the event generator. To estimate the systematic uncertainties of these sources for jets, a *fast simulation* is used. PYTHIA events are generated and the charged final-state particles are extracted. Instead of a full detector simulation, the efficiency and resolution are parametrised and applied track-by-track. The anti- k_T algorithm is used to reconstruct jets of the original “particle level” and of the “detector level” (i.e. after applying efficiency parametrisations, rejecting/adding tracks, etc.) individually. The generated and reconstructed particle spectra are then used to extract the correction factors. The parameters of the fast simulation (efficiency, resolution, generator tune, etc.) are varied within their estimated systematic uncertainties. The spread of the resulting correction factors is then assigned as the systematic uncertainty of the bin-by-bin correction.

5.3.1. Tracking Efficiency and Momentum Resolution

According to the charged particle analysis [34], the systematic uncertainty of the efficiency correction is 5% and the momentum resolution has an uncertainty of 20%. These estimates are adopted for this work. In the inclusive case, the uncertainty of the efficiency correction factor is just set to 5%. For the corresponding uncertainty for jets, the fast simulation will be used as explained in the following. Furthermore, the uncertainty of the momentum resolution needs to be propagated to the uncertainties of the bin-by-bin correction factors, which is again achieved by the fast simulation.

In Fig. 5.14, the single track efficiencies for positive particles are shown, which are obtained from the full simulation. A parametrisation — a piecewise multipolynomial fit — is fitted to each histogram to get a smooth and continuous description, which is used for the fast simulation. The same is done for negative particles, which yields similar plots. The single track efficiencies are the ratios of the generated and reconstructed spectra (after all cuts) for primaries, without

5. Monte Carlo Corrections

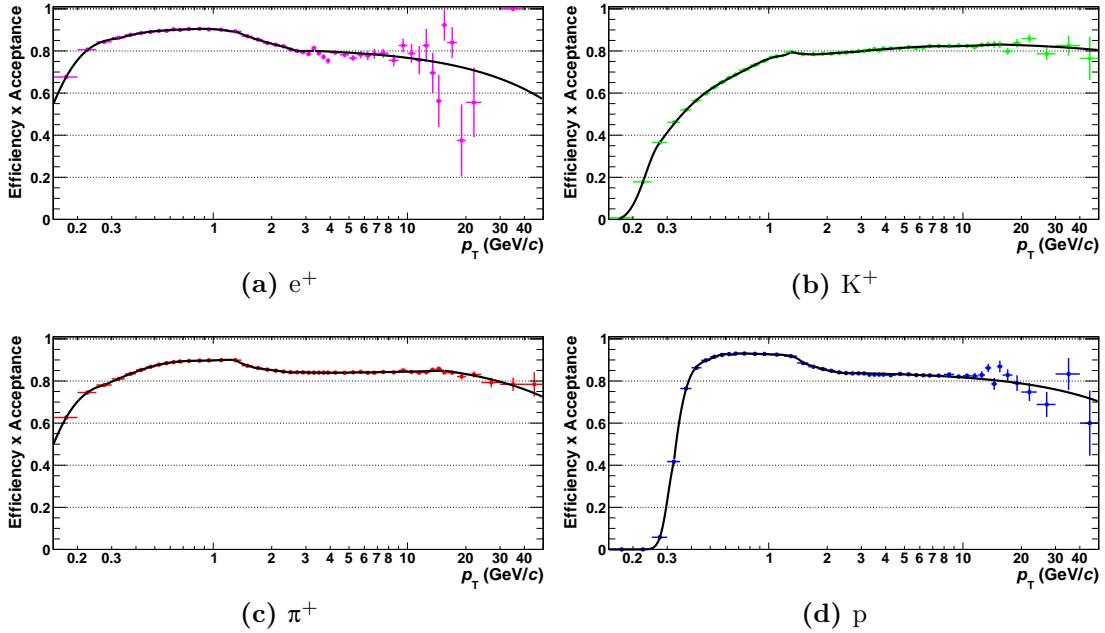


Figure 5.14.: The single track efficiencies for positive particles are shown. The black lines are a parametrisation of the data points that allow for a smooth description.

taking into account resolution effects and without normalisation. In the fast simulation, the inverse p_T of each track is smeared with a Gaussian with width $\sigma = 0.002 (\text{GeV}/c)^{-1}$, which is the Kalman filter estimate for the resolution during the tracking procedure.

The systematic errors of efficiency and resolution are conservatively estimated via varying the tracking efficiency by $\pm 5\%$ and the momentum resolution by $\pm 20\%$. The bin-by-bin correction factors of full and fast simulations are compared in Fig. 5.15 for $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$ jets as a function of p_T^{track} and z . The corrections for all charged particles, π , K and p are shown. The cyan error bands indicate the variation of the correction factors from the fast simulation, when the efficiency is changed by $\pm 5\%$. The overall agreement between full and fast simulations is good. Note that it is not necessary to have perfect agreement, since the systematic error is taken as the relative change of the correction factors for the variations in the fast simulation. The efficiency variation leads to typically 5% uncertainty of the correction factors at low z . The uncertainty decreases towards intermediate z , when the correction factors get close to unity, and increase again for higher z because p_T^{jet} bin migration effects become important.

Similarly, the effect of a variation of the resolution by $\pm 20\%$ is studied. The transverse momentum spectra in jets are relatively flat. Hence, the variation of the resolution has only small impact on the correction factors. The resulting uncertainties are below 1% in most bins. From these jet results, the resolution systematic uncertainty of the bin-by-bin correction factors for the inclusive case

5.3. Systematic Uncertainty of the MC Correction Factors

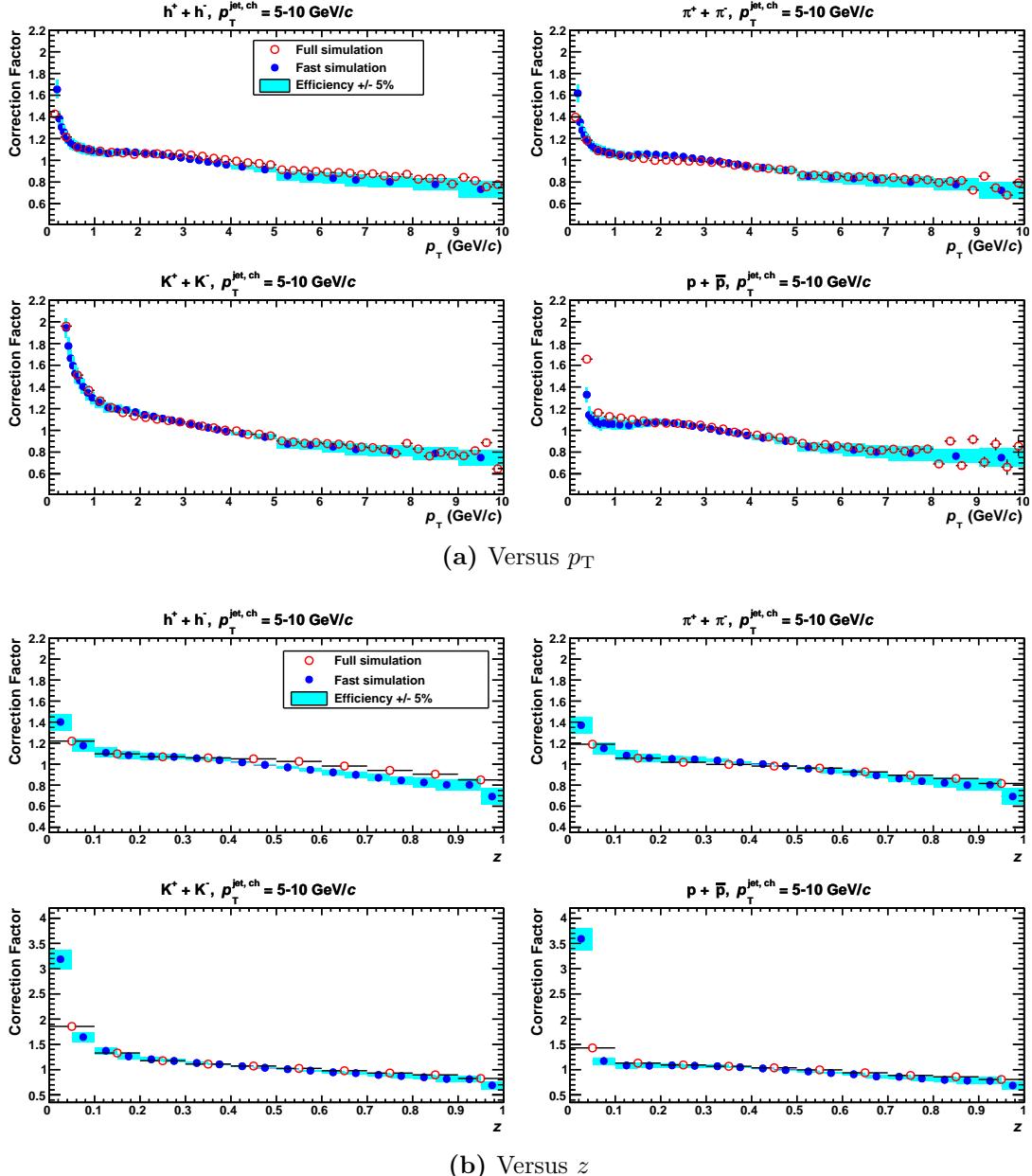


Figure 5.15.: Tracking efficiency systematics. The bin-by-bin correction factors as a function of **a)** p_T and **b)** z are shown for all charged particles, π , K and p for $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$. The full simulation (open red points) is compared to the fast simulation (full blue points). The cyan bands indicate variations of the toy MC correction factors with efficiency.

5. Monte Carlo Corrections

is estimated to be 2%, independent of p_T . On the one hand, the corresponding error in the jet case is typically smaller than this value in all p_T and all p_T^{track} bins². On the other hand, the inclusive p_T spectra are much softer than spectra in jets and, thus, have a more critical p_T dependence. The presented value tries to take into account these competing effects.

5.3.2. Shape Dependence

The bin-by-bin correction factors rely on the correct description of the jet spectrum and the jet shapes, like the (longitudinal) fragmentation distributions shown in Fig. 5.16. The systematic error of this shape dependence is estimated by comparing the correction factors obtained from simulations with modified fragmentation patterns:

- PYTHIA tune Perugia-0.
- PYTHIA with low- p_T depletion: the leading track is kept, but the other jet constituents are randomly discarded with a survival probability of 75%.
- PYTHIA with low- p_T enhancement: the jet constituents (except for the leading track) are artificially duplicated. The additional constituents are then randomly discarded with a momentum dependent survival probability, which is 75% for the lowest and 25% for the highest momentum.

For the low- p_T enhancement and depletion, the jet reconstruction is re-run after the duplication and removal of jet constituents. The duplicated particles are rotated by 180 degrees in azimuth with respect to the jet axis. Note that in [34], also other event generators like HERWIG or PYTHIA tune Perugia-2011 have been included into the comparison. It turned out that the manual low- p_T enhancement and depletion typically provides the most extreme variations. Hence, the systematic error study for this analysis was restricted to these variations.

The impact on the fragmentation distributions is illustrated in Fig. 5.16 for $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$ as a function of z . The comparison to the unmodified spectra is shown for all charged particles, π , K and p. As expected, removing (duplicating) low- p_T jet particles depletes (enhances) the fragmentation distribution at low z . It is instructive to look at the reduction of the fragmentation distribution at intermediate z in case of enhanced low- p_T . It can be explained by the fact that the transverse momenta of all constituents (also the duplicated ones) sum up to the jet transverse momentum, which can introduce p_T^{jet} bin migration. The resulting systematic uncertainty of the correction factors from this source increases with p_T (z) and typically amounts to $\pm 5\%$.

²In Section 5.3.9, a detailed look will reveal that kaons and protons have larger systematic errors from this source at very low p_T . However, these bins are below a cut-off and will not be used for the analysis. The cut-off will be discussed in Section 5.5.

5.3. Systematic Uncertainty of the MC Correction Factors

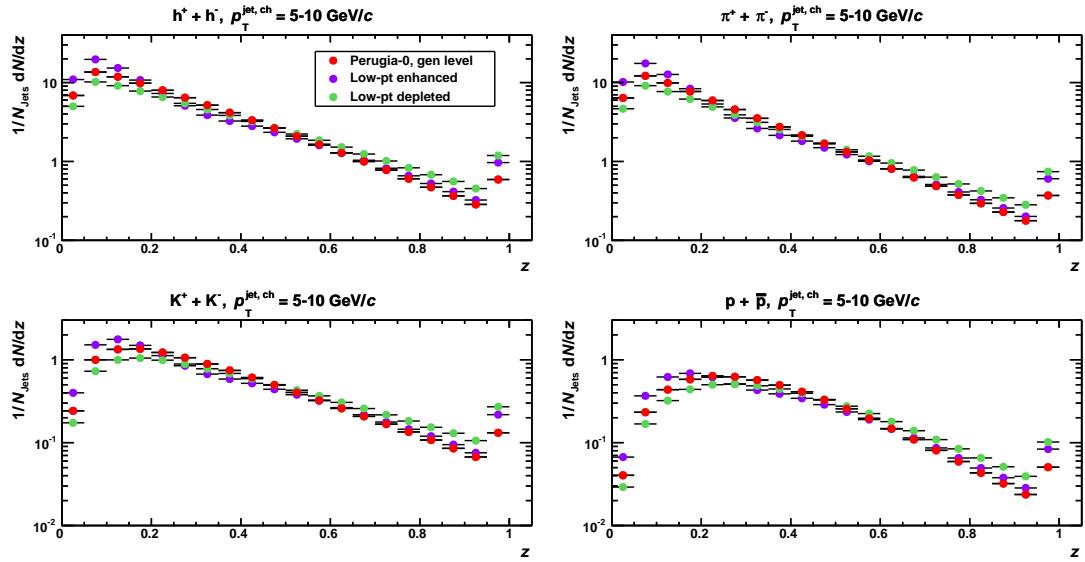


Figure 5.16.: Modified fragmentation distributions (green and violet) are compared to the case without modifications (red). The comparison is shown for all charged particles, π , K and p in jets with $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$.

5.3.3. Secondary Particle Contamination

The fully corrected results of the analysis must not depend on the chosen track cuts. Varying the track cuts (and consistently the MC correction factors) and investigating the spread of the fully corrected results is therefore a means to estimate the systematic uncertainties of the MC correction factors. A cut on the Distance of Closest Approach (DCA) provides a handle on the amount of secondaries that pass the track selection. The comparison of loose and tight DCA cuts (details about these cuts can be found in Appendix C.5) is used in this section to estimate the systematic uncertainty of the secondary correction. The change of the DCA cut primarily alters the secondary contamination and minimises the impact on other quantities like the efficiency. This is important in order to separate the systematic uncertainties of individual error sources. It turns out that the procedure is not as straightforward as it may appear at first glance. The considered cuts are different from what is used for the analysis in the end. Therefore, the systematic uncertainty of the secondary correction needs to be estimated relative to the corresponding correction factor, which can then be evaluated for any cut. It has to be noted that the cut for which the uncertainty is obtained should be as close as possible to what is used in the analysis in order to minimise possible correlations of other sources with the secondary contamination.

5. Monte Carlo Corrections

Mathematical Considerations

Some mathematical considerations are necessary to figure out from which observables the systematic uncertainties can be estimated and how they are related to the cut variations. Starting point is Eq. 5.2, which is written in the following again for convenience. For better legibility, the superscript “rec”, the additional subscript “*s*” and the dependence on *obs* will be dropped:

$$f_{\text{prim}} \equiv \frac{N_{\text{prim}}}{N_{\text{prim}} + N_{\text{sec}}} = \frac{N_{\text{prim}}}{N}, \quad (5.6)$$

$$N \equiv N_{\text{prim}} + N_{\text{sec}}. \quad (5.7)$$

From these relations, it follows:

$$1 - f_{\text{prim}} = \frac{N_{\text{sec}}}{N_{\text{prim}} + N_{\text{sec}}} \iff N_{\text{prim}} + N_{\text{sec}} = \frac{N_{\text{sec}}}{1 - f_{\text{prim}}} \quad (5.8)$$

and

$$N_{\text{prim}} = f_{\text{prim}} \cdot (N_{\text{prim}} + N_{\text{sec}}) = N_{\text{sec}} \cdot \frac{f_{\text{prim}}}{1 - f_{\text{prim}}}, \quad (5.9)$$

where Eq. 5.8 has been used in the second step.

The systematic uncertainty due to the tracking efficiency has already been estimated in Section 5.3.1 and must not enter here again. Therefore, it is assumed in the following that the efficiency correction has already been applied and is perfect, i.e. there is no uncertainty from that source and the error of N_{prim} vanishes by definition. From Eq. 5.6, the systematic uncertainty of the primary fraction f_{prim} then reads:

$$\begin{aligned} \Delta f_{\text{prim}} &= \Delta N_{\text{sec}} \cdot \frac{N_{\text{prim}}}{(N_{\text{prim}} + N_{\text{sec}})^2} = \Delta N_{\text{sec}} \cdot \frac{f_{\text{prim}}}{N_{\text{prim}} + N_{\text{sec}}} \\ \Rightarrow \frac{\Delta f_{\text{prim}}}{f_{\text{prim}}} &= \frac{\Delta N_{\text{sec}}}{N_{\text{prim}} + N_{\text{sec}}} = \frac{\Delta N_{\text{sec}}}{N_{\text{sec}}} \cdot (1 - f_{\text{prim}}), \end{aligned} \quad (5.10)$$

where Eq. 5.8 was used in the last step. The relative uncertainty of the primary fraction is proportional to the relative uncertainty of the secondary yield and the actual correction $(1 - f_{\text{prim}})$ — bearing in mind that $f_{\text{prim}} = 1$ corresponds to no correction.

The quantities in the above equations depend on the considered track cuts. From now on, the cuts will be denoted by a superscript. The loose and tight DCA cuts will be indicated by *l* and *t*, respectively, whereas *x* will be a wild-card for any cut. The systematic uncertainty of the secondary correction is estimated from the ratio between the fully corrected yields from the loose and tight DCA

5.3. Systematic Uncertainty of the MC Correction Factors

cuts. The tilde indicates that the corresponding quantity is fully corrected, such that the fully corrected (primary) yield reads:

$$\widetilde{N}^x \equiv \widetilde{N}_{\text{prim}}^x + \widetilde{N}_{\text{sec}}^x = N_{\text{prim}} + \widetilde{N}_{\text{sec}}^x. \quad (5.11)$$

The second equality holds true under the assumption of a perfect efficiency correction. $\widetilde{N}_{\text{sec}}^x$ is the *residual* secondary yield after all corrections, which can in general also be negative in case of overcorrection. If both secondary and efficiency correction were perfect and had zero systematic uncertainty, $\widetilde{N}_{\text{sec}}^x$ would vanish and $\widetilde{N}^x = N_{\text{prim}}$ would not depend on the cut. Hence, assuming that the systematic uncertainty of the secondary correction dominates, $\widetilde{N}_{\text{sec}}^x$ can be used as a measure of this uncertainty.

The deviation of the loose over tight ratio (l/t ratio) from unity can be written as:

$$1 - \frac{\widetilde{N}^l}{\widetilde{N}^t} = \frac{\widetilde{N}_{\text{sec}}^t - \widetilde{N}_{\text{sec}}^l}{\widetilde{N}^t} \equiv \frac{\Delta \widetilde{N}_{\text{sec}}}{\widetilde{N}^t} \approx \frac{\Delta \widetilde{N}_{\text{sec}}}{N_{\text{prim}}}, \quad (5.12)$$

using Eq. 5.11 and $\widetilde{N}^t \approx N_{\text{prim}}$ for $\widetilde{N}_{\text{sec}}^t \ll N_{\text{prim}}$. This approximation is valid, if the systematic uncertainty of the secondary correction is rather small. The results of this section will justify this approximation. Note that the primary yields in the numerator drop out due to the assumption of a perfect efficiency correction in Eq. 5.11.

These quantities need to be related to the systematic uncertainty of f_{prim}^x in Eq. 5.10. In particular, the uncertainty of f_{prim}^x depends on the chosen track cuts and needs to be converted correspondingly. The Eq. 5.12 may be re-written as:

$$1 - \frac{\widetilde{N}^l}{\widetilde{N}^t} \approx \frac{\Delta \widetilde{N}_{\text{sec}}}{N_{\text{prim}}} = \frac{\Delta \widetilde{N}_{\text{sec}}}{N_{\text{sec}}^x} \cdot \frac{1 - f_{\text{prim}}^x}{f_{\text{prim}}^x} = \frac{\Delta N_{\text{sec}}^x}{N_{\text{sec}}^x} \cdot (1 - f_{\text{prim}}^x), \quad (5.13)$$

where Eq. 5.9 has been used. Since $\Delta \widetilde{N}_{\text{sec}}$ is related to the corrected (residual) secondary yield³, the error of the secondary yield for cut x becomes $\Delta N_{\text{sec}}^x = \Delta \widetilde{N}_{\text{sec}} / f_{\text{prim}}^x$. This was used in the last step. The result can be re-written as a double ratio:

$$\left(1 - \frac{\widetilde{N}^l}{\widetilde{N}^t}\right) / (1 - f_{\text{prim}}^x) \approx \frac{\Delta N_{\text{sec}}^x}{N_{\text{sec}}^x}, \quad (5.14)$$

which directly corresponds to the relative uncertainty of the secondary yield. With Eq. 5.10, this finally yields the relative systematic uncertainty of the secondary correction.

Estimation Procedure

For technical reasons, the following procedure is carried out using the minimum bias MC production. The jet finding with resolution parameter R is run on

³By assumption, a perfect efficiency correction is applied to all quantities. Thus, the tilde indicates that only the secondary correction is applied in addition.

5. Monte Carlo Corrections

the default hybrid tracks sample. The tracks for the following procedure are then selected from ideal cones with radius R around the reconstructed jet axis (a further discussion follows in Section 6.6) and demanded to satisfy the corresponding cut criteria. Two sets of global tracks are considered, one with a tight and another with a loose DCA cut. A more detailed discussion of these cuts is postponed to Section 6.3.10. What is important here is that this pair of cuts maximises the difference of the secondary particles, while other differences are minimal.

Individually for both cuts, the MC correction factors are obtained and used to correct the raw particle yields extracted with the MTF. As discussed before, the systematic uncertainty of the secondary correction is related to the loose/tight DCA ratio of the fully corrected spectra. The latter will be denoted as l/t ratio in the following and is shown for pions and kaons in jets with $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$ in Fig. 5.17. Note that some data points for kaons at very low p_T have been removed, since the efficiency is too low. The ratio is close to unity, which implies that the correction procedure is pretty accurate. However, both pions and kaons show the largest deviation from unity at high p_T — a similar deviation is observed for protons (not shown). Since the secondary contamination decreases with p_T and the effect is observed for all species, this suggests that the efficiency correction introduces a small bias.

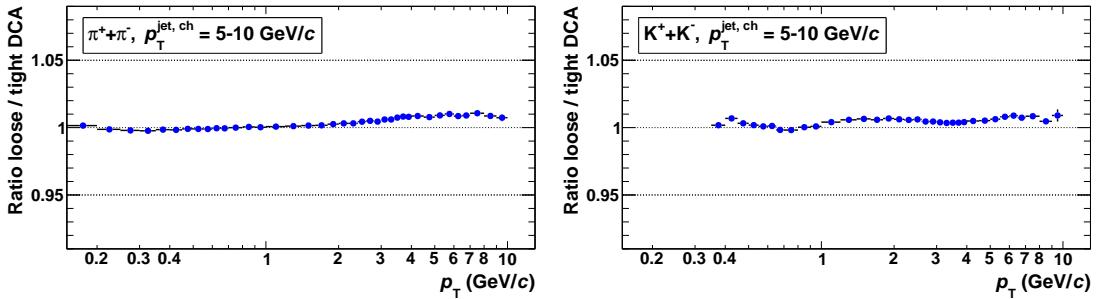


Figure 5.17.: Ratio loose/tight DCA for pions (left) and kaons (right) in jets with $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$.

The efficiency bias needs to be corrected to avoid an overestimation of the systematic uncertainty. Kaons have in good approximation zero secondary contamination. It is assumed that the deviations of their l/t ratio from unity are only due to systematic effects of the efficiency correction. To leading order, these effects should be species independent. It is assumed further that the change of the DCA cut does not introduce a PID bias in regions without dE/dx crossings. An approximate solution of the above issue is then to divide the l/t ratios of any species by that of the kaons. This procedure will be called *kaon correction* in the following. The resulting l/t ratios for pions and protons after the kaon correction are shown in Fig. 5.18. Obviously, the ratio for pions is closer to unity at high p_T than in Fig. 5.17, which is the desired outcome. Note, however, that

5.3. Systematic Uncertainty of the MC Correction Factors

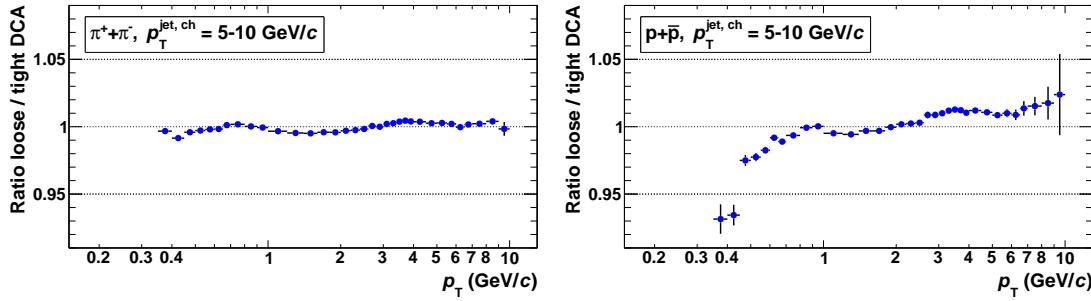


Figure 5.18.: Ratio loose/tight DCA for pions (left) and protons (right) in jets with $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$ after applying the kaon correction.

the correction seems to introduce a wiggle around $p_T \approx 2 \text{ GeV}/c$. This is close to the dE/dx crossing of kaons and protons, where systematics of the MTF significantly increase. Although the influence of the DCA cut on the PID should be very small, this may still lead to some bias in the dE/dx crossing regions. For that reason, the crossing regions are excluded in the further procedure. Similarly, the high- p_T region ($p_T > 2 - 3 \text{ GeV}/c$, say) is excluded. The argument is that even after the kaon correction there may be small biases (either from PID or still from efficiency), which can be problematic in view of the very small correction factors in that region.

A glance back to Eq. 5.10 reveals that the relative uncertainty of the secondary contamination, $\Delta N_{\text{sec}}^x/N_{\text{sec}}^x$, needs to be determined. According to Eq. 5.14, the next step is, thus, to evaluate the double ratio $(1 - \frac{\tilde{N}^l}{\tilde{N}^t})/(1 - f_{\text{prim}}^x)$. Pictorially, the numerator of the double ratio for protons is one minus the right panel of Fig. 5.18 and the denominator is basically one minus the blue histogram in Fig. 5.3b⁴. The resulting double ratio for $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$ is shown in Fig. 5.19 and discussed in the following.

The double ratio exhibits considerable statistical fluctuations. These become much stronger for higher p_T^{jet} . It is therefore desirable to fit a parametrisation to the data points. Naively, a constant double ratio could be expected, since it amounts to the relative uncertainty of the secondary contamination, which should be rather independent of p_T . However, residual efficiency effects, impacts from p_T^{jet} bin migration and PID biases may still enter. It turns out that a linear fit as a function of $\ln(p_T)$ (black line in Fig. 5.19) describes the data reasonably well, if the signed double ratio and not its absolute value is considered. As mentioned before, the dE/dx crossing regions may cause PID biases and are excluded from the fit (red points). The dashed line is just an extrapolation to get a feeling how the double ratio could behave towards even lower p_T . For the inclusive case, an additional degree of freedom is added to the fit via allowing for a constant value

⁴Note that the figure is for a different p_T^{jet} . This is just to get a feeling about the values. The actual procedure takes the correct p_T^{jet} bins, of course.

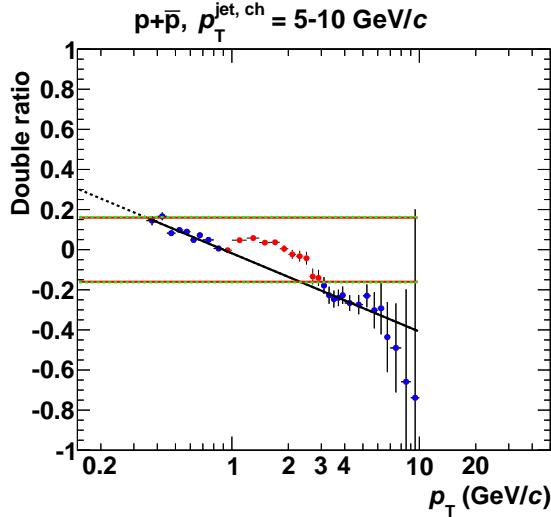


Figure 5.19.: Double ratio for protons in jets with $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$. The red points have been excluded from the fit. See discussion in text for further details.

above a floating threshold. The difference to the jet case is most likely due to effects from p_T^{jet} bin migration, which turn the constant efficiency bias at high p_T in the inclusive case into a p_T dependent one for jets. The corresponding plots for the inclusive case and the discussion can be found in Appendix E.1.

With so many other biases entering, it is rather difficult to estimate the relative uncertainty of the secondary contamination. Obviously, the fit describes the blue points in the full p_T range reasonably well (see Fig. 5.19). Since other effects dominate the high- p_T part (above $p_T = 2 - 3 \text{ GeV}/c$), this region only enters the fit, but will not enter the error estimate. The latter is restricted to the blue points in the $p_T < 1 \text{ GeV}/c$ regime, where secondary contamination is largest. A p_T independent error estimate is obtained via taking the extreme (absolute) value from the fit in that region — starting from the first blue point; the dashed line does not enter. The red horizontal lines indicate the region with plus-minus the error estimate around zero. The green dashed lines similarly show the average systematic error estimated from all p_T^{jet} bins (the inclusive case is treated separately). The average uses as weight the uncertainties of the utilised extreme point of the fit in each p_T^{jet} bin, which is evaluated via the errors of the fit parameters.

The double ratios for all p_T^{jet} bins for protons and pions are shown in Figs. 5.20 and 5.21, respectively. Comparing the weighted average (green dashed lines) with the error estimate of each bin (red lines), it can be seen that the p_T^{jet} dependence for protons is small. In particular, the weighted average reasonably well describes the spread at low p_T .

Pions have only small secondary contamination at the order of 1%. Therefore, large fluctuations are observed and only for the smallest p_T^{jet} the fit is meaningful in the sense that the uncertainty of the red lines (which is not shown) is not too

5.3. Systematic Uncertainty of the MC Correction Factors

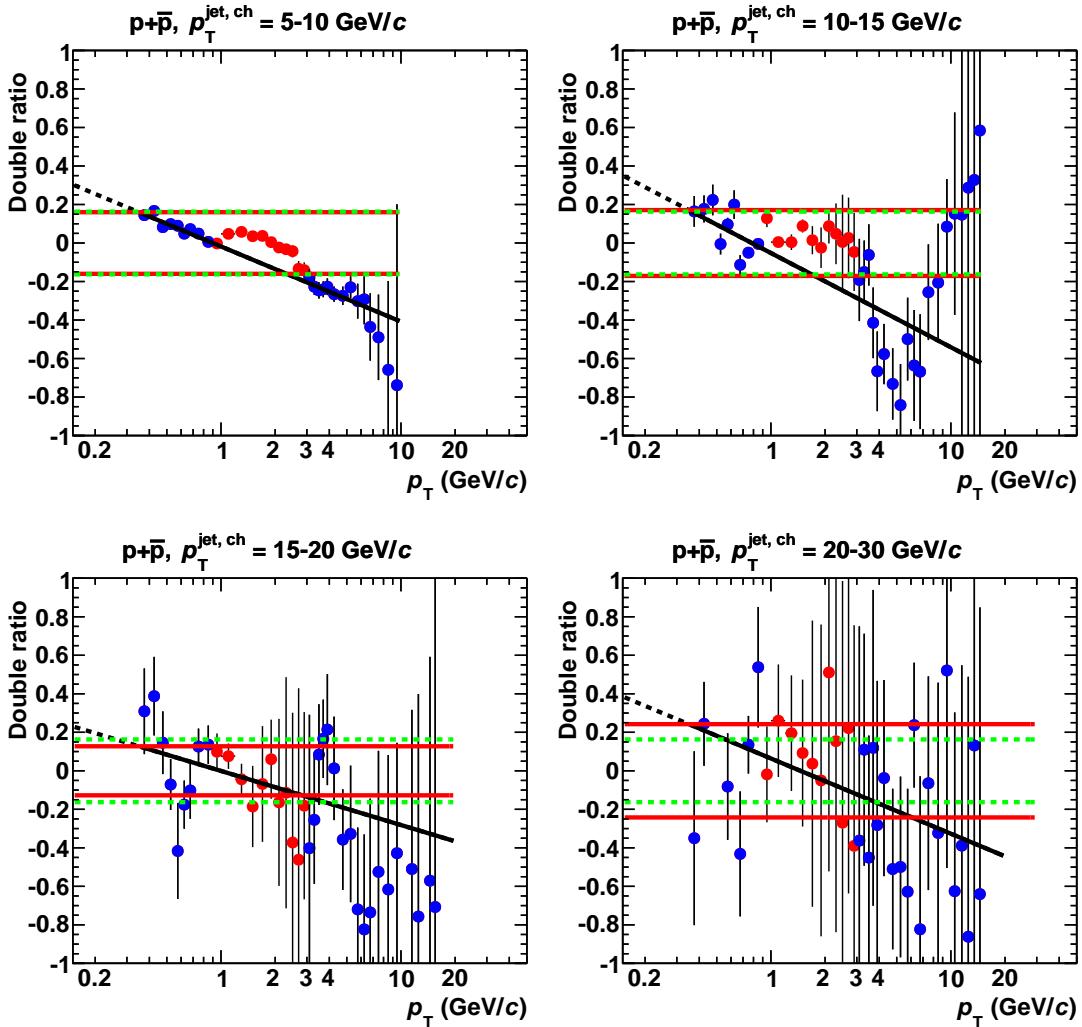


Figure 5.20.: Double ratio for protons in jets. The red points have been excluded from the fit. See discussion in the text for further details.

large. The average uncertainty (green dashed lines) is therefore driven by the low- p_T^{jet} bins, but also reasonably well describes the spread for higher p_T^{jet} within the huge error bars.

Note that the kaons cannot be fitted. First of all, they have already been used to divide out efficiency biases, i.e. their l/t ratio is unity by construction. Secondly, their secondary contamination is basically zero. Due to this, the actual choice of systematic error estimate is rather irrelevant for them. Formally, the error from the pion fits is assigned to kaons as well.

5. Monte Carlo Corrections

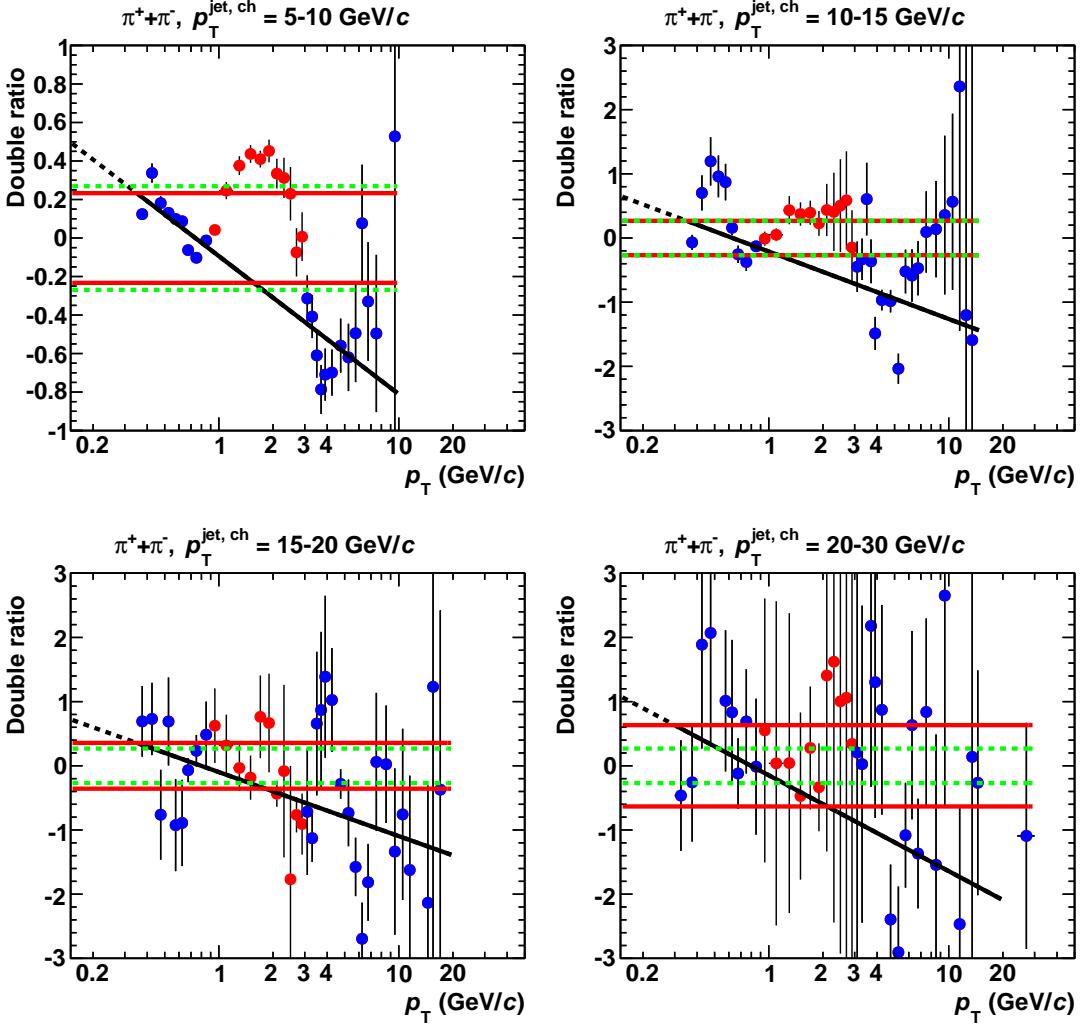


Figure 5.21.: Double ratio for pions in jets. The red points have been excluded from the fit. See discussion in the text for further details. Note the different ranges for the vertical axis.

Results

Finally, having obtained $\Delta N_{\text{sec}}^x / N_{\text{sec}}^x$, the relative systematic uncertainty of the secondary correction can be calculated via Eq. 5.10. A summary plot for all the “ingredients” is presented in Fig. 5.22 for protons and in Fig. 5.23 for pions as a function of p_T . The corresponding plots as a function of z and for the inclusive analysis can be found in Appendix E.1. The full blue points are the l/t ratio and the full red points indicate the secondary correction factor, which is just the primary fraction. The open red points are the final estimated error and should be compared to the open blue points, which correspond to the observed error. Obviously, the former reasonably well describe the low- p_T region and give

5.3. Systematic Uncertainty of the MC Correction Factors

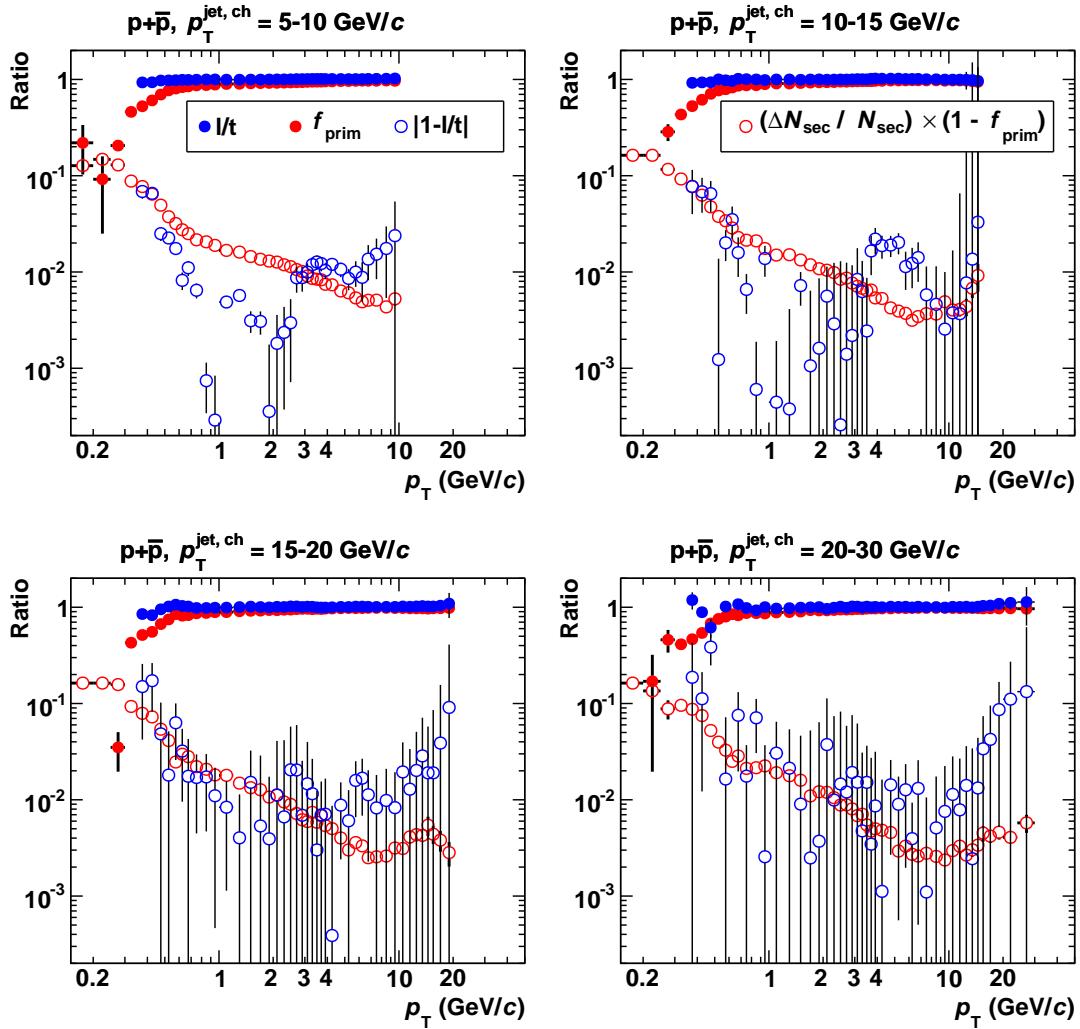


Figure 5.22.: Summary plot for the systematic uncertainty estimation of the secondary correction for protons in jets as a function of p_T . The open red points show the relative systematic uncertainty estimate and should be compared to the open blue points, which show the relative change of spectra due to DCA cut variations.

a conservative estimate up to about $2 \text{ GeV}/c$ for the protons. The error bars for pions are large for high p_T^{jet} , but the error estimate successfully describes the trend within uncertainties. As can be seen, the pion primary fraction is close to unity and the actual value of $\Delta N_{\text{sec}}^x / N_{\text{sec}}^x$ does not matter too much. The uncertainty of the secondary correction is about 1% and below, which is anyway completely dominated by other error sources.

Note that the dip for the protons around $1 - 2 \text{ GeV}/c$ is due to the fact that the l/t ratio crosses unity. This does not necessarily imply that the MC description is perfect in this region. On the one hand, it could mean that there is a p_T dependent systematic bias of the correction, which switches an under- to an overcorrection at

5. Monte Carlo Corrections

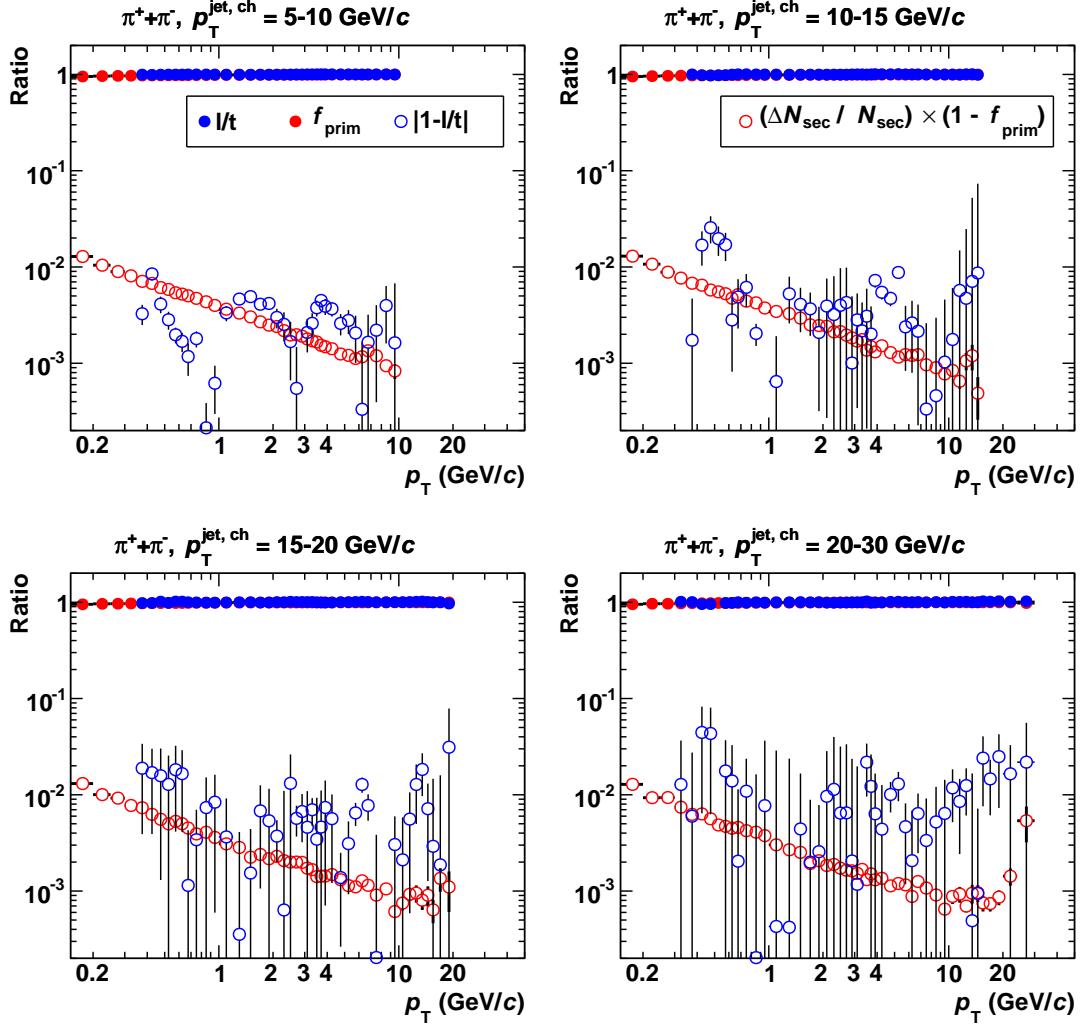


Figure 5.23.: Summary plot for the systematic uncertainty estimation of the secondary correction for pions in jets as a function of p_T . The open red points show the relative systematic uncertainty estimate and should be compared to the open blue points, which show the relative change of spectra due to DCA cut variations.

this p_T . On the other hand, it may be that residual efficiency biases start playing a dominant role. In any case, the assigned systematic uncertainty conservatively covers this region.

In conclusion, it appears justified to assume a constant relative systematic uncertainty of the secondary contamination. The p_T dependence of the secondary correction is then propagated to its uncertainty and correctly describes the observed trends (leaving apart the proton dip). In principle, the same procedure could be carried out as a function of z (or ξ). However, the z fits are problematic, since the dE/dx crossing regions contribute to many bins, such that basically only 1 – 2 data points are left for the error estimate (and the fit) at low z (cor-

5.3. Systematic Uncertainty of the MC Correction Factors

responding to low p_T). This makes the procedure highly sensitive to statistical fluctuation in these bins. Hence, the fit results from p_T are adopted for z (and ξ). This is valid, since the uncertainties are anyway relative to the correction factor. The resulting summary plots can be found in Appendix E.1. Here, it should only be mentioned that the error estimate adequately describes the spread due to the cut variations.

It should be emphasized at this point that the procedure does not yield reasonable results, if the kaon correction is not applied. For example, the l/t ratio for pions increases in that case at high p_T , in contrast to the decreasing secondary contamination. Eventually, this introduces strange trends in the double ratio making a reliable uncertainty estimation impossible. A more detailed discussion is presented in Appendix E.1.

Finally, to boil the findings of these section down to a few numbers, the relative systematic uncertainties of the secondary contamination for the given set of track cuts are summarised in Tab. 5.1 and hold true for all observables (p_T , z , ξ). From these values, the relative uncertainty of the secondary correction factor is obtained via Eq. 5.10 by multiplying with one minus the correction factor (a correction factor of unity corresponds to no correction). This means that if e.g. the relative uncertainty from the table is 10% and $f_{\text{prim},s}(p_T^{\text{track}})$ is 95% for some species s at some p_T^{track} , then the correction is 5% and the resulting relative systematic uncertainty of $f_{\text{prim},s}(p_T^{\text{track}})$ amounts to $10\% \cdot 5\% = 0.5\%$.

Table 5.1.: Relative systematic uncertainties of the secondary contamination. The values for the modified hybrid track cuts (mod. hybrid) and the default cuts (default) are presented.

Species	Relative systematic uncertainty (%)		
	Inclusive (default)	Inclusive (mod. hybrid)	Jets (mod. hybrid)
p	15.7	10.1	16.3
π, K	38.1	28.2	26.9

5.3.4. GEANT-FLUKA Correction

The GEANT-FLUKA correction relies on the parametrisation in Eq. 5.5, which might slightly change for different cuts. Also, it is only an effective correction and does not remedy the origin of the problem, namely the wrong reaction cross-section in the detector material for anti-protons and negative kaons. To be conservative, the difference of the correction factors with and without this correction is used as systematic uncertainty.

Note that the MC production for the jet case was carried out with a newer GEANT version that already uses the corrected cross-section. Hence, this systematic uncertainty estimate only contributes in the inclusive case.

5.3.5. Strangeness Scaling

The strangeness scaling is based on MB (CMS) measurements and it is not clear whether it can be applied to jet spectra also. In addition, the CMS results have been obtained for a different rapidity range. As a consequence, 100% of the difference of the correction factors with and without strangeness scaling is taken as systematic uncertainty.

5.3.6. Muon Contamination

Finally, the correction of the muon contamination strictly relies on the MC model. The correction is small (less than 2%) and the systematic uncertainty is set to 100% of the difference of the correction factors with and without taking into account the muon contamination. This should cover the model dependencies of the correction factors and still contributes only a small fraction to the total systematic uncertainty.

5.3.7. Multiplicity Dependence

The 5% uncertainty assigned to the efficiency correction completely covers all deviations among multiplicity bins for the reference multiplicity and no multiplicity dependence is observed for the V0M estimator. Therefore, no additional systematic uncertainty is assigned with respect to the multiplicity dependence of the efficiency.

For the secondary correction, the observations for both multiplicity estimators are similar. Only the correction factors for protons exhibit a multiplicity dependence for $p_T \lesssim 0.6 \text{ GeV}/c$, with the spread rising gradually to roughly $\pm 4 - 8\%$ at $p_T = 0.35 \text{ GeV}/c$ with decreasing p_T . This spread is covered by the conservatively estimated error bars of the secondary correction (and the strangeness scaling) and it should be noted that bins with intermediate multiplicity have approximately equal primary fractions, i.e. a smaller spread. Therefore, no additional systematic uncertainty is assigned with respect to the multiplicity dependence of the secondary correction.

5.3.8. Error Propagation to the To-Pion Ratios

Efficiency, Resolution and Shape

The systematic uncertainties of efficiency, resolution and shape are expected to partially cancel for the to-pion ratios. The correction factor for the to-pion ratio of species s is (cf. Eq. 5.1):

$$corr_{s/\pi,\varepsilon}(obs) \equiv \frac{corr_{s,\varepsilon}(obs)}{corr_{\pi,\varepsilon}(obs)}. \quad (5.15)$$

5.3. Systematic Uncertainty of the MC Correction Factors

The correlation between the systematic uncertainties $\delta corr_{s,\varepsilon}(obs)$ and $\delta corr_{\pi,\varepsilon}(obs)$ is assumed to be 100% positive. The systematic uncertainty of the to-pion ratio, $\delta corr_{s/\pi,\varepsilon}(obs)$, is then estimated from the spread of the variations:

$$corr_{s/\pi,\varepsilon}^{\text{variation}}(obs) \equiv corr_{s/\pi,\varepsilon}(obs) \frac{1 + \frac{\delta corr_{s,\varepsilon}(obs)}{corr_{s,\varepsilon}(obs)}}{1 + \frac{\delta corr_{\pi,\varepsilon}(obs)}{corr_{\pi,\varepsilon}(obs)}}, \quad (5.16)$$

where the uncertainties in the numerator and the denominator are varied in the same directions. In the above equations, an implicit dependence on p_T^{jet} is understood for all variables.

Secondaries

The systematic uncertainty of the secondary correction for each species is propagated to that of the to-pion ratios by assuming that the correction factors of the two considered species are completely uncorrelated. Since the sources of the secondary particles can be different (pions from decays of neutral kaons and hadronic interactions, protons from material and Λ decays), the systematic uncertainties are at least partially uncorrelated. Treating them as fully uncorrelated gives a more conservative estimate.

Other Sources

All remaining systematic uncertainty sources are treated as fully correlated. This means the correction factors for the to-pion ratios are calculated from those of the corresponding yields for each variation (e.g. with and without applying the GEANT-FLUKA correction). The systematic uncertainty is then estimated from these variations completely analogous to the estimate for the yield correction factors.

5.3.9. Total Uncertainties

The discussion about the systematic uncertainty estimation of the MC correction factors is summarised in Tab. 5.2. It should be emphasized that the estimates for the inclusive and the jet case are different.

In the following figures, the individual contributions of each error source are shown in a stacked representation, i.e. each histogram adds (linearly) one error source to the error of all sources considered to that point. The total error is the square root of the quadratic sum of all error sources and is shown as the black dotted line.

The errors for the inclusive case using the modified hybrid cuts are shown in Fig. 5.24. Likewise, Fig. 5.25 illustrates the errors for $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ for the yields and to-pion ratios, respectively. The corresponding figures for all other p_T^{jet} bins are collected in Appendix E.2.

5. Monte Carlo Corrections

Table 5.2.: Overview of the systematic uncertainty estimation of the MC correction factors.

Error source	Systematic uncertainty estimate	
	Inclusive	Jets
Tracking efficiency	5%	Vary tracking efficiency by $\pm 5\%$
Momentum resolution	2%	Vary momentum resolution by $\pm 20\%$
Jet shape dependence of MC generator	–	Compare to PYTHIA with high- z enhancement/depletion
GEANT-FLUKA correction	100% of corr. factor	–
Strangeness scaling	100% of difference of corr. factors with and without strangeness scaling	
Secondary particle contamination	Run full analysis with corrections for loose and tight DCA cuts	
Muon contamination	100% of corr. factor	

As can be seen, the errors from efficiency, resolution and shape largely cancel for the to-pion ratios as expected⁵. Note that for protons in the very first p_T^{track} bins, the contribution from individual error sources is zero or even the total systematic uncertainty is zero. This happens if the corresponding correction factors are zero and the relative error is not defined. It is caused by the almost zero efficiency in these bins, such that they have to be excluded from the analysis.

The systematic errors depend only weakly on the p_T^{jet} as a function of p_T^{track} (cf. Appendix E.2). However, since the strangeness scaling is solely a function of p_T^{track} , its error in bins of z does depend on the p_T^{jet} . Typically, the systematic error of pions and kaons is around 5%, but increases above 10% as p_T^{track} approaches p_T^{jet} and jet bin migration effects become important. For protons, the error is 10 – 20% at low p_T^{track} (z) due to the large uncertainty from the strangeness scaling. Depending on p_T^{jet} , it reaches a minimum around 3 – 9 GeV/c, where the strangeness scaling is of minor importance and bin migration effects are still marginal, and increases for higher p_T^{track} as for the other species.

⁵The cancellation is by construction in the inclusive analysis, where the same fixed errors are assigned to each species. The observation in jets, in particular for $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$, where statistical fluctuations are small, corroborates this. Of course, the regions with bin migration, i.e. p_T^{track} close to p_T^{jet} , should not be considered for this purpose.

5.3. Systematic Uncertainty of the MC Correction Factors

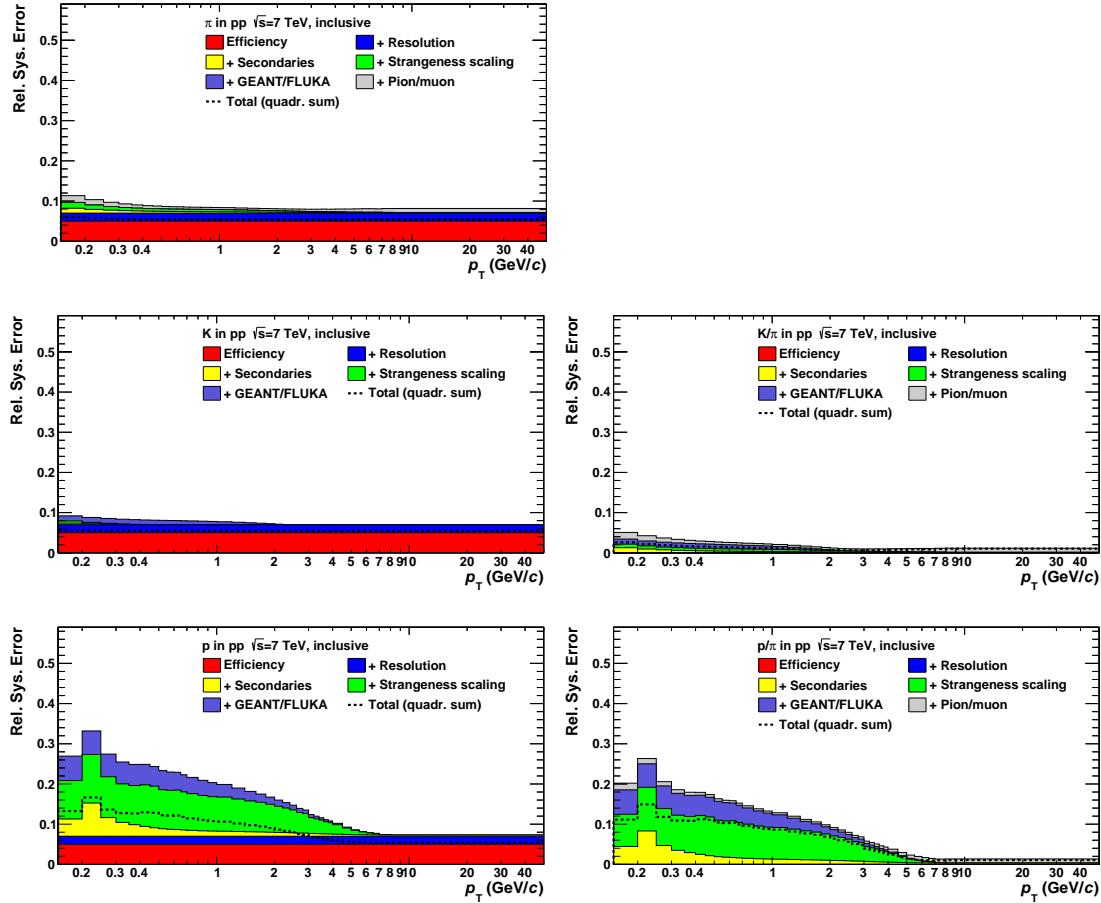


Figure 5.24.: Relative systematic uncertainties of the MC correction factors for the inclusive yields (left column) and to-pion ratios (right column) as a function of p_T^{track} . The rows show (from top to bottom) π , K and p , respectively.

Due to the cancellation, the systematic uncertainties of the kaon-to-pion ratios are typically below 5%. For the proton-to-pion ratio, they are similar to the proton yields because the uncertainty of the strangeness scaling does not cancel. Only above $3 - 5 \text{ GeV}/c$, they also drop below 5%.

Since the multiplicity dependent analysis uses a different set of cuts, the relative systematic uncertainties for the yields are slightly different. They are shown in Fig. 5.26 for the MB correction factors, that are used for the V0M analysis. In order not to just repeat what was already presented for the inclusive case of the jet analysis, the figure shows the uncertainties separately for positively and negatively charged particles. Especially, the errors for protons are significantly different from those of anti-protons. The latter suffer from uncertainties of the GEANT-FLUKA correction, whereas those originating from secondary contamination are smaller than for protons at low p_T . The relative systematic uncertainties of the to-pion ratios are shown in Fig. 5.26 as well. As for the jet analysis, they are completely dominated by the uncertainties of the kaons and protons.

5. Monte Carlo Corrections

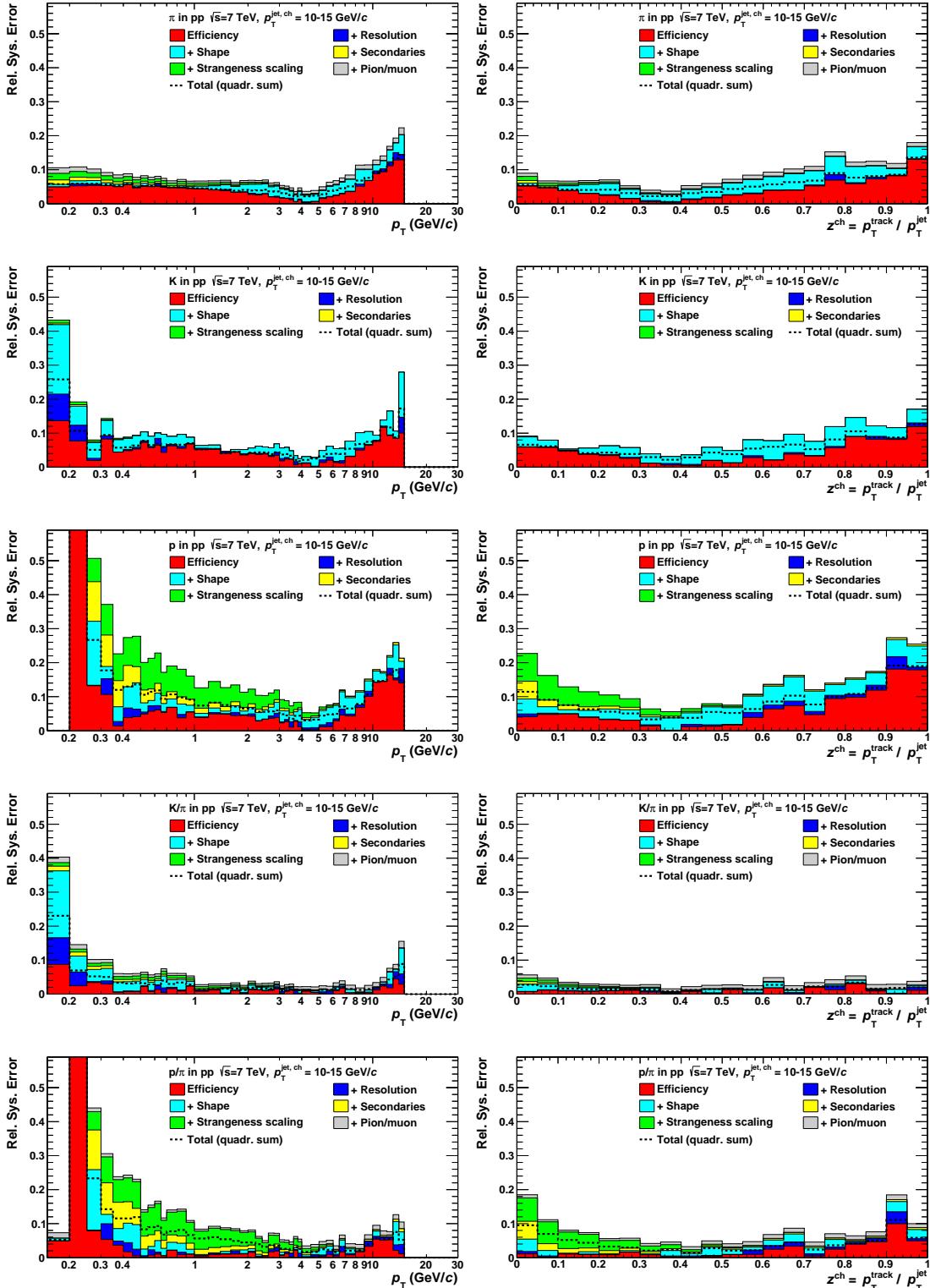


Figure 5.25.: Relative systematic uncertainties of the MC correction factors for $p_T^{jet} = 10 - 15 \text{ GeV}/c$. The rows show (from top to bottom) pions, kaons, protons, K/π and p/π , respectively, as a function of p_T^{track} (left column) and z (right column).

5.3. Systematic Uncertainty of the MC Correction Factors

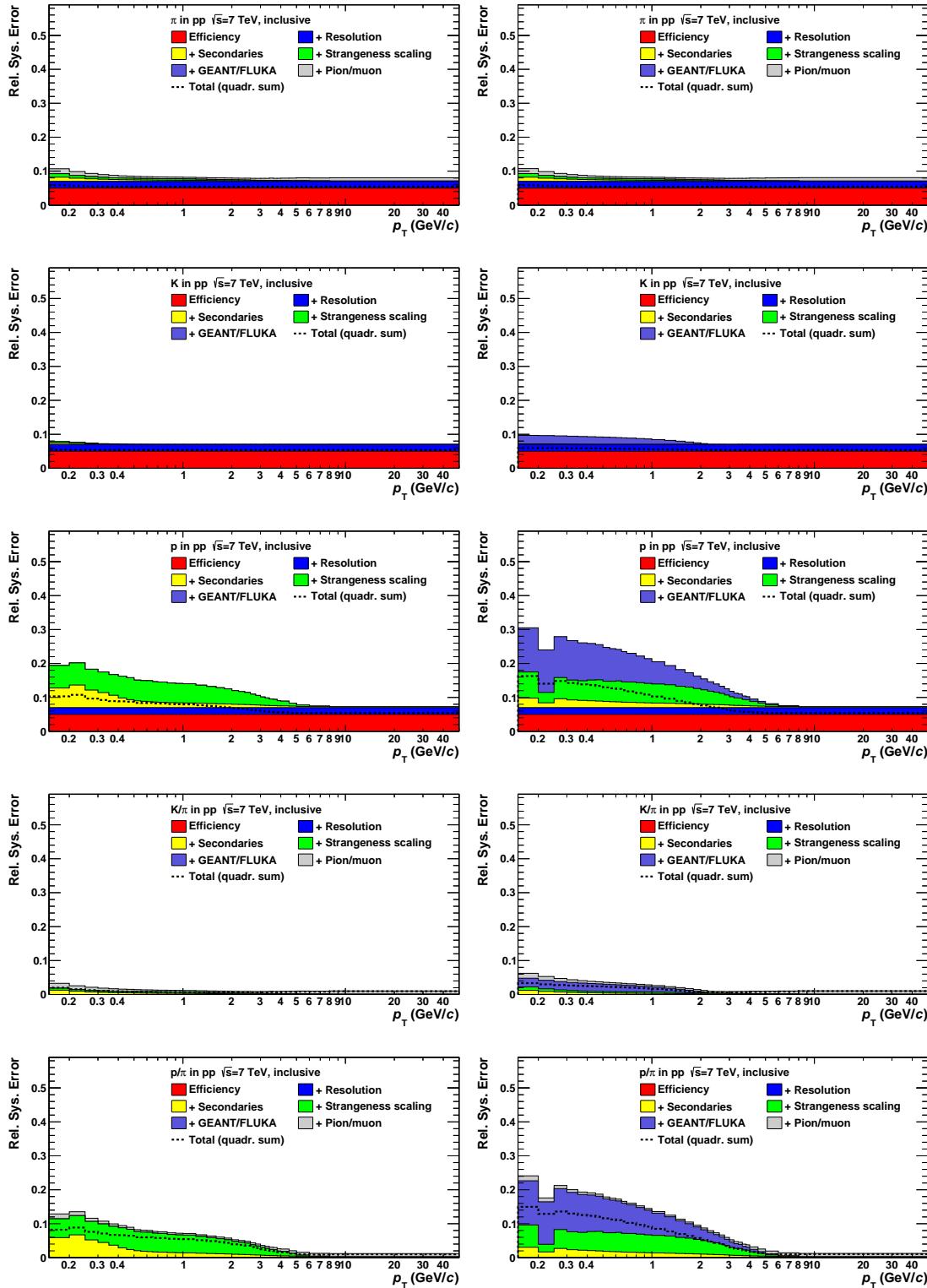


Figure 5.26.: Relative systematic uncertainties of the MB MC correction factors in the multiplicity analysis. The left column is for positively, the right column for negatively charged particles. The rows show (from top to bottom): pions, kaons, protons, K/π and p/π .

5. Monte Carlo Corrections

The multiplicity bin used to extract the corrections for the reference multiplicity yields very similar correction factors. The difference to the relative systematic uncertainties of the MB correction is immaterial and the corresponding plots are not shown for that reason.

5.4. MC Closure Check

The correction procedure is validated via an MC closure check using the minimum bias MC production (see Appendix B.1 for the run list). To distinguish possible biases originating from the fitting and the correction procedure, the corrections are applied to the MC truth at the detector level, i.e. the “identification” is simply the MC ID. To avoid correlations, the MC sample is split into two disjunct samples A and B. The correction factors are extracted from sample A, whereas the raw yields at detector level and also the MC true yields at particle level are taken from sample B.

All corrections mentioned in this chapter are then applied to the raw yields, except for the GEANT-FLUKA correction and the strangeness scaling. Those two must not be applied in order to have consistency between particle and detector level. Since the muon correction is used, the raw yield of MC identified muons is added to the pions, which is consistent with the treatment of muons for the fits in data.

The comparison of corrected detector level to particle level is shown in Fig. 5.27. The ratios are consistent with unity for all species in case of inclusive and also for jets with the yields being a function of p_T^{track} or z . Note that some data points at low p_T^{track} may be missing for the detector level. This happens in case of no statistics for the corrections or the raw yield.

The figure proves that in case of perfect particle identification at the detector level, the particle level can be recovered within statistical uncertainties. The systematic error of the correction procedure is then due to uncertainties of how adequate the simulation describes the real world.

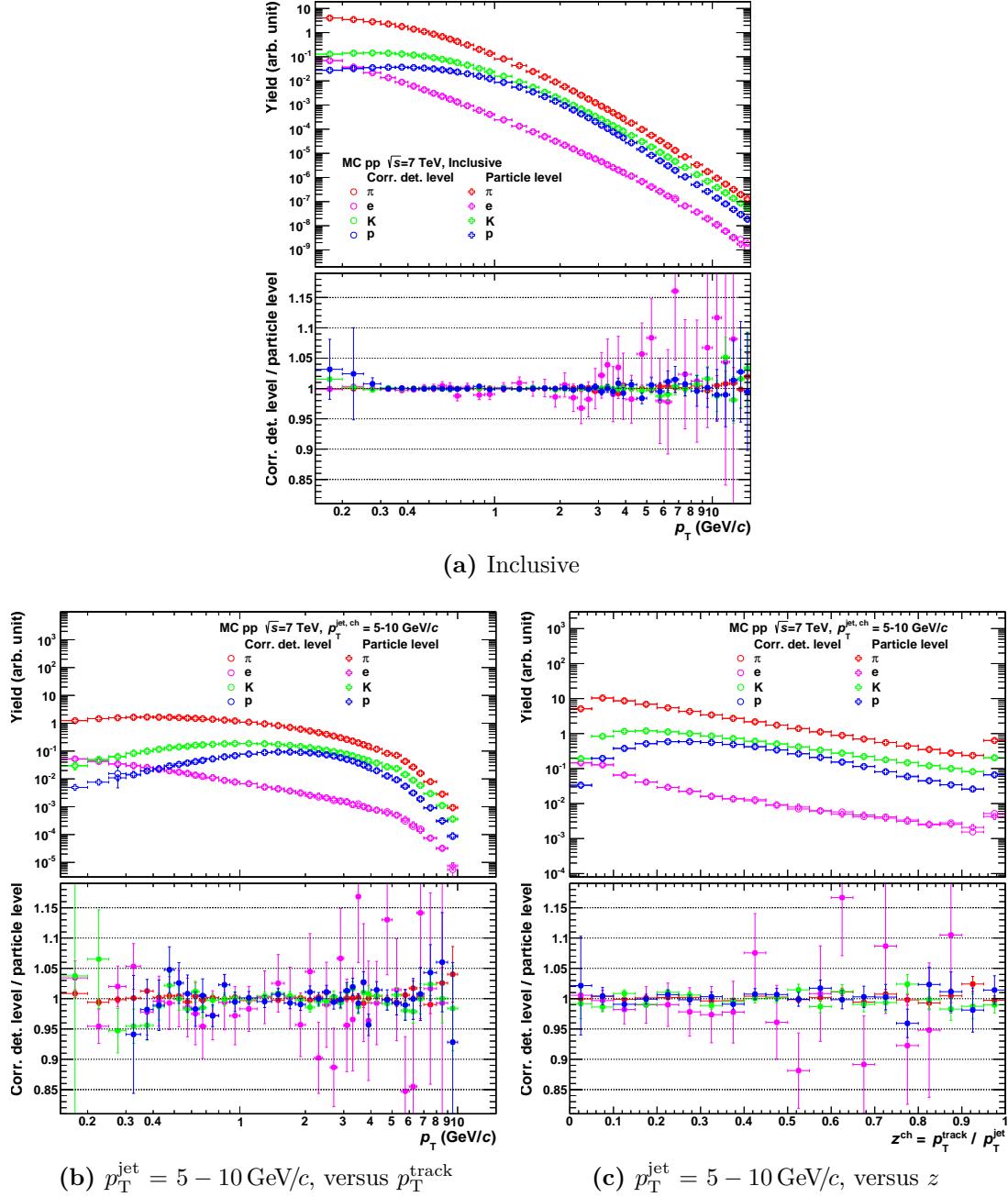


Figure 5.27.: MC closure check of the corrections for **a)** inclusive and for $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$ as a function of **b)** p_T^{track} and **c)** z . The corrected yields of the MC truth at detector level (open dots) are compared to the MC truth at particle level (open crosses).

5.5. Limitations of the MC Correction Factors

In some bins, the MC correction factors are rather unreliable. This happens if the total systematic uncertainty in such a bin becomes very large or if the efficiency is too small — a typical example are protons at low p_T , see Figs. 5.25 and 5.1. A small efficiency implies a huge correction factor and is highly sensitive to a proper detector simulation. Hence, bins are removed from the results if at least one of the following conditions is fulfilled:

- MC bin-by-bin correction factor for efficiency (including acceptance and p_T resolution), CF , is too large: $CF \geq 9/4$,
- relative total systematic uncertainty of MC bin-by-bin correction factor, ΔCF , is too large: $\Delta CF \geq 1/3$.

The chosen thresholds exhibit some degree of arbitrariness. The presented values lead to rather stable rejection cut-offs for each species as a function of the considered data sample (inclusive and different p_T^{jet}). The cut-offs in p_T and z are listed in Tab. 5.3. The kaon p_T cut-offs are slightly smaller for the inclusive reference of the jet analysis than in the multiplicity analysis due to the higher efficiency of the modified hybrid tracks.

Table 5.3.: p_T and z cut-offs due to limitations of the MC correction factors. In the column “charge”, the value “both” means that all charged particles are considered, whereas “all” indicates that the cut-offs hold for positive particles (“ > 0 ”), negative particles (“ < 0 ”) and for “both”. The case “jets, $p_T^{\text{jet}} \geq 10 \text{ GeV}/c$ ” for the data sample indicates that the values hold for $p_T^{\text{jet}} = 10 - 15, 15 - 20$ and $20 - 30 \text{ GeV}/c$.

Species	Data sample	Charge	p_T cut-off (GeV/c)	z cut-off
π	inclusive (jets & mult.)	all	0.15	—
	jets, all p_T^{jet}	all	0.15	0
K	inclusive (jets) & jets, $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$	$\begin{cases} \text{both} \\ > 0 \\ < 0 \end{cases}$	$\begin{cases} 0.35 \\ 0.3 \\ 0.35 \end{cases}$	$\begin{cases} 0.05 \\ 0.05 \\ 0.05 \end{cases}$
	jets, $p_T^{\text{jet}} \geq 10 \text{ GeV}/c$	all	0.35	0
	inclusive (mult.)	$\begin{cases} \text{both} \\ > 0 \\ < 0 \end{cases}$	$\begin{cases} 0.35 \\ 0.35 \\ 0.4 \end{cases}$	$\begin{cases} — \\ — \\ — \end{cases}$
p	inclusive (jets & mult.)	all	0.35	—
	jets, $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$	all	0.35	0.05
	jets, $p_T^{\text{jet}} \geq 10 \text{ GeV}/c$	all	0.35	0

6. Identified Particle Production in pp Collisions in ALICE

In this chapter, results on identified particle production in pp collisions at a centre-of-mass energy of $\sqrt{s} = 7\text{ TeV}$ are presented. Three different data samples are analysed to determine:

- the charged particle production in jets
- and the inclusive charged particle production in bins of multiplicity, the multiplicity being estimated at
 - forward rapidity
 - and mid-rapidity.

Firstly, the event and track selection is described in Section 6.1, followed by the presentation of the raw results in Section 6.2 and a discussion of the systematic uncertainties in Section 6.3. The inclusive minimum bias particle production and that in multiplicity bins is evaluated in Sections 6.4 and 6.5, respectively. Finally, the particle production in jets is discussed in Section 6.6, followed by a discussion of all obtained results.

This work only deals with charged particles and charged jets, without information about neutral particles. In the following, all (jet) momenta are implicitly understood to be the charged momenta, e.g. p_T^{jet} denotes the transverse momentum of a jet that was reconstructed from charged particle momenta only. The same applies to the fragmentation variables z and ξ , defined in Eqs. 4.8 and 4.9, respectively. As mentioned at the end of Section 4.4, the results as a function of ξ will not be shown in this thesis, but the corresponding procedure is also described for that case. If it turned out as beneficial, it would be easy to just run the analysis as a function of ξ .

6.1. Event and Track Selection

The analysis is run on data from pp collision at $\sqrt{s} = 7\text{ TeV}$ recorded in 2010. The run lists are available in Appendix B. Events fulfilling the minimum bias trigger condition have been selected for this analysis. The minimum bias trigger requires at least one hit in either the V0 forward scintillators or in the two innermost SPD layers of the ITS, in coincidence with an LHC bunch crossing.

6. Identified Particle Production in pp Collisions in ALICE

The event selection is carried out by the standard ALICE physics selection which re-evaluates the hardware trigger condition on the reconstructed data and suppresses beam-gas and other machine-induced background triggers by demanding the event to coincide with an LHC bunch crossing [118]. The event sample after trigger/physics selection has been restricted to events passing the following vertex cuts:

- the event has a reconstructed primary vertex,
- the primary vertex of the event has at least one contributor and
- the vertex is not a TPC primary vertex.

Moreover, the position of the reconstructed vertex is demanded to be within $|z| < 10$ cm along the beam axis (Z_{vtx} cut). It can happen that multiple collisions occur in the same LHC bunch leading to separate interaction vertices and, thus, “overlapping” events. This situation is called *pile-up*. In such cases, the analysis result can be biased. For this reason, a pile-up rejection is applied, which uses the SPD to detect multiple interaction vertices. The resulting statistics is listed in Tab. 6.1.

Table 6.1.: Statistics of various analysis data samples after physics selection, event cuts and pile-up rejection.

Part of analysis	Number of events (10^6)
V0M multiplicity	194
Reference multiplicity	216
Jets (+ inclusive reference)	216

Each part of the analysis uses an individual set of cuts as is listed in Tab. 6.2. The splines and η -maps are created from tracks after the ITS-TPC standard track cut 2010 selection (see Appendix C.2 for details) with an additional requirement of at least 60 PID clusters, a restriction to the fiducial TPC acceptance $|\eta| < 0.9$ and a cut on the lower track transverse momentum $p_T > 0.15 \text{ GeV}/c$. The cut is configured to select primary particles via the Distance of Closest Approach (DCA) — the latter is not used for tracks with V0 topology. These cuts ensure a good PID quality. The same cuts are used for the inclusive analysis in multiplicity bins, but with $|\eta| < 0.8$.

However, the jet reconstruction requires a flat efficiency/acceptance in azimuthal angle. Due to dead zones in the ITS, the standard cuts do not have a flat efficiency. The (default) *hybrid track cuts* [119] have been developed for a uniform azimuthal acceptance. They also accept tracks with no or too few ITS hits, but constrain such tracks to the primary vertex to improve the p_T resolution. In this work, the jet analysis uses *modified hybrid track cuts* with the same additional

Table 6.2.: Track cuts used in the individual parts of the analysis.

Part of analysis	Track cuts	η range
Extraction of PID parametrisations V0M multiplicity Reference multiplicity	ITS-TPC	$ \eta < 0.9$
	standard	$ \eta < 0.8$
	tracks 2010	$ \eta < 0.8$
Jets (+ inclusive reference)	Modified hybrid tracks	$ \eta < 0.9$

cuts as mentioned in the previous paragraph. Compared to the default hybrid tracks, they have a smaller secondary contamination. More details can be found in Appendix C.3.

Both hybrid track cuts and standard cuts are rather similar with respect to the TPC dE/dx quality. Moreover, secondary contamination and detector effects are corrected for via MC, such that it is valid to use these different cuts for the extraction of the dE/dx parametrisations and the yield extraction. A further validation is discussed in Section 6.3.10, where it will be shown that even for extremely different PID quality cut variations, the results are still consistent within systematic uncertainties.

6.2. Discussion of Uncorrected Results for Jets

To get a feeling about the results from the fitting without any impact from the MC corrections, a selection of typical raw results is presented briefly in this section. The systematic uncertainties are shown as well, but their discussion is reserved for Section 6.3. The overall trends, as well as the fit quality are rather similar for all parts of the analysis. Hence, this section focusses on the analysis of jets with $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$.

6.2.1. Typical Uncorrected Results

The uncorrected particle fractions in jets with $p_T^{\text{jet}} = 5 - 10$ and $10 - 15 \text{ GeV}/c$ are shown in Fig. 6.1. The overall trends are similar for both p_T^{jet} bins. The electrons contribute more than 10% to the total yield for $p_T < 0.3 \text{ GeV}/c$, but are mainly secondary particles from photon conversions (cf. Fig. 5.3). The pion fraction becomes rather flat at high p_T , whereas the kaon fraction continuously increases. Correspondingly, the proton fraction exhibits a maximum and decreases towards higher p_T . The systematic uncertainties significantly increase in regions with dE/dx crossings. The increased error at high p_T is due to the larger uncertainty of the mean dE/dx parametrisation.

6. Identified Particle Production in pp Collisions in ALICE

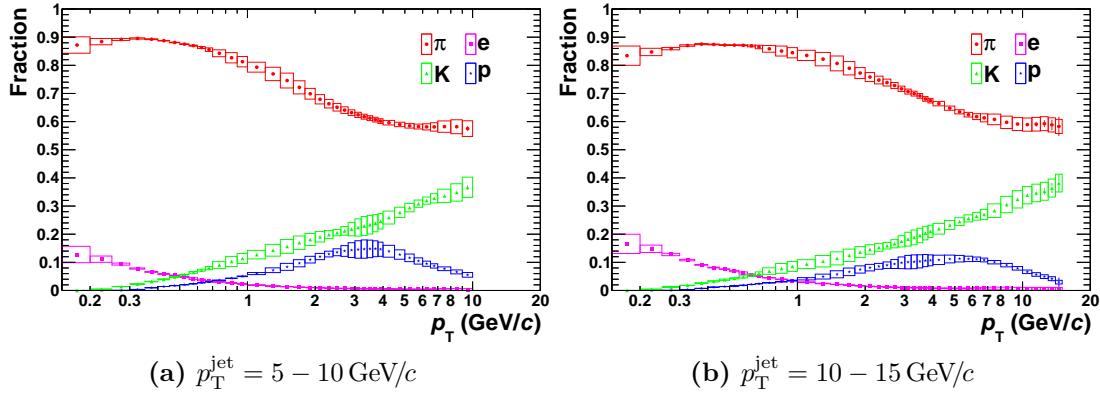


Figure 6.1.: The uncorrected particle fractions are shown for particles in jets with $p_T^{\text{jet}} = 5 - 10$ (a) and $10 - 15 \text{ GeV}/c$ (b) in data. The boxes show the total systematic uncertainty.

6.2.2. Fit Quality Checks

To judge the quality of the fit, the fitted dE/dx distributions are shown in Fig. 6.2 for typical bins at low, intermediate and high p_T , respectively. In addition, the relative deviation between fit and data is presented, for which the error bars are solely the statistical errors of the data points, i.e. the uncertainty of the fit in the corresponding bin is not shown. The plots are for jets with $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ as a function of p_T^{track} in data. The same plots with a linear vertical axis and similar plots for templates in z are shown in Appendix F.2. As can be seen, the templates describe the data with good precision, the relative deviation between fit and data being typically smaller than 10%.

The particle fractions are changing significantly as a function of p_T with increasing p_T^{jet} , as is visible in Fig. 6.1¹. That the TPC Multi-Template Fit is sensitive to these changes, is demonstrated in Fig. 6.3. It shows the fit of the dE/dx distribution for $p_T^{\text{track}} = 4 - 4.5 \text{ GeV}/c$ in the individual p_T^{jet} bins and the inclusive case. The fit quality is good in all cases despite some deviations for the tails with very low statistics, which are mainly visible for the inclusive case with the highest statistics and, thus, the largest dE/dx range with non-empty bins.

¹This observation will become even more evident in Fig. 6.34.

6.2. Discussion of Uncorrected Results for Jets

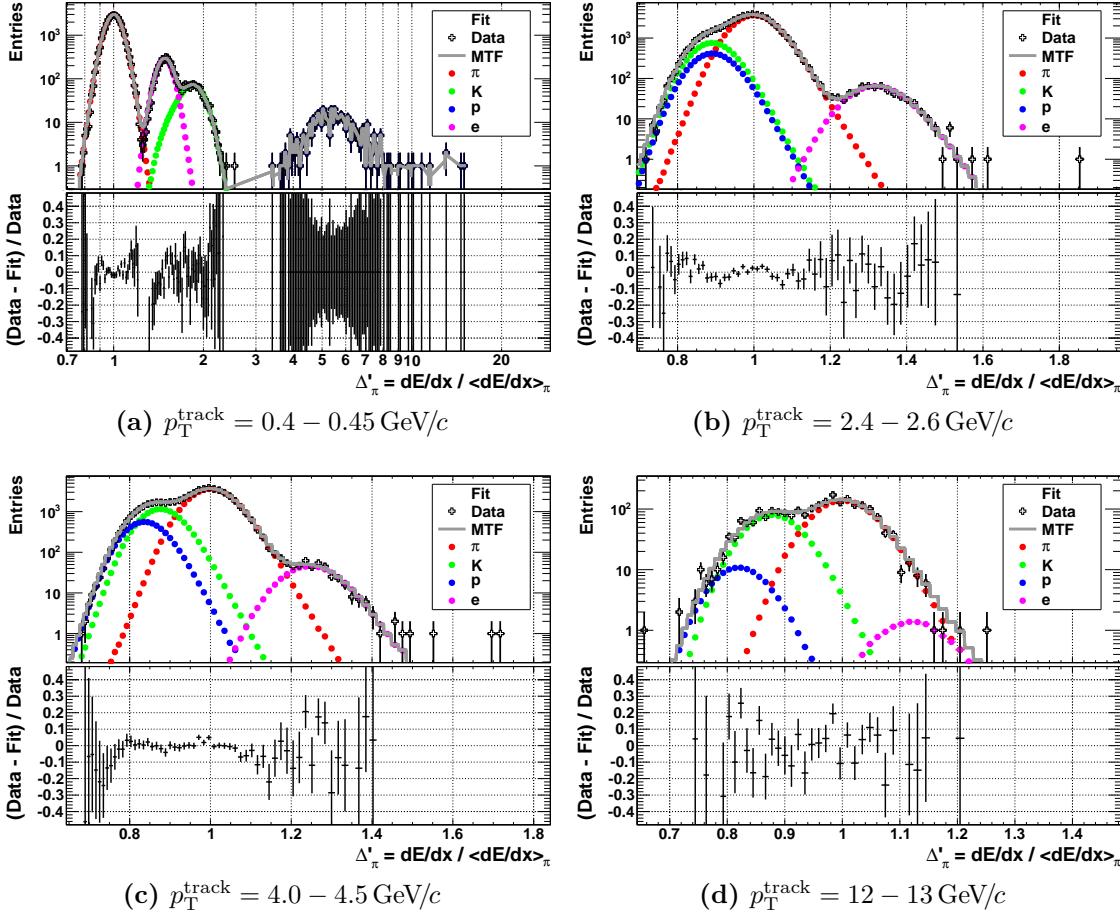


Figure 6.2.: Template fits to data in different p_T^{track} intervals for $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$. The panels a) – d) show the following p_T^{track} ranges in GeV/c : 0.4 – 0.45, 2.4 – 2.6, 4.0 – 4.5 and 12 – 13. Note the different ranges of the abscissa and its logarithmic scale in a).

6. Identified Particle Production in pp Collisions in ALICE

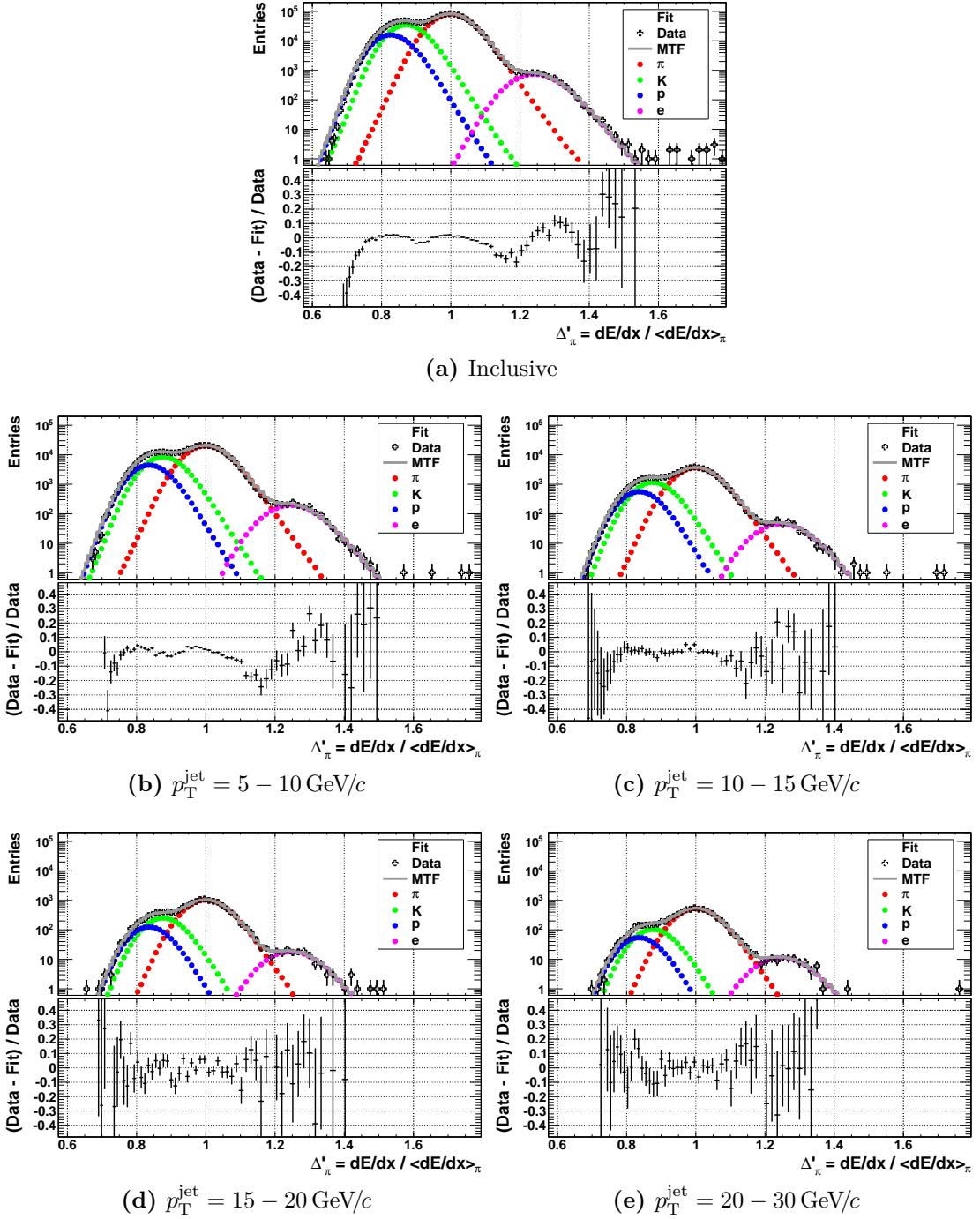


Figure 6.3.: Comparison of the template fits for minimum bias data and jet constituents. The fits are shown for $p_T^{\text{track}} = 4 - 4.5 \text{ GeV}/c$.

6.2.3. Comparison of Uncorrected Results to the TPC Coherent Fit and MC

Independently of the TPC Multi-Template Fit (MTF), the TPC Coherent Fit (TCF) was developed to extract particle fractions and yields (and, within the same fit, also the TPC PID response parameters) in jets [112, 120]. At the time of writing this thesis, the TCF analysis has only been run on default hybrid tracks and the minimum bias MC production (LHC10f6a) was used for correction [121]. To directly compare the uncorrected particle fractions obtained with the two methods, the same data sets and cuts — namely the default hybrid cuts — are used in the following. To illustrate the discriminating power with respect to MC simulations, the MC prediction of PYTHIA Perugia-0 propagated through a full detector simulation with GEANT will also be shown. The MC fractions are determined by counting MC labels, i.e. no fitting is used.

In Figs. 6.4, 6.5 and 6.6, the uncorrected particle fractions of pions, kaons and protons, respectively, are shown as a function of p_T . The corresponding plots as a function of z can be found in Appendix H.1, for which similar observations hold. In all cases, good agreement within uncertainties is observed for both methods². This comparison confirms the obtained particle fractions: two independent methods with different systematic uncertainty estimations yield consistent results.

The shape in MC is very different from that observed in data for the lowest p_T^{jet} bin and the situation is similar for the inclusive particle production. The discrepancy can reach up to 20 – 30%. For higher p_T^{jet} , the agreement gradually becomes better and is typically within 10 – 20%. However, this does not hold for protons. Their agreement gets better from $p_T^{\text{jet}} = 5 – 10$ to $10 – 15 \text{ GeV}/c$, but seems to get worse again for $p_T^{\text{jet}} = 15 – 20 \text{ GeV}/c$.

²Note that there are no uncorrected results from the TCF available for the inclusive case with this set of cuts. Furthermore, the TCF has not been run for $p_T^{\text{jet}} = 20 – 30 \text{ GeV}/c$ yet.

6. Identified Particle Production in pp Collisions in ALICE

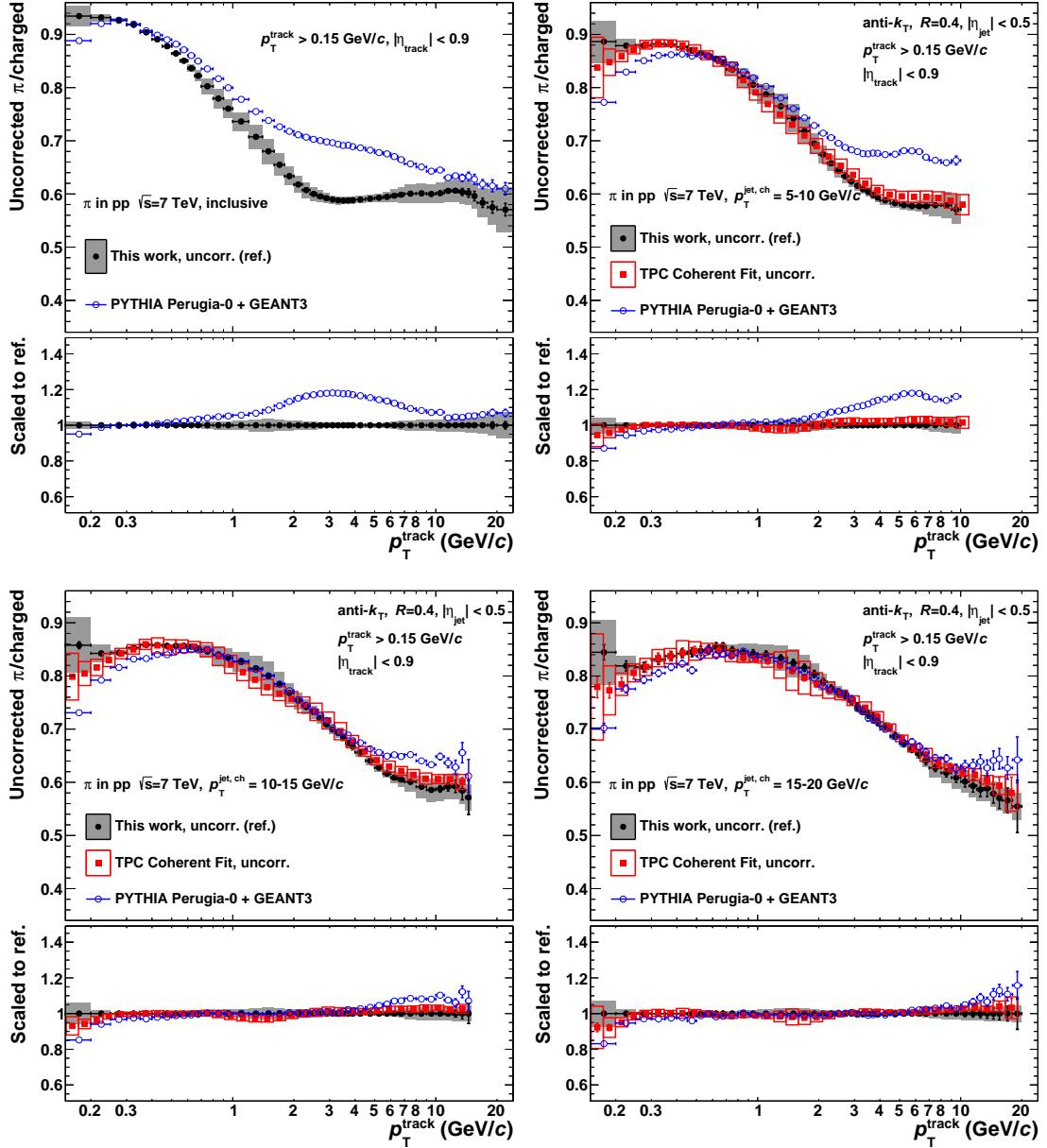


Figure 6.4.: Comparison of uncorrected pion fractions as a function of p_T to the TPC Coherent Fit. In addition, the MC prediction from PYTHIA Perugia-0 is overlaid. Note that there are no results from the TPC Coherent Fit for the inclusive case.

6.2. Discussion of Uncorrected Results for Jets

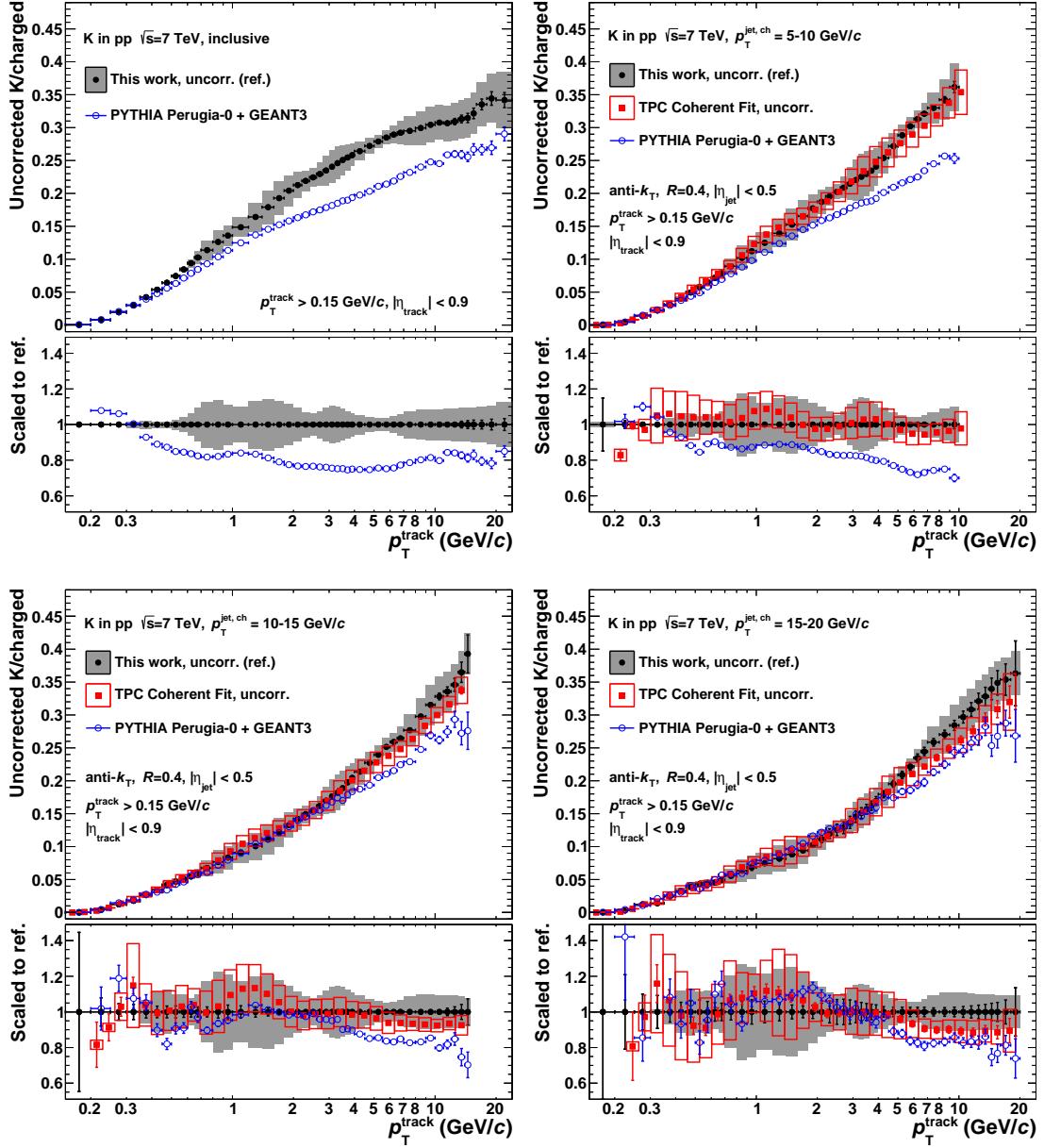


Figure 6.5.: Comparison of uncorrected kaon fractions as a function of p_T to the TPC Coherent Fit. In addition, the MC prediction from PYTHIA Perugia-0 is overlaid. Note that there are no results from the TPC Coherent Fit for the inclusive case.

6. Identified Particle Production in pp Collisions in ALICE

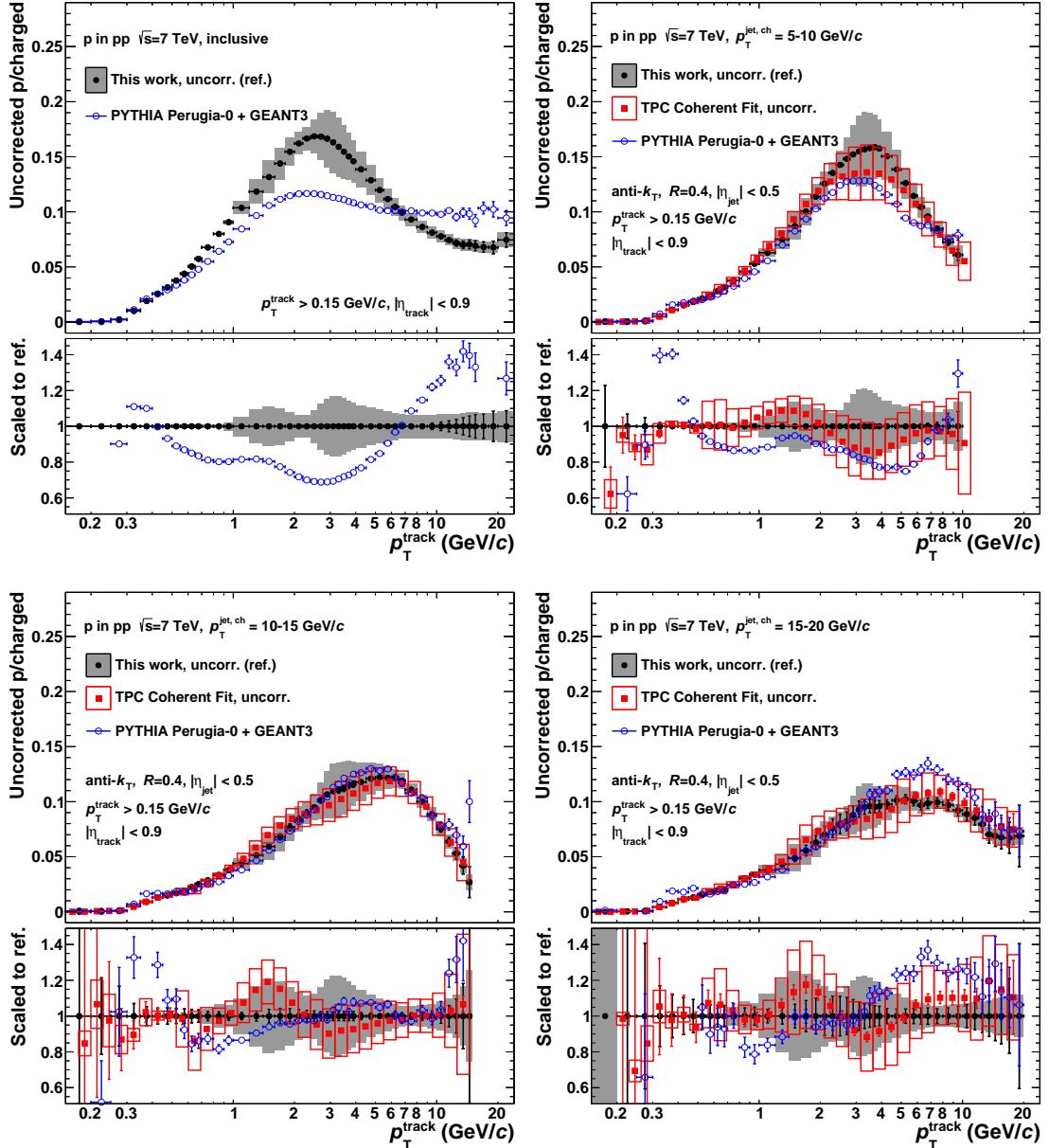


Figure 6.6.: Comparison of uncorrected proton fractions as a function of p_{T} to the TPC Coherent Fit. In addition, the MC prediction from PYTHIA Perugia-0 is overlaid. Note that there are no results from the TPC Coherent Fit for the inclusive case.

6.3. Estimation of Systematic Uncertainties

The systematic uncertainties of the MC corrections have already been evaluated in Section 5.3. This section discusses the systematic uncertainties arising from the fitting procedure, as well as those from the pile-up rejection. The latter corresponds to an uncertainty of the normalisation and will be shown separately in the plots with the final results. The uncertainties from the MC corrections and the fitting procedure are independent of each other and are added in quadrature.

The shape of the templates is varied to estimate the systematic uncertainty of the raw fractions and yields. An overview of the variations can be found in Tab. 6.3. In cases in which there are different variation sizes below and above a threshold, the variation is smoothly changed around this threshold via the error function.

Table 6.3.: Variations for systematic uncertainty estimation (cf. discussion in the text). If variations are quoted in per cent, the default values are understood to be included. In all other cases, the default case is stated explicitly.

Template input	Variations
Mean dE/dx (splines)	$\pm 0.2\%$ for $\beta\gamma \lesssim 50$ $+1.3\%$ and -0.6% for $\beta\gamma \gtrsim 50$
$\sigma(dE/dx)$	$\pm 3\%$ for $dE/dx \lesssim 250^{\dagger}$ $\pm 50\%$ for $dE/dx \gtrsim 250$
η dependence of mean dE/dx	$\pm 3\%$ for $p \lesssim 0.45 \text{ GeV}/c$ $\pm 0.5\%$ for $p \gtrsim 0.45 \text{ GeV}/c$
PID weighting	Default: Combined PID from ITS, TPC, TOF with default priors Flat
Shape of detector response	Default: Asymmetric shape (cf. Eq. 4.5) Pure Gaussian

The systematic uncertainty of the results is extracted from the spread of the variations. The method follows [122], the statistical errors of the variations are fully correlated and not taken into account⁴, such that the systematic error becomes the sigma of the result distribution from the variations. The default parameters describe the data best. For this reason, the mean of the result is the same as for the default settings and not the mean of the systematic variations.

[†]The dE/dx is measured in arbitrary units throughout in this work, see, e.g., Fig. 3.1.

⁴Each variation is obtained from the same data set with equal statistics. Hence, the statistical errors do not contribute to the spread of the results. In contrast, uncorrelated statistical errors are assumed in [122], which then also need to be taken into account.

6. Identified Particle Production in pp Collisions in ALICE

In the following, the variations for each type of systematic error source are discussed. For every type, the systematic error of the raw fraction is derived as described before. The size and trending of the systematic errors are rather similar for the inclusive and the jet case. Hence, the results for $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ are discussed as a typical example. Finally, a variation of the track cuts is carried out in Section 6.3.10 and will verify the estimated (total) systematic uncertainties.

6.3.1. Splines

The systematic uncertainty due to the splines is estimated by varying the splines within their uncertainties (estimated in Section 3.4.5) and looking at the spread of the results. For each variation, the change (scaling) of the splines is done globally (except for the $\beta\gamma$ threshold) for all bins and all species. A typical result is shown in Fig. 6.7 for jets with $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ in data. The uncertainty of the splines has mainly an impact on the dE/dx crossing regions, where the dE/dx shapes of two species are very similar and a small shift of the mean has a large effect, and at high momenta, where the Bethe-Bloch parametrisation approaches the plateau and the uncertainties are large due to a lack of data constraints. The relative systematic error is up to 10% for kaons and protons and more like 4% for pions.

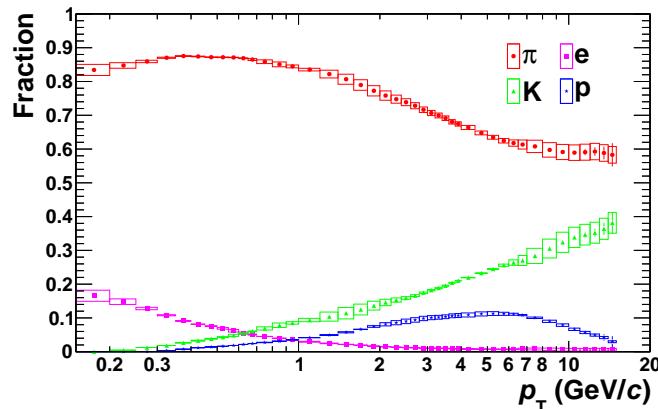


Figure 6.7.: The uncorrected particle fractions are shown for particles in jets with $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ in data. The boxes show the systematic error due to the spline uncertainties.

6.3.2. η Correction

The systematic error of the η correction is derived from the variations $\delta\widetilde{\Delta}' = 0, \pm 0.5\% (\pm 3\%)$ for p greater (less) than $0.45 \text{ GeV}/c$, respectively, in Eq. 3.20 — cf. Section 3.5.1. For each variation, the change (scaling) is done globally (except for the momentum threshold) for all bins and all species. As depicted in Fig. 6.8, the

impact on the extracted fractions is predominantly around the dE/dx crossing of kaons and protons. This crossing is much flatter as a function of momentum than for any other pair of species, such that details of the shape become important. Similar large errors are observed for the dE/dx crossing of pions and electrons in the very first bins. Although it is rather steep, the higher uncertainty of the η correction at such low momenta propagates to the fractions.

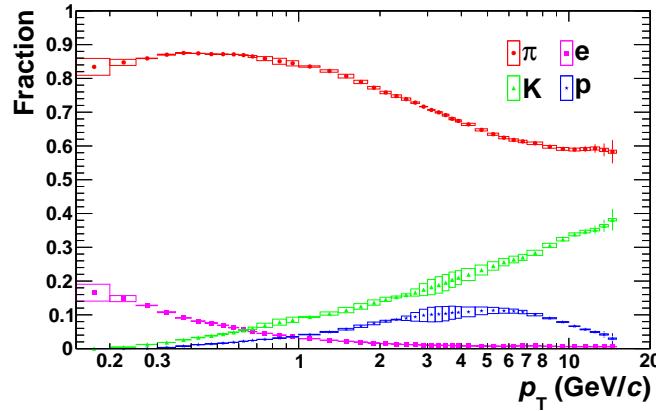


Figure 6.8.: The uncorrected particle fractions are shown for particles in jets with $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ in data. The boxes show the systematic error due to uncertainties of the η correction.

6.3.3. Resolution Map

The impact of the systematic uncertainty of the resolution map on the fit results is studied by looking at the variation between default resolution parametrisation and scaling the resolution globally (except for the dE/dx threshold) for all bins and all species by $\pm 3\%$ ($\pm 50\%$) for dE/dx smaller (larger) than 250 — cf. Section 3.6.1. The large uncertainty in the high- dE/dx region has little to no impact on the analysis, since special templates are used in that region (cf. Section 4.2.2).

In Fig. 6.9, the systematic error of the fractions due to uncertainties of the resolution are illustrated. The systematic error is largest in the region, where the dE/dx of all three hadrons ($\pi/K/p$) are similar and one species is “sandwiched” between the others. In that case, the width of the shape matters most to disentangle the particle content. Obviously, the width of the shape has only a minor impact at the dE/dx crossing of two species, which is mainly sensitive to the mean of the shape.

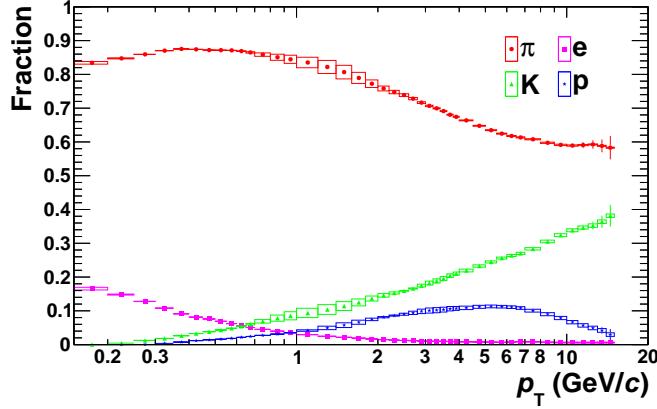


Figure 6.9.: The uncorrected particle fractions are shown for particles in jets with $p_T^{\text{jet}} = 10-15 \text{ GeV}/c$ in data. The boxes show the systematic error due to uncertainties of the resolution.

6.3.4. Shape of the Detector Response

The templates are created with the asymmetric dE/dx response (see Eq. 4.5) per default. This is compared to the case where a pure Gaussian is used for the response. The shape uncertainties cause systematic errors of the fractions mainly in the dE/dx crossing regions, see Fig. 6.10.

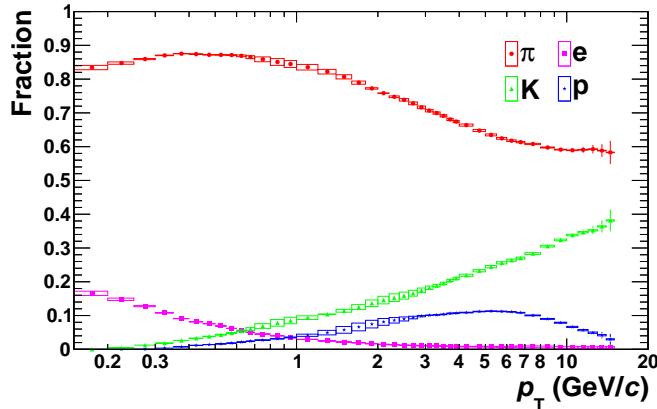


Figure 6.10.: The uncorrected particle fractions are shown for particles in jets with $p_T^{\text{jet}} = 10-15 \text{ GeV}/c$ in data. The boxes show the systematic error due to uncertainties of the shape of the dE/dx response.

6.3.5. PID Weighting

The default case uses Bayesian probabilities from ITS, TPC and TOF and default priors for the weighting of the template generation. The result is compared to that for template generation with flat weighting, i.e. each track contributes equally to

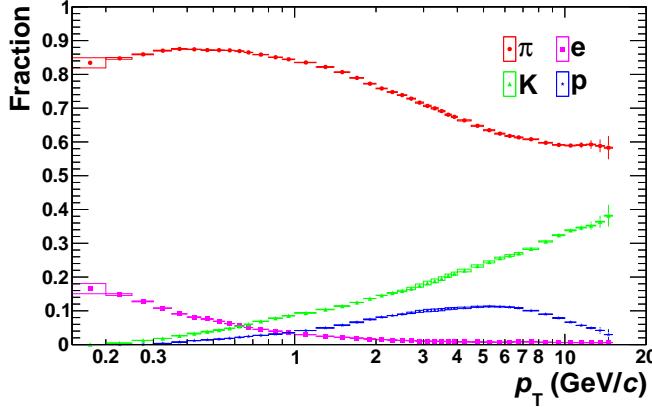


Figure 6.11.: The uncorrected particle fractions are shown for particles in jets with $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ in data. The boxes show the systematic error due to uncertainties of the PID weighting.

the templates of all species. Fig. 6.11 shows the impact on the extracted fractions, which is found to be negligible compared to other error sources in almost the full p_T range.

It is instructive to look at $p_T = 0.15 \text{ GeV}/c$, where the relative error of the pions is about 2%, whereas it is much smaller for the subsequent bins. At such low momentum, the detection efficiency for kaons and protons is essentially zero, i.e. the only relevant species are pions, electrons and muons. Muon templates are not included, since they are typically completely covered by the pion templates. In this region however, the muon template starts to deviate from that of the pions. The effect is not big enough to reliably fit muons separately, but it already distorts the shape. Also, the systematics of the η correction may cause further shape distortions, so that after all, the templates do not perfectly reproduce the observed shape. It can happen that the fit converges to a local minimum, which, since the regularisation has only constraints from higher p_T , causes a small jump of the pion and electron fractions.

In Fig. 6.12, the template shapes are shown. A close look at the upper panel reveals that the pion template at $\Delta' = 1$ is slightly lower in case of equal weights (compare red and black histograms), the electron template being shifted towards lower Δ' at the same time⁵. The ratio in the lower panel looks pretty similar demonstrating that the change of the templates can be compensated by a change of the particle fractions. The template difference is mainly due to the priors — this has been checked explicitly by comparing the templates with PID weighting with and without priors, i.e. only using the detector probabilities in the latter case. The electrons have a prior probability of $O(1\%)$, whereas the pions are close to 100%. The priors become important for the weighting if the distributions of the

⁵The maximum is at $\Delta' \approx 1.14$ for the pre-PID weighted templates and at $\Delta' \approx 1.1$ in case of equal weights.

6. Identified Particle Production in pp Collisions in ALICE

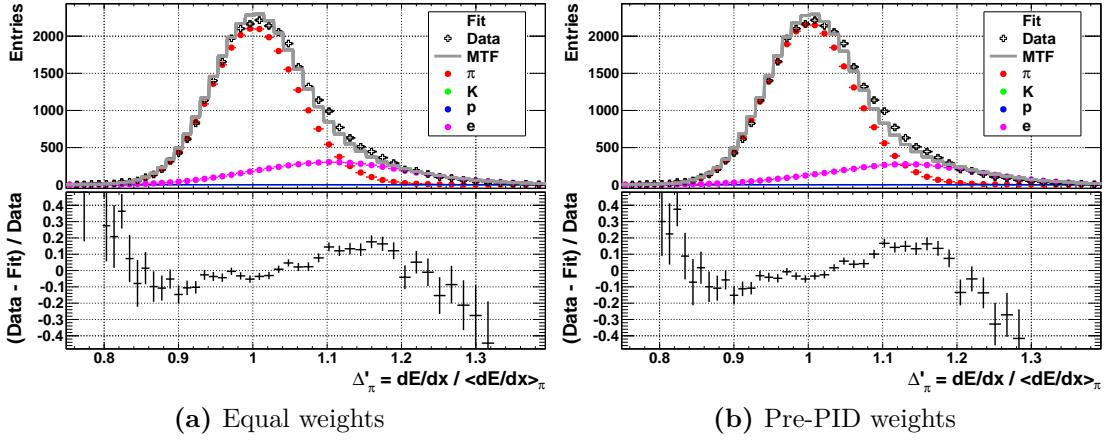


Figure 6.12.: Shape uncertainty of pre-PID weighting for $p_T^{\text{track}} = 0.15 - 0.2 \text{ GeV}/c$ and $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ in data. Effectively, there are no kaons and protons due to low detection efficiency. **a)** shows the template shapes created with equal weights, while in **b)**, the pre-PID weights with priors have been used.

detector responses overlap, which is the case in the lowest p_T bin for electrons and pions. Since the separation depends on η , the by almost two orders of magnitude different priors cause that the phase space is sampled completely differently by the electron templates than for a flat weighting. For some η , there might be good separation, such that the detector probability is zero for pions and the input of all tracks from this pseudo-rapidities are attributed to the electron template. On the contrary, there might be a finite probability for pions at other η ranges, which causes together with the priors that the weight for the electron templates is highly suppressed — which is desired, since pions are predominantly contributing in that case.

To discuss this in detail, three methods can now be compared: equal weighting (1), weighting with detector probabilities without priors (2) and weighting with detector probabilities with priors (3). Method (1) can only be used, if the input parameters of the tracks for the template generation are similar. At such low p_T , the electrons mainly stem from photon conversion and may have a completely different topology than primary pions. So, it is not clear whether method (1) can be safely used. For method (2), the input parameters are assigned correctly to the corresponding templates, if the separation between the species is large. Once there is overlap, the assignment is shared. Consider the case that the detector probabilities are 50% for both species, but one species is much more abundant than the other. Without priors, the less abundant species will get a too large share from that region, which distorts the total shape once other regions are added. If priors are used, method (3), the abundances are taken into account correctly.

To sum up, this explains the varying electron template shapes in Fig. 6.12 for the methods (1) and (3). Method (3) theoretically gives the most realistic

weighting. However, if the expected detector response used for the probabilities is slightly off (or the priors are much off), the very different priors can bias the electrons by strongly reducing the contribution to the templates from some phase space regions by mistake. As a result, the two variations (1) and (3) give a reasonable estimate of the systematic error from the pre-PID.

Note that the large systematic error is merely visible for the crossing of two particles with priors that vary by at least one order of magnitude. This is only fulfilled for the electron-pion crossing. For instance, for the crossing between electrons and kaons or protons around $p = 0.5$ and $1 \text{ GeV}/c$, respectively, the priors are of order 1% for both species. In those cases, the exact values for the detector probabilities do not change the template shape significantly. Also, at momenta above $10 \text{ GeV}/c$, where electrons and pions approach each other again in dE/dx , there is no large error observed, since the electron fraction is fixed for the fit — the contribution of the electron template becomes anyway negligible.

6.3.6. Pile-up Rejection

The pile-up rejection can introduce biases in a twofold way. It may either reject events that are not affected by pile-up (false positives) or it may accept events that are affected (false negatives). The minimum bias MC production LHC10f6a, that was used to extract the correction factors for the inclusive analyses, does not simulate pile-up. Hence, all rejected events are false positives. Their fraction is found to be completely negligible. Consequently, any systematic uncertainty of the pile-up rejection arises from false negatives.

In general, the pile-up rejection has very different impact on the individual parts of the analysis. For the jet and the inclusive MB analysis, less than 1% of the events are rejected due to pile-up after all other event cuts. The impact on the identified particle yields is below 1% and considered as negligible compared to other systematic uncertainties. In the multiplicity analysis, the bins with highest multiplicities are affected most by pile-up/the pile-up rejection, i.e. the fraction of events tagged as pile-up is largest. The systematic uncertainty due to (residual) pile-up is estimated as follows. All events that are tagged as pile-up are not counted for the number of analysed events N_{evt} . The relative systematic uncertainty of N_{evt} due to pile-up is estimated as 100% of the fraction of pile-up rejected events. Since the spectra are proportional to $1/N_{\text{evt}}$ (by choice of the normalisation), the error can simply be propagated from N_{evt} to the spectra.

According to some first more detailed studies [123], this is a valid conservative estimate. It may be realistic to decrease the estimated uncertainties by roughly a factor 3 – 10, which is desired with respect to currently estimated errors beyond 30% for the highest multiplicity bins. However, this requires further thorough checks and has to be done for all analysed runs.

6.3.7. Special Treatment of Errors for the Multiplicity-to-Minimum Bias Ratios

As will become evident in Section 6.5.3 and the following discussion, it is crucial for drawing conclusions to calculate the uncorrelated errors for the multiplicity evolution. Hence, the “real ratios” of the multiplicity bins with respect to MB will be shown instead of the “scaled to reference” results used for the other analysis parts (cf. discussion in Appendix A).

To eliminate the correlated errors, the following method is used. It is based on the observation that the dE/dx parametrisation does not depend on the multiplicity for the considered multiplicity classes, which will be discussed in detail in Section 6.5.2. The dE/dx template variation is done *simultaneously* for the multiplicity bin *mult* and MB, i.e. if e.g. the mean value of the splines is increased for *mult*, it is increased at the same time by the same amount for MB. It is then checked how the variation changes the “ $\text{result}(\text{mult}) / \text{result}(\text{MB})$ ” ratio. In the final step, the systematic error of the ratio is estimated from this variation in the same way it is done for the result of a single bin.

The (utilised) MC correction factors are identical for all multiplicity bins. It is assumed that their uncertainties completely cancel in the ratio. As discussed in Section 5.3.7, this assumption is fulfilled in good approximation, except for protons at $p_T \lesssim 0.6 \text{ GeV}/c$, for which these uncertainties of the ratio gradually increase to 4 – 8% for the most extreme multiplicity bins with decreasing p_T , until the cut-off is reached.

Finally, it should be noted that the statistical errors of the “*mult*/MB” ratio are calculated by assuming uncorrelated errors. It turns out that the extent of correlation does not matter, since the statistical error of the *mult* bin dominates by far that of the MB result.

6.3.8. Total Systematic Errors

The error sources are considered to be independent of each other, i.e. the systematic errors are summed in quadrature. The only exception is the pile-up rejection uncertainty, which will be plotted separately, but anyway drops out in the particle fractions, that are shown in the following. Note that although some variations are done globally, there is only correlation between neighbouring p_T (z) bins, such that the errors can be considered as being derived for each bin individually.

The results with total systematic uncertainty are shown in Fig. 6.13. The relative errors are typically below 10% for kaons and protons and below 5% for pions. In the dE/dx crossing regions, the relative uncertainties can be up to 20% for kaons and protons, but stay below 5% for pions.

In the relativistic rise region at momenta around $4 - 10 \text{ GeV}/c$, some interesting features are visible: the systematic errors of the pion and kaon fractions first decrease, then stay constant and finally increase again. This is because the

crossing regions lead to larger uncertainties first, then the separation becomes larger and constant in the relativistic rise. At some point, the uncertainties of the splines increase (the “gap” in the Bethe-Bloch fit, cf. Fig. 3.2b) for pions — and at higher momenta also for kaons and (for much higher momenta) for protons. This also effects the fraction of the kaons, which are next to the pions in dE/dx , but hardly the protons, which still have the kaons in between.

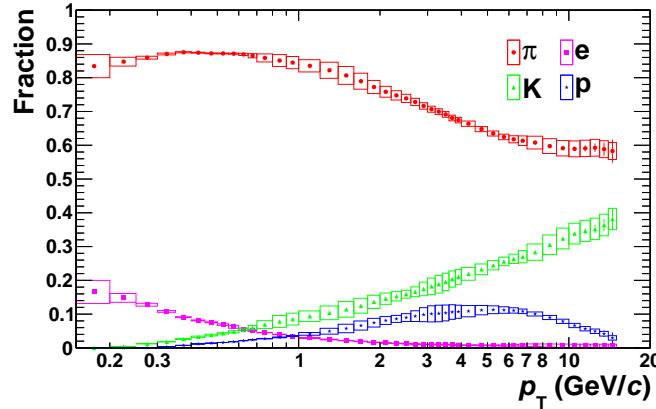


Figure 6.13.: The uncorrected particle fractions are shown for particles in jets with $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ in data. The boxes show the total systematic error as the quadratic sum of all sources.

A glance back to Section 4.4 shows that the total systematic errors obviously are reasonable when applying the procedure to MC (“semi-blind test”) and comparing to the MC truth. The difference between MC truth and fit is largely described by the systematic error bands.

6.3.9. Systematic Uncertainty Verification via η -Differential Jet Analysis

The detector performance changes significantly with the polar angle (cf. Chapter 3). The comparison of the results in different η slices serves as a good cross-check of the robustness of the analysis and the systematic error estimation. In particular, the dE/dx crossings of two species are at different p_T for fixed η . The detector efficiencies are rather similar within $|\eta| < 0.6$, such that the raw particle fractions of different η slices in this range can be compared directly. While the results at sufficiently high p_T should be consistent for all slices, physical deviations can be expected at low p_T . If the particle momentum becomes similar to its mass, the rapidity y and the pseudo-rapidity η differ. The particle fractions are measured and compared for fixed η ranges, but physics-wise the fractions should be identical for a fixed y range.

In Fig. 6.14, the uncorrected particle fractions in data for $|\eta| < 0.2$, $0.2 < |\eta| < 0.4$ and $0.4 < |\eta| < 0.6$ are compared to the full range $|\eta| < 0.9$. The

6. Identified Particle Production in pp Collisions in ALICE

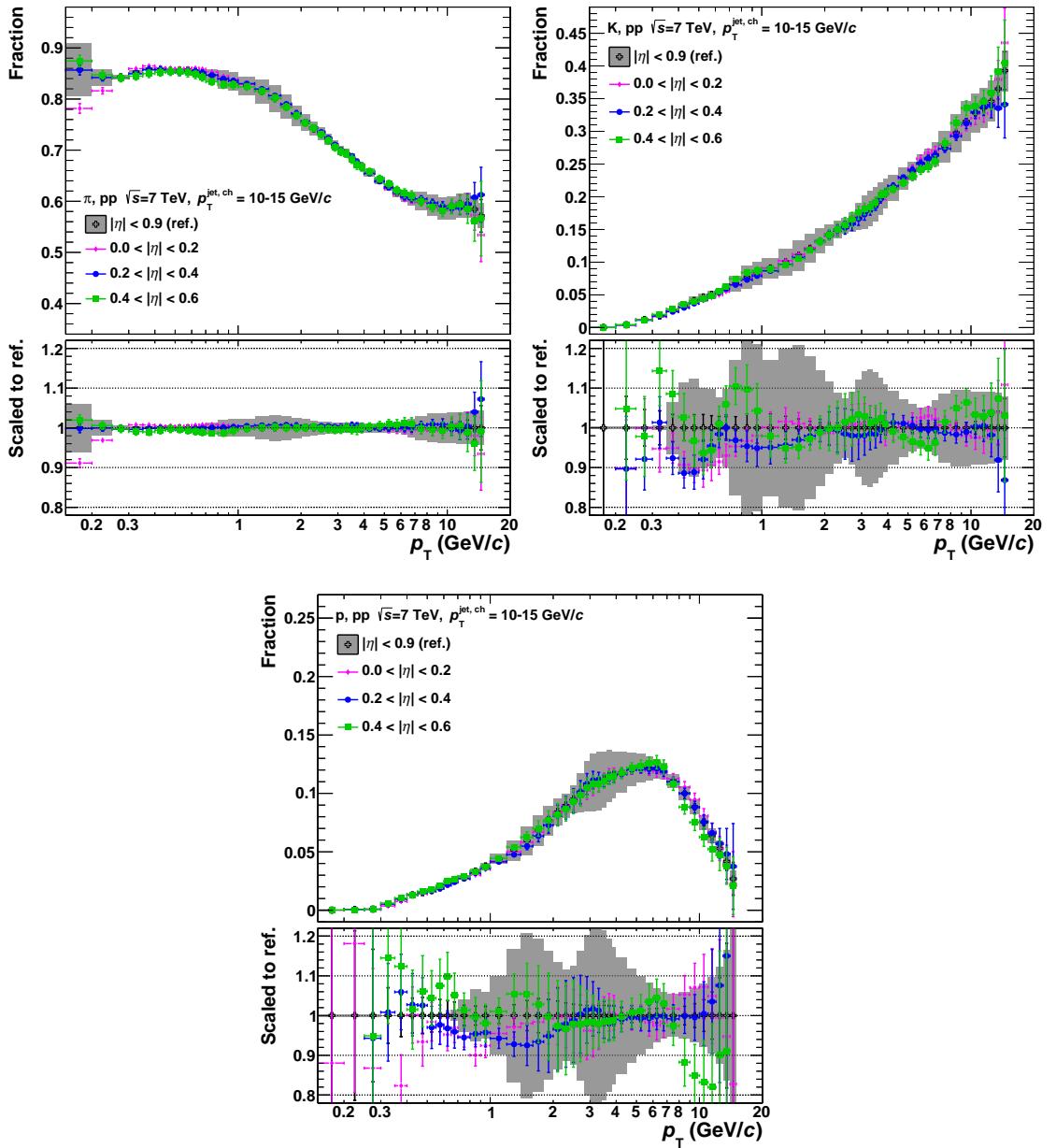


Figure 6.14.: Comparison of the uncorrected particle fractions in η slices (coloured symbols) and in the full η range (black open crosses) versus p_T^{track} for $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ in data. The results are shown for (from upper left to lower right) π , K and p . The grey area indicates the systematic uncertainties for the full η range.

results are presented for $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ and are similar for other p_T^{jet} bins and the inclusive case (cf. Appendix G). Typically, the results of the single η slices agree within $\pm 10\%$ with the result of the full range and are especially in agreement within the errors. At low p_T , larger deviations and an ordering of the η slices is visible as expected for the non-highly-relativistic regime. Note that the

systematic errors of the individual η slices have not been evaluated explicitly.

For the most forward bin (green rectangles), protons have very little statistics at highest p_T , such that the statistical error bars become large and so would the systematic error bars. The slice $0.6 < |\eta| < 0.9$ is not shown because its statistics is too small to allow for a clean extraction of the fractions (in particular around the dE/dx crossings). The relative track abundances in jets for $|\eta| < 0.2 \leftrightarrow 0.2 < |\eta| < 0.4 \leftrightarrow 0.4 < |\eta| < 0.6 \leftrightarrow 0.6 < |\eta| < 0.9$ are typically $9 \leftrightarrow 8 \leftrightarrow 5 \leftrightarrow 1$. Furthermore, the efficiency slightly changes for tracks with η being close to the edge of the detector acceptance.

In summary, the good consistency of the results of the η slices with that of the full range proves the robustness of the analysis. In particular, the estimate of the systematic uncertainty due to the fitting procedure is confirmed.

6.3.10. Robustness Against Cut Variations

The fully corrected results must not depend on the choice of track cuts, i.e. changing the track cuts for the analysis and the MC corrections at the same time should not change the final result within estimated uncertainties. Cut variations demonstrate the robustness of the analysis method including the estimation of the systematic uncertainties. Such checks are performed in this section.

Change of the PID Quality

In Section 6.1, it was mentioned that the track cuts used to extract the TPC dE/dx response are different from those used in the jet analysis (including the template generation). Since there is no impact on the PID quality from the difference of those cuts expected, this choice is justified. A further justification is presented in this section by studying extreme cut variations with large impact on the PID quality and showing that the results are still consistent within the systematic uncertainties.

The TPC dE/dx response is extracted with the ITS-TPC standard track cuts 2010 (see Appendix C.2 for details) with an additional cut on at least 60 PID clusters as before. The templates and particle yields are obtained from a default hybrid track sample⁶ with one of the following additional cuts:

- cut on at least 60 PID clusters (*ncl cut*),
- “TPC geometrical cut” (*geo cut*, see Appendix C.4 for details).

The geo cut mainly rejects short tracks or tracks with large overlap with the TPC sector boundaries. Especially, this removes tracks with worse resolution, mainly

⁶Note that the final results are extracted from the modified hybrid tracks, which already removes shorter secondary tracks with respect to the default hybrid sample. In this sense, the cut variation presented in this section is even more extreme.

6. Identified Particle Production in pp Collisions in ALICE

resulting from the lower number of PID clusters. Also, the geo cut may reject more tracks compared to the ncl cut in some η regions relative to others. This can change the mean dE/dx due to its η dependence. The bottom-line is that the measured dE/dx shape of a geo cut track sample can be different from that of an ncl cut sample. It is expected that taking such different cuts for the dE/dx response extraction compared to the analysed track sample biases the results.

The single track efficiencies for the geo cut compared to those for the ncl cut are shown in Fig. 6.15. Compared to the ncl cut, the geo cut reduces the efficiency by about 30% around $p_T = 1 \text{ GeV}/c$ and by about 5% for $p_T > 5 \text{ GeV}/c$ (see Appendix C.4 for further discussion). The efficiency decrease is basically species independent, except for the kaons. They can decay inside the TPC and the geo cut partially removes them due to the cut on the track length. This implies that the raw particle fractions will slightly change and the corrected results must be compared for both cuts instead.

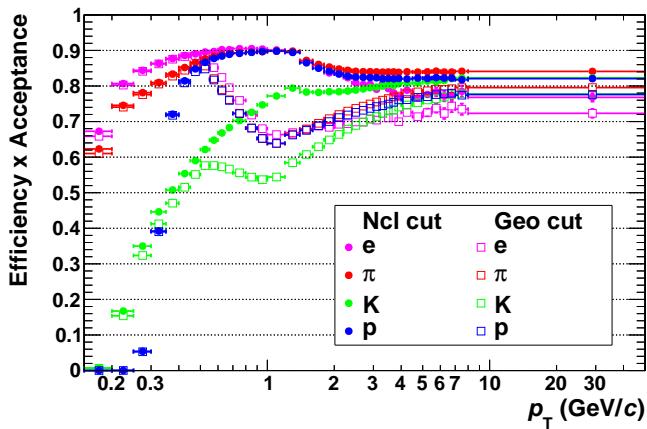


Figure 6.15.: The MC single track efficiencies of the ncl cut (dots) are compared to those of the geo cut (open boxes). Compared to the ncl cut, the geo cut is most effective around $p_T = 1 \text{ GeV}/c$ and saturates at a 5% lower efficiency for $p_T > 5 \text{ GeV}/c$.

To get an impression how much the dE/dx parametrisation will change for the geo cut, the splines have been extracted for MC (using the particle level truth for “identification”) for the geo cut. The result is compared to the default splines for the ncl cut in Fig. 6.16. The difference of the mean dE/dx is at the per mille level and becomes smaller than 0.5% for momenta beyond 4 GeV/c . Although this is a very small change, it is as large as (or at least a good fraction of) the estimated spline accuracy of about 2% (cf. Section 3.4.4) and may lead to considerable systematic biases in the dE/dx crossing regions. Similar or even larger impacts can be expected for the resolution maps. Consequently, strong variations of the results for the geo cut in the order of the systematic error may occur.

The corrected particle yields for $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ in data are shown in Fig. 6.17 for the ncl cut and the geo cut. Each set of cuts was corrected with its own MC correction factors. The black boxes are the systematic errors of the ncl

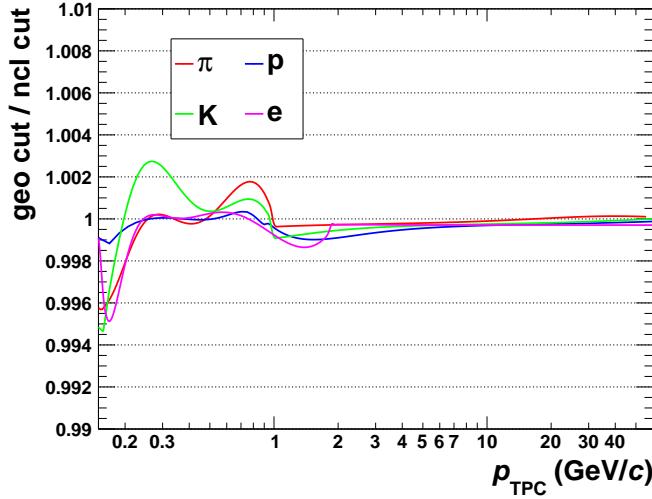


Figure 6.16.: The splines extracted from the geometrical cut sample (*geo cut*) are compared to the ones of the default ncl cut sample (*ncl cut*) in MC. Their ratio is shown as a function of momentum for the individual species. In both cases, the particle level truth was used to “identify” particles.

cut results, which include the systematic uncertainties of the MC corrections. The yields for the geo cut are consistent with those of the ncl cut within systematic errors. The biggest effect is visible for kaons around $p_T = 1.5 \text{ GeV}/c$, where the geo cut causes the largest difference in the track sample. On the one hand, the overall resolution of the remaining tracks should be improved after the geo cut. On the other hand, strongly reducing the statistics hides details in the dE/dx distributions and makes the PID more difficult. Therefore, there is a trade-off for the geo cut. The results for other p_T^{jet} bins and inclusive show similar consistency and can be found in Appendix I.

It should be noted again that for this comparison the geo cut is *not* treated the same way as the ncl cut, since the track cuts for the TPC dE/dx response determination are close to the ncl cut and are used for both, ncl and geo cut. The consistency of the results proves the robustness of the analysis against cut variations: even if the PID quality of the cuts for the dE/dx response determination is very different than that for the yield extraction, the systematic error bars still cover the resulting spread in such a situation. As a result, it is legitimate to use the cut combination described in Section 6.1 for this analysis because these cuts are expected to even have approximately the same PID quality.

Change of Efficiency and Secondary Contamination

For the cut variation, global tracks are considered besides modified and default hybrid tracks. Global tracks are tracks which have SPD hits and an ITS refit. They are a subset of the modified hybrid tracks, which are in turn a subset of the default hybrid tracks. 4 different sets of global tracks have been evaluated:

6. Identified Particle Production in pp Collisions in ALICE

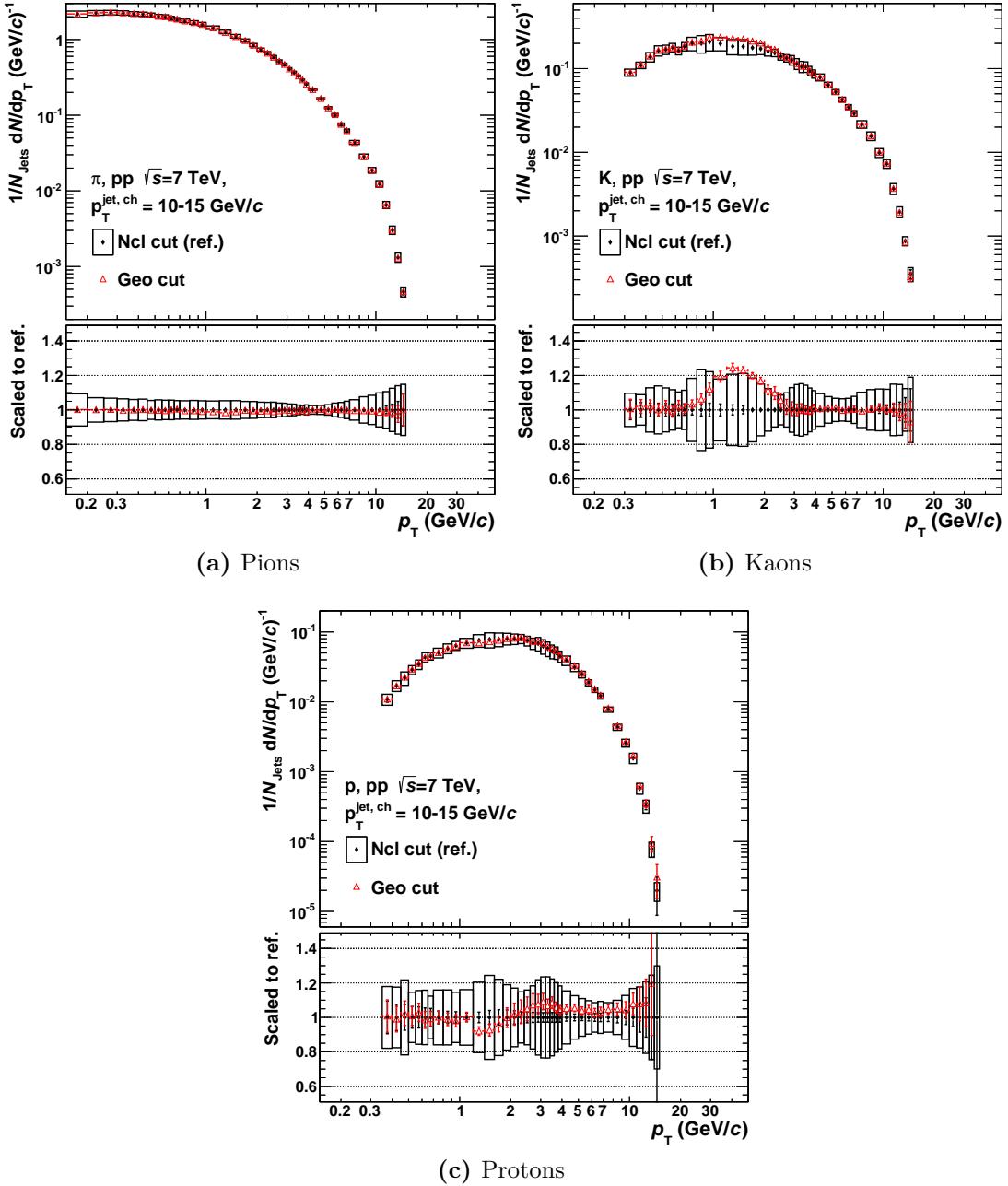


Figure 6.17.: Comparison of the corrected particle yields for the TPC geometrical cut (red open triangles) and the ncl cut (black diamonds) versus p_T^{track} for $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ in data. Each set of cuts is corrected with its own MC correction factors. The boxes correspond to the systematic errors for the ncl cut, including the uncertainties of the corresponding MC correction factors.

- tighter cut on the Distance of Closest Approach (DCA),
- looser cut on the DCA,
- reduction of the demanded number of tracking clusters in the TPC, N_{cl} , and
- replacing the N_{cl} requirement by a cut on the number of crossed TPC pad rows.

The change is to be understood relative to the default value that is used for the global tracks subset of the modified hybrid tracks. Details about the global track cut variations can be found in Appendix C.5.

The ratio of the cut variations to the modified hybrid cuts is shown for pions, kaons and protons in jets in Figs. 6.18, 6.19 and 6.20, respectively. The systematic errors, which include both fitting and MC correction uncertainties, typically cover the observed spread of the results. Obviously, particles without SPD hit(s) cause some bias, which is obvious from the difference between global and modified hybrid tracks. On top of this, tracks without ITS refit cause a bias between modified and default hybrid tracks. It is important to note that the trends for $p_T > 2 - 3 \text{ GeV}/c$ go into the same direction for all species. This implies that the trend is not due to PID effects, but is related to the total yield. Indeed, it was checked that repeating the same procedure for the total yields, without any PID fitting, produces results similar to those of the pions. The change of the total yield can be traced to the DCA distributions, which are slightly different in MC than in data. The largest deviations are observed for particles without ITS refit [124]. As a consequence, the efficiency and secondary contamination are biased in MC. The cuts may also influence to some extent the PID quality and, hence, the fitting. This is taken into account in the systematic error bands.

Similar plots as a function of z and for the inclusive case are collected in Appendix I. In general, good agreement is observed. The only mentionable deviation is visible for protons in the inclusive case. The yield for the default hybrid cuts is slightly above the error band for $p_T \approx 5 - 10 \text{ GeV}/c$. Most likely this is again related to the difference of the DCA distributions in MC and data, with largest impact on the default hybrid cuts. Hence, the latter have additional systematic uncertainties, which should also enter in the ratio to possibly cover the remaining deviation.

It should be noted at this point that the results for modified hybrid cuts in $|\eta| < 0.9$ agree within uncertainties with those from the standard track cuts in $|\eta| < 0.8$. The latter results are very similar to the global track cut variations and are not shown for this reason. Nevertheless, to compare apples to apples, the inclusive reference for the jet analysis is obtained with the same cuts (and range in η) as the jet results.

6. Identified Particle Production in pp Collisions in ALICE

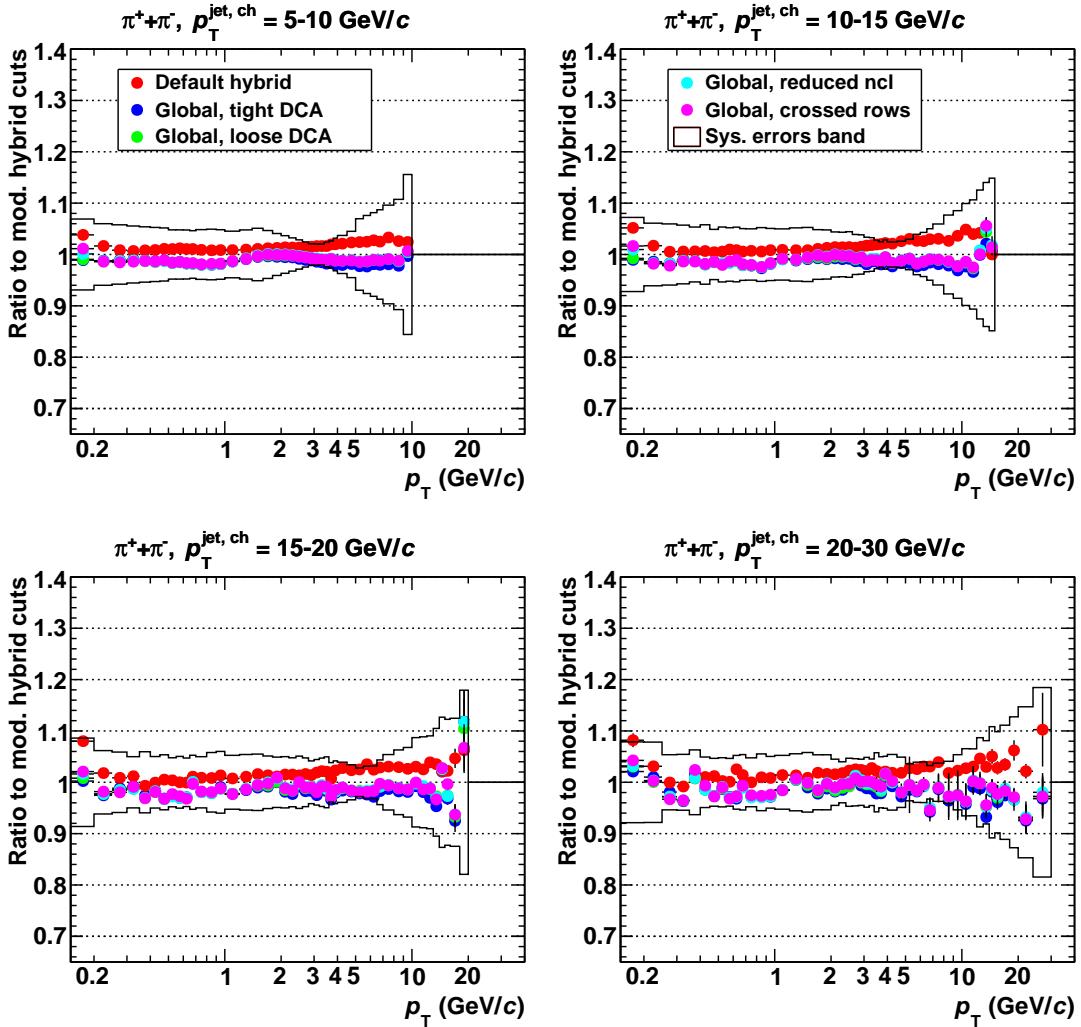


Figure 6.18.: Cut variations for pions in jets. The ratio to the modified hybrid track cuts is shown. Each set of cuts is corrected with its own MC correction factors. The systematic error band shows the total uncertainty from fit and MC corrections.

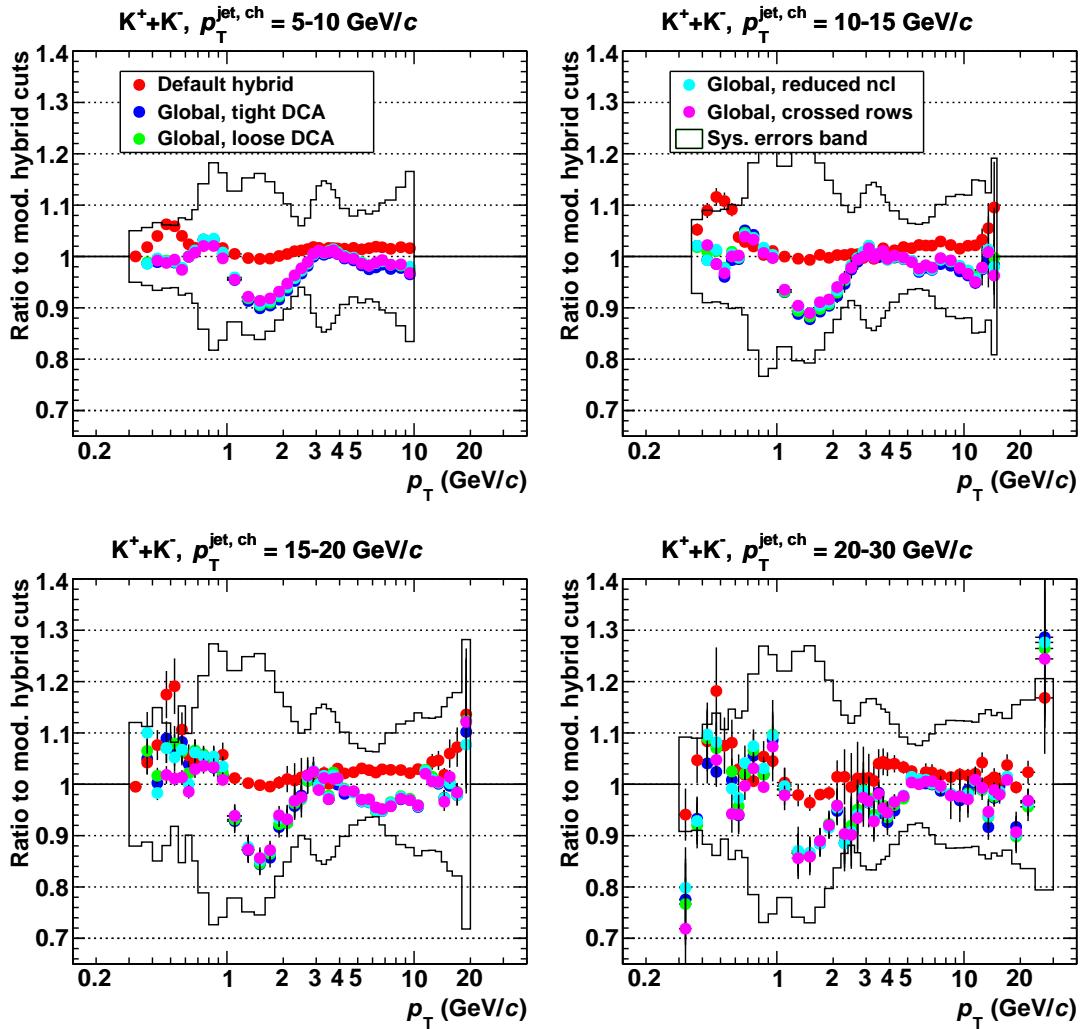


Figure 6.19.: Cut variations for kaons in jets. The ratio to the modified hybrid track cuts is shown. Each set of cuts is corrected with its own MC correction factors. The systematic error band shows the total uncertainty from fit and MC corrections.

6. Identified Particle Production in pp Collisions in ALICE

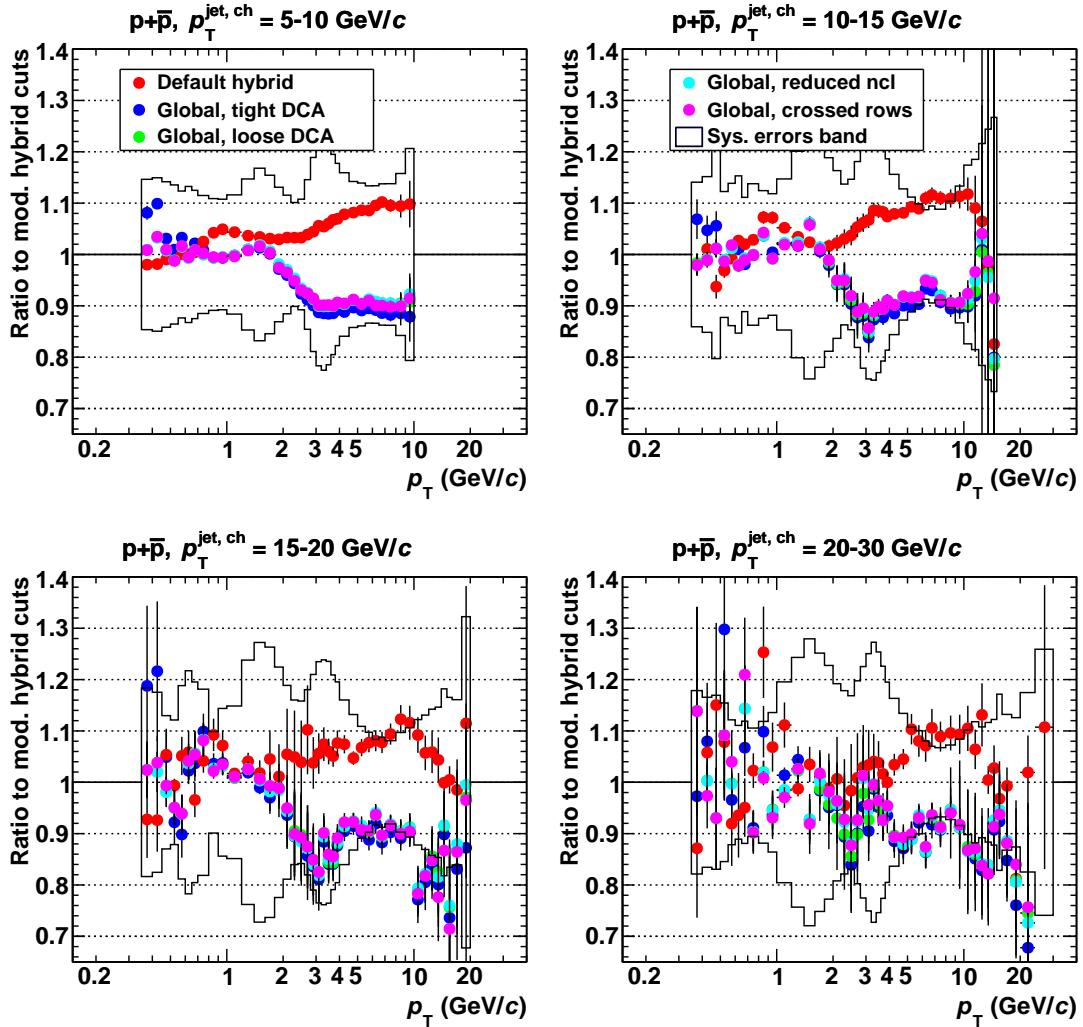


Figure 6.20.: Cut variations for protons in jets. The ratio to the modified hybrid track cuts is shown. Each set of cuts is corrected with its own MC correction factors. The systematic error band shows the total uncertainty from fit and MC corrections.

As a final remark, it should be mentioned that the above check was done with the minimum bias MC production (LHC10f6a) for historical reasons. It was checked that the conclusion of this study is the same for the new jet MC production (LHC14b6).

6.4. Inclusive Minimum Bias Particle Production

To (further) validate the analysis, the minimum bias results are compared to published and preliminary inclusive particle yields from other analysis. For $p_T < 3 - 6 \text{ GeV}/c$, results are available from a combined PID analysis [125], which identifies particles using ITS, TPC, TPC+TOF or HMPID or via kink topologies. The high- p_T region with $p_T > 3 \text{ GeV}/c$ was originally analysed by a group from the Lund University (Sweden) identifying particles with the TPC [126]. The latter method is different from the one presented in this thesis. In particular, it is not based on templates, but purely relies on parametrisations and extracts the particle yields via a multi-Gauss fit of the TPC dE/dx (also see [106]).

The minimum bias yields obtained in this work must be normalised accordingly. In particular, the number of inelastic collision events, $N_{\text{evt}}(\text{INEL})$, occurs in the normalisation and requires some care due to the so-called “bin-0” issue: event cuts can change the shape and normalisation of the track p_T spectra, which is particle type independent. This happens due to events that trigger, but have no reconstructed vertex and are, thus, rejected by the event cuts.

An MC-based approach is used to correct for this effect [112]. The inclusive p_T spectra of charged primaries at the particle level normalised to $1/N_{\text{evt}} \cdot dN/dp_T$ are evaluated after Minimum Bias (MB) trigger selection (“MB”, $N_{\text{evt}} = N_{\text{evt}|\text{MB}}$), after additional vertex cut (“vtx”, $N_{\text{evt}} = N_{\text{evt}|(\text{MB} \& \text{Vtx})}$) and after additional Z_{vtx} cut and pile-up rejection⁷ (“z vtx”, $N_{\text{evt}} = N_{\text{evt}|(\text{MB} \& \text{Vtx} \& Z_{\text{vtx}})}$) (see Section 6.1 for the details of these cuts). The spectra ratios for these sets are shown in Fig. 6.21. The p_T dependence of the ratios is at the per mille level and negligible. While the Z_{vtx} cut does not change the spectra, the vertex cut alters the normalisation by about 11%.

The Table 6.4 shows that the ratio of the normalised spectra before and after the vertex cut is on (p_T -)average the same as the ratio of the number of events. For completeness, the corresponding values for the Z_{vtx} cut are given in Tab. 6.5. As can be seen, the spectra ratio scales with the number of events in that case.

⁷The considered MC production does not include pile-up. Since the fraction of events wrongly tagged as pile-up is negligible, the pile-up rejection has no effect in this consideration. The presented method will be applicable to the MB analysis, where the pile-up fraction is negligible. However, it will not be applicable to the multiplicity analysis, where pile-up may contribute significantly and needs to be subtracted.

6. Identified Particle Production in pp Collisions in ALICE

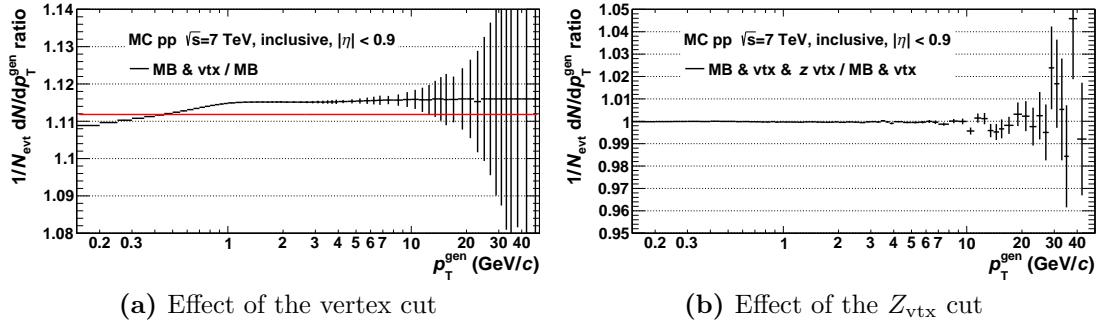


Figure 6.21.: Impact of vertex cuts on the inclusive charged primary p_T spectra in MC. The $1/N_{\text{evt}} \cdot dN/dp_T$ spectra at the particle level are compared for **a)** MB trigger selection (“MB”) before and after an additional vertex cut (“vtx”) and **b)** for MB selection with vertex cut before and after an additional Z_{vtx} cut (“ z vtx”). The solid red line in **a)** is the fit of a constant to the ratio.

Table 6.4.: Comparison of the normalised p_T spectra ratios with the N_{evt} ratios for the vertex cut. The former value is obtained from a fit of a constant to the ratio.

$R_{\text{spec}} \equiv \frac{(1/N_{\text{evt}} \cdot dN/dp_T) _{(MB \& Vtx)}}{(1/N_{\text{evt}} \cdot dN/dp_T) _{MB}}$	$R_{\text{evt}} \equiv \frac{N_{\text{evt} MB}}{N_{\text{evt} (MB \& Vtx)}}$	$\frac{R_{\text{evt}}}{R_{\text{spec}}} - 1$
1.1118	1.1160	0.38%

Table 6.5.: Comparison of the normalised p_T spectra ratios with the N_{evt} ratios for the Z_{vtx} cut. The former value is obtained from a fit of a constant to the ratio.

$R_{\text{spec}} \equiv \frac{(1/N_{\text{evt}} \cdot dN/dp_T) _{(MB \& Vtx \& Z_{\text{vtx}})}}{(1/N_{\text{evt}} \cdot dN/dp_T) _{(MB \& Vtx)}}$	$R_{\text{evt}} \equiv \frac{N_{\text{evt} (MB \& Vtx)}}{N_{\text{evt} (MB \& Vtx \& Z_{\text{vtx}})}}$
0.9998	1.1144

It has been checked that varying the pseudo-rapidity (η) range yields consistent results for both cuts at the per mille level.

In conclusion, the “bin-0” issue can be corrected by simply scaling the normalised spectra with the ratio of the number of events after and before vertex cut, $\frac{N_{\text{evt}|(MB \& Vtx)}}{N_{\text{evt}|MB}}$. Note that the corresponding event numbers should be taken from the considered data sample, i.e. not from MC in case of data.

Taking into account all corrections, the normalised MB (charged) yield of species s reads:

$$\frac{1}{N_{\text{evt}}(\text{INEL})} \frac{1}{2\pi p_T} \frac{d^2 N_s}{dy dp_T} = \frac{1}{N_{\text{evt}|(\text{MB \& Vtx \& } Z_{\text{vtx}})}} \frac{N_{\text{evt}|(\text{MB \& Vtx})}}{N_{\text{evt}|\text{MB}}} \varepsilon_{\text{trig}} \\ \cdot R_s^{\text{G-F}}(p_T) \frac{f_{\text{prim},s}(p_T) f_{\pi-\mu,s}(p_T)}{\varepsilon_s(p_T)} , \quad (6.1) \\ \cdot \frac{1}{2\pi p_T} \frac{\Delta N(p_T)}{\Delta y_s(p_T, \eta) \Delta p_T} A_s(p_T)$$

with the following definitions:

- $N_{\text{evt}}(\text{INEL})$ is the number of inelastic collision events,
- $N_{\text{evt}|\text{MB}}$, $N_{\text{evt}|(\text{MB \& Vtx})}$ and $N_{\text{evt}|(\text{MB \& Vtx \& } Z_{\text{vtx}})}$ are the number of events after Minimum Bias selection (MB), after MB and vertex cut, and after MB, vertex cut, Z_{vtx} cut and pile-up rejection (cf. Section 6.1), respectively, with $N_{\text{evt}|(\text{MB \& Vtx})}/N_{\text{evt}|\text{MB}} \approx 0.9009$,
- $\varepsilon_{\text{trig}}$ is the minimum bias trigger efficiency estimated as 0.852 for $\sqrt{s} = 7 \text{ TeV}$ pp collisions [127],
- $R^{\text{G-F}}(p_T)$ is the GEANT-FLUKA correction factor (see Eq. 5.5),
- $f_{\text{prim},s}(p_T)$ is the fraction of primary particles, including strangeness scaling (cf. Section 5.1.2),
- $f_{\pi-\mu,s}(p_T)$ is the pion fraction in the “muon + pion” sample, including strangeness scaling (cf. Section 5.1.3),
- $\varepsilon_s(p_T)$ is the tracking efficiency, including the p_T resolution correction (cf. Section 5.1.1),
- $\Delta N(p_T)$ is the total number of charged tracks of all species for the considered sample in the corresponding p_T bin with width Δp_T ,
- $\Delta y_s(p_T, \eta)$ is the rapidity range calculated from the pseudo-rapidity range via Eq. 6.3 as discussed below and
- $A_s(p_T)$ is the particle fraction of species s in the considered sample.

Note that from now on $N_{\text{evt}|(\text{MB \& Vtx \& } Z_{\text{vtx}})}$ implicitly includes the pile-up rejection, i.e. events tagged as pile-up are subtracted. The factor in the first line of the right-hand side of Eq. 6.1 normalises the spectrum to $N_{\text{evt}}(\text{INEL})$ and takes into account the “bin-0” correction factor.

Strictly speaking, the conversion from $dN/d\eta$ to dN/dy normally requires the Jacobian⁸ $dy/d\eta = p/E$, where p is the total momentum and E the energy of a

⁸It is discussed in [128] whether the Jacobian exhibits some additional factors.

6. Identified Particle Production in pp Collisions in ALICE

particle. However, the particle identification is not done track-by-track, but on a statistical basis, such that the energy is unknown. Also, p/E is not constant in a given p_T and η bin. Hence, the results are just normalised to rapidity, but obtained for a fixed η cut. The pseudo-rapidity η has already been defined in Eq. 3.11. The rapidity y is defined as:

$$y \equiv 0.5 \ln \frac{E + p_z}{E - p_z}, \quad (6.2)$$

where E is the total energy of the particle and p_z its momentum along the z -direction. In the high-energy limit, η and y are identical. For a particle from species s with mass m_s and transverse momentum p_T , Eqs. 3.11 and 6.2 lead to:

$$y_s(p_T, \eta) = 0.5 \ln \frac{\sqrt{m_s^2 + p_T^2} \cosh^2 \eta + p_T \sinh \eta}{\sqrt{m_s^2 + p_T^2} \cosh^2 \eta - p_T \sinh \eta}. \quad (6.3)$$

The rapidity range the results are normalised to is then defined as:

$$\Delta y_s(p_T, \eta) \equiv y_s(p_T, \eta_2) - y_s(p_T, \eta_1), \quad (6.4)$$

with $\eta \equiv \eta_2 = -\eta_1 = 0.8$ in this case. In Fig. 6.22, the normalisation factors for pseudo-rapidity (taken to be 1.6) and rapidity are compared. They differ most at low p_T and for heavy species. For protons at $p_T = 0.15 \text{ GeV}/c$, the factors differ by about a factor 5.5. The difference strongly decreases with rising p_T and becomes negligible around $p_T = 6 \text{ GeV}/c$ for all species.

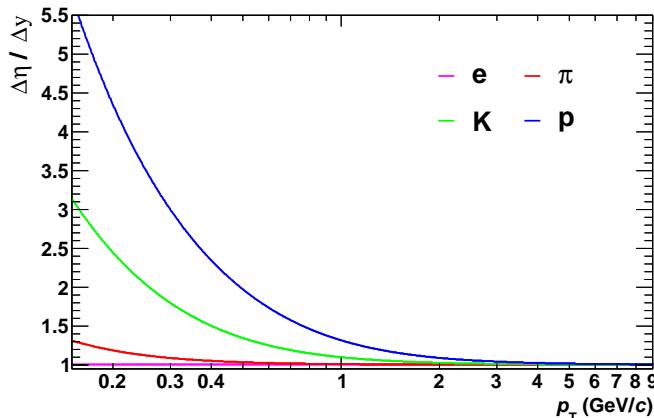


Figure 6.22.: Comparison of the normalisation factors for rapidity (Δy) and pseudo-rapidity ($\Delta\eta = 2 \cdot 0.8 = 1.6$). The ratio $\Delta\eta/\Delta y$ is the conversion factor to change the normalisation of the yields from pseudo-rapidity to rapidity. In the relevant p_T range, $p_T > 0.15 \text{ GeV}/c$, the conversion factor is unity for electrons, smaller than 1.4 for pions and can go up to 5.5 for protons. It drops rapidly with increasing p_T and approaches unity for all species around $p_T = 6 \text{ GeV}/c$.

6.4.1. Comparison of Results to Other Analyses

The corrected inclusive yields of negatively charged particles are shown and compared to the results of the combined PID and the Lund analysis in Fig. 6.23. The comparison of positively charged particles yields similar observations and can be found in Appendix H.2. The overall agreement is good, all three analyses agree within uncertainties in the whole p_T range. The pion results typically vary by less than 5%, which is smaller than the systematic uncertainties. Only the first p_T bin is off by almost 20%. This is just an artefact from the calculation of the result “ratio”. Since the binning is different, interpolation is used to obtain the value for the denominator. If the bin centre of a bin in the numerator is smaller (larger) than the centre of the first (last) bin in the denominator, no extrapolation is performed, but the value of the first (last) available bin is taken.

The proton yields agree with better than 10% with the combined PID results, whereas the deviations to the Lund results are typically smaller than 15% (and the errors are considerably large). The situation seems to be inverted for the kaons. Here, the agreement with Lund is around 5%, but the combined PID results show deviations by up to 15% around $p_T = 1 - 2 \text{ GeV}/c$. This coincides with the dE/dx crossing region. Although the deviation is still covered by the systematic uncertainties, it corresponds to the largest tension between the results. It is clear that the TPC only analysis in this work has larger uncertainties in the dE/dx crossing region than the combined PID analysis, which uses e.g. TOF information in addition. Deviations are, thus, expected and are indeed described by the larger systematic uncertainty.

It should be noted that the MB results presented in this section have a rather fine binning at high p_T , which coincides with that of the Lund analysis. For the multiplicity-binned analysis, the statistics does not suffice for such a fine high- p_T binning. To have equal binning for multiplicity bins and MB, the MB yields will be extracted again with correspondingly rougher binning.

In conclusion, the observed consistency of the different analysis results are reassuring. In particular, it is worth noting that the three analyses were run on different data sets. Moreover, the combined PID analysis is directly done in $|y| < 0.5$, i.e. no conversion from pseudo-rapidity to rapidity is necessary. Also, that analysis corrects for the secondary contamination by fitting DCA MC templates to the measured distribution, which does not require a scaling of the strangeness content in MC. Eventually, this further verifies the methods presented in this thesis.

6. Identified Particle Production in pp Collisions in ALICE

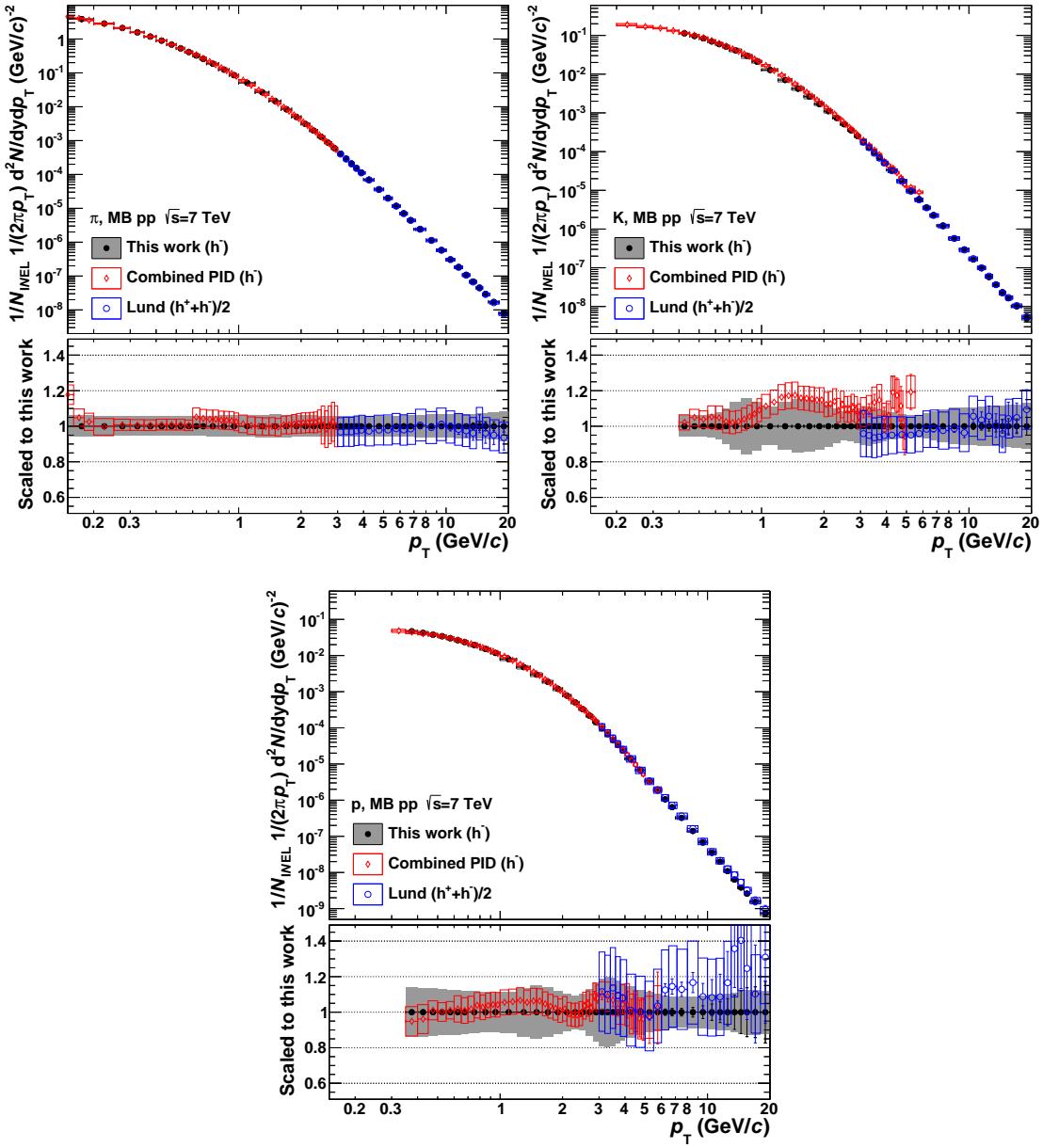


Figure 6.23.: The corrected MB yields of negatively charged particles are compared to the combined PID [125] and the Lund analysis [126]. The error bars indicate the statistical and the boxes the systematic uncertainties. Note that the Lund result (blue) shows the yield of both charges divided by two.

6.5. Inclusive Particle Production in Multiplicity Bins

To study possible similarities between high-multiplicity pp collisions and p–Pb and Pb–Pb collisions, the analysis is carried out in various bins of multiplicity in minimum bias pp events.

6.5.1. Multiplicity Estimators

The multiplicity is a measure for the hadronic activity of an event. Various choices for multiplicity estimators exist. This work focuses on two of them which have been proposed for such studies by the collaboration: the reference multiplicity and the V0M multiplicity estimators. Comparing the results obtained with those estimators may help to better understand the biases of each type of multiplicity selection.

The *reference multiplicity* estimator (RM) counts tracks and tracklets at mid-rapidity. It was proposed to provide a good correlation between estimated and true multiplicity, as well as to have the least bias among mid-rapidity estimators. It counts the number of global ITS-TPC tracks within $|\eta| < 0.8$, adds ITS standalone tracks not associated to global tracks and finally adds ITS tracklets that are neither associated to ITS standalone tracks, nor to global tracks. The η window was chosen as wide as possible while retaining full TPC coverage to get the best handle on the hadronic activity.

The *V0M multiplicity* estimator (V0M) makes use of the V0 detectors at forward-rapidity and is similar to what is done for the analysis of Pb–Pb collisions. The sum of the signal amplitudes from V0A and V0C is utilised to calculate multiplicity percentiles.

The analysis is carried out at mid-rapidity. While the analysed tracks contribute to the RM estimate, they are separated by a gap in η from any event activity contributing to the V0M estimate. Such correlations have to be taken into account for the interpretation of the results.

To get a feeling about the available statistics for the analysis, the number of events after all selections is shown in Fig. 6.24a as a function of multiplicity for both estimators. The analysis needs about 10^4 events to yield spectra with reasonably small uncertainties. This requirement is matched for the presented binning for V0M. For RM, the threshold is reached around multiplicity 70. The distributions of the number of events per multiplicity (percentile) are shown in Fig. 6.24b. The distribution is largely flat for the V0M percentile. It is not perfectly flat, since the quantiling of the V0 amplitudes is based on the minimum bias sample after pile-up rejection, but without applying vertex cuts [123].

6. Identified Particle Production in pp Collisions in ALICE

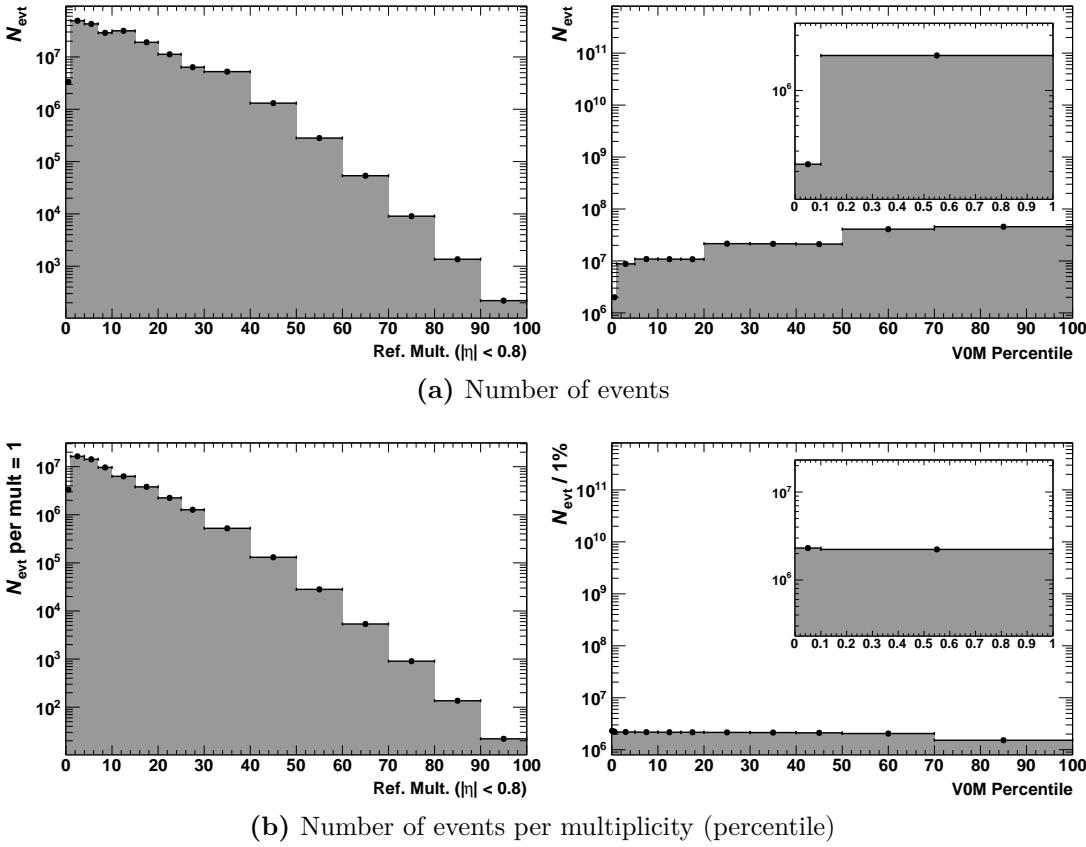


Figure 6.24.: Number of events after all selections as a function of the reference multiplicity (left) and the VOM percentile (right). Note that no normalisation to the bin width is applied in (a), which causes jumps in the histograms. After normalising to the bin width (b), the distribution for the VOM percentile is largely flat. Also note that the horizontal axis in the left panel shows the multiplicity and not a percentile, i.e. low percentiles correspond to high multiplicities.

6.5.2. Quality Checks

The comparison of the spectra in different multiplicity bins can only be meaningful as long as the utilised dE/dx parametrisations are valid for all multiplicity bins. To check, whether the mean dE/dx varies with multiplicity, V0 pions⁹ as a clean sample have been selected from events in different multiplicity classes. In Fig. 6.25, their Δ' as a function of momentum is shown for various multiplicity classes for both the V0M and the reference multiplicity estimator. Obviously, there is no systematic trend when changing the multiplicity. In particular, the data points are typically within the estimated intrinsic spline uncertainty of $\pm 0.2\%$ around unity. The same observation holds true for V0 protons and V0

⁹It is important to note that “V0” relates to the track topology in this case and has nothing to do with the V0 detector/multiplicity estimator.

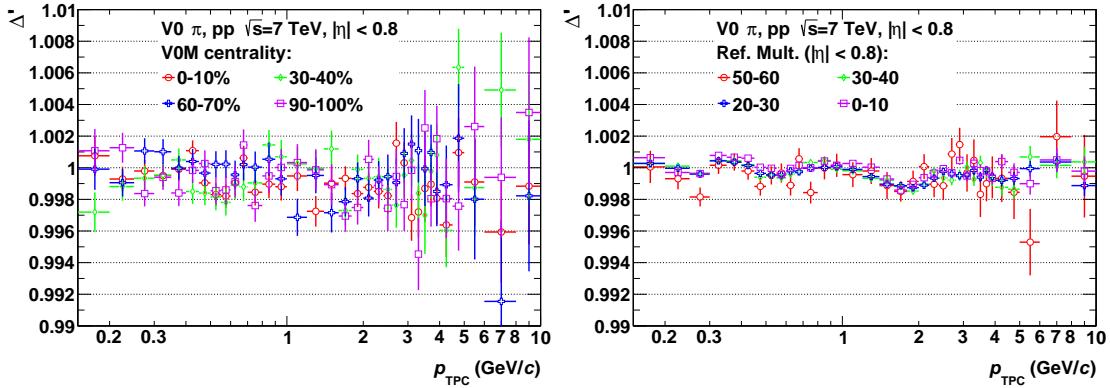


Figure 6.25.: Multiplicity dependence of the mean dE/dx for V0 pions. The left and right panels compare the Δ' in different bins of V0M multiplicity and reference multiplicity, respectively.

electrons with considerably larger statistical uncertainties.

As a result, there is no indication for a multiplicity dependence of the mean dE/dx in pp collisions in the considered multiplicity bins. The intrinsic spline uncertainty is sufficient to account for possible deviations of Δ' from unity in the individual multiplicity classes. Such a check needs to be re-done in case triggered samples would allow to investigate much higher multiplicities, e.g. a reference multiplicity around 500 or higher.

This observation is supported by the comparison of the fit quality in different multiplicity bins. An example is shown in Fig. 6.26 for RM for tracks with $p_T = 4 - 4.5$ GeV/c. This is a typical example for a momentum bin in the region of the dE/dx relativistic rise. Evidently, the fit describes the data similarly well in all multiplicity bins, although the statistics is different by about a factor 30 between some of these bins. Similar observations hold true for V0M and other p_T ranges. The corresponding plots are collected in Appendix F.1.

It should be mentioned at this point that cross-checks to three other multiplicity-binned analyses have been carried out. One of these analyses uses the ITS dE/dx for PID and extracts the particle spectra in the low- p_T region ($p_T \lesssim 0.6$ GeV/c). Good agreement with the MTF within systematic uncertainties is observed in the overlap region [129]. A similarly good agreement is found for the comparison with two independent (and technically different) analyses using TPC and TOF for PID up to $p_T \approx 2 - 3$ GeV/c [130, 131]. It is planned to combine the results of all analyses to obtain particle spectra with small uncertainties in the full p_T range.

6. Identified Particle Production in pp Collisions in ALICE

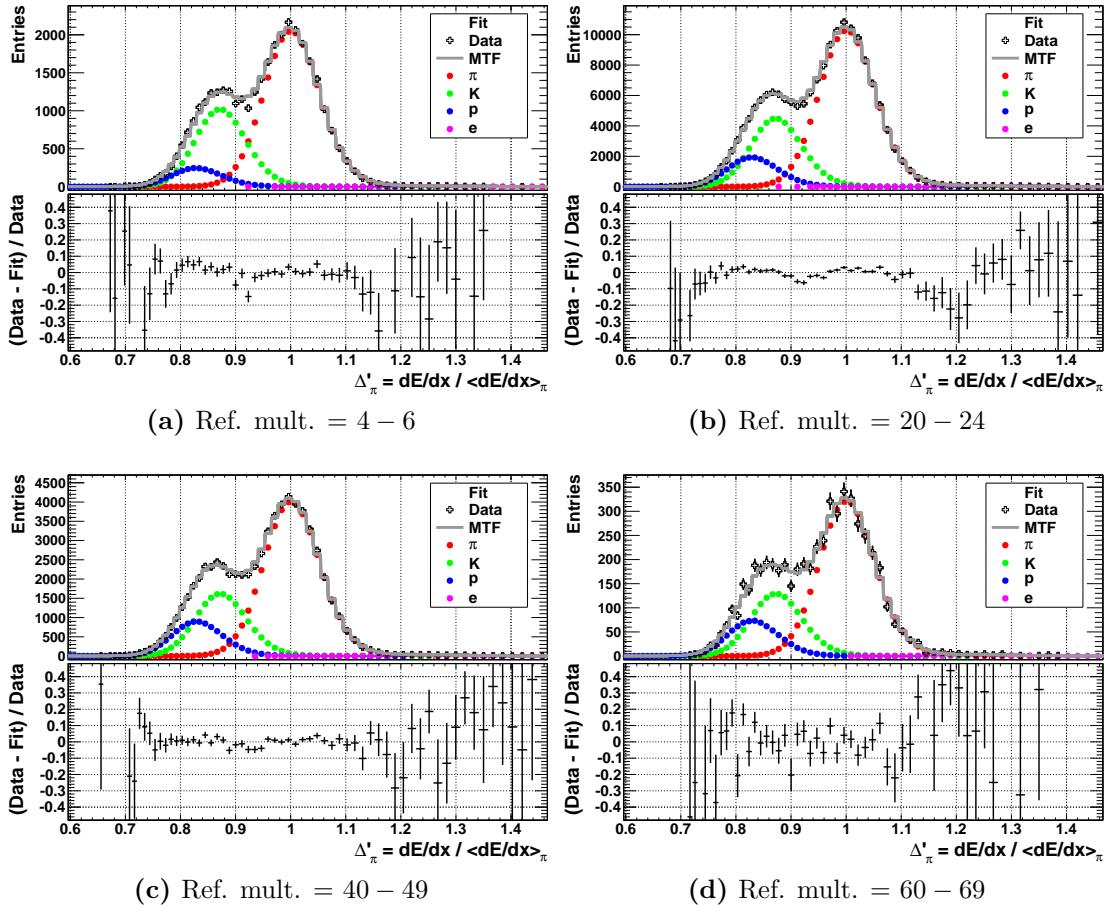


Figure 6.26.: Template fits to data for different reference multiplicity bins for $p_T = 4 - 4.5 \text{ GeV}/c$. The reference multiplicity increases from the upper left to the lower right plot.

6.5.3. Results

The normalised (charged) yield for a given multiplicity class is defined as:

$$\frac{1}{N_{\text{evt}}} \frac{1}{2\pi p_{\text{T}}} \frac{d^2 N_s}{dy dp_{\text{T}}} = \frac{1}{N_{\text{evt}}} \frac{1}{2\pi p_{\text{T}}} \frac{\Delta N(p_{\text{T}})}{\Delta y_s(p_{\text{T}}, \eta = 0.8) \Delta p_{\text{T}}} A_s(p_{\text{T}}) \\ \cdot R_s^{\text{G-F}}(p_{\text{T}}) \frac{f_{\text{prim},s}(p_{\text{T}}) f_{\pi-\mu,s}(p_{\text{T}})}{\varepsilon_s(p_{\text{T}})}, \quad (6.5)$$

where the abbreviation $N_{\text{evt}} \equiv N_{\text{evt}|(\text{MB} \& \text{Vtx} \& Z_{\text{vtx}})}$ was used and all other quantities are defined similarly as in Eq. 6.1, but here have an implicit dependence on the considered multiplicity class. In general, this class can be different for the MC corrections and the fitted quantities (cf. Section 5.2.5). It turns out that the “bin-0” correction cannot be applied as straightforward as for the minimum bias case. A proper treatment is still being discussed in the collaboration.

For this reason, the normalisation of the yields in multiplicity bins is changed in this thesis from the number of inelastic collisions $N_{\text{evt}}(\text{INEL})$ — as used for the minimum bias results — to the number of events after all cuts and selections $N_{\text{evt}|\{\text{MB} \& \text{Vtx} \& Z_{\text{vtx}}\}}$. The different normalisation should be kept in mind, when the minimum bias results and those for multiplicity bins are compared. Though, it is only an overall normalisation factor $N_{\text{evt}|\{\text{MB} \& \text{Vtx} \& Z_{\text{vtx}}\}}/N_{\text{evt}}(\text{INEL})$, which includes the “bin-0” correction and the minimum bias trigger efficiency. It does not affect the cancellation of correlated errors in the ratio of multiplicity bin and minimum bias results. Furthermore, the goal of this thesis is to compare the multiplicity evolution in pp with that in p–Pb and Pb–Pb collisions. The “bin-0” effects should cancel to a good extent for the ratios between multiplicity bin and minimum bias and anyway drop out for the to-pion ratios. In that sense the main results of this work are not harmed by such a normalisation issue.

As for the yields in Eq. 6.5, the to-pion ratio for a given multiplicity class also takes into account the conversion from pseudo-rapidity to rapidity. It reads for species s (different from π):

$$\begin{aligned} & \left(\frac{1}{N_{\text{evt}}} \frac{1}{2\pi p_{\text{T}}} \frac{d^2 N_s}{dydp_{\text{T}}} \right) / \left(\frac{1}{N_{\text{evt}}} \frac{1}{2\pi p_{\text{T}}} \frac{d^2 N_{\pi}}{dydp_{\text{T}}} \right) = \frac{d^2 N_s}{dydp_{\text{T}}} / \frac{d^2 N_{\pi}}{dydp_{\text{T}}} \\ & = \frac{A_s(p_{\text{T}})}{A_{\pi}(p_{\text{T}})} \frac{\Delta y_{\pi}(p_{\text{T}}, \eta = 0.8)}{\Delta y_s(p_{\text{T}}, \eta = 0.8)} R_s^{\text{G-F}}(p_{\text{T}}) \frac{f_{\text{prim},s}(p_{\text{T}})}{f_{\text{prim},\pi}(p_{\text{T}}) f_{\pi-\mu,\pi}(p_{\text{T}})} \frac{\varepsilon_{\pi}(p_{\text{T}})}{\varepsilon_s(p_{\text{T}})}, \end{aligned} \quad (6.6)$$

where the covariance matrix is used for the error calculation of the ratio $\frac{A_s(p_{\text{T}})}{A_{\pi}(p_{\text{T}})}$. Again, the variables have an implicit dependence on the considered multiplicity class.

The fully corrected spectra in bins of V0M and RM are shown in Figs. 6.27 and 6.28 for pions, kaons and protons, respectively. As visible from the ratio to MB, the particle yields increase by several orders of magnitude with multiplicity. Note that the lower panel shows a real ratio taking into account the correlation of errors between multiplicity bin and MB (cf. Section 6.3.7). The different spread at low and high p_{T} indicates that the spectra become harder with multiplicity. The results are roughly similar for both estimators. A closer look reveals that the spectra rise by about one order of magnitude at low p_{T} from the lowest to the highest multiplicity bin for both estimators. At high p_{T} , however, the increase is up to four orders of magnitude for RM, but only about 2 – 3 orders for V0M. This would indicate that the spectra hardening is more pronounced for RM. Yet, it should be noted that there is no one-to-one correspondence between the bins of the two multiplicity estimators.

The species dependence of the multiplicity evolution can be studied by looking at the corresponding evolutions of the to-pion ratios. The K/π ratio is shown in Fig. 6.29. Evidently, there is basically no dependence on multiplicity for V0M, whereas a significant change of the ratios is observed versus RM. In Fig. 6.30, the p/π ratios are shown. Both estimators show similar trends. The peak of the distribution increases in height with multiplicity and is shifted towards higher p_{T} .

6. Identified Particle Production in pp Collisions in ALICE

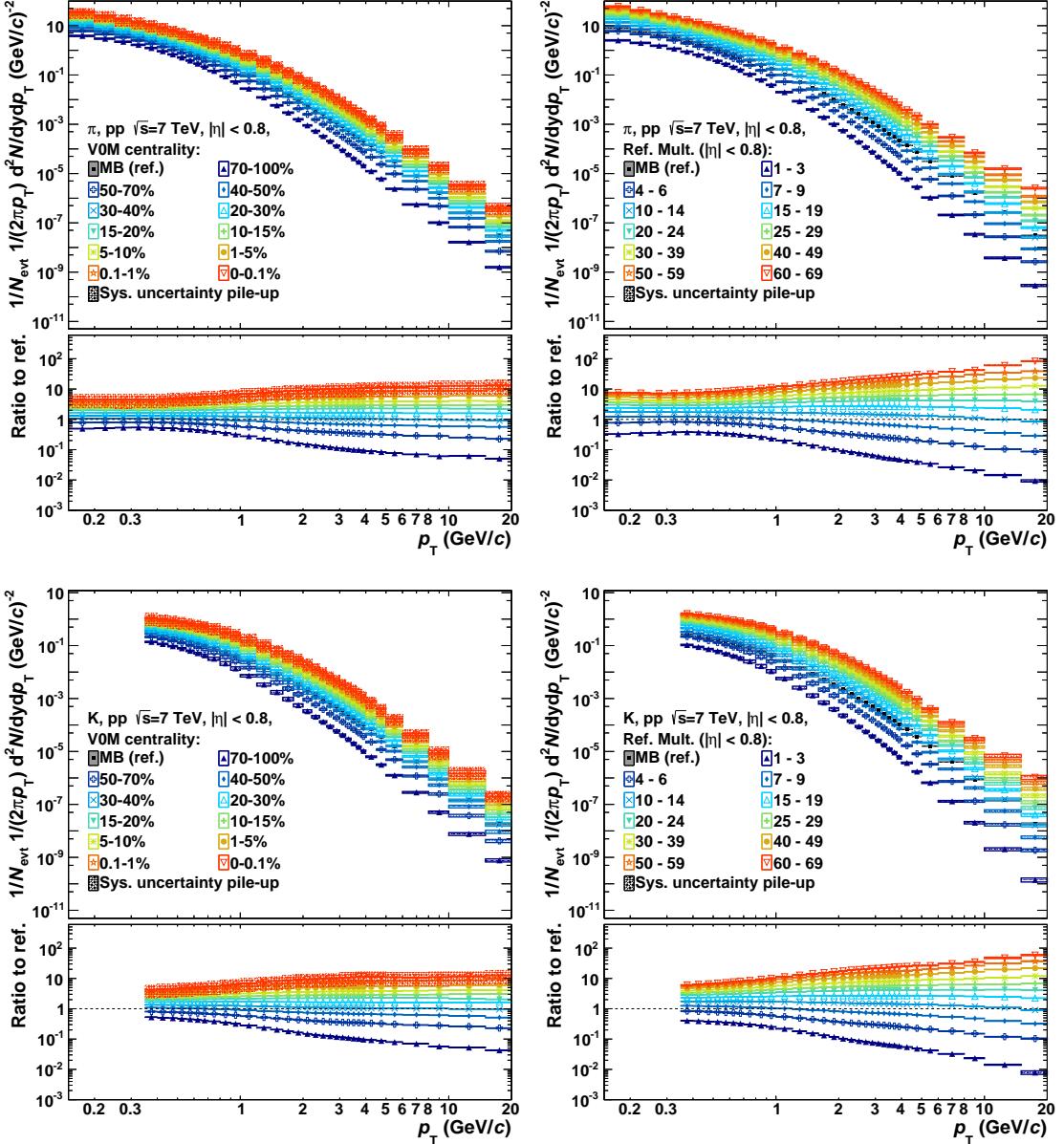


Figure 6.27.: π (top) and K (bottom) spectra for different bins of V0M (left) and reference multiplicity (right). The lower panels show the ratio to MB. Note that the MB data points are normalised to $N_{\text{evt}}(\text{INEL})$ instead of $N_{\text{evt}}|(\text{MB} \& \text{Vtx} \& Z_{\text{vtx}})$.

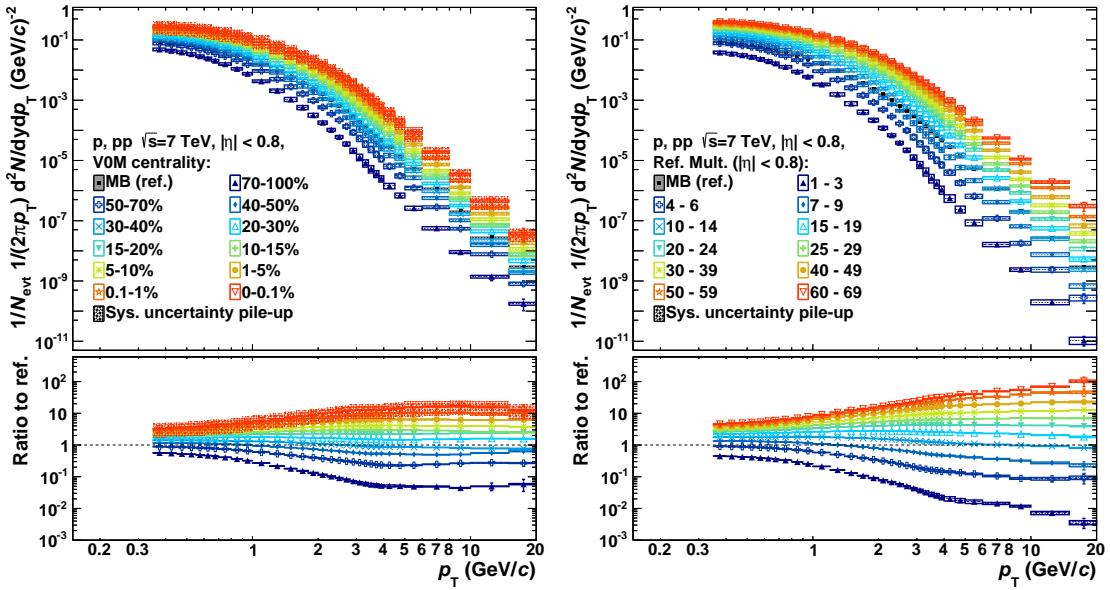


Figure 6.28.: p_T spectra for different bins of V0M (left) and reference multiplicity (right). The lower panels show the ratio to MB. Note that the MB data points are normalised to $N_{\text{evt}}(\text{INEL})$ instead of $N_{\text{evt}}|(\text{MB} \& \text{Vtx} \& Z_{\text{vtx}})$.

Furthermore, there is a crossing point in p_T at which all multiplicity bins have the same or at least a very similar to-pion ratio. Obviously, this point is at higher p_T for RM ($p_T \approx 2.3 \text{ GeV}/c$) than for V0M ($p_T \approx 1.7 \text{ GeV}/c$). For both estimators, the p/π ratio of the bin with lowest multiplicity looks shifted downwards by about 10% relative to MB with respect to the general multiplicity evolution.

It was observed in Section 6.5.2 that the dE/dx response is rather independent of the multiplicity. Indeed, this manifests itself in the almost complete cancellation of the systematic errors of the ratios to MB.

6.5.4. Charge Dependence

The analysis is run separately on positively and negatively charged particles. This allows to check whether there is an asymmetry between the particle and anti-particle production. Earlier studies [117, 132] of minimum bias pp collisions with varying beam energy showed that the mid-rapidity production of e.g. protons and anti-protons becomes equal within uncertainties for $\sqrt{s} = 7 \text{ TeV}$, whereas anti-protons are less abundant by a few per cent at lower energies. The similar cross-section for particles and anti-particles originates from the fact that the production of single-inclusive hadrons at mid-rapidity is driven by gluon-induced processes and fragmentation, cf. discussion in [33].

In this section, the particle and anti-particle production at low and high multiplicities is compared. Due to limited statistics, the most extreme multiplic-

6. Identified Particle Production in pp Collisions in ALICE

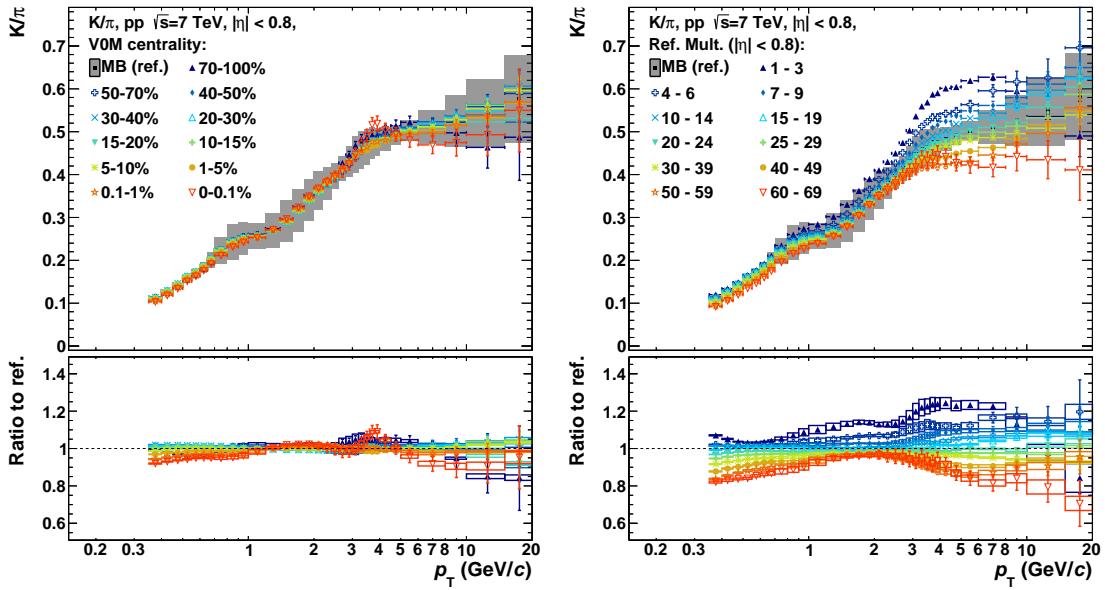


Figure 6.29.: K/π ratios for different bins of V0M (left) and reference multiplicity (right). For better visibility, only the systematic uncertainties of the MB data points are plotted in the upper panel. The lower panel shows the ratio to MB including the corresponding systematic uncertainties.

ity bins cannot be used. It should be stressed at this point that instead of a particle/anti-particle ratio only the “scaled to reference” (see Appendix A) result will be presented. In the end, this is already sufficient to judge possible trends as a function of multiplicity, such that a detailed evaluation of uncertainty correlations is dispensable. Although the statistical uncertainties of positively and negatively charged particles are uncorrelated, their systematic uncertainties are partially, but not fully, correlated. On the one hand, the fit uncertainties are largely the same. On the other hand, the MC correction factors, especially for protons and anti-protons, are significantly different. So, mainly the statistical error bars in the lower panels of the following figures should be considered.

The particle and anti-particle production in the V0M multiplicity bins 70 – 100% and 0.1 – 1% is compared in Fig. 6.31. The spectra of both charges typically agree within a few per cent as could be expected from the earlier minimum bias results. The scaled results in the bottom panels are basically the same for both multiplicity bins. In fact, this holds true for the other multiplicity bins, as well as for bins of the reference multiplicity estimator, which is presented in Appendix J.1. It is noteworthy that the negatively charged kaon yield is systematically below that of the positively charged ones. Furthermore, the anti-proton yield grows beyond the proton yield in the very first p_T bins. However, it is not clear to which extent the uncorrelated systematic uncertainties cover these deviations. In any case, taking into account the large statistical uncertainties at high p_T , the

6.5. Inclusive Particle Production in Multiplicity Bins

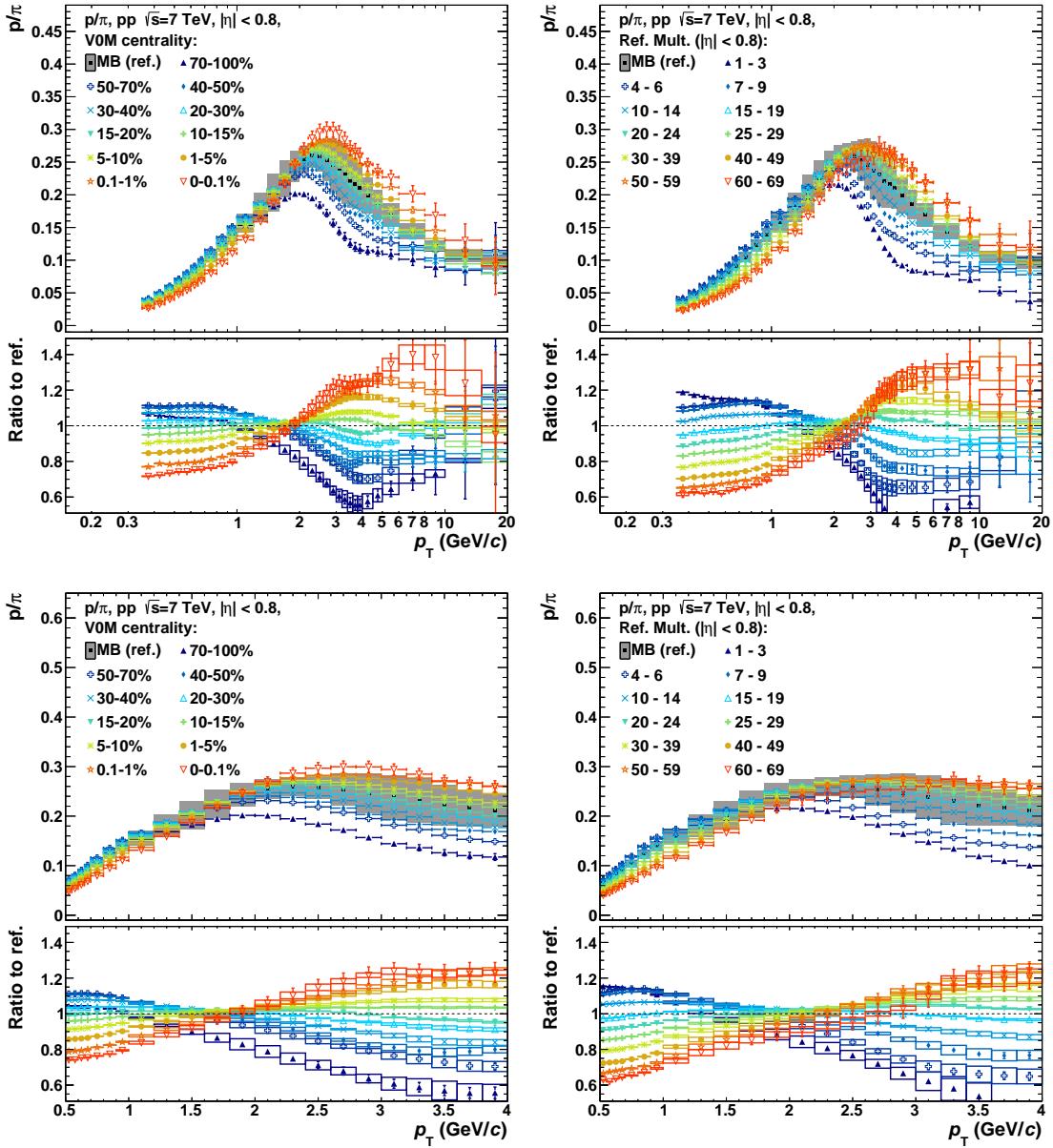


Figure 6.30.: p/π ratios for different bins of V0M (left) and reference multiplicity (right). For better visibility, only the systematic uncertainties of the MB data points are plotted in the upper panel of each row. The lower panel of each row shows the ratio to MB including the corresponding systematic uncertainties. The bottom row shows a zoom to the crossing point with linear p_T scale.

deviation of the scaled results, if any, is at the level of a few per cent in the vast majority of p_T bins. There is no indication that the charge dependence of the particle production evolves as a function of multiplicity.

6. Identified Particle Production in pp Collisions in ALICE

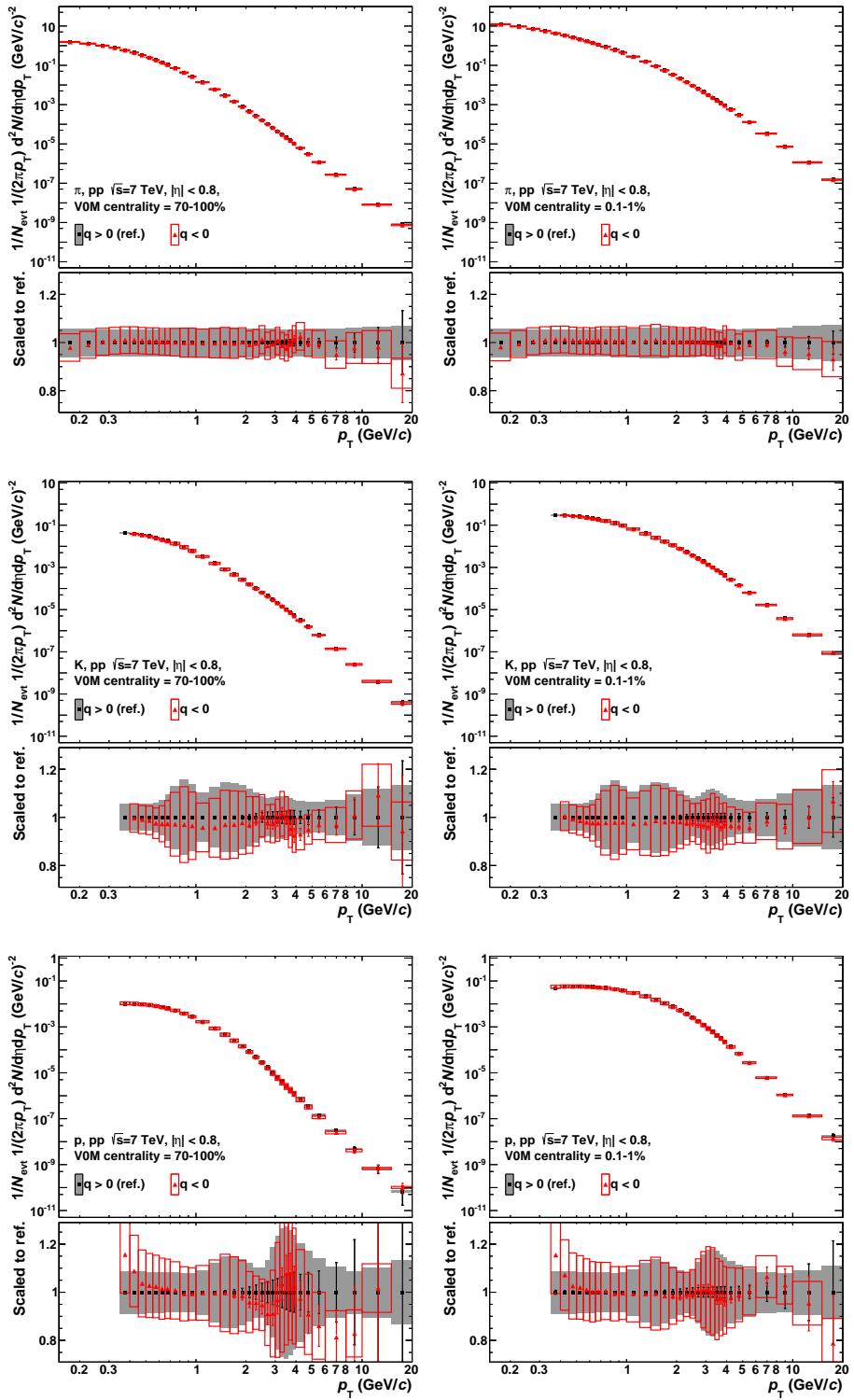


Figure 6.31.: Comparison of positively ($q > 0$) and negatively ($q < 0$) charged particles for the VOM multiplicity bins 70 – 100% (left) and 0.1 – 1% (right). The rows show (from top to bottom): π , K and p .

6.6. Particle Production in Jets

The charged jet reconstruction is carried out with the FastJet package [133] using the anti- k_T algorithm with resolution parameter $R = 0.4$ (cone size), cf. Section 2.3. Only jets fully contained in the TPC acceptance are considered in this analysis, i.e. $|\eta_{\text{jet}}| < 0.9 - R = 0.5$.

The reconstructed charged jet spectrum is shown in Fig. 6.32. To extract the identified particle spectra for a given p_T^{jet} bin, rather high statistics is necessary because the PID requires an additional, high-granular dimension. It turns out that several 10^4 jets are needed for the PID fits. From the figure, it can be read off that this is only possible up to the p_T^{jet} bins 15 – 20 or at most (with larger statistical uncertainties) 20 – 30 GeV/ c .

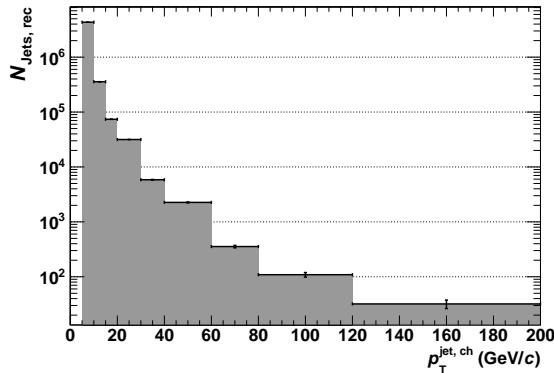


Figure 6.32.: Reconstructed (charged) jet spectrum.

Note that the dE/dx response is extracted from minimum bias data to have the full statistics. It is also used for jet particles, since the dE/dx response of the latter is expected to be the same as the average of all particles. One possible exception is the local track density, which is higher in jets. First estimates [120] show that there might be some small impact on the dE/dx , which shifts the mean by a few per mille, if the p_T^{jet} is increased from 5 – 10 GeV/ c to 15 – 20 GeV/ c . In any case, the possible shift is covered by the systematic errors discussed in Section 6.3.

Since the jet finding is run on a sample of hybrid tracks, but the analysis is done with modified hybrid tracks (see Appendix C.3 for details), the track references from the jet finding cannot be used directly to select the tracks for the analysis¹⁰. Instead, a cone with the same radius R as used for the jet finding is

¹⁰If the analysis tracks are a true sub-sample of the tracks from the jet reconstruction, it is possible to take the track references and apply the harder cut. However, the method discussed in the following is more general. The difference between these two alternatives for a true sub-sample is found to be negligible.

6. Identified Particle Production in pp Collisions in ALICE

put around the reconstructed jet axis and all tracks within this cone satisfying the modified hybrid track cuts are selected for analysis.

The charged particle yield of species s in jets is normalised as:

$$\frac{1}{N_{\text{Jets}}} \frac{dN_s}{dobs} = \frac{1}{N_{\text{Jets}}} \frac{f_{\text{prim},s}(obs)}{\varepsilon_s(obs)} \frac{f_{\pi-\mu,s}(obs)}{\Delta obs} \frac{\Delta N(obs)}{\Delta obs} A_s(obs), \quad (6.7)$$

where obs is either p_T , z or ξ and N_{Jets} is the number of (reconstructed) charged jets in the considered sample. All other quantities are defined analogously to Eq. 6.1, but in general have different values in jets than in the inclusive case. It is worth mentioning that the particle yields in jets are evaluated in a fixed η range and are not differential in η or y .

The to-pion yield ratio of species s (different from π) in jets is defined as:

$$\begin{aligned} \left(\frac{1}{N_{\text{Jets}}} \frac{dN_s}{dobs} \right) / \left(\frac{1}{N_{\text{Jets}}} \frac{dN_\pi}{dobs} \right) &= \frac{dN_s}{dobs} / \frac{dN_\pi}{dobs} \\ &= \frac{A_s(obs)}{A_\pi(obs)} \frac{f_{\text{prim},s}(obs)}{f_{\text{prim},\pi}(obs) f_{\pi-\mu,\pi}(obs)} \frac{\varepsilon_\pi(obs)}{\varepsilon_s(obs)}, \end{aligned} \quad (6.8)$$

where the covariance matrix is used for the error calculation of the ratio $\frac{A_s(obs)}{A_\pi(obs)}$.

To be able to compare with the corresponding inclusive to-pion yield ratios, the latter also needs to be evaluated with respect to the pseudo-rapidity range $\Delta\eta$, i.e. $\Delta y_s(p_T, \eta)$ in Eq. 6.1 is replaced by $\Delta\eta$ and drops out in the ratio. It reads:

$$\begin{aligned} \left(\frac{1}{N_{\text{evt}}(\text{INEL})} \frac{1}{2\pi p_T} \frac{d^2 N_s}{d\eta dp_T} \right) / \left(\frac{1}{N_{\text{evt}}(\text{INEL})} \frac{1}{2\pi p_T} \frac{d^2 N_\pi}{d\eta dp_T} \right) &= \frac{d^2 N_s}{d\eta dp_T} / \frac{d^2 N_\pi}{d\eta dp_T} \\ &= \frac{A_s(p_T)}{A_\pi(p_T)} R_s^{\text{G-F}}(p_T) \frac{f_{\text{prim},s}(p_T)}{f_{\text{prim},\pi}(p_T) f_{\pi-\mu,\pi}(p_T)} \frac{\varepsilon_\pi(p_T)}{\varepsilon_s(p_T)}, \end{aligned} \quad (6.9)$$

where again the covariance matrix is used for the error calculation of the ratio $\frac{A_s(p_T)}{A_\pi(p_T)}$. With these definitions, the right-hand sides of Eqs. 6.8 and 6.9 are identical for $obs = p_T$, except for the GEANT-FLUKA correction factor $R_s^{\text{G-F}}(p_T)$. It needs to be included for the inclusive case, since the corresponding minimum bias MC production requires it — in contrast to the (newer) jet MC production.

The fitting procedure for z and ξ is exactly as for the p_T case, only the binning and the observable (p_T , z , ξ) are different. There are two further exceptions:

- Similar to the p_T case described in Section 4.3.4, a z (ξ) threshold is chosen to apply a special handling to the electrons. A function is fitted to the electron fraction below (above) this threshold and then used to fix the electron fraction beyond the threshold. It turns out that the same functional form of the fit can be used and it is convenient to take the middle of the p_T^{jet} bin as an effective p_T^{jet} , which is then used to calculate the threshold in z and ξ from the corresponding threshold in p_T via Eqs. 4.8 and 4.9.

- The situation is similar for the special templates, which use the TPC dE/dx response of particles that can be identified (uniquely) track-by-track via the TPC (see Section 4.2.2). The method stays the same, but the upper (lower) thresholds need to be converted from p_T to z (ξ). To be sure that the species are well separated when the special templates are used, the most extreme p_T^{jet} of the considered p_T^{jet} range is chosen for the conversion. For z and ξ , this is the largest p_T^{jet} , i.e. all tracks below (above) the z (ξ) threshold have a p_T below the corresponding p_T threshold.

Note that in both cases the actual position of the threshold is not critical, since many p_T bins are contributing to each z or ξ bin, such that the threshold is anyway smeared completely. Also, the impact of the special templates on the result is small, since the lowest p_T bins, in which they are important, are mixed with higher p_T bins. The shape in the higher p_T bins is reliable and the efficiency for kaons and protons steeply rises with p_T at low momenta. Therefore, the region with less reliable templates is statistically dominated by the regions with reliable templates.

6.6.1. Results

This section presents the final results including the corrections to charged particle level discussed in Chapter 5 and their systematic uncertainties, which have been added in quadrature. The p_T^{jet} scale evolution of the corrected yields is shown in Fig. 6.33 as a function of particle p_T and z , respectively. The p_T spectra span 3 – 4 orders of magnitude and become harder with increasing p_T^{jet} . A clear p_T^{jet} ordering of the spectra is observed; the ordering inverts at $p_T^{\text{track}} = 0.4 \text{ GeV}/c$ for pions and at $p_T^{\text{track}} = 2 \text{ GeV}/c$ for kaons and protons. Note that the logarithm of the yield is parabolic as a function of $\ln(p_T^{\text{track}})$. This implies that the ξ spectra have an approximately Gaussian shape. The z spectra span only 1 – 2 orders of magnitude and exhibit an opposite p_T^{jet} ordering. The ordering is inverted around $z = 0.3 – 0.4$ for all species.

In most cases, the last z bin has a higher yield than the neighbouring bin, which is best visible for $p_T^{\text{jet}} = 5 – 10 \text{ GeV}/c$. This happens due to the contribution of single-track jets, i.e. (charged) jets consisting of only one track, to the last z bin. Single-track jets may consist of several neutral jet constituents (that are not reconstructed in this analysis), accompanied by a single charged one. It can also happen that all but one of the charged jet constituents have transverse momenta below the threshold of $150 \text{ MeV}/c$ used for the jet reconstruction and/or are outside of the reconstructed jet cone with radius $R = 0.4$. In any case, the same feature is observed in MC simulations at the particle level (see Fig. K.7 and compare discussion in Section 6.6.2).

The onset of a scaling with p_T^{jet} is observed versus z (extrapolating the trends towards higher p_T^{jet}). The kaon and proton yields do not change within uncertainties above $p_T^{\text{jet}} = 15 – 20 \text{ GeV}/c$. Yet, the pion yield does change beyond uncertain-

6. Identified Particle Production in pp Collisions in ALICE

ties at least up to $p_T^{\text{jet}} = 20 - 30 \text{ GeV}/c$, but also here the ratio between subsequent p_T^{jet} bins seems to get closer to unity for higher p_T^{jet} . The bin $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ is chosen as reference in the lower panels of Fig. 6.33 for two reasons. Firstly, the statistics is sufficiently high, such that statistical fluctuations are small. Secondly, the spectrum as a function of z is similar to that of higher p_T^{jet} bins, such that the onset of the scaling is well visible in the ratio panels.

The Figs. 6.34 and 6.35 show the corrected to-pion ratios as a function of particle p_T and z , respectively. The inclusive result is chosen as reference in the lower panels of Fig. 6.34 in order to directly show how the to-pion ratios in jets compare to the inclusive ones. It may occur at first glance that the K/π ratio has much larger systematic uncertainties at $z \gtrsim 0.5$ for high p_T^{jet} than the p/π ratio. In fact this observation is to some extent a drawing artefact because the absolute values of the K/π ratio are about a factor of 3 – 4 larger in the upper panel than those for p/π . However, this argument holds true only for the upper panels. The lower panels give the relative uncertainties times the mean value of the ratio to the reference, the mean being rather close to unity in most bins. Eventually, the larger uncertainties of the K/π ratio are due to the increase of the spline uncertainty at high $\beta\gamma$ (cf. Tab. 6.3), which is around $p_T = 7 \text{ GeV}/c$ for pions. This has a larger impact on the species neighbouring in dE/dx , namely the kaons, than on the protons. The increased uncertainty is clearly visible in Fig. 6.34 and just propagates to the to-pion ratios as a function of z .

Also for the to-pion ratios, the onset of a scaling with p_T^{jet} is observed versus z . The K/π and p/π ratios do not change within uncertainties above $p_T^{\text{jet}} = 10 - 15$ and $15 - 20 \text{ GeV}/c$, respectively. The K/π ratios show a monotonic increase as a function of z , i.e. the strangeness fraction increases with z . For highest z , the ratio has a pretty large value of about 0.5 – 0.7. For the p/π ratio, there is a maximum at intermediate z and the ratio drops for $z \rightarrow 1$. This means that leading baryons are suppressed.

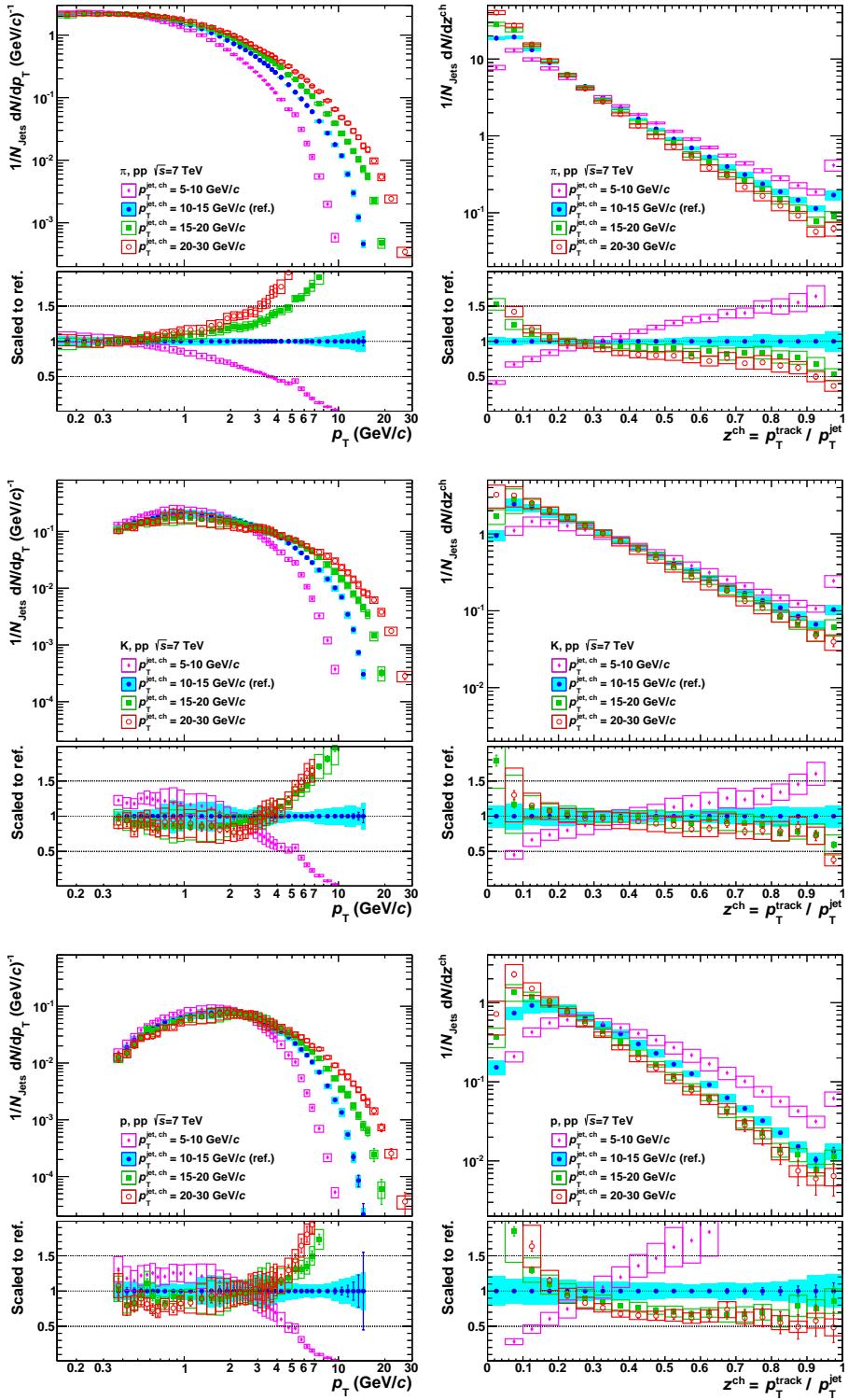


Figure 6.33.: Comparison of the corrected charged $\pi/K/p$ yields for different p_T^{jet} in data as a function of particle p_T (left) and z (right). The rows show (from top to bottom) pions, kaons and protons.

6. Identified Particle Production in pp Collisions in ALICE

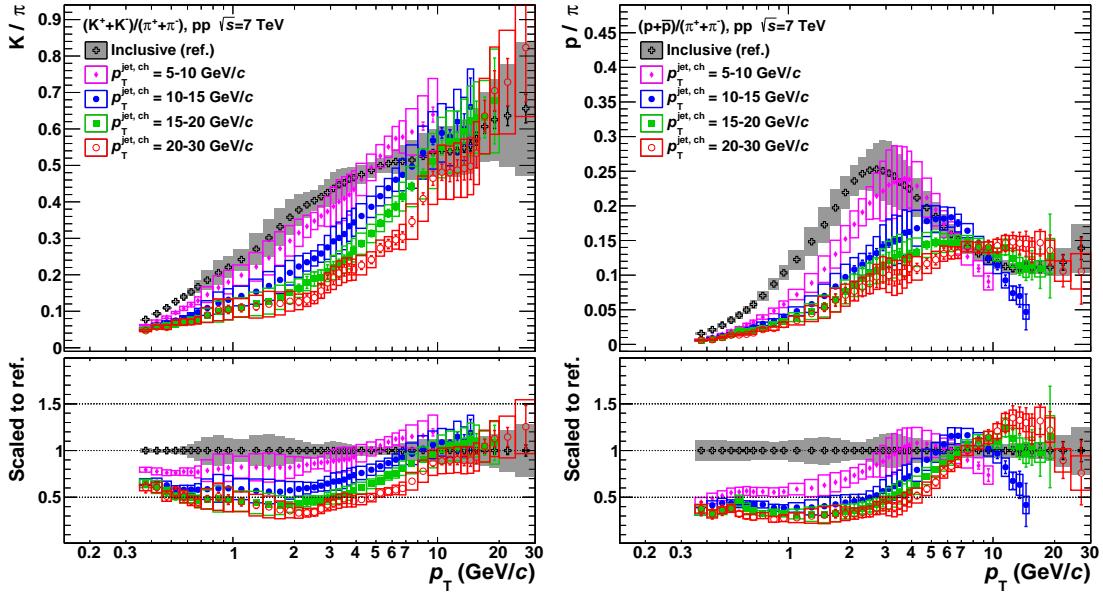


Figure 6.34.: Comparison of the corrected to-pion ratios for different p_T^{jet} in data as a function of particle p_T . The inclusive case is shown in black (open crosses). The left panel shows the kaon-to-pion and the right panel the proton-to-pion ratio.

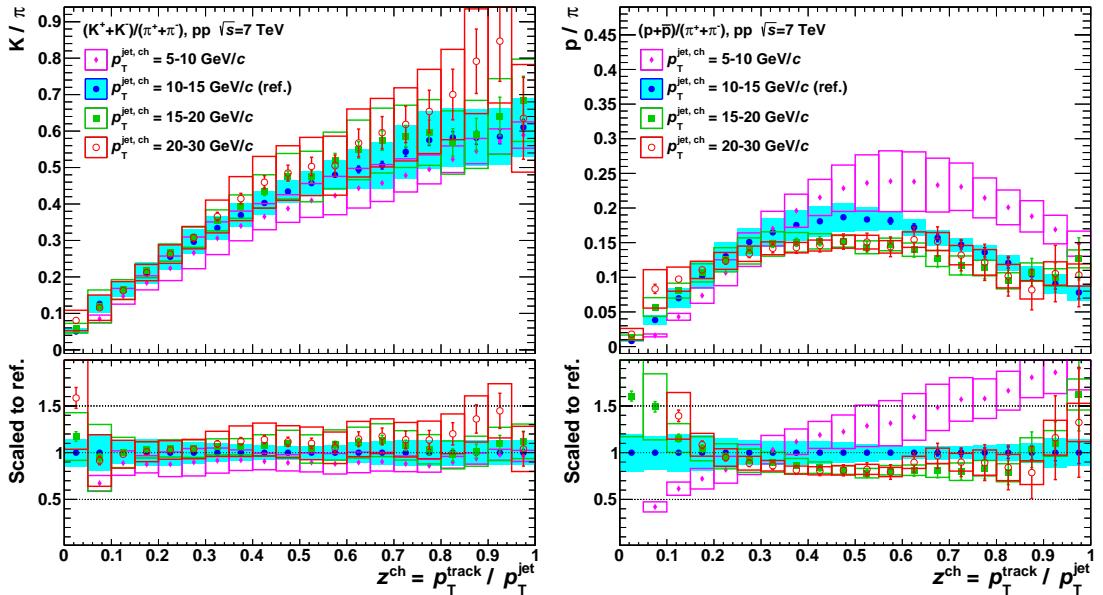


Figure 6.35.: Comparison of the corrected to-pion ratios for different p_T^{jet} in data as a function of z . The left panel shows the kaon-to-pion and the right panel the proton-to-pion ratio.

6.6.2. Comparison to MC Models

In Figs. 6.36, 6.37 and 6.38, the corrected yields of charged $\pi/K/p$ as a function of particle p_T are confronted with different MC models, namely the PYTHIA [87] tunes Perugia-0 (blue), Perugia-0 without colour reconnection (red) and Perugia-2011 (green) [114]. Some important details of these tunes are discussed in Section 6.7.1. For completeness, the corresponding plots as a function of z and the same plots in logarithmic scale can be found in Appendix K. Since the features of the ratio between MC and data are basically the same in bins of z and p_T — except that in the former case the axis is scaled differently —, this section discusses the p_T case only.

All models describe the basic trends, but have a strong tension to the measurement. In particular, all models fail to describe the pions in the lowest p_T^{jet} bin and underestimate the pion yield below $p_T = 1 \text{ GeV}/c$ for all p_T^{jet} . The kaon yield is underestimated by PYTHIA as well, but better agreement is found at high p_T^{jet} . The measurement favours the tune Perugia-0 without colour reconnection (red open triangles). The maximum position of the proton yield around $p_T = 2 \text{ GeV}/c$ is reproduced by the MC models, but the models fail to describe the width of this peak and also the high- p_T slope. In addition, Perugia-2011 predicts a higher yield at the maximum, which is at the edge of the systematic uncertainties of the data points.

The to-pion ratios are compared to MC models in Figs. 6.39 and 6.40 as a function of particle p_T . In Appendix K, the to-pion ratios as a function of z are collected. The trends observed in data (strangeness increase with p_T (or z), leading baryon suppression) are qualitatively described by the MC models. Quantitatively, the deviations observed for the yields are just propagated to the to-pion ratios. Hence, the results versus z are not discussed explicitly in this section.

The observation that the measured kaon yield favours the Perugia-0 model without colour reconnection, is clearly visible in the K/π ratio in Fig. 6.39, in particular in the upper right panel. For the p/π ratio, the MC predictions roughly describe the p_T dependence of the maximum, but its height is predicted to be rather constant in contradiction to the measurement.

Overall, none of the considered PYTHIA tunes gives a perfect description of the data. Nevertheless, the typical agreement for most of the bins is better than 30% for pions and kaons and around 50% for protons, which holds true for all tunes.

6. Identified Particle Production in pp Collisions in ALICE

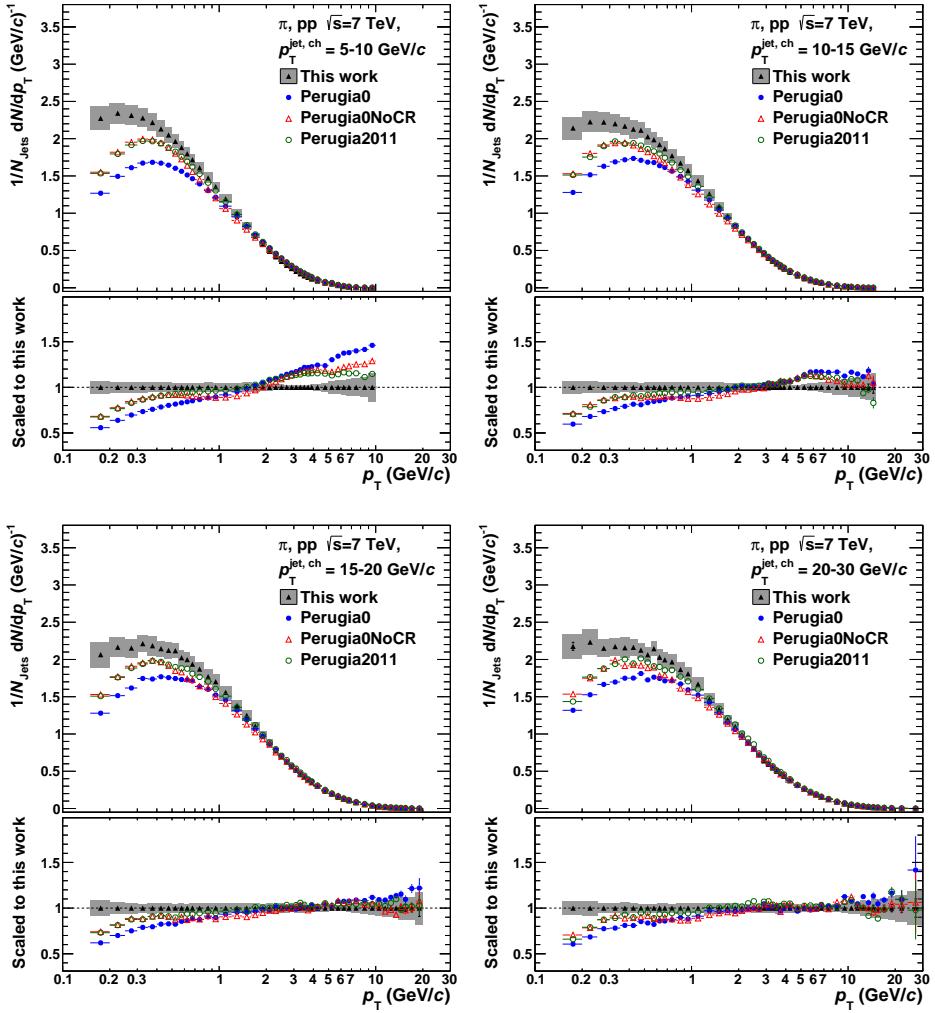


Figure 6.36.: Comparison of corrected pion yields in jets to different MC predictions as a function of p_T . The p_T^{jet} increases from the upper left to the lower right panel.

6.6.3. Charge Differential Analysis

A charge differential analysis contains even more information, since the quark content of the produced particles is different for each charge, but it suffers from lower statistics by about a factor 2. The same motivation and comments about uncertainty correlations as mentioned in Section 6.5.4 apply also in this case. In particular, the systematic uncertainties in the ratios should partly cancel, whereas the statistical uncertainties are independent for the different charges. For jets, the variable of interest is the momentum fraction z , since, in general, the contribution from gluon- and quark-induced processes depends on z [134] (also see discussion in Section 6.5.4). Unfortunately, the large systematic uncertainties of the currently available Fragmentation Functions (FFs), that enter Next-to-Leading

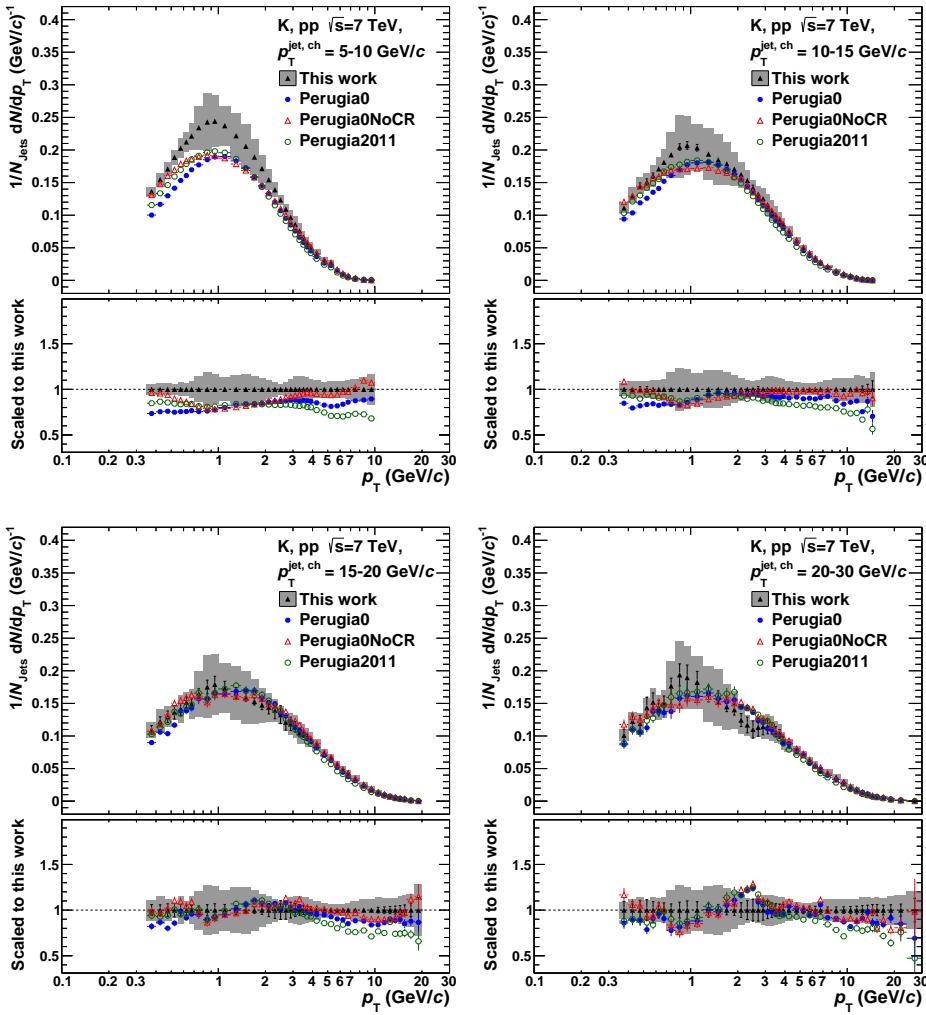


Figure 6.37.: Comparison of corrected kaon yields in jets to different MC predictions as a function of p_T . The p_T^{jet} increases from the upper left to the lower right panel.

Order (NLO) calculations via Eq. 1.2, do not allow for a conclusive prediction of the ratio between particles and anti-particles as a function of z [134, 135].

The spectra of positively and negatively charged particles in jets with $p_T^{\text{jet}} = 5 - 10$ and $10 - 15 \text{ GeV}/c$ are shown in Fig. 6.41 as a function of z . The bin $p_T^{\text{jet}} = 15 - 20 \text{ GeV}/c$ can be found in Appendix J.2 and shows similar trends, but suffers from considerable statistical uncertainties. For $p_T^{\text{jet}} = 20 - 30 \text{ GeV}/c$, these uncertainties are too large to draw conclusions. Typically, both charges are equally abundant within 10% for pions and kaons and within 20% for protons. For the latter, both statistical and systematic uncertainties are large for $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ (and beyond). The yield of negatively charged particles seems to systematically decrease towards higher z for all species. To check whether this trend is covered by the uncorrelated systematic uncertainties would require a more

6. Identified Particle Production in pp Collisions in ALICE

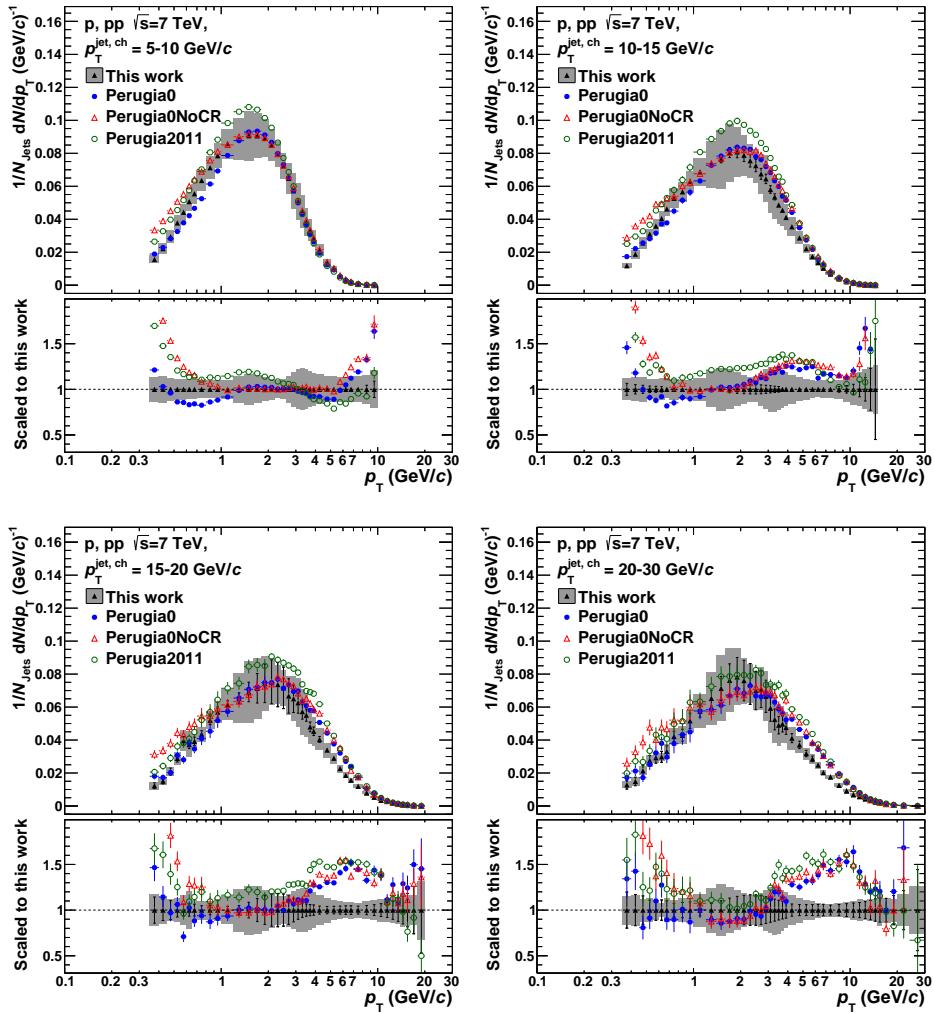


Figure 6.38.: Comparison of corrected proton yields in jets to different MC predictions as a function of p_T . The p_{T}^{jet} increases from the upper left to the lower right panel.

detailed evaluation of the latter for the ratio. Likewise, the anti-proton yield is systematically above the proton yield in the first z bin. It has to be noted that the systematic uncertainty of the MC correction dominates in this bin. Since the correction factors for protons and anti-protons are significantly different for p_T around a few hundred MeV/ c , which corresponds to the low z bins¹¹, it can be expected that the systematic uncertainty of the ratio is less correlated.

In conclusion, the z yields of particles and anti-particles are found to be very similar. Further conclusions require a careful evaluation of the systematic uncertainties of the yield ratios and higher statistics is needed for drawing conclusions for protons in p_T^{jet} bins above 10 GeV/ c .

¹¹To get a feeling about the difference of the correction factors, Fig. 5.7 can be considered, which compares the efficiencies of the individual charges for the inclusive particle production.

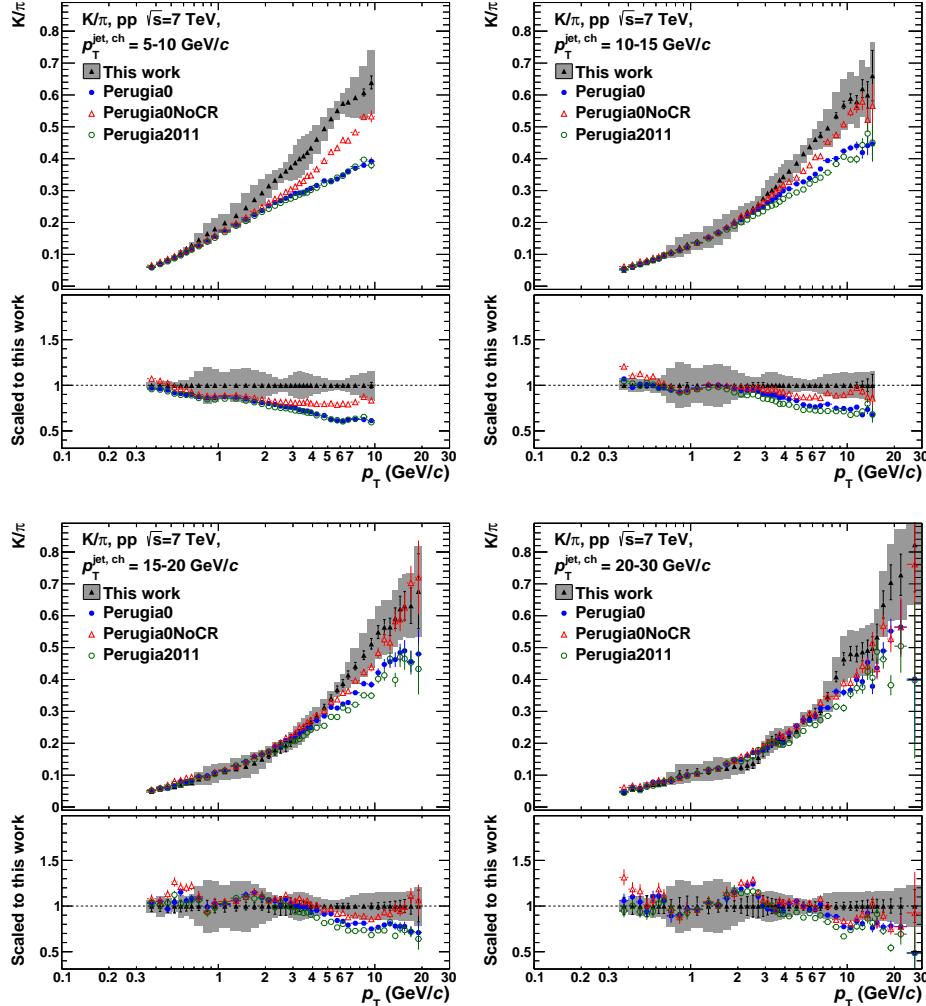


Figure 6.39.: Comparison of corrected K/π ratios in jets to different MC predictions as a function of p_T . The p_T^{jet} increases from the upper left to the lower right panel.

6. Identified Particle Production in pp Collisions in ALICE

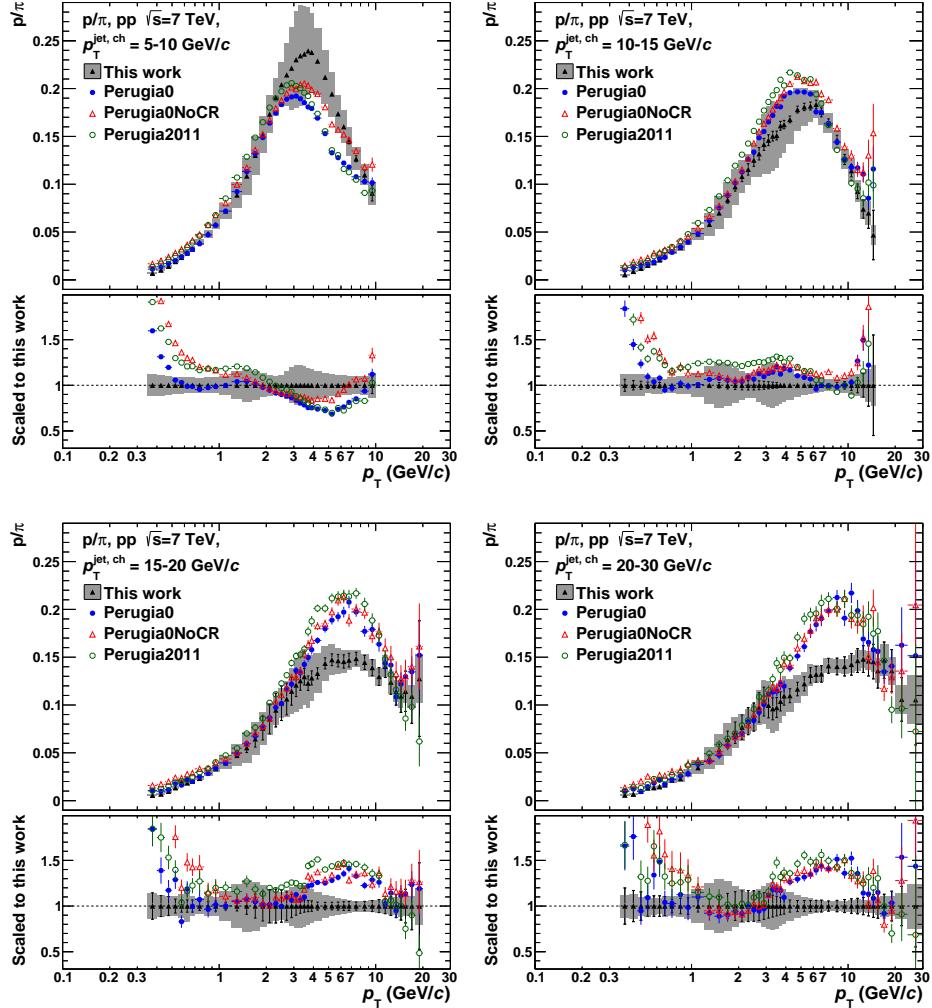


Figure 6.40.: Comparison of corrected p/π ratios in jets to different MC predictions as a function of p_T . The p_T^{jet} increases from the upper left to the lower right panel.

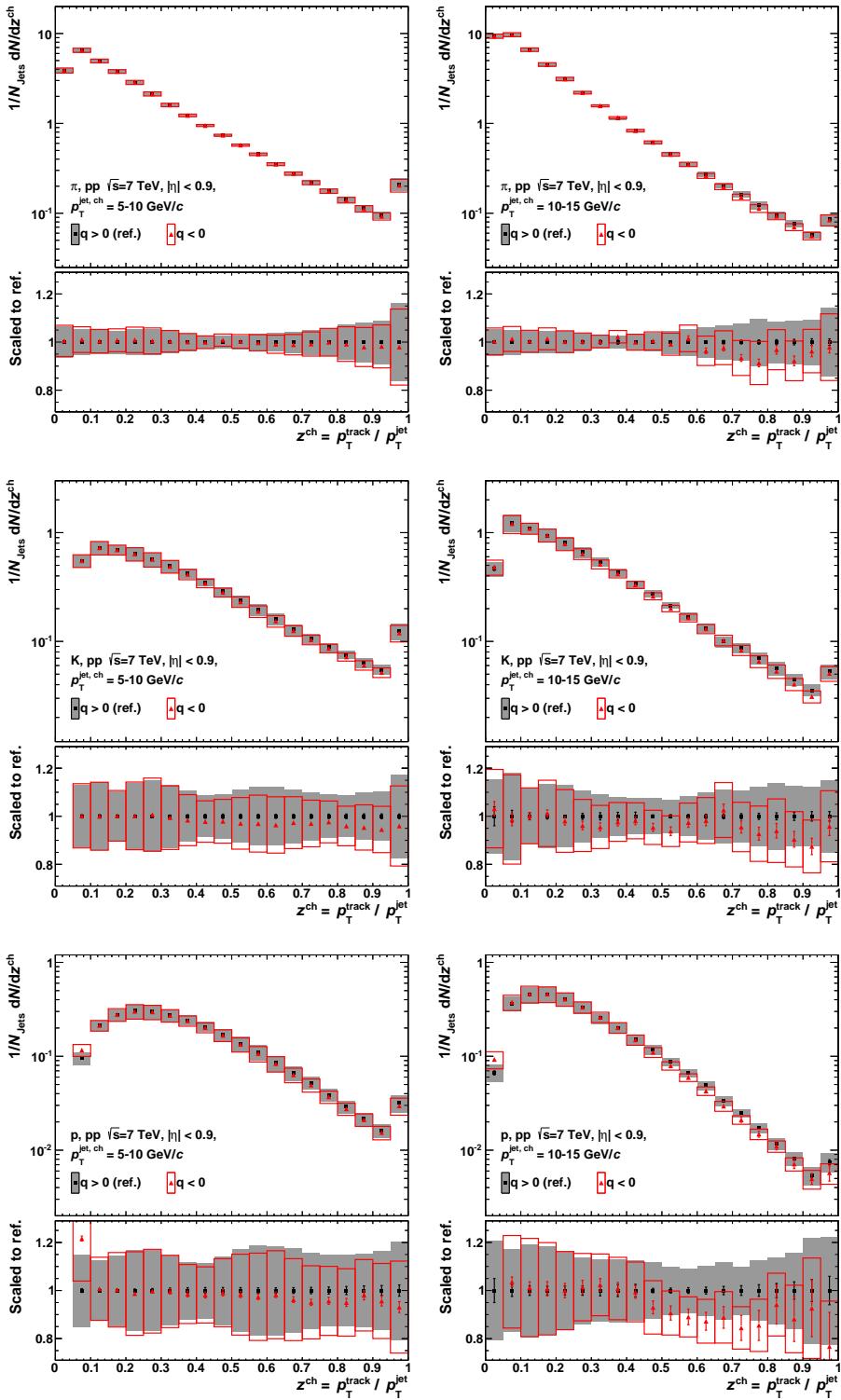


Figure 6.41.: Comparison of positively ($q > 0$) and negatively ($q < 0$) charged particles in jets with $p_{\text{T}}^{\text{jet}} = 5 - 10$ (left) and $10 - 15 \text{ GeV}/c$ (right). The rows show (from top to bottom): π , K and p .

6.7. Discussion

This section discusses the results of the jet and multiplicity analysis parts and demonstrates that the observations may be linked to each other.

6.7.1. Jets

The onset of p_T^{jet} scaling of the identified yields as a function of z (cf. Fig. 6.33) corroborates a similar observation for unidentified charged particles shown in Fig. 6.42. The charged yields are still changing at least up to $p_T^{\text{jet}} = 15 - 20 \text{ GeV}/c$, then becoming largely independent of p_T^{jet} at not too low z . Note that the results in the right panel of the figure are not directly comparable to the left panel and the results of this thesis because the underlying event has been subtracted and the yields are obtained for leading jets.

The comparison to MC PYTHIA Perugia tunes (cf. figures in Section 6.6.2) shows that the agreement between MC models and data typically gets better when moving from low to high p_T^{jet} (except for the protons). Obviously, the MC models better describe the hard processes with large momentum transfer than the softer ones. The ratio between MC prediction and data is rather constant for $p_T^{\text{jet}} = 15 - 20$ and $20 - 30 \text{ GeV}/c$ (cf. Figs. 6.36 – 6.38). This could mean that the rather well-described processes with large momentum transfer start dominating at these jet momenta, such that there is no further improvement of the description. In particular, the bin $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$ seems to be somewhere between the inclusive case and higher p_T^{jet} with respect to the contribution of soft and hard processes and/or the momentum transfer. Firstly, the trends of the MC/data ratio in the bin $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$ are rather similar to the inclusive case, see Figs. 6.4 – 6.6. Secondly, the yields as a function of z are rather different for the bin $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$ and those for higher p_T^{jet} bins (cf. right column of Fig. 6.33). Increasing the p_T^{jet} therefore means to look at samples with higher “jettiness”. It should be kept in mind that the underlying event of the collision may also contribute at low p_T^{track} or z .

The PYTHIA tunes Perugia-0 without colour reconnection and Perugia-2011 give a very similar description of the measured pion yields (see Fig. 6.36), but the latter shows the largest deviation of all considered tunes for the kaon yields (Fig. 6.37) and also, to some extent, for the proton yields (Fig. 6.38). This is a surprising observation because the difference between the tunes Perugia-0 and Perugia-2011 is that the former is only based on LEP, Tevatron and SPS data, whereas the latter also uses LHC pp data at $\sqrt{s} = 0.9$ and 7 TeV (with slight variations of the parameters controlling the modelling of the underlying event and of the fragmentation). Furthermore, the measured kaon-to-pion ratio and also the kaon yield (see Figs. 6.37 and 6.39) favour the Perugia-0 model without

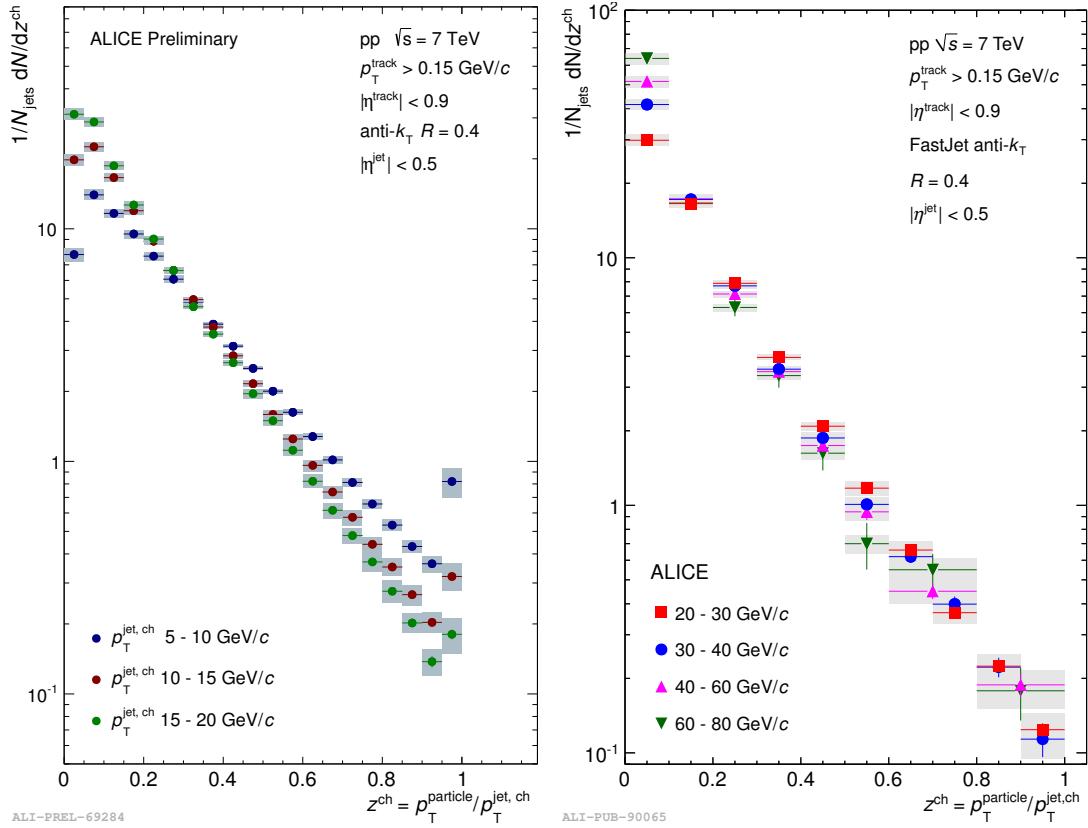


Figure 6.42.: ALICE measurement of charged particle yields in charged jets as a function of z . The left panel shows lower p_T^{jet} bins complementary to the right panel. Note that the left plot is for inclusive jets without subtraction of the underlying event [124], whereas the right plot shows leading jet results after underlying event subtraction [34].

colour reconnection. This is an interesting observation, since colour reconnection is one candidate to explain collective phenomena in high-multiplicity pp events, which will be further discussed below. Obviously, there is room for improvement and further tuning of the event generators seems to be required.

6.7.2. pp Versus Multiplicity

The identified particle spectra show a considerable evolution as a function of multiplicity. As was shown in Figs. 6.27 and 6.28, the particle yields of every species change by more than one order of magnitude from lowest to highest multiplicities. From the ratio to the Minimum Bias (MB) spectra it can be seen that the slope versus p_T is also changing with multiplicity. The origin of the high multiplicity events and this behaviour are not understood yet. A possible contribution could stem e.g. from multi-parton interactions [136].

6. Identified Particle Production in pp Collisions in ALICE

In addition to this “first order” effect, a species dependence of the multiplicity evolution is observed, which can be studied via the to-pion ratios. It was discussed in Section 1.4 that from the evolution of the to-pion ratios in Pb–Pb and p–Pb collisions as a function of multiplicity some features can be naively anticipated for the evolution in pp. In particular, the K/π ratio is expected not to change with multiplicity and p/π should still exhibit a crossing point for different multiplicities around $p_T = 1.6 \text{ GeV}/c$. Early studies of pp collisions [137] as a function of mid-rapidity multiplicity, which have been confirmed and extended in p_T by (the reference multiplicity part of) this work, showed a deviating behaviour. The p/π crossing point is at a different position (more like $p_T \approx 2.3 \text{ GeV}/c$) and the K/π ratio is shifted towards higher p_T with increasing multiplicity. The V0M results of this thesis show that for a multiplicity estimator at forward rapidity, the results are close to the picture that would be expected if the physics (like hydrodynamics) taking place in small systems was the same as in large systems. The K/π ratio is largely independent of multiplicity and the p/π crossing point is at $p_T \approx 1.7 \text{ GeV}/c$.

The question is what causes the difference between these estimators. Is there a (larger) jet bias on the mid-rapidity multiplicity estimator? High values for this estimator may originate from high- p_T jets fragmenting into many particles. The V0M multiplicity estimator at forward rapidity can mitigate this effect, since there is an η gap between the region for the multiplicity estimate and the analysis region.

As a first check, scatter plots for the reconstructed p_T^{jet} versus the multiplicity of each estimator are shown in Fig. 6.43. In the scatter plots, every (charged) jet corresponds to one entry, i.e. an event with n jets contributes with n entries. Note that due to the threshold $p_T^{\text{jet}} \geq 5 \text{ GeV}/c$ in the jet reconstruction, events with no jets or “jets” with smaller momentum do not contribute to these plots. From the plots it can be seen that the p_T^{jet} distribution in each multiplicity bin is continuous and monotonous. Hence, it is reasonable to take these distributions and estimate the mean p_T^{jet} (for jets above the threshold) per multiplicity class and search for trends.

To better judge possible correlations between p_T^{jet} and the multiplicity, the lower panel in the figure shows the profiles of the scatter plots with the mean p_T^{jet} in each multiplicity bin. Obviously, the mean p_T^{jet} is uncorrelated from the V0M multiplicity (constant at $\sim 8.2 \text{ GeV}/c$), but increases approximately linearly with reference multiplicity above multiplicity ≈ 6 . In the latter case, the change of the mean p_T^{jet} is about $1 \text{ GeV}/c$ for the multiplicity range from 10 to 100. The drop for reference multiplicity $\approx 2 - 5$ could be due to mini-jets. If the latter fragment into many particles, this increases the reference multiplicity on the one hand. On the other hand, such soft fragmentation results in a broad jet profile and particles may be emitted outside of the (cone) radius used for the jet reconstruction and, thus, do not contribute to p_T^{jet} . As a consequence, the mean p_T^{jet} might drop for mini-jets with increasing reference multiplicity.

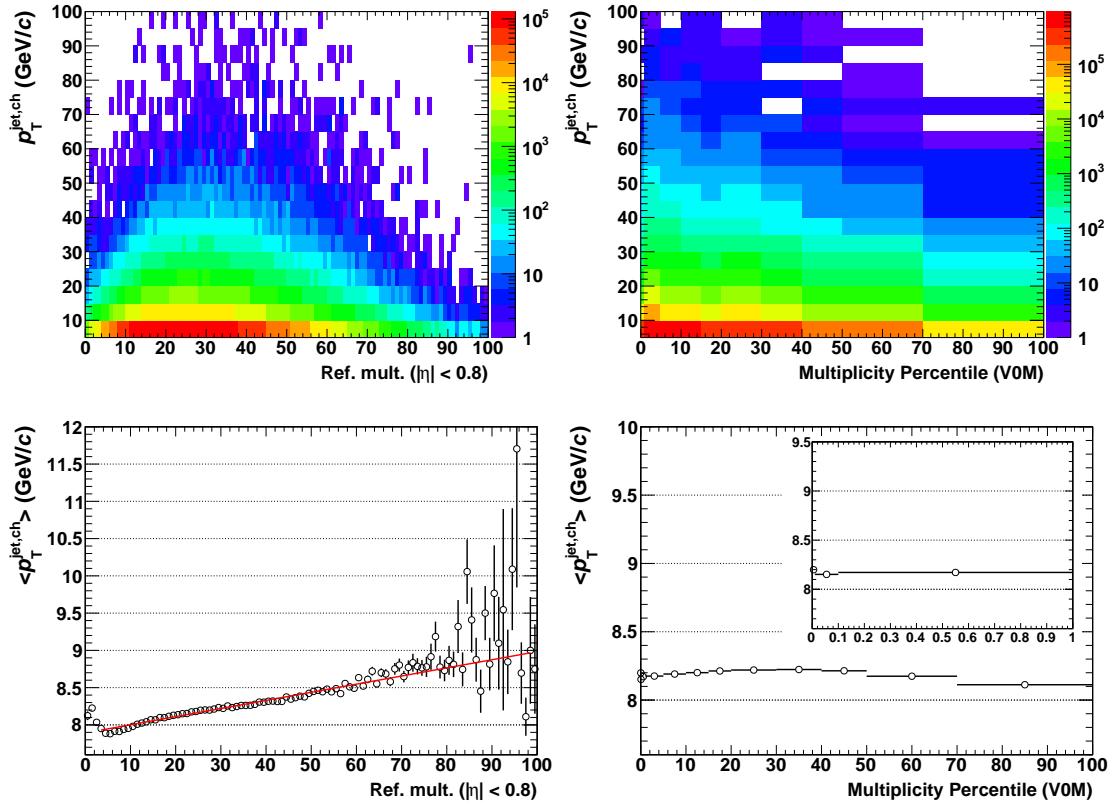


Figure 6.43.: Correlation between p_T^{jet} and multiplicity for the reference multiplicity (left) and the V0M multiplicity (right). The upper panel shows the scatter plots, the lower panel depicts the profiles of the corresponding mean p_T^{jet} in each multiplicity bin. The colour code gives the number of jets. The red line in the lower left panel is a linear fit.

Turning back to high reference multiplicities, they are evidently caused by more energetic jets to some extent. Selecting high reference multiplicities, thus, may include a bias to a larger admixture of (high- p_T) jets, whereas low reference multiplicities are similar to a large V0M percentile — maybe with even a slightly suppressed jet contribution for the reference multiplicity. Of course, there is no clear one-to-one correspondence between a bin in reference multiplicity and another one in V0M. Furthermore, different p_T^{jet} may contribute and their weight could change as a function of p_T^{track} .

A further hint to such a jet bias comes from the observation in Figs. 6.27 and 6.28: the hardening of the spectra seems to be more pronounced for high reference than for high V0M multiplicities. Since jets exhibit a harder p_T spectrum, a jet bias could explain this trend.

Looking at the p/π ratios in jets compared to MB, Fig. 6.34, it seems as if the ratio to MB only changes little with p_T^{jet} for high p_T^{jet} and p/π is suppressed in jets. Keeping in mind that a mixture of all p_T^{jet} contributes to the inclusive

6. Identified Particle Production in pp Collisions in ALICE

measurement, the naive expectation is to see a rather constant suppression of p/π for jet particles with respect to MB at low p_T^{track} , which becomes gradually smaller supposedly somewhere around $p_T^{\text{track}} = 4 - 10 \text{ GeV}/c$, finally approaching the MB value at high p_T^{track} . Similarly, for K/π , there should be a rather constant suppression at low p_T^{track} in jets. The shift of the K/π shape with p_T^{jet} seems to continue even at high p_T^{jet} . Hence, the K/π suppression in jets with respect to MB may last up to highest p_T^{track} . One important difference between the evolution with p_T^{jet} for K/π and p/π is that basically the whole distribution is just shifted for K/π . In contrast, the peak of p/π is not only shifted, but also decreases in height. Hence, even if the shift of both to-pion ratios may be similar, the change is much smaller for p/π because of the rather flat shape.

If the difference between reference multiplicity (RM) and V0M was due to a (larger) jet bias in the reference multiplicity, this would be reflected in the observation that the ratio between reference and V0M multiplicity shows similar trends as that of jets and MB. From the above deliberation it is reasonable to find bin pairs for both multiplicity estimators for which the p/π ratio roughly matches at high p_T^{track} . In Fig. 6.44, two such bin pairs at high (V0M 0.1 – 1%, RM 40 – 49) and low multiplicities (V0M 30 – 40%, RM 10 – 14), respectively, are compared with MB as a common reference. For convenience, two different p_T^{jet} bins have been added to the plot to better estimate how jet particles change the to-pion ratios in comparison to MB. Indeed, the low multiplicity bin pair is found to have a similar p/π all over p_T^{track} . In contrast, for the high-multiplicity bin pair, the RM result is biased towards lower p/π at low p_T^{track} , finally (within large statistical uncertainties) gradually becoming identical with the V0M result around $p_T^{\text{track}} = 4 - 5 \text{ GeV}/c$. Taking the same bin pairs for K/π , also here the low multiplicity bin pair gives very similar results. For the high-multiplicity bin pair, the RM result is systematically below that of V0M, the ratio being rather independent of p_T^{track} . Overall, this demonstrates at least qualitatively that a (stronger) jet bias of the reference multiplicity estimator could account for the difference to the V0M estimator or at least for some part of it. It should be emphasized that this is not a proof or a quantitative statement, but rather a plausibility consideration with some (reasonable) assumptions.

There is yet another intriguing observation. The p/π ratios for different V0M multiplicities seem to reunite around $p_T = 15 - 20 \text{ GeV}/c$ (see left panel of Fig. 6.30), whereas they look much more separated for the reference multiplicity (right panel of that figure). It is not obvious in the latter case, whether there is a reunion at higher p_T . Unfortunately, the error bars are pretty large and do not allow for a clear judgement.

In summary, using the V0M multiplicity estimator at forward rapidity shifts the p/π crossing point towards lower p_T compared to the reference multiplicity estimate at mid-rapidity, namely to about $p_T \approx 1.7 \text{ GeV}/c$. This is close to what is observed in p–Pb and Pb–Pb collisions ($p_T \approx 1.5 - 1.6 \text{ GeV}/c$). There may still be a residual jet bias, which is suppressed in larger systems by $1/N_{\text{part}}$. But the

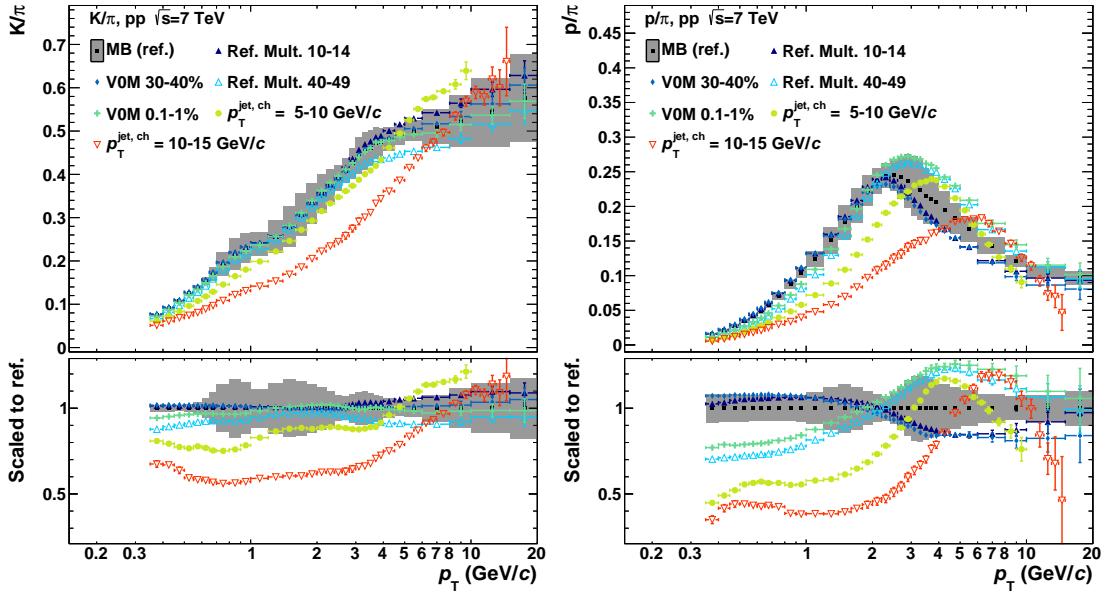


Figure 6.44.: Comparison of K/π (left) and p/π (right) for different multiplicity estimators and jets, see discussion in the text. Note that for this comparison, the conversion from pseudo-rapidity to rapidity was not applied, neither to the jet, nor to the multiplicity results. The jet results are for $|\eta| < 0.9$, those for multiplicity for $|\eta| < 0.8$, which should have negligible impact on this comparison.

comparison of the VOM and the reference multiplicity estimators clearly shows a trend in the expected direction. Evidently, this might imply that in a careful analysis, the same physics of big systems is discovered in small systems.

Taking together this observation and other discoveries like the ridge structure in two-particle correlations (cf. Section 1.4), there are hints for the presence of flow in small systems. It remains to be investigated what the origin of the flow is. Is a medium formed also in small systems, which allows for a hydrodynamical description? If yes, how can an object as small as a few fm^3 behave like a macroscopic fluid? Or are the flow-like patterns generated by other mechanisms like Colour Reconnection (CR) [138] (also see [139])? Or can the observation be explained by initial-state effects from gluon saturation described by Colour Glass Condensate (CGC) models [140, 141]? The answers to these questions have to be provided by future measurements and theoretical model calculations.

6.8. Conclusion About Analysis Method and Outlook

The TPC Multi-Template Fit (MTF) is a powerful tool that allows to extract identified particle spectra in a wide momentum range extending from $p_T = 0.15$ to about $20 \text{ GeV}/c$. It relies on a detailed modelling of the TPC dE/dx response developed in this thesis. As was discussed in Section 3.4.5, there are no data points from clean particle samples available for pp collisions to constrain the Bethe-Bloch fit of the dE/dx in the region with $\beta\gamma = O(100)$. This leads to large systematic uncertainties of the extracted particle yields at high momenta. On the one hand, a unique identification of primary particles is not possible in that region. On the other hand, the available statistics from secondary particles identified via track topologies is too small. A possible alternative are muons from cosmic rays [142]. They offer sufficient statistics for this $\beta\gamma$ region. A pure sample could be achieved by using a cosmic trigger and demanding that no beam collisions take place at the same time. The framework developed in this work is designed for general purposes. It is therefore simple to just add data points from cosmic muons to the Bethe-Bloch fit and leave anything else untouched. In particular, there is no modification needed for the MTF.

The LHC run 2 has just started. It will offer higher statistics and, in particular, data takings with high-multiplicity and jet triggers are planned. This will allow to extract identified particle spectra beyond $p_T = 20 \text{ GeV}/c$ and will give access to even higher event multiplicities. It remains to be seen what the ultimate high- p_T limit will be, since the statistical PID method does not work anymore as soon as all particles reach the Fermi plateau of the Bethe-Bloch curve (cf. discussion in Section 3.8).

It is planned to use the MTF for the extraction of identified particle spectra in jets from p–Pb and Pb–Pb collisions. A comparison to the pp results of this work then provides information about the medium created in such collisions (see Section 1.3.1). However, this requires some further investigations, in particular for the study of Pb–Pb collisions. Firstly, the TPC dE/dx parametrisations need to take into account the large occupancy effects in heavy-ion collisions. Secondly, the jet spectra must be corrected for the underlying event, which gives a sizeable contribution to the jet particle spectra in Pb–Pb collision. This is a common problem for analyses of jets in a heavy-ion environment and different approaches already exist. A possible procedure could be to extract the identified particle spectra in cones perpendicular to the reconstructed jet axis and to subtract them from the jet particle spectra. No matter which method is chosen in the end, the correction for the underlying event has to be carried out with great care, since it can be affected by various effects. For instance, the hydrodynamical flow leads to correlations of the angular particle distribution.

6.8. Conclusion About Analysis Method and Outlook

The Dijet Calorimeter (DCal), that was installed during the shutdown after LHC run 1, extends the calorimeter coverage and enables the study of dijets. This is in particular interesting with respect to Pb–Pb collisions, where the back-to-back partons from the hard-scattering travel a different distance through the medium and are, thus, modified in various ways. Dijet asymmetries contain information about the medium, that can e.g. induce gluon radiation. In general, a trigger on jets in the calorimeter acceptance will allow for the study of full jets. The charged tracks could be identified as done for the charged track analysis in this work. By combining them with calorimeter hits then gives access to the neutral energy and with that to the full jet energy. Such a measurement will facilitate the comparison to theoretical calculations.

In conclusion, the PID calibration procedures developed in this work are an important basis for future studies. The MTF is basically ready to be used for the study of other collision systems. It was demonstrated in Section 4.4 that the MTF is capable of extracting particle spectra as a function of various observables, like p_T , z or ξ , and only minor modifications are required to switch the observable. In that sense, it is conceivable to utilise even further observables in future.

7. Summary

In the framework of this thesis, the TPC Multi-Template Fit (MTF) was developed as a powerful tool for statistical Particle IDentification (PID). It fits templates with the expected dE/dx response of individual species to the measured TPC dE/dx distribution. The MTF is capable of extracting charged $\pi/K/p$ spectra in a wide momentum range, which extends from $p_T = 0.15$ to at least $20\text{ GeV}/c$. It allows for a consistent treatment of the whole p_T range, which is an advantage over other methods commonly used in ALICE that are either limited to low or to high p_T . Good agreement between the MTF and the latter is found for the extracted minimum bias spectra of pions, kaons and protons.

The MTF relies on a detailed modelling of the TPC dE/dx response. In this work, a framework was developed which models the TPC dE/dx response, extracts its parameters and allows to evaluate PID-relevant quantities. It was integrated into the common PID framework of the ALICE software package.

The goal of this thesis was to study identified charged particle spectra and in particular the charged hadron composition in jets. Such an analysis requires various Monte Carlo (MC) corrections, e.g. to take into account detector effects. The correction procedure for jets is not straightforward, since detector inefficiencies, secondary particles, etc. also influence the jet reconstruction. The correction procedure developed in this work is based on that for the charged particle jet analysis carried out by the ALICE collaboration earlier. It was extended to be applicable to individual species and especially takes into account effects that are important for a specific species, but were negligible in the analysis of unidentified charged jet particles. Based on ideas from the latter analysis, an estimation procedure for the systematic uncertainty of the MC corrections for pions, kaons and protons was developed.

The main outcome of this thesis is the application of the MTF in two different studies of pp collisions at $\sqrt{s} = 7\text{ TeV}$ recorded by ALICE in 2010. The first analysis investigated the charged $\pi/K/p$ spectra and K/π and p/π ratios at mid-rapidity in bins of event multiplicity. The second analysis studied the identified charged particle composition in charged jets at mid-rapidity.

The multiplicity-binned analysis is motivated by the observation of intriguing similarities between high-multiplicity pp collisions and p–Pb and Pb–Pb collisions. Besides the observation of long-range correlations in all three systems, also the K/π and p/π ratios were found to exhibit strikingly similar patterns in p–Pb and Pb–Pb collisions for different multiplicity classes. While hydrodynamical models can explain these features for Pb–Pb, their application to smaller systems seems questionable. A variety of other models, such as Colour Reconnection (CR) [138]

7. Summary

or Colour Glass Condensates (CGCs) [140, 141], has been suggested that could account for these observations. The K/π and p/π ratios in pp collisions obtained in this work may help to distinguish between these models. Two complementary multiplicity estimators have been utilised: the reference multiplicity (RM) at mid-rapidity and the V0M multiplicity (V0M) at forward rapidity. The former is strongly correlated with the analysis region, whereas the latter is separated by a pseudo-rapidity gap. Potential biases by these estimators can be inferred from the comparison of the two. Indeed, the results hint to a possible jet bias of the RM. For V0M, it is found that the multiplicity evolution of the K/π and p/π ratios exhibits noticeable similarities to those in $p\text{-Pb}$ and Pb-Pb collisions. Together with the observed long-range correlations, this raises the question about the origin of these effects. The ultimate answer may be found by comparing theoretical models with all currently available data.

The final topic of this thesis was the study of identified charged particles in charged jets. The corresponding spectra can either be extracted as a function of the transverse momentum of the constituent, p_T^{track} , or the reduced momentum $z^{\text{charged}} \equiv p_T^{\text{track}}/p_T^{\text{jet, charged}}$ of the constituent. While the former can be compared to the inclusive charged particle production, the latter is closely related to the Fragmentation Functions (FFs) in theoretical calculations. In particular, the reconstructed jet momentum approximates the momentum of the original parton emerging from the hard-scattering, thus, allowing for a rather direct comparison of theory and measurement. The objective of this study was threefold. Firstly, the measurement allows to further constrain FFs, which currently exhibit large uncertainties that lead to a very limited accuracy of cross-section calculations. Secondly, it allows to distinguish between different models for the physics of jets. Thirdly, it provides a baseline for the corresponding measurements in $p\text{-Pb}$ and Pb-Pb collisions. The comparison of the results in all three collision systems finally provides information about the medium that is created in Pb-Pb collisions. For the z^{charged} spectra and the corresponding K/π and p/π ratios, the onset of a scaling with p_T^{jet} is observed. The K/π ratio is found to increase with z^{charged} and becomes as large as $0.5 - 0.7$, which implies that the strangeness fraction rises with z^{charged} . The p/π ratio exhibits a maximum at intermediate z^{charged} , which implies that leading baryons are suppressed. These results were confronted with predictions from various PYTHIA Perugia tunes. None of the tunes gives a perfect description of the data, but all tunes roughly describe the observed trends with an accuracy of about $30 - 50\%$.

In conclusion, ALICE is perfectly suited to study identified particles, which is unique at the LHC. The MTF can be used for any analysis that requires statistical PID of charged particles. Higher statistics, dedicated triggers, detector and accelerator upgrades in LHC run 2 will allow to study the physics of pp , $p\text{-Pb}$ and Pb-Pb collisions in more detail.

Appendix A.

Remarks About the Ratios of two Results

In this thesis, various results are usually compared in an extra panel in each figure containing all the graphs “scaled to” some reference, see e.g. Fig. 6.34, instead of showing the real “ratio to”. In general, showing just the scaled results avoids problems with possible correlations of errors between two results. It may be difficult to figure out to which extent the errors are correlated in order to eventually calculate the true errors of the ratio. In addition, the binning of the results may be different, e.g. when two analyses are compared. While the mean values of the ratio can be calculated via interpolating the value in the denominator, this is not trivially possible for correlated errors with the correlation changing from one bin to the other.

In most of the cases shown in this thesis, the conclusion of the “scaled to reference” plots is already clear and there is no need to evaluate the actual ratios and errors. However, there are a few cases in which the error size is crucial to interpret the result and it is worth to study the errors of the ratios in detail. An example is the comparison of various multiplicity bins with minimum bias (MB) (see Section 6.5.3). If just the systematic errors of MB and all multiplicity bins were plotted for the “scaled to reference” case, it wouldn’t be possible to judge whether the shift of the crossing point in the p/π ratio is significant.

Appendix B.

Run Lists for Analysis

The individual parts of the analysis are run on the following lists of runs for the pp collisions at $\sqrt{s} = 7 \text{ TeV}$ data recorded in 2010:

LHC10d.pass2:

122374, 125085, 125097, 125101, 125134, 125296, 125630, 125632, 125633, 125842, 125847, 125848, 125849, 125850, 125851, 125855, 126004, 126007, 126008, 126073, 126078, 126081, 126082, 126088, 126090, 126097, 126158, 126160, 126168, 126284, 126351, 126352, 126359, 126404, 126406, 126407, 126408, 126409, 126422, 126424, 126425, 126432.

LHC10e.pass2:

127712, 127714, 127718, 127822, 127933, 127935, 127936, 127937, 127941, 127942, 128185, 128186, 128189, 128191, 128192, 128260, 128366, 128452, 128483, 128486, 128494, 128503, 128504, 128507, 128582, 128605, 128609, 128611, 128615, 128777, 128778, 128820, 128824, 128836, 128843, 128850, 128853, 128855, 128913, 129514, 129520, 129523, 129528, 129540, 129586, 129587, 129641, 129647, 129650, 129652, 129653, 129654, 129666, 129667, 129726, 129729, 129735, 129736, 129738, 129742, 129744, 129959, 129960, 129961, 129983, 130149, 130157, 130158, 130172, 130178, 130375, 130480, 130517, 130519, 130696, 130704, 130793, 130795, 130799, 130834, 130844, 130847, 130848.

The selected runs have good quality according to the ALICE Run Condition Table (RCT) and $\mu < 0.15$, i.e. a relatively low pile-up. The underlined run numbers have only been used for the jet and reference multiplicity analysis parts, but not for V0M. The reason is that the amplitudes of the V0 detectors show problems for those runs and cannot be used for a multiplicity estimate. Yet, the events have good quality to be used for the other analysis parts.

B.1. List of Monte Carlo Runs

To extract correction factors for the inclusive case (as well as for the Monte Carlo (MC) blind test and MC closure check), the MC minimum bias production **LHC10f6a** with the following list of runs has been used:

Appendix B. Run Lists for Analysis

122374, 122375, 124751, 125023, 125085, 125097, 125100, 125101, 125134, 125296, 125628, 125630, 125632, 125633, 125842, 125843, 125844, 125847, 125848, 125849, 125850, 125851, 125855, 126004, 126007, 126008, 126073, 126078, 126081, 126082, 126088, 126090, 126097, 126158, 126160, 126168, 126283, 126284, 126285, 126359, 126403, 126404, 126405, 126406, 126407, 126408, 126409, 126422, 126424, 126425, 126432, 126437.

The selected runs largely have anchor runs used for the analysis (see lists above). Since no significant run dependence for the correction factors is observed, some additional runs which have anchor runs with good quality according to the ALICE RCT have been added. The underlined run is excluded from the analysis using the V0M multiplicity because it shows problems with the V0 amplitude (but the run itself can be used for the other analysis parts).

For the correction factor extraction for the jet analysis, the following list of runs from the MC jet productions **LHC14b6[b-e]** is used:

122374, 125101, 125296, 125633, 125848, 125851, 126007, 126073, 126082, 126097, 126168, 126351, 126404, 126408, 126424.

Appendix C.

Track Selection and Cuts

This appendix comprises details about the track selection and cuts in the individual analysis parts.

C.1. V0 Candidate Selection

In Section 3.4, it was discussed that the extraction of the TPC PID parametrisation requires clean identified particle samples. Besides selecting primary tracks with PID information from TPC and TOF, track topologies provide a means for clean particle identification. Such V0 candidates are selected via cuts on the invariant mass and on the Armenteros-Podolanski variables. In detail, the following code is used for the V0 selection:

```
1 AliESDv0KineCuts* fV0Cuts = new AliESDv0KineCuts;
2 fV0Cuts->SetGammaCutChi2NDF(5.);
3 Float_t gammaProdVertexRadiusCuts[2] = { 3.0, 45. };
4 fV0Cuts->SetGammaCutVertexR(&gammaProdVertexRadiusCuts[0]);
5 fV0Cuts->SetMode(AliESDv0KineCuts::kPurity, AliESDv0KineCuts::kPP);
```

Firstly, the reduced χ^2 of the γ candidates fitted from the daughter tracks is demanded to be smaller than 5. Secondly, the γ production radius must lie within 3 and 45 cm (cf. discussion in Section 3.4.1). The last line enables the purity mode for γ candidates, such that tight cuts on the Armenteros-Podolanski variables are applied. Besides the cut on the γ production radius, the extraction of the TPC PID parametrisation requires an additional TOF selection for high-momentum V0 electrons, which is an even stronger criterion.

C.2. ITS-TPC Standard Track Cuts 2010

A set of standard track cuts using ITS and TPC information has been defined for the data taking in 2010. These cuts ensure a reasonable tracking quality, but further cuts are required to improve e.g. the TPC PID performance. The ITS-TPC standard cuts for 2010 only accept tracks reconstructed with both TPC and ITS information and which have at least one SPD hit. Daughters from kinks are

rejected. Furthermore, the track is required to have at least 70 tracking clusters in the TPC. The χ^2 of the track fit needs to be smaller than 4 for each TPC cluster and less than 36 per ITS cluster. The Distance of Closest Approach (DCA) to the primary vertex is required to be smaller than 2 cm in beam direction. If primary particle selection is demanded, a p_T dependent DCA cut in the plane transverse to the beam direction is added: $DCA_{xy} < \left(0.0182 + \frac{0.0350}{(p_T/\text{GeV}/c)^{1.01}}\right)$ cm. Also, the χ^2 of the track fit with constraint to the primary vertex needs to be below 36 in that case.

C.3. Hybrid Tracks

C.3.1. Default Hybrid Tracks

Charged tracks are reconstructed with the combined information of ITS and TPC. A hybrid reconstruction technique [119] is used to obtain a uniform azimuthal acceptance in $|\eta| < 0.9$. The hybrid tracks consist of two diverse track classes: (i) tracks containing at least three (out of six) hits in the ITS, including at least one hit in the SPD, and (ii) tracks with fewer than three hits in the ITS or no SPD hit. To improve the determination of the transverse momentum, tracks of class (ii) are constrained to the primary vertex. The momentum resolution $\delta p_T/p_T$ is about 1% at $p_T = 1 \text{ GeV}/c$ for all accepted tracks and approximately 4% at $p_T = 40 \text{ GeV}/c$ for 95% of the tracks [34]. The remaining 5% are tracks without a hit in the ITS. Their momentum resolution is 7% at $p_T = 40 \text{ GeV}/c$.

Only tracks with a Distance of Closest Approach (DCA) to the vertex smaller than 2.4 cm in the plane transverse to the beam and less than 3.2 cm in beam direction are accepted by the hybrid cuts. These cuts suppress the contribution of secondary particles stemming from weak decays or from interaction with detector material and beam pipe.

The jet finding algorithms require a homogeneous acceptance because inhomogeneities can bias the results. Hence, the default hybrid tracks are used for the jet finding.

C.3.2. Modified Hybrid Tracks

The particle spectra and fragmentation functions are not as sensitive to a uniform azimuthal acceptance as the jet finding, but they can suffer more from secondary particle contamination. The default hybrid track cuts are changed accordingly to accept only tracks that have an ITS refit, called *modified hybrid tracks* in this work. In particular, this requirement removes secondary tracks produced outside the ITS. The reduction of secondaries from weak decays also removes a good fraction of daughter tracks from strange mothers. This makes the analysis less sensitive to the strangeness scaling (cf. Section 5.1.2) and reduces the corresponding systematic uncertainty.

C.4. Geometrical Cut

The geometrical cut was proposed [142] to reject short tracks or tracks with large overlap with the TPC sector boundaries. The cut removes tracks with worse dE/dx resolution or that have potentially biased dE/dx . The former is due to a lower number of PID clusters. The latter may arise from a bad calibration of pads close to the boundaries. This happens, if the energy of Krypton decays used for the calibration is partially deposited in the dead zones, such that a reconstructed cluster at the sector boundaries does not contain the full energy.

Basically, the length of the track in the active zone, the number of PID clusters and the number of crossed pad rows are evaluated for this cut. Since it is largely a purely geometrical cut, the effect of the cut is well reproduced in MC simulations [142].

To the time of writing this thesis, the cut parameters have not been optimised yet. For the current parameters, the track survival rate around $p_T = 1 \text{ GeV}/c$ is reduced considerably by about 30% for the geometrical cut compared to the default cut, namely requiring at least 60 PID clusters (cf. Section 6.3.10). The reduction is around 5% at higher p_T . That is why the cut has not been used as a default, but only as a cross-check in this work. The considerable efficiency reduction at $p_T = 1 \text{ GeV}/c$ happens because the tracks start to become straight and can largely overlap with the TPC sector boundaries. Such tracks may still survive the default cuts, but are rejected by the (stricter) geo cut. With increasing p_T , the curvature of the tracks decreases further and both cuts become more similar.

C.5. Global Tracks and Cut Variations

Global tracks are tracks which have SPD hits and an ITS refit. They are a subset of the modified hybrid tracks, which are in turn a subset of the default hybrid tracks. In Section 6.3.10, 4 different sets of global tracks have been considered for cut variations. Their details are presented in the following list:

1. Global tracks with loose cut on the Distance of Closest Approach (DCA):

This cut set is basically the same as the ITS-TPC standard track cuts 2010 described in Appendix C.2. The most important difference is the cut on the DCA to the vertex, which is demanded to be smaller than 2.4 cm in the plane transverse to the beam and less than 3.2 cm in beam direction. In addition, the cut on the number of tracking clusters in the TPC is p_T dependent for global tracks: $N_{\text{cl}} > 70 + \frac{30}{20} p_T / (\text{GeV}/c)$.

Appendix C. Track Selection and Cuts

2. Global tracks with tight cut on the DCA:

This cut set is identical to that in the previous point, but with a p_T dependent DCA cut in the plane transverse to the beam direction:

$$DCA_{xy} < \left(0.0105 + \frac{0.0350}{(p_T/(\text{GeV}/c))^{1.01}} \right) \text{ cm}.$$

3. Global tracks with fewer demanded TPC tracking clusters:

This cut set corresponds to that in the first point, but the p_T dependent cut on the number of tracking clusters in the TPC has a lower offset: $N_{\text{cl}} > 60 + \frac{30}{20} p_T/(\text{GeV}/c)$.

4. Cut on the number of crossed TPC rows:

This cut set corresponds to that in the first point, but the cut on the number of TPC tracking clusters is replaced by the requirement that the track has crossed at least 70 TPC pad rows and the ratio between crossed rows and findable clusters is at least 80%. The latter is a measure of the track length fraction that does not end up in dead zones of the TPC sector boundaries.

Appendix D.

MC Blind Test Results

Additional plots for the MC blind test (see Section 4.4) are shown in this appendix for the inclusive particle production (Fig. D.1) and the results fitted versus z (Fig. D.2) and ξ (Fig. D.3) in jets. It is demonstrated in Fig. D.4 that the fit is also capable of reproducing the electron fraction of the MC truth in jets. The upper panels of each plot show the uncorrected yield fractions, while the lower panels contain the fractions scaled (means and errors) to the fit, i.e. the fit is at unity. In all cases, the fit result is in agreement with the MC truth within uncertainties. In particular, the results of Fig. D.3 demonstrate that the Multi-Template Fit (MTF) is capable of extracting the particle content in bins of ξ .

Appendix D. MC Blind Test Results

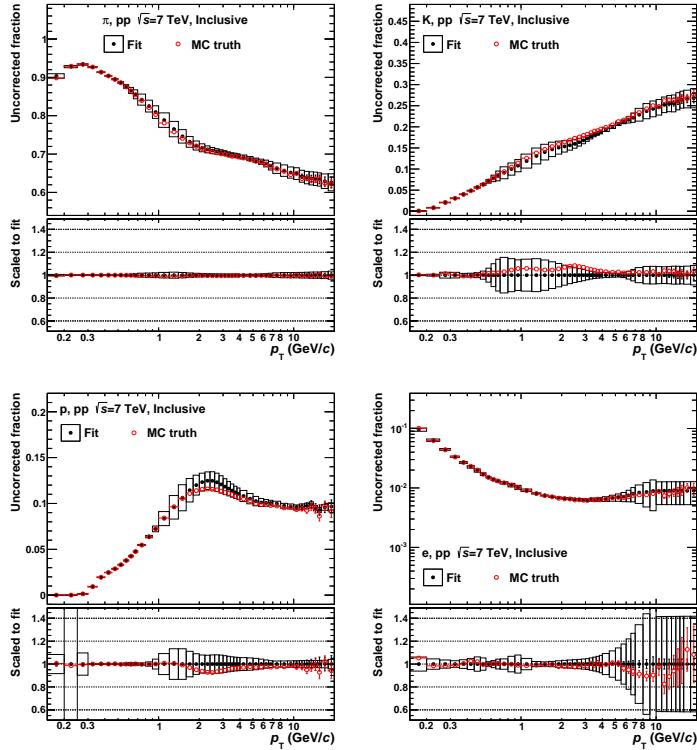


Figure D.1.: MC blind test for the inclusive particle production. The raw fractions from fitting the MC output (black) are compared to the MC truth (red) for (from upper left to lower right): pions, kaons, protons and electrons.

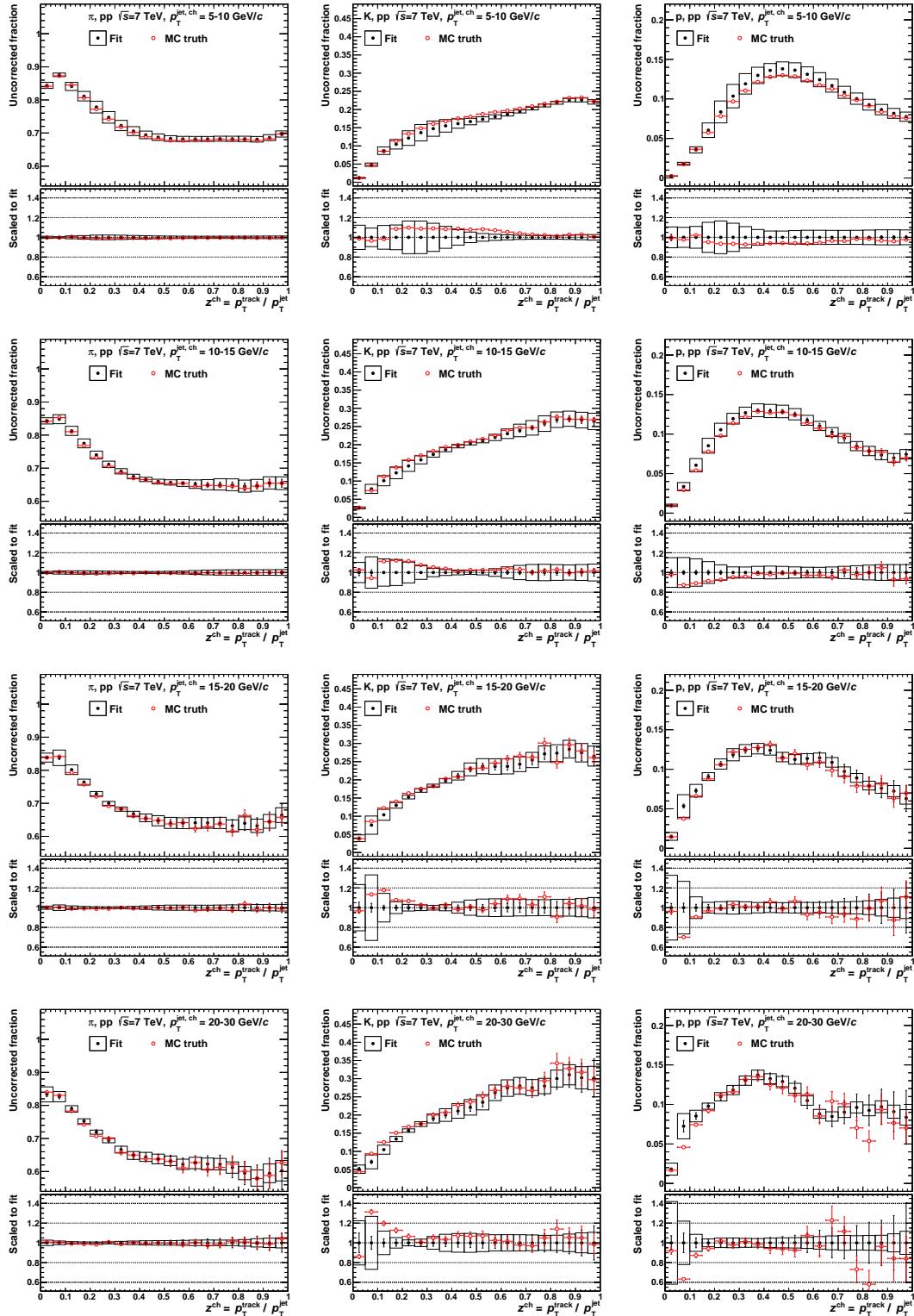


Figure D.2.: MC blind test for fitting versus z . The raw fractions from fitting the MC output (black) are compared to the MC truth (red) for (from left to right): pions, kaons and protons. The rows show (from top to bottom): $p_T^{\text{jet}} = 5 - 10$, $10 - 15$, $15 - 20$ and $20 - 30 \text{ GeV}/c$.

Appendix D. MC Blind Test Results

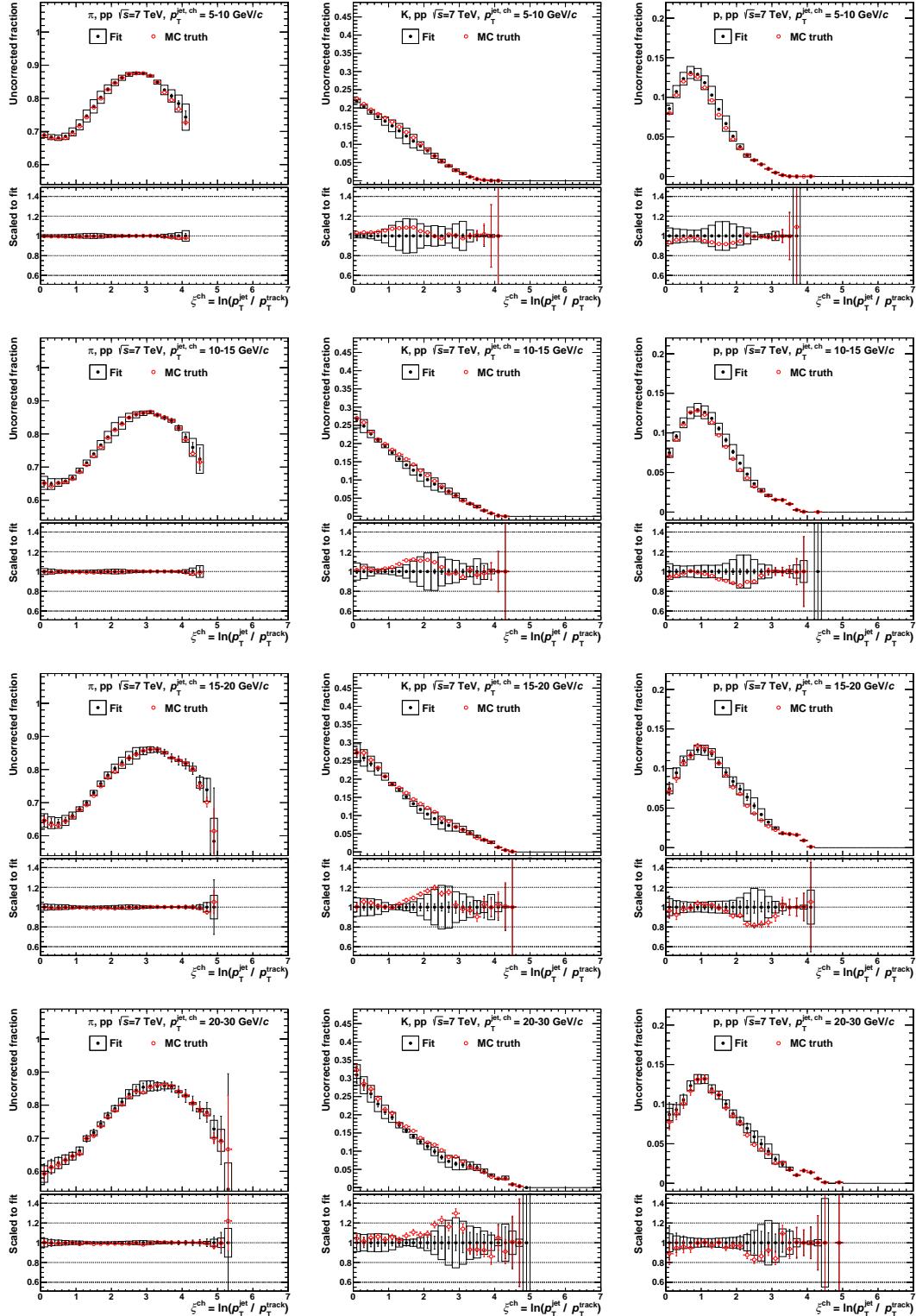


Figure D.3.: MC blind test for fitting versus ξ . The raw fractions from fitting the MC output (black) are compared to the MC truth (red) for (from left to right): pions, kaons and protons. The rows show (from top to bottom): $p_T^{\text{jet}} = 5 - 10$, $10 - 15$, $15 - 20$ and $20 - 30$ GeV/ c .

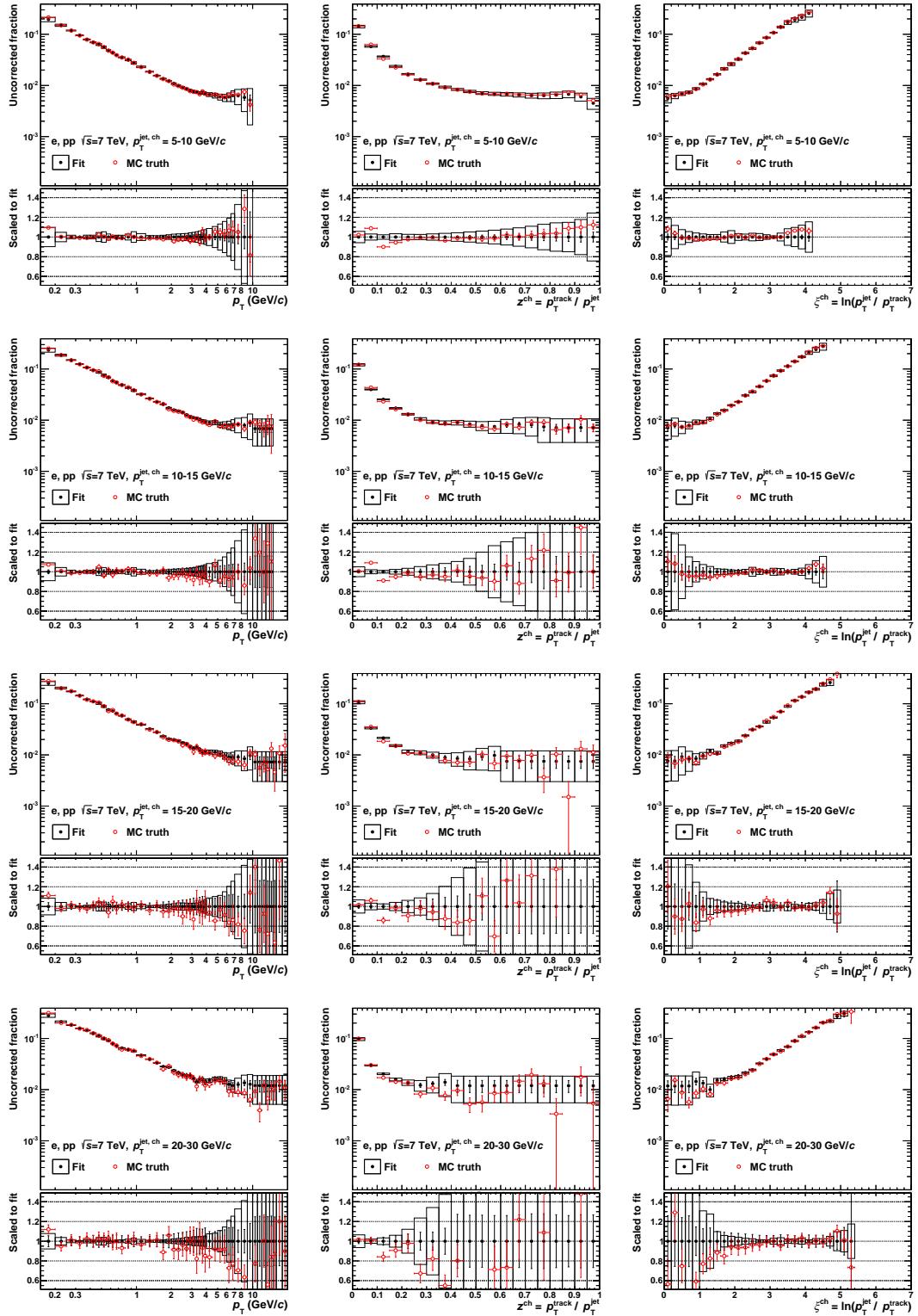


Figure D.4.: MC blind test for electrons in jets. The raw fractions from fitting the MC output (black) are compared to the MC truth (red) for fitting versus (from left to right): p_T , z and ξ . The rows show (from top to bottom): $p_T^{\text{jet}} = 5 - 10$, $10 - 15$, $15 - 20$ and $20 - 30 \text{ GeV}/c$.

Appendix E.

Systematic Uncertainties of MC Corrections

In Section 5.3, the systematic uncertainties of the MC correction factors have been discussed. This appendix comprises further discussions on the estimation of the systematic uncertainties of the secondary correction, as well as the full set of plots showing the contributions from all error sources for all considered p_T^{jet} bins.

E.1. Secondary Correction

The procedure to estimate the relative systematic uncertainty of the secondary correction has been discussed in Section 5.3.3. The *kaon correction* was introduced in order to remove residual biases from the efficiency correction. Indeed, without this correction, the estimation procedure does not work properly. This becomes evident when looking at Fig. E.1, which shows the double ratio for pions without application of the kaon correction. The horizontal red and green lines are not important in this consideration. Above $p_T = 1 \text{ GeV}/c$, the double ratio drops steeply. This happens due to the decreasing secondary correction factor with p_T , whereas the efficiency bias is rather constant or even increasing. More dangerous with respect to the estimation procedure is the change of slope at low p_T . Obviously, efficiency and secondary contamination effects are mixed up to some extent at low p_T .

On the one hand, for protons, the secondary contamination is considerably larger at low p_T and dominates possible efficiency effects. On the other hand, the precise value from the fit does not matter much for the pions, since the overall correction factor is anyway small. Altogether, this makes the estimation procedure sufficiently accurate, if the kaon correction is utilised. This will be the case from now on.

The double ratios for the inclusive case are shown in Fig. E.2. The standard cuts are compared to the modified hybrid track cuts. The overall trends are similar, but the size of the spread differs due to the individual correction factors entering the denominator. Note that the fits exhibit an additional degree of freedom compared to the jet case. The linear fit is allowed to stay at a constant

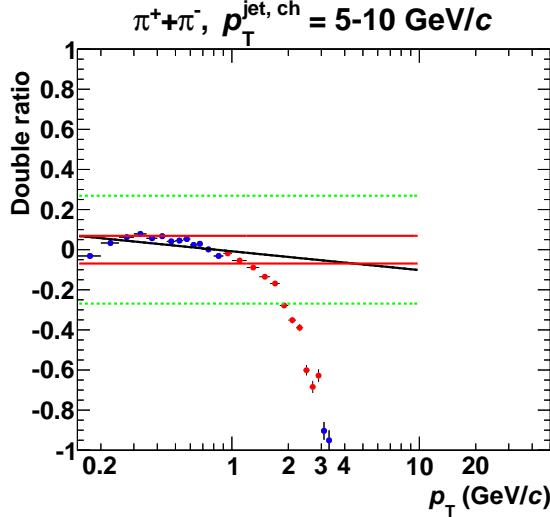


Figure E.1.: Double ratio for pions in jets with $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$ without applying the kaon correction. The red points have been excluded from the fit.

value above a floating threshold, which evidently improves the description of the data.

The fit results eventually enter the summary plots for the inclusive case, depicted in Fig. E.3. The error estimate (open red points) covers the spread due to DCA cut variations (open blue points) up to $p_T = 1 - 2 \text{ GeV}/c$. Above this threshold, efficiency effects slowly start dominating, such that it is not an issue to see the red points dropping below the blue ones.

As was discussed in Section 5.3.3, the values from the fit versus p_T are used to estimate the uncertainties of the correction factors as a function of z . The corresponding summary plots are shown in Figs. E.4 and E.5 for protons and pions, respectively. As is visible in the plots, the errors in the low- z region are reasonably well described. Note that, depending on p_T^{jet} , the dE/dx crossing regions and also the high- p_T region are reached already at rather moderate z . Overall, this supports that the fit values from the estimation versus p_T can simply be adopted for the one versus z .

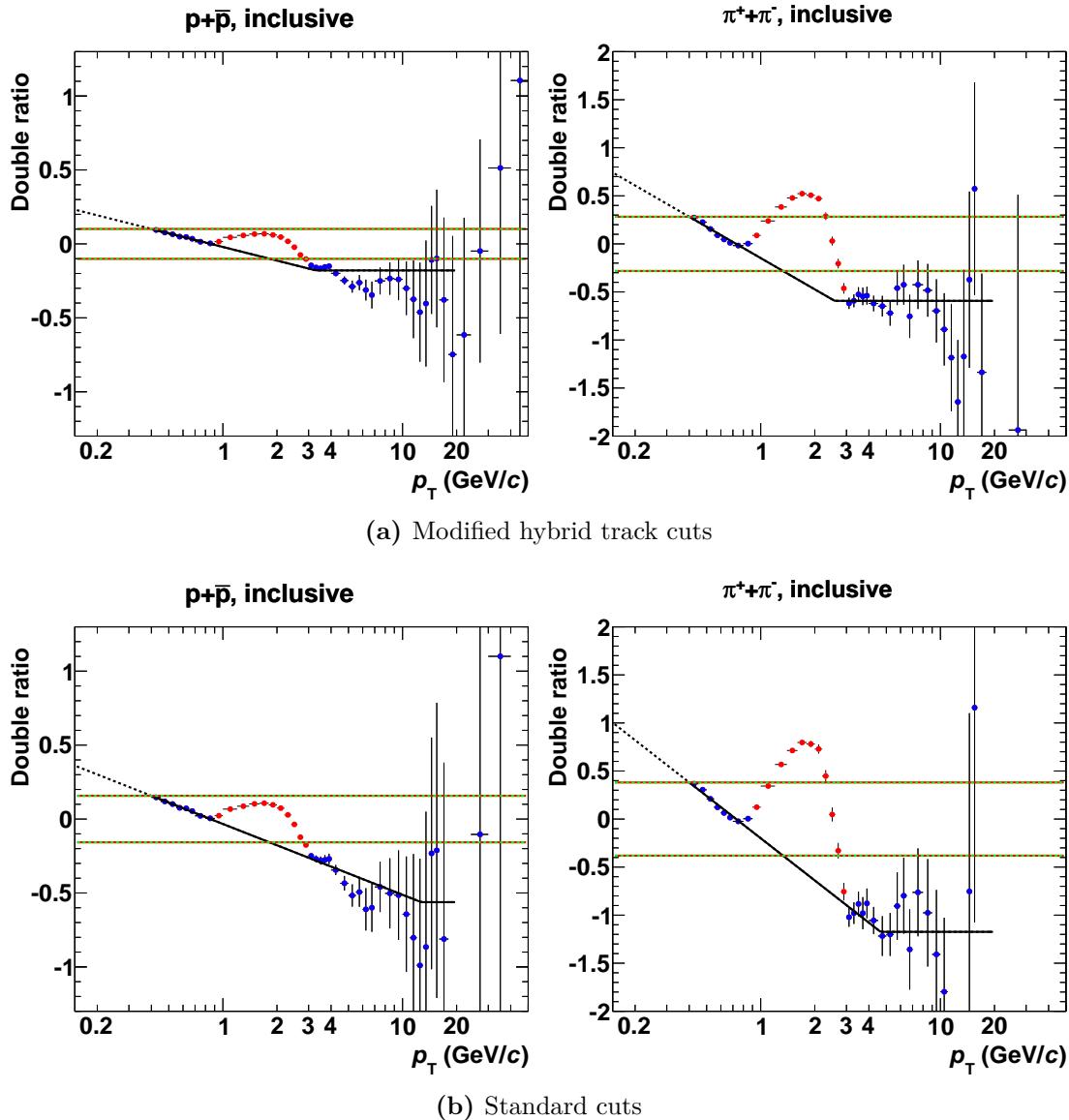


Figure E.2.: Double ratio for protons (left) and pions (right) for the inclusive case. The results for the modified hybrid track cuts are shown in **a**), those for the standard cuts in **b**). The red points have been excluded from the fit.

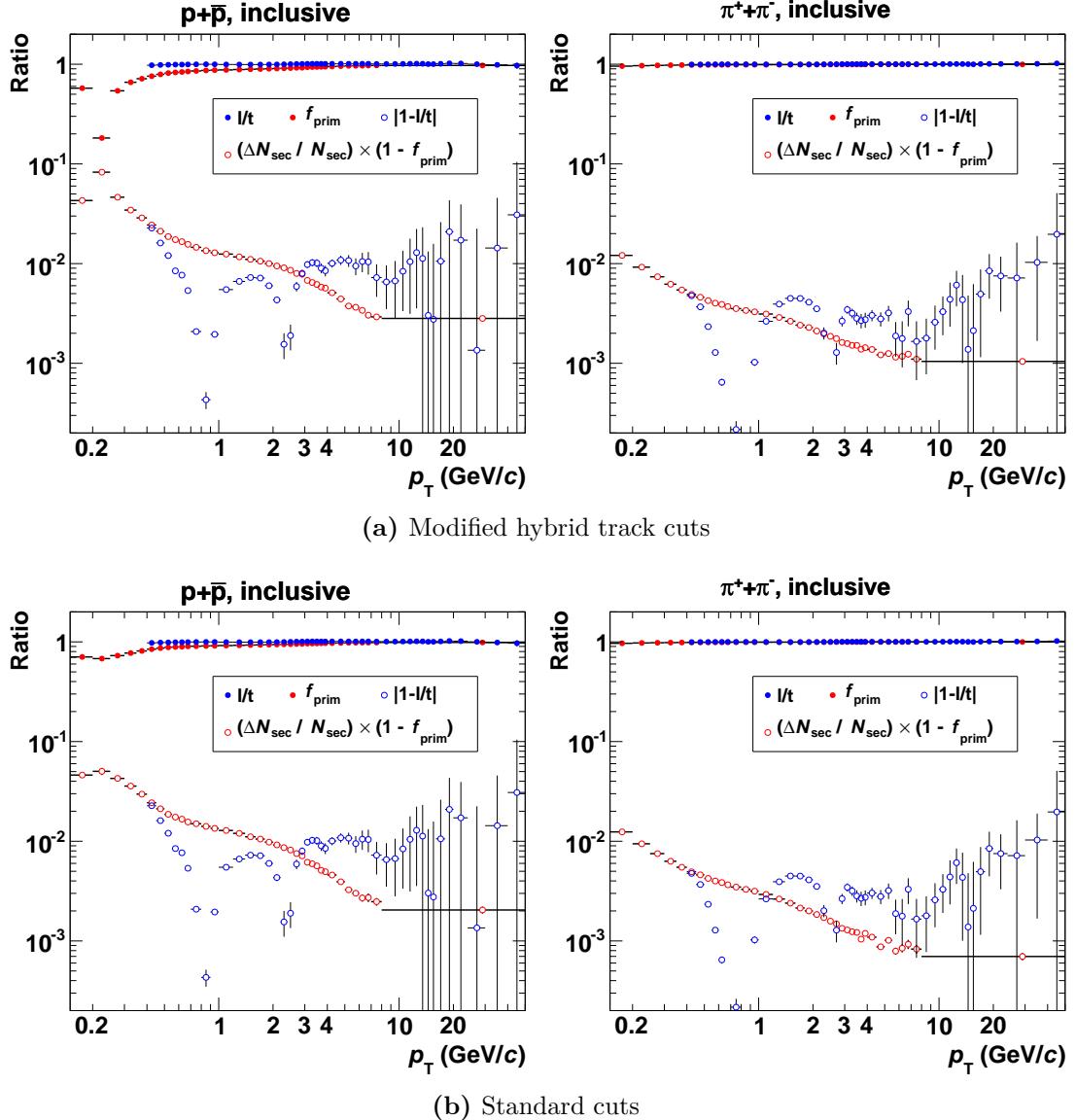


Figure E.3.: Summary plot for the systematic uncertainty estimation of the secondary correction for protons (left) and pions (right) in the inclusive case. The modified hybrid track cuts (**a**) are compared to the standard cuts (**b**). The open red points show the relative systematic uncertainty estimate and should be compared to the open blue points, which show the relative change of spectra due to DCA cut variations.

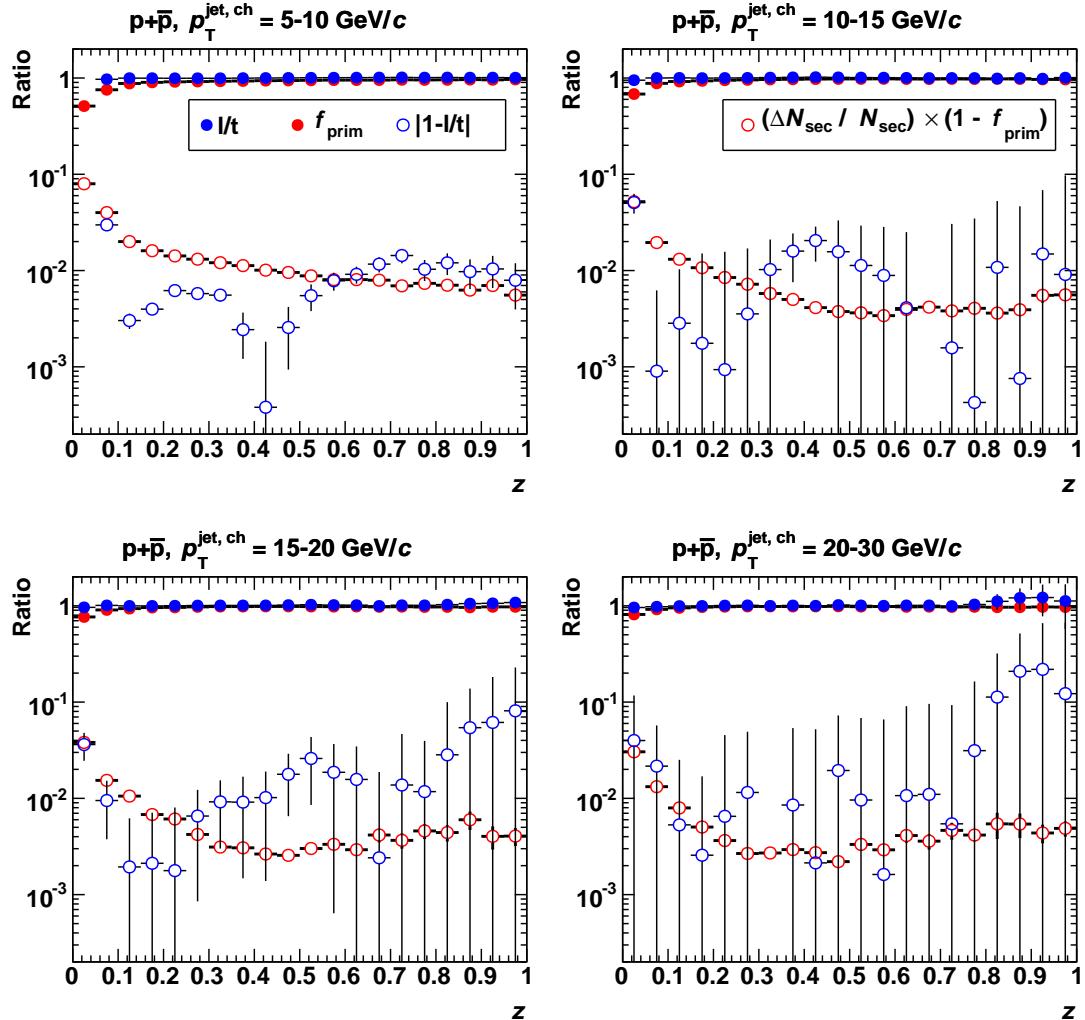


Figure E.4.: Summary plot for the systematic uncertainty estimation of the secondary correction for protons in jets as a function of z . The open red points show the relative systematic uncertainty estimate and should be compared to the open blue points, which show the relative change of spectra due to DCA cut variations.

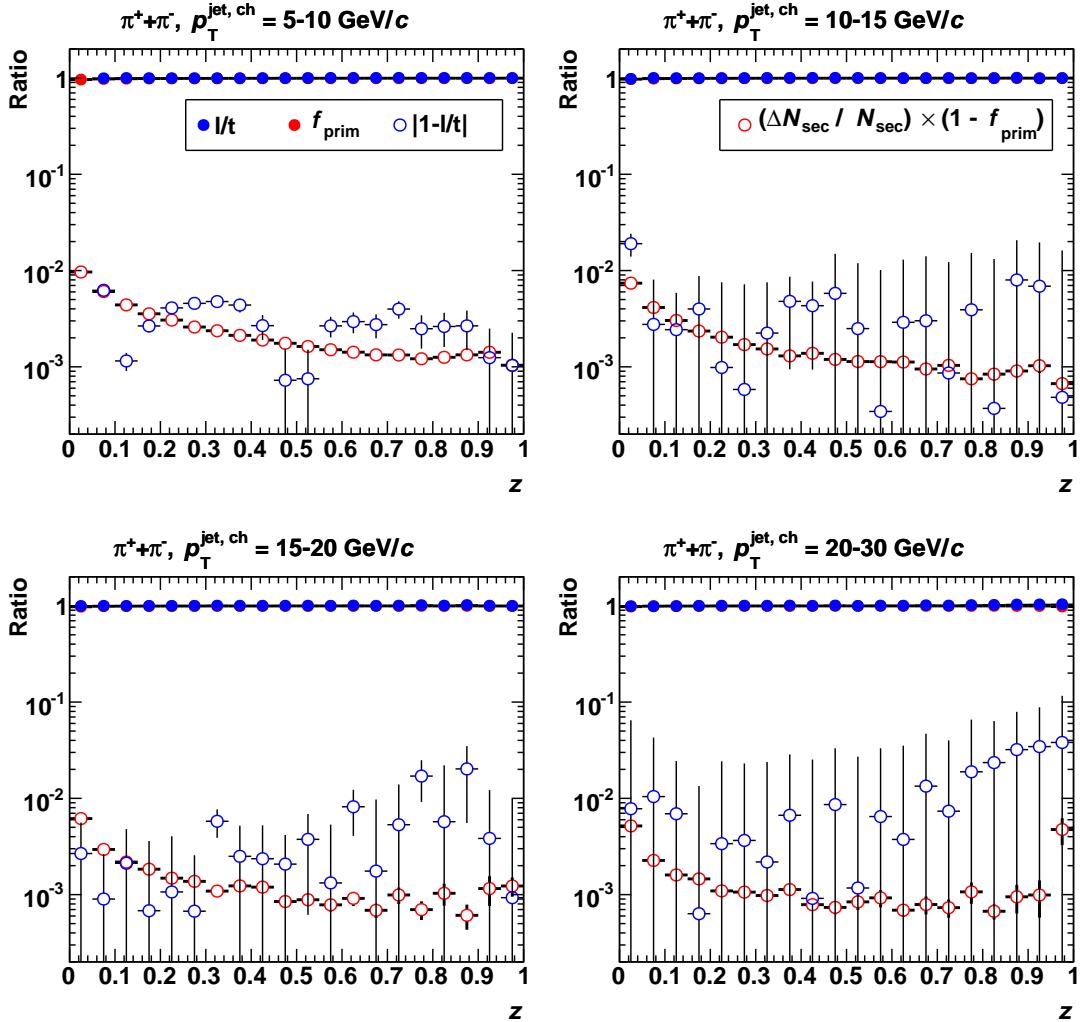


Figure E.5.: Summary plot for the systematic uncertainty estimation of the secondary correction for pions in jets as a function of z . The open red points show the relative systematic uncertainty estimate and should be compared to the open blue points, which show the relative change of spectra due to DCA cut variations.

E.2. Relative Systematic Uncertainties from Various Sources

It has been discussed in Section 5.3.9 that the MC correction factors have systematic uncertainties from various sources. The relative systematic uncertainties for the pion, kaon and proton yields are shown in Figs. E.6, E.7 and E.8, respectively, and for K/π and p/π in Figs. E.9 and E.10. The corresponding discussion is included in Section 5.3.9. As for the figures in that section, the coloured histograms are the linear sum of all errors from the red histogram up to the corresponding error source. The dashed black curve is the total systematic uncertainty defined as the square root of the quadratic sum of all individual errors.

Appendix E. Systematic Uncertainties of MC Corrections

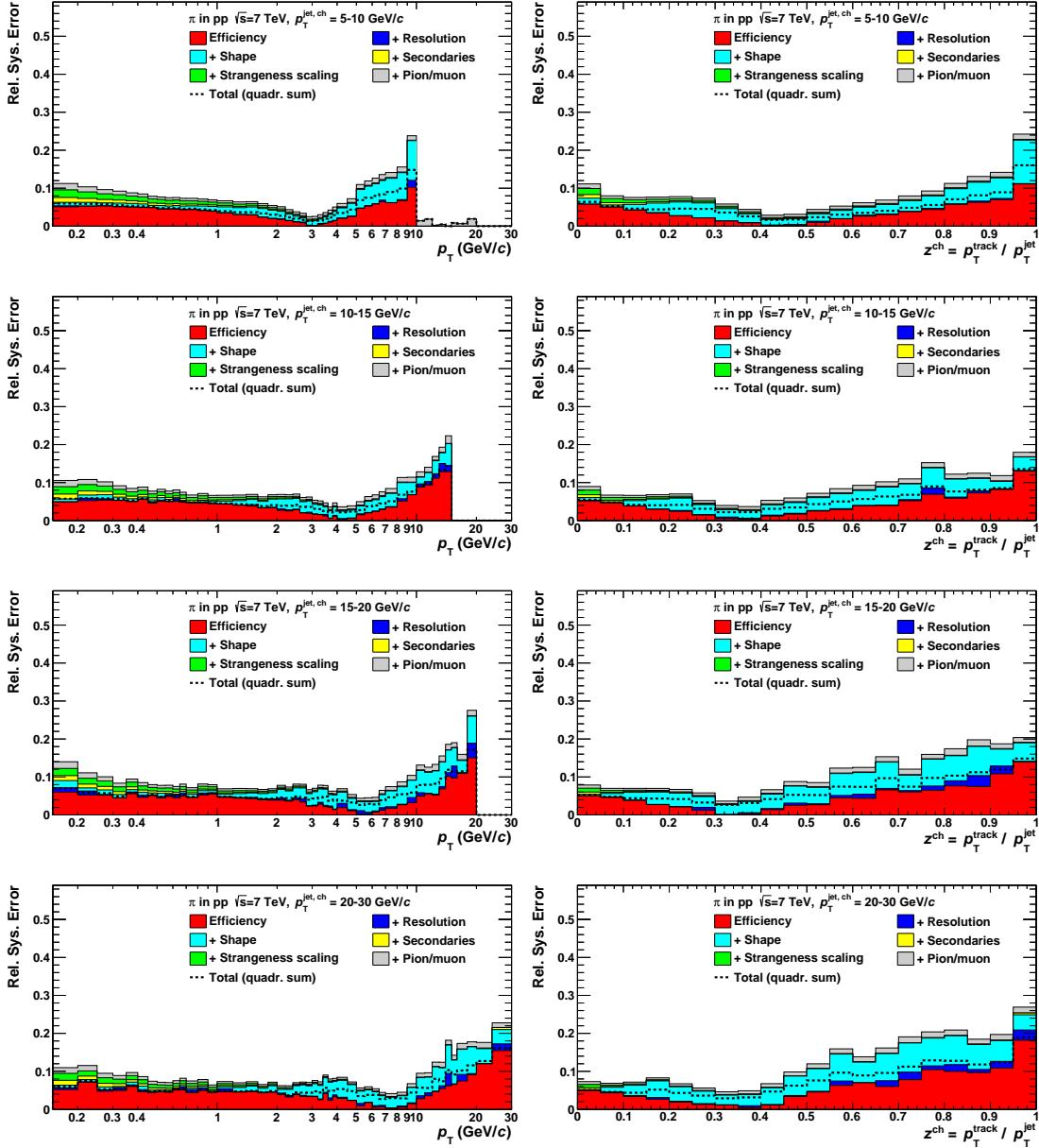


Figure E.6.: Relative systematic uncertainties of the MC correction factors for the pion yields. The rows show (from top to bottom) $p_T^{\text{jet}} = 5 - 10, 10 - 15, 15 - 20$ and $20 - 30 \text{ GeV}/c$, respectively, as a function of p_T^{track} (left column) and z (right column).

E.2. Relative Systematic Uncertainties from Various Sources

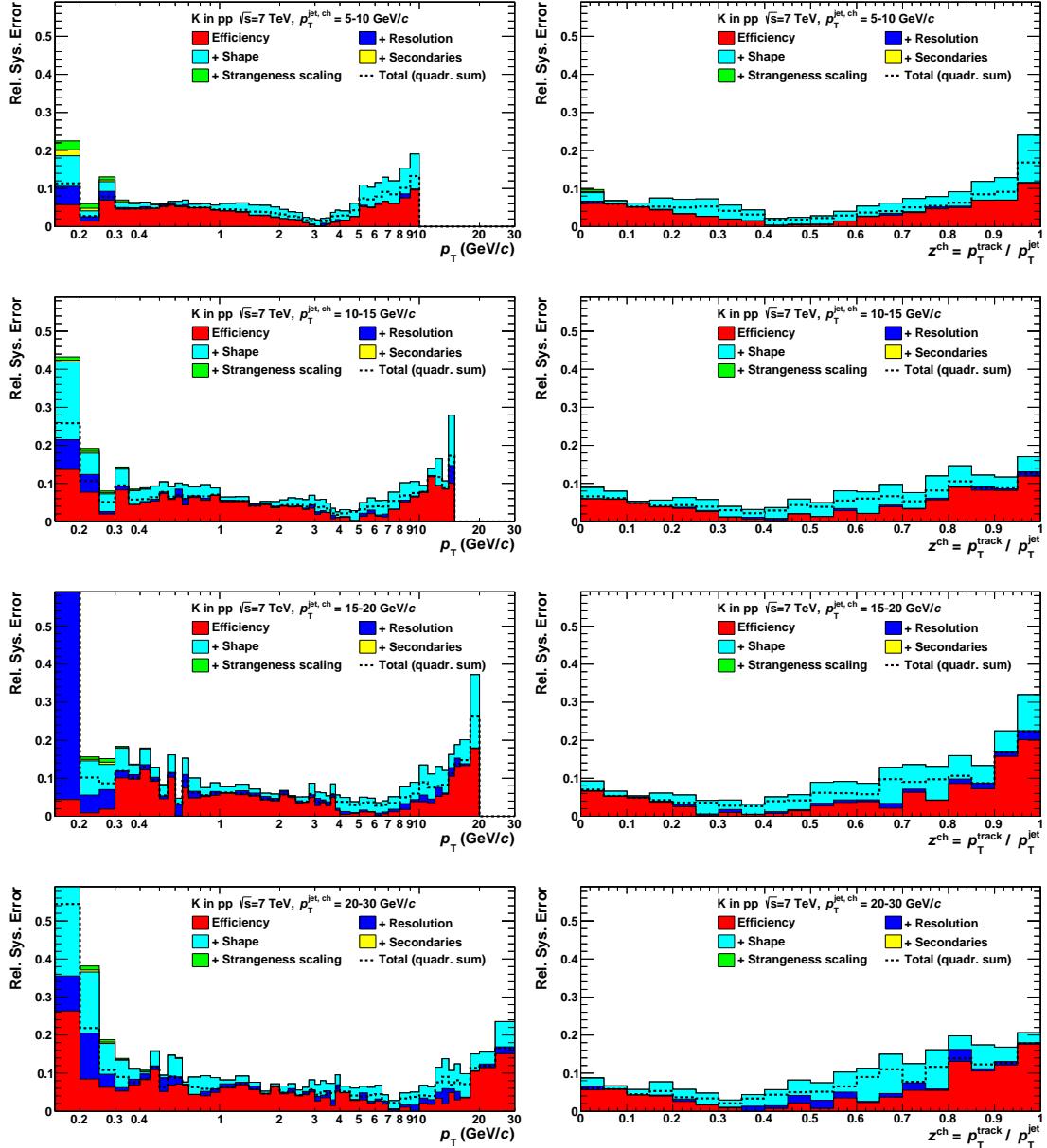


Figure E.7.: Relative systematic uncertainties of the MC correction factors for the kaon yields. The rows show (from top to bottom) $p_T^{\text{jet}} = 5 - 10, 10 - 15, 15 - 20$ and $20 - 30 \text{ GeV}/c$, respectively, as a function of p_T^{track} (left column) and z (right column).

Appendix E. Systematic Uncertainties of MC Corrections

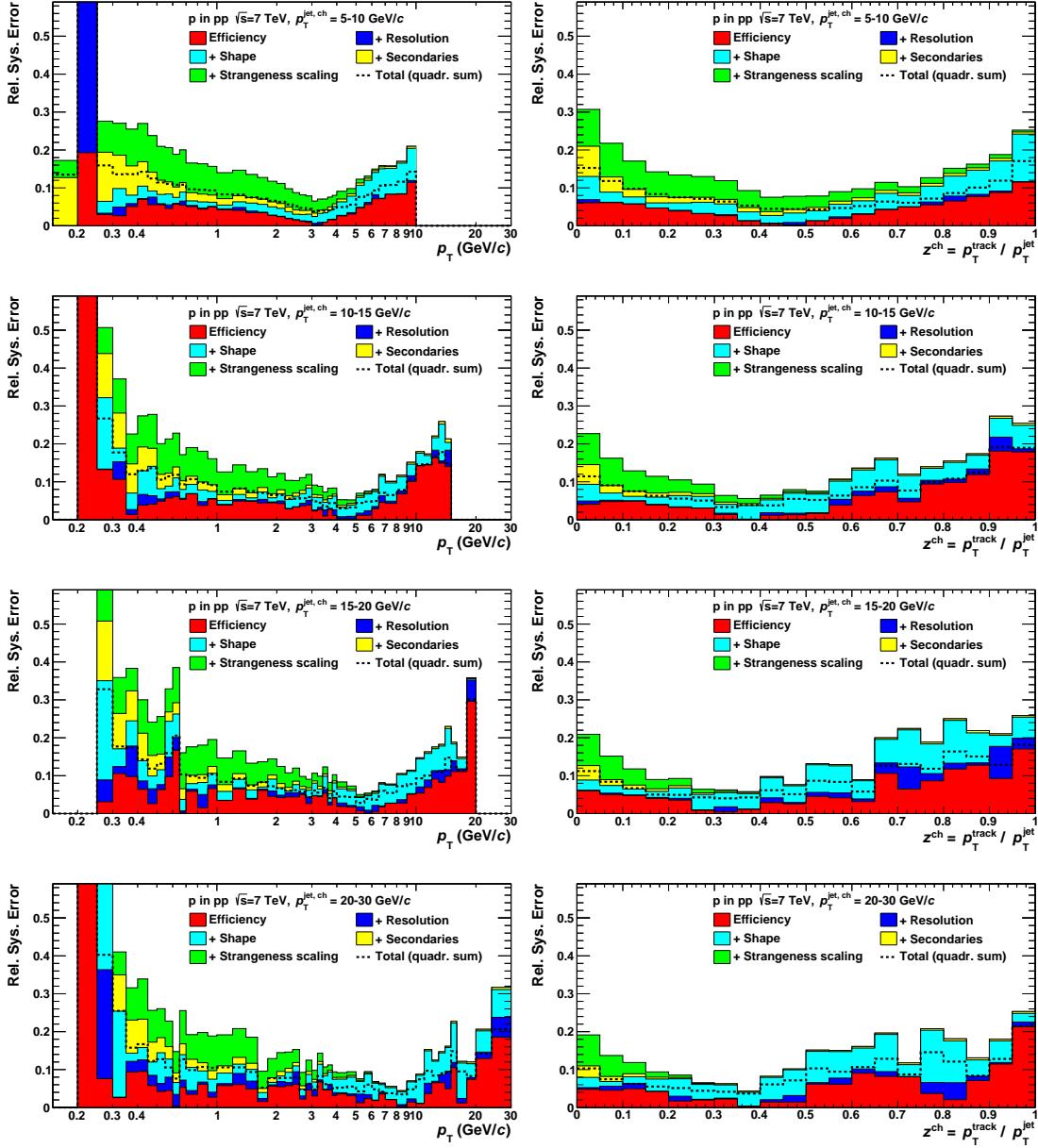


Figure E.8.: Relative systematic uncertainties of the MC correction factors for the proton yields. The rows show (from top to bottom) $p_T^{\text{jet}} = 5 - 10, 10 - 15, 15 - 20$ and $20 - 30 \text{ GeV}/c$, respectively, as a function of p_T^{track} (left column) and z (right column).

E.2. Relative Systematic Uncertainties from Various Sources

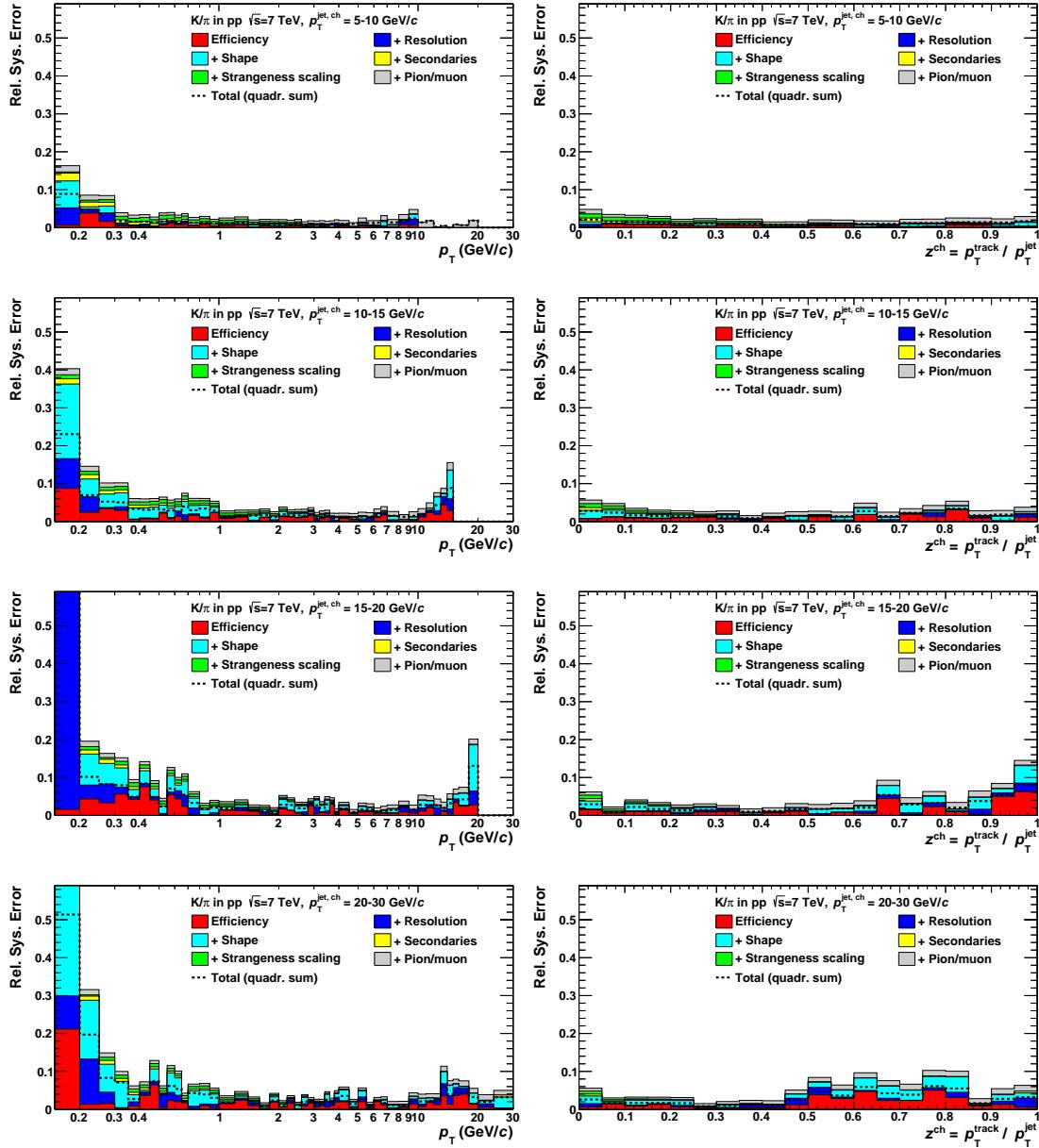


Figure E.9.: Relative systematic uncertainties of the MC correction factors for the top-pion ratios of kaons. The rows show (from top to bottom) $p_T^{\text{jet}} = 5-10, 10-15, 15-20$ and $20-30 \text{ GeV}/c$, respectively, as a function of p_T^{track} (left column) and z (right column).

Appendix E. Systematic Uncertainties of MC Corrections

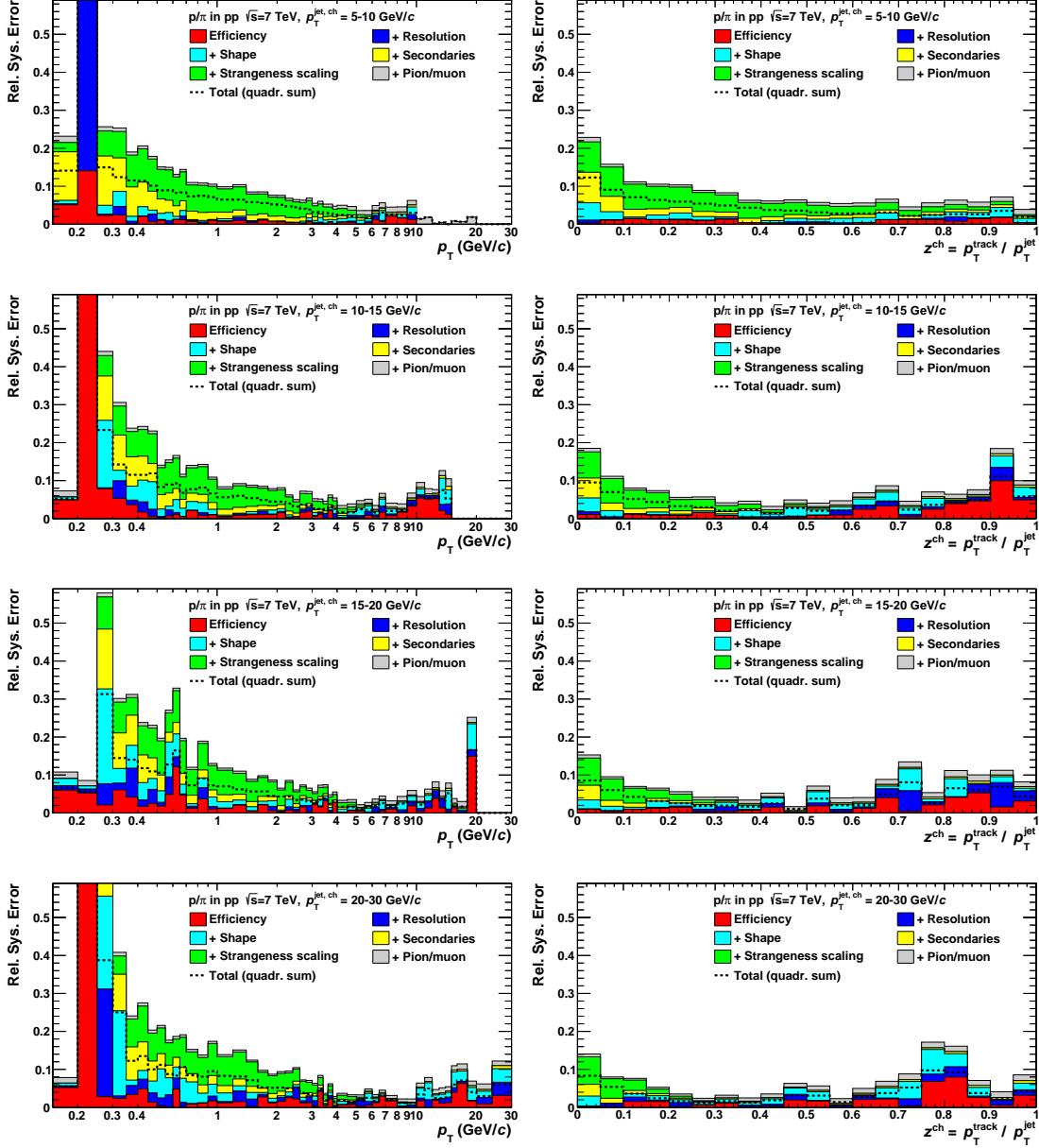


Figure E.10.: Relative systematic uncertainties of the MC correction factors for the to-pion ratios of protons. The rows show (from top to bottom) $p_T^{\text{jet}} = 5 - 10, 10 - 15, 15 - 20$ and $20 - 30 \text{ GeV}/c$, respectively, as a function of p_T^{track} (left column) and z (right column).

Appendix F.

Fit Quality Plots

In this appendix, further plots showing the fit quality in various bins and data samples are collected.

F.1. Multiplicity Analysis

In Section 6.5.2, it has been demonstrated with a few example plots that the fit quality is similar in all multiplicity bins. Further fit quality plots are collected in this section. The Figs. F.1 and F.2 are for tracks with $p_T = 0.8 - 0.9$ and $8 - 10 \text{ GeV}/c$, respectively, in bins of reference multiplicity. Similarly, the Figs. F.3, F.4 and F.5 show the fit quality in bins of V0M percentile for tracks with $p_T = 0.8 - 0.9, 4 - 4.5$ and $8 - 10 \text{ GeV}/c$, respectively. Note that the lower the percentile becomes, the higher is the event multiplicity. In general, the fit provides a good description of the data and the quality is largely independent of the multiplicity (for both estimators). Yet, the deviation at low Δ' for $p_T = 0.8 - 0.9 \text{ GeV}/c$ may appear to be dangerous at first glance. In the end, this deviation is not an issue because it has little statistical weight (with respect to the total integrals of a given species).

Appendix F. Fit Quality Plots

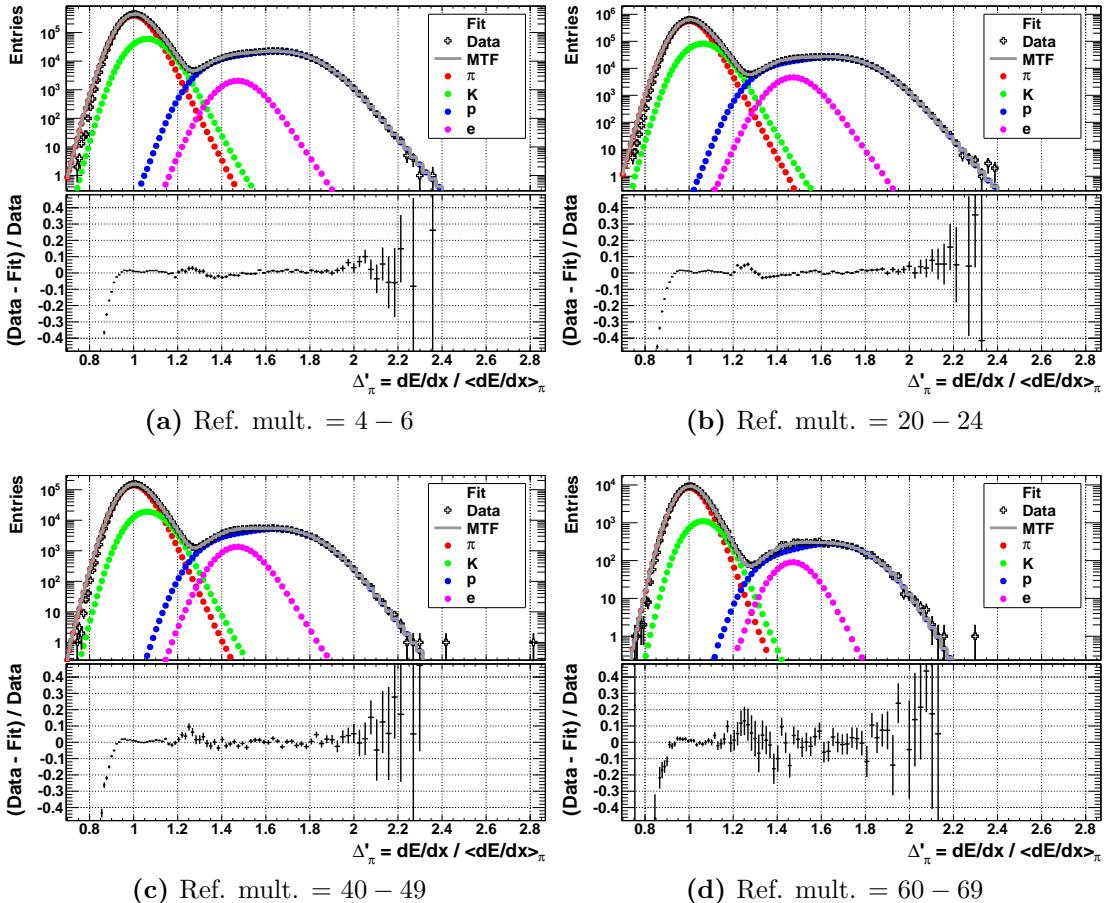


Figure F.1.: Template fits to data for different reference multiplicity bins for $p_T = 0.8 - 0.9 \text{ GeV}/c$. The multiplicity increases from the upper left to the lower right plot.

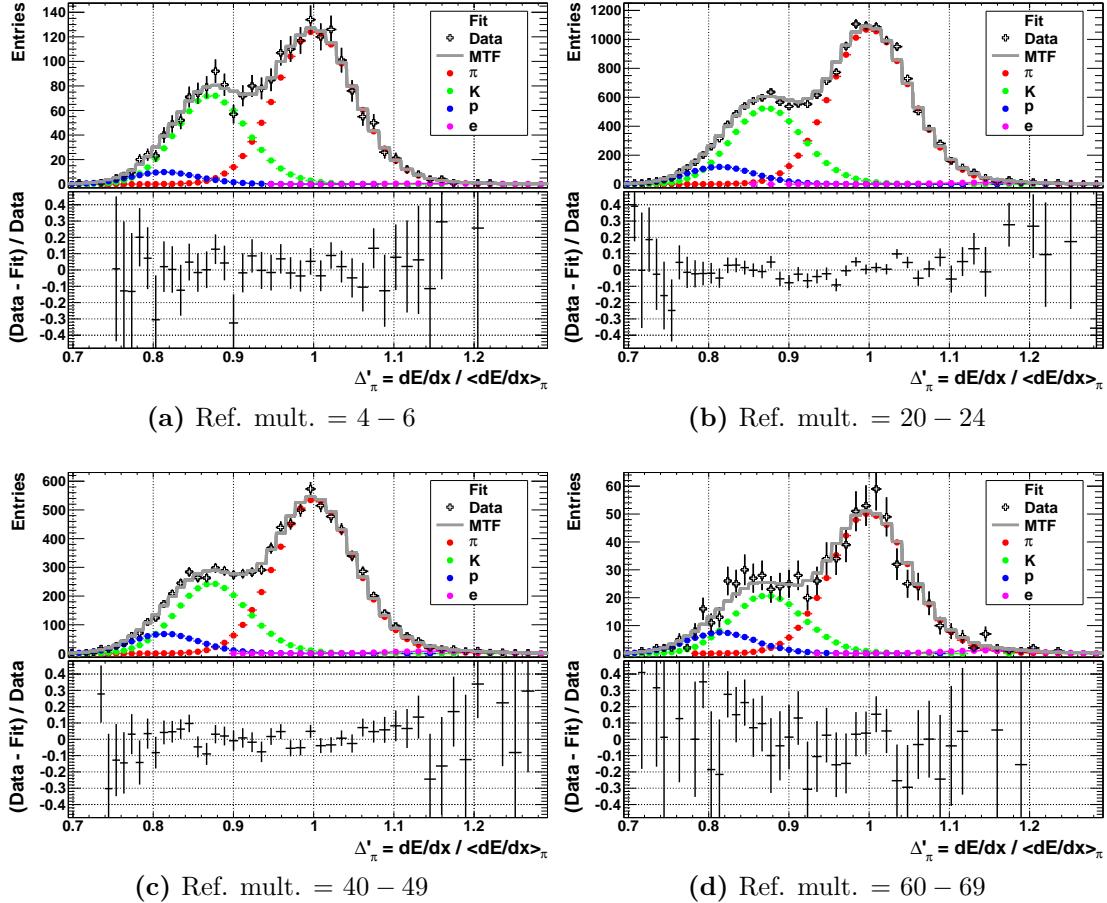


Figure F.2.: Template fits to data for different reference multiplicity bins for $p_T = 8 - 10 \text{ GeV}/c$. The multiplicity increases from the upper left to the lower right plot.

Appendix F. Fit Quality Plots

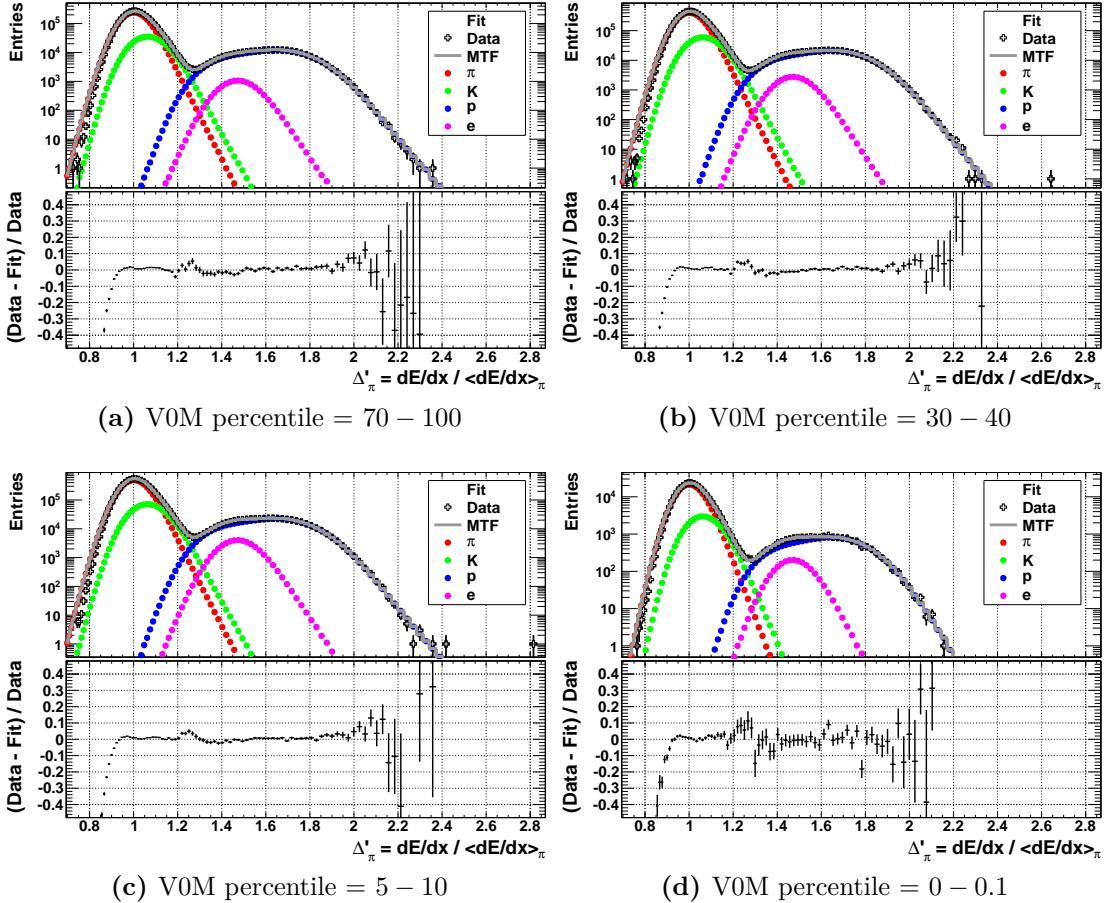


Figure F.3.: Template fits to data for different V0M percentile bins for $p_T = 0.8 - 0.9 \text{ GeV}/c$. The V0M percentile decreases from the upper left to the lower right plot.

F.1. Multiplicity Analysis

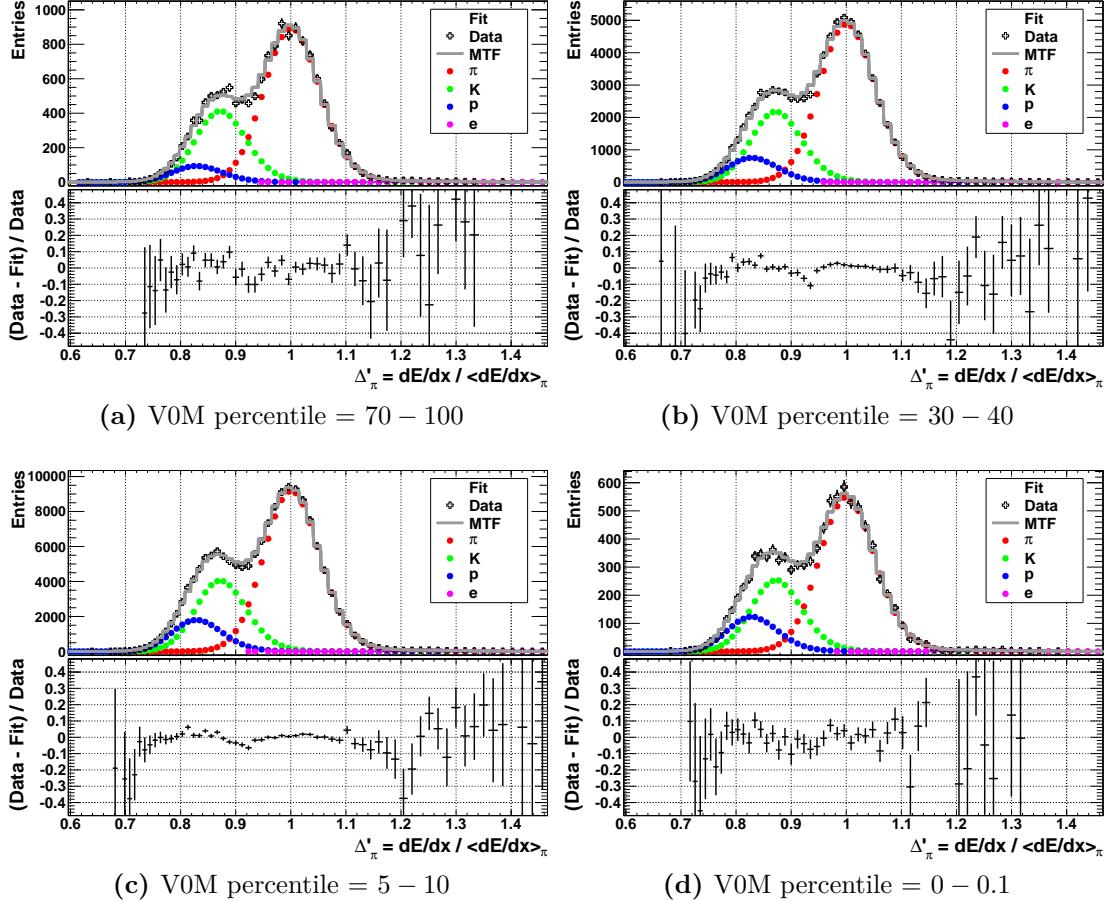


Figure F.4.: Template fits to data for different V0M percentile bins for $p_T = 4 - 4.5 \text{ GeV}/c$. The V0M percentile decreases from the upper left to the lower right plot.

Appendix F. Fit Quality Plots

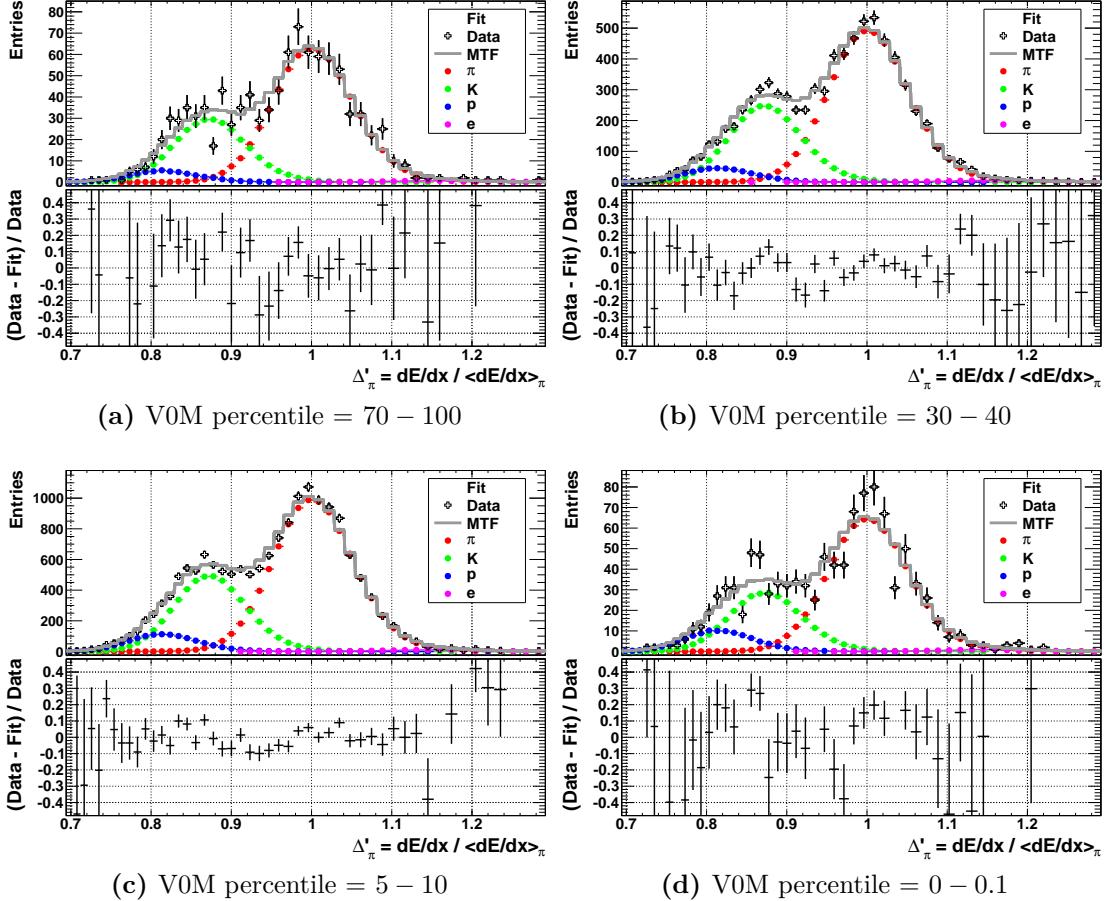


Figure F.5.: Template fits to data for different V0M percentile bins for $p_T = 8 - 10 \text{ GeV}/c$. The V0M percentile decreases from the upper left to the lower right plot.

F.2. Jet Analysis

The fit quality plots for particles in jets with $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ in data are shown in Fig. F.6 for different p_T^{track} bins. The figure shows the same results as Fig. 6.2, but with a linear vertical axis. The corresponding fit quality plots in bins of z are shown in Fig. F.7. The fit describes the data well in all presented z bins.

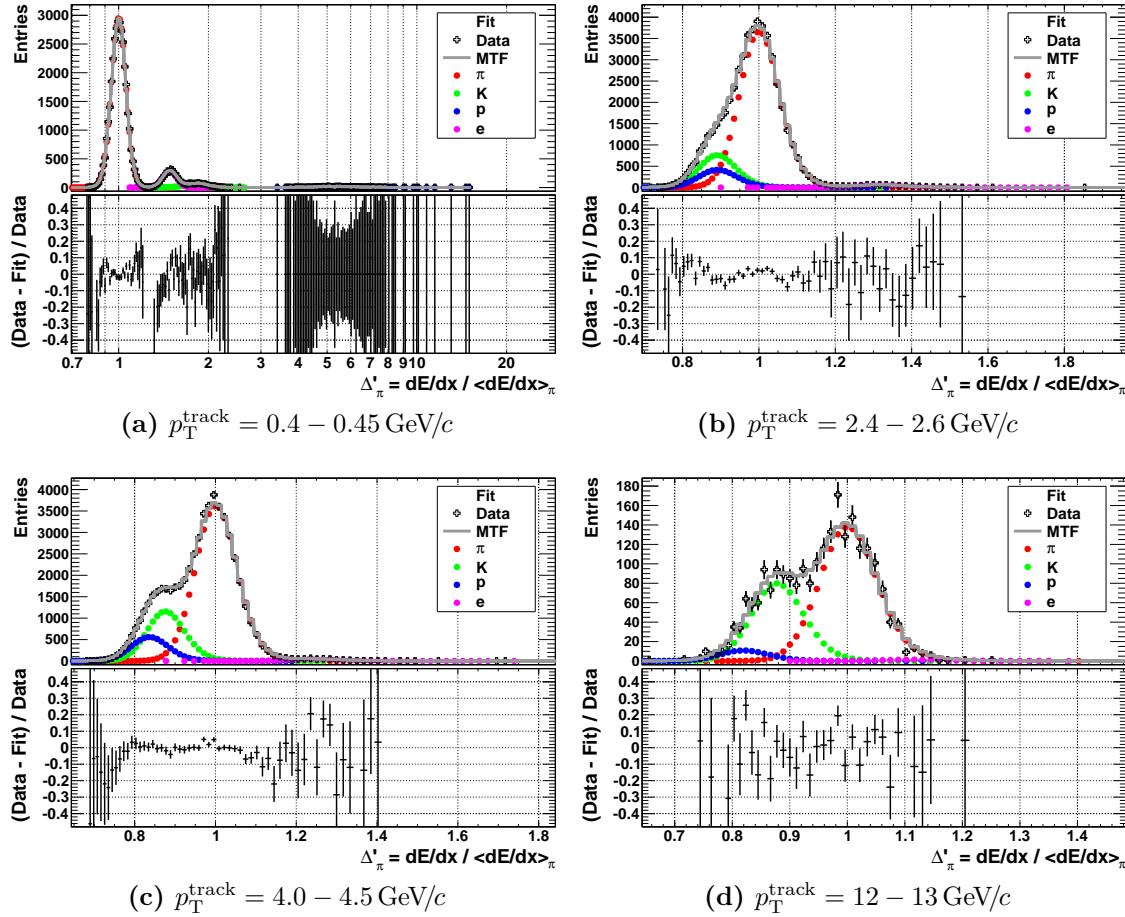


Figure F.6.: Template fits to data in different p_T^{track} intervals for $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$. The panels a) – d) show the following p_T^{track} ranges in GeV/c : $0.4 - 0.45$, $2.4 - 2.6$, $4.0 - 4.5$ and $12 - 13$. Note the different ranges of the abscissa and its logarithmic scale in a).

Appendix F. Fit Quality Plots

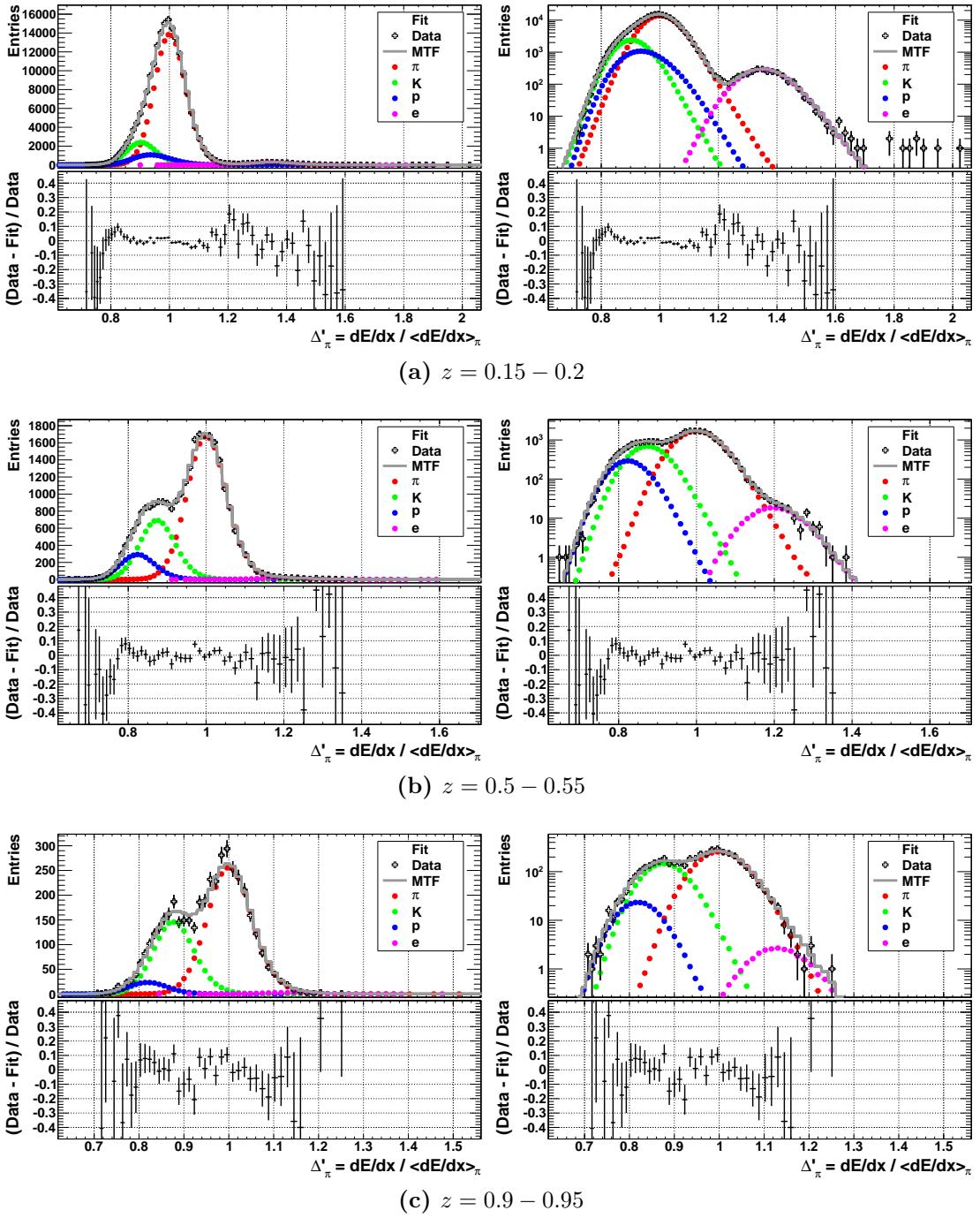


Figure F.7.: Template fits to data in different z intervals for $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$. The left panel shows the plots with linear vertical axis, the right panel with logarithmic axis. The rows show (starting from the top) the following z ranges: $0.15 - 0.2$, $0.5 - 0.55$ and $0.9 - 0.95$. Note the different ranges of the abscissa.

Appendix G.

Plots for the η -differential Jet Analysis

In this section, the uncorrected particle fractions in data for $|\eta| < 0.2$, $0.2 < |\eta| < 0.4$ and $0.4 < |\eta| < 0.6$ are compared to the full range $|\eta| < 0.9$, similar to what was done in Section 6.3.9. The Fig. G.1 shows the results for the inclusive case and Figs. G.2 and G.3 those for $p_T^{\text{jet}} = 5 - 10$ and $15 - 20 \text{ GeV}/c$, respectively. Good agreement is observed in all cases, except for the expected deviations in the non-highly-relativistic regime (best visible for protons), where the difference between rapidity y and pseudo-rapidity η becomes important. The statistical uncertainties for $p_T^{\text{jet}} = 15 - 20 \text{ GeV}/c$ are considerable, which is why the bin $p_T^{\text{jet}} = 20 - 30 \text{ GeV}/c$, that has even less statistics, is not shown.

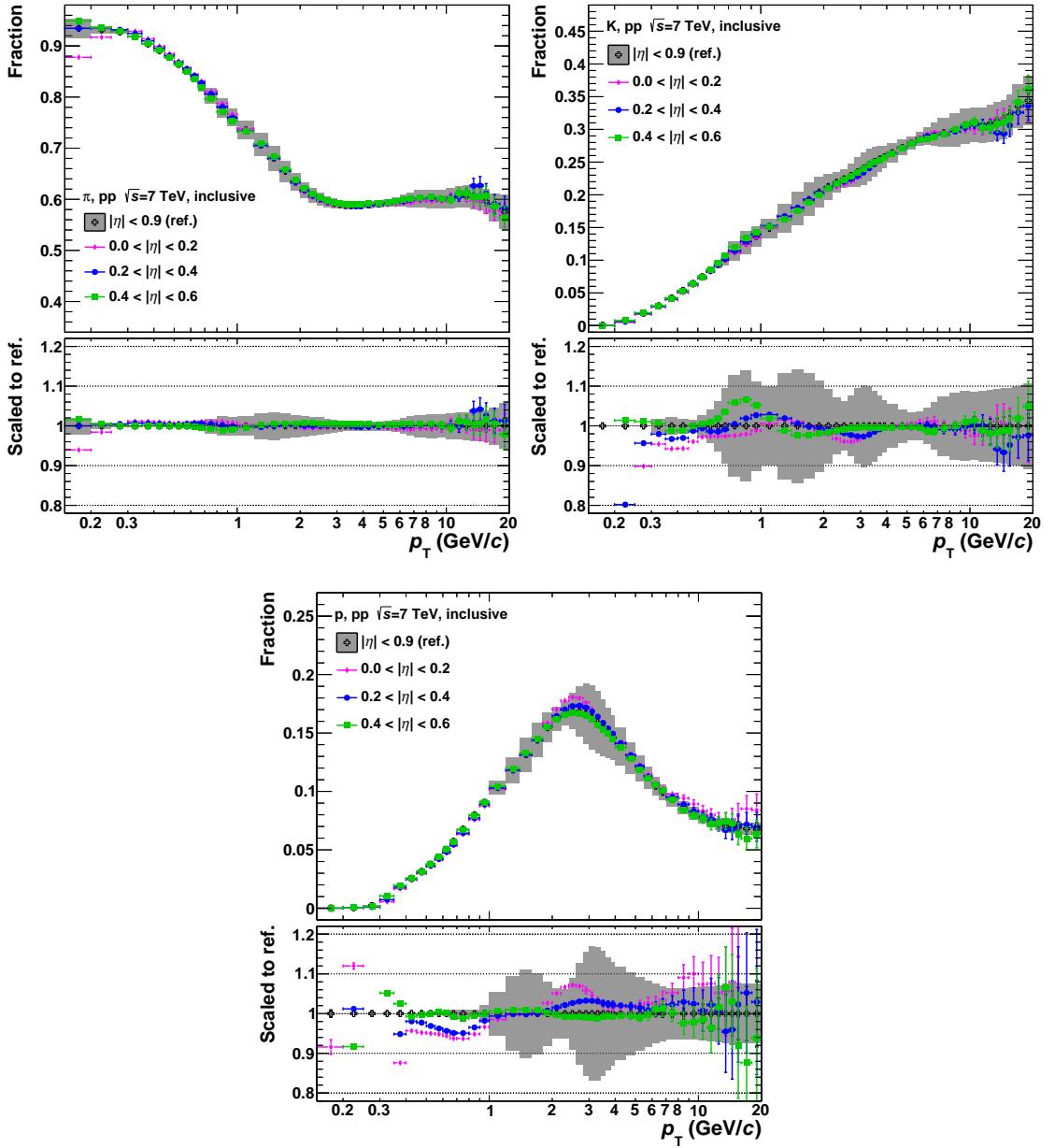


Figure G.1.: Comparison of the uncorrected particle fractions in η slices (coloured symbols) and in the full η range (black open crosses) versus p_T^{track} for the inclusive case in data. The results are shown for (from upper left to lower right) π , K and p . The grey area indicates the systematic uncertainties for the full η range.

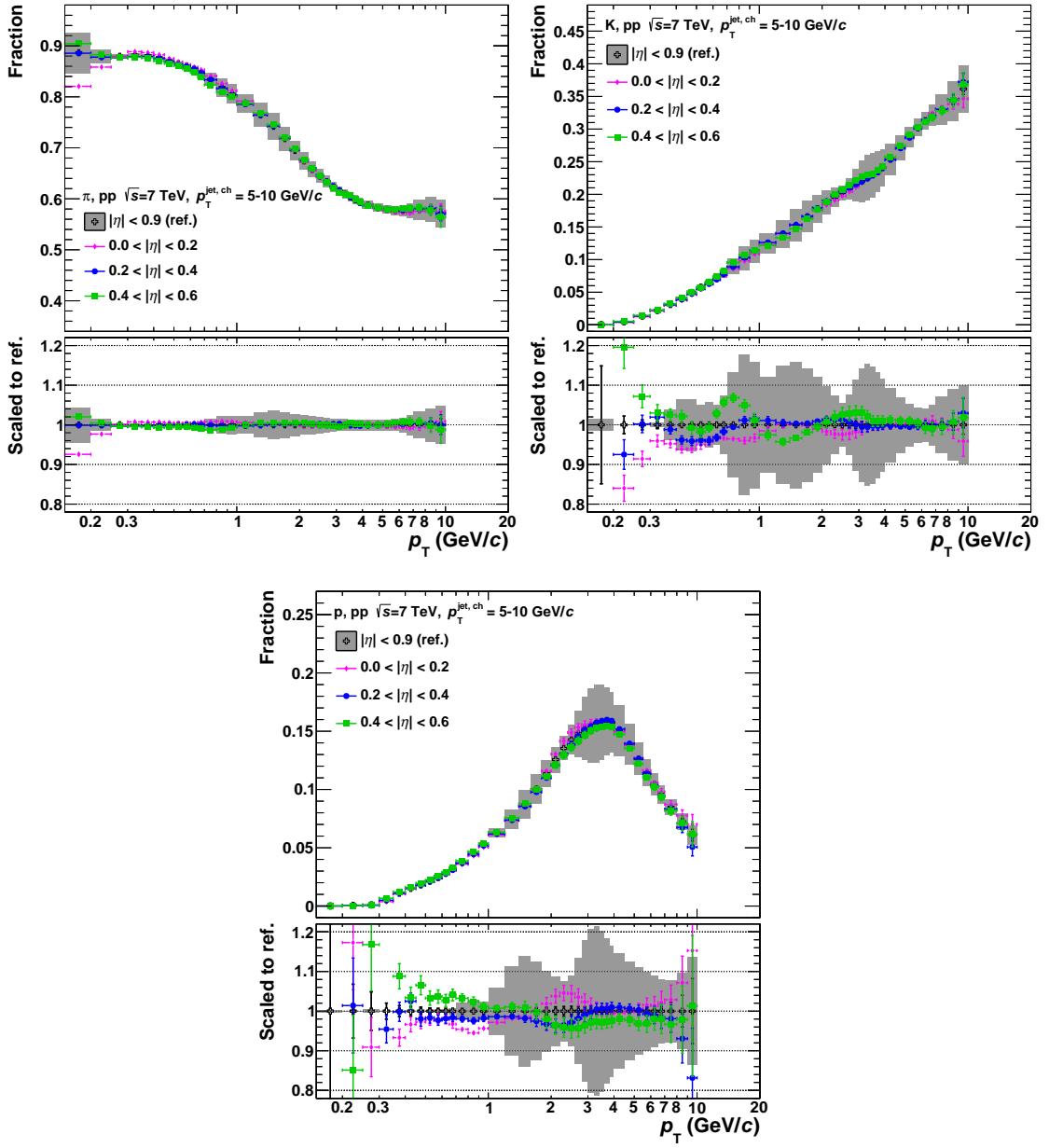


Figure G.2.: Comparison of the uncorrected particle fractions in η slices (coloured symbols) and in the full η range (black open crosses) versus p_T^{track} for $p_T^{\text{jet}} = 5 - 10 \text{ GeV}/c$ in data. The results are shown for (from upper left to lower right) π , K and p . The grey area indicates the systematic uncertainties for the full η range.

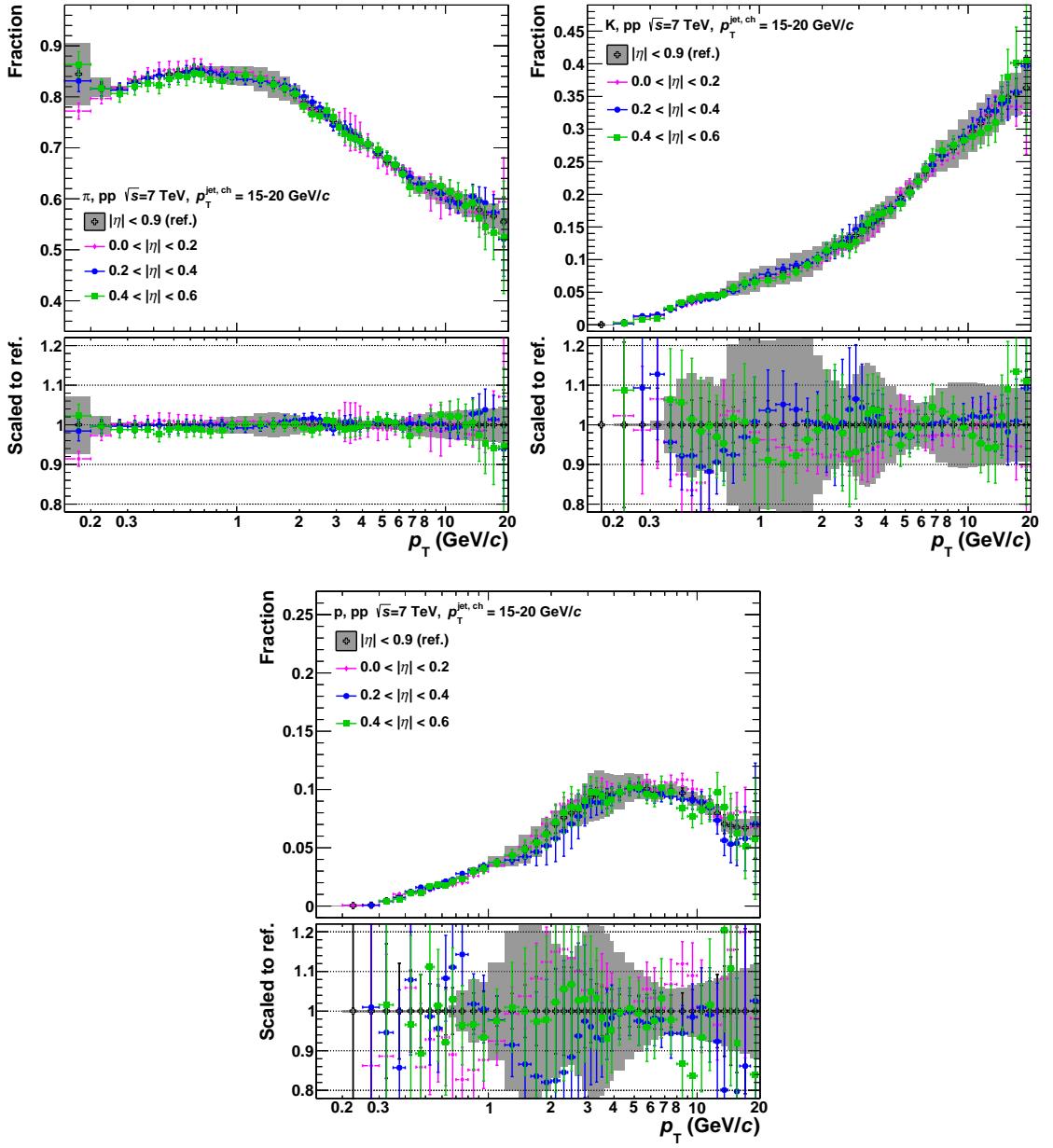


Figure G.3.: Comparison of the uncorrected particle fractions in η slices (coloured symbols) and in the full η range (black open crosses) versus p_T^{track} for $p_T^{\text{jet}} = 15 - 20 \text{ GeV}/c$ in data. The results are shown for (from upper left to lower right) π , K and p . The grey area indicates the systematic uncertainties for the full η range.

Appendix H.

Cross-Checks

This appendix comprises further plots with cross-checks.

H.1. TPC Coherent Fit

The uncorrected particle fractions as a function of z for pions, kaons and protons are compared to the results of the TPC Coherent Fit (using the same set of cuts and running on the same data set) and MC PYTHIA Perugia-0 in Figs. H.1, H.2 and H.3, respectively. The MC results are from the full detector simulation with GEANT and the fractions are determined by counting MC labels. The observations are similar to those made in Section 6.2.3 for the results as a function of p_T .

Appendix H. Cross-Checks

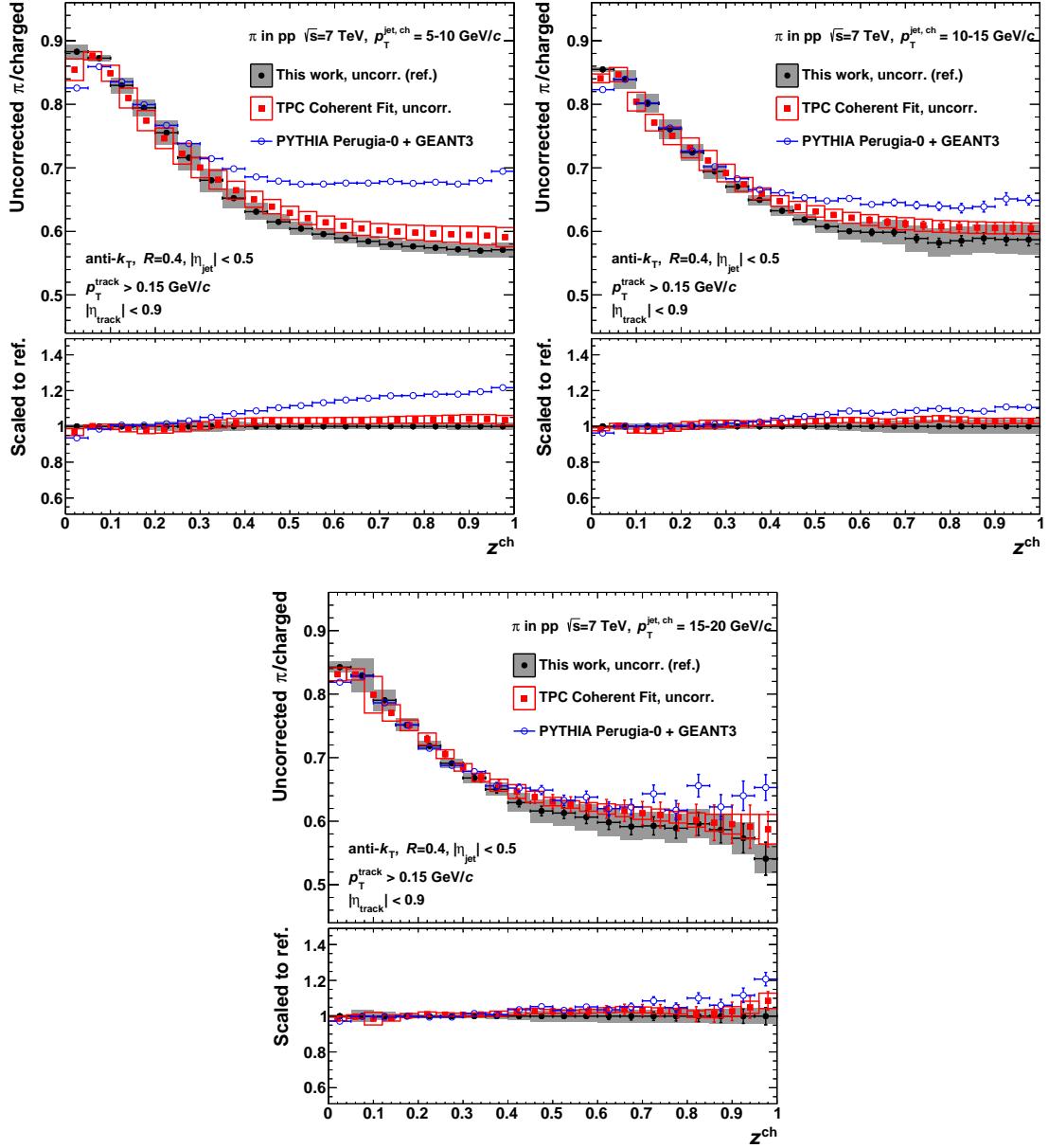


Figure H.1.: Comparison of uncorrected pion fractions as a function of z to the TPC Coherent Fit. In addition, the MC prediction from PYTHIA Perugia-0 is overlaid.

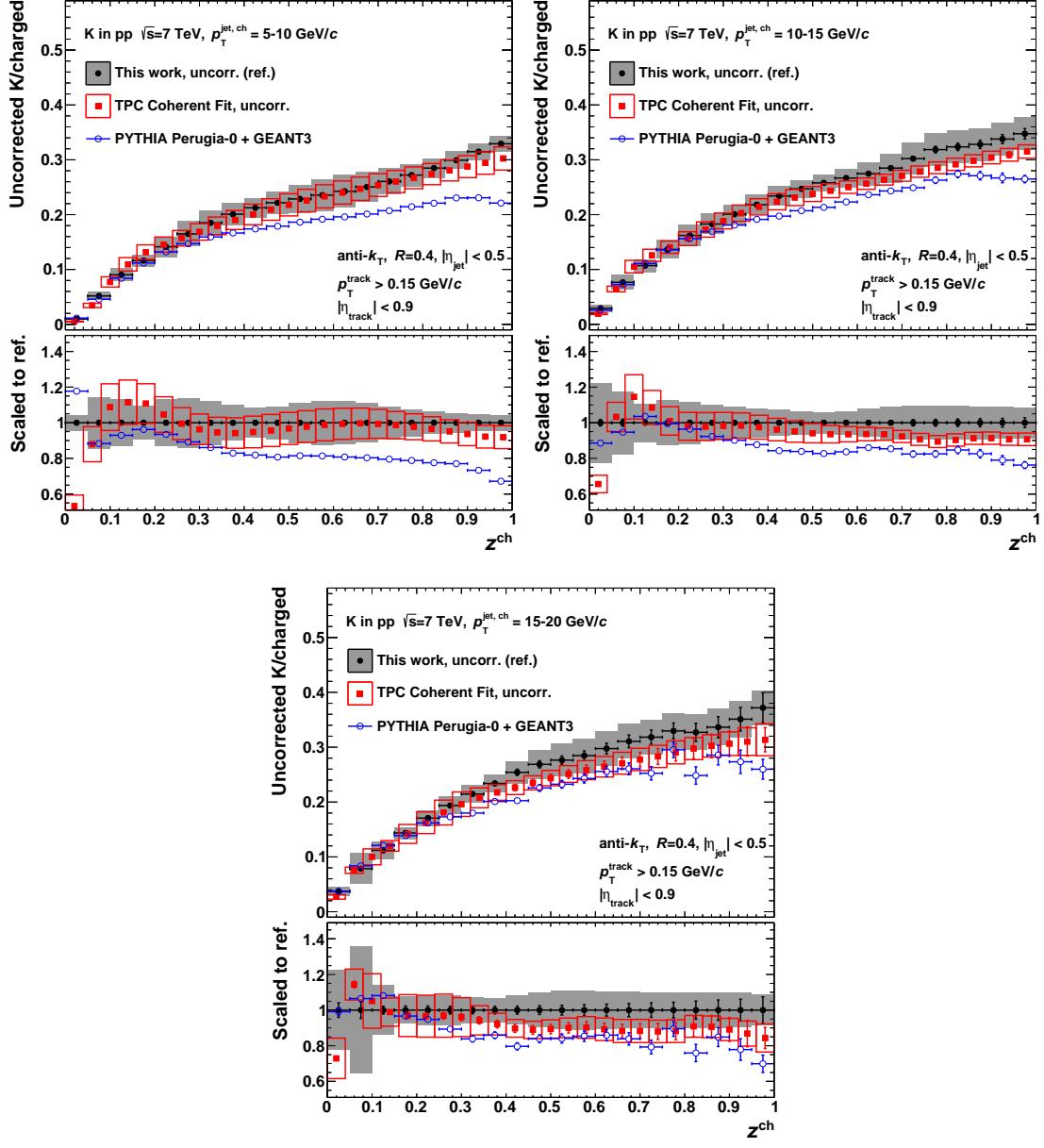


Figure H.2.: Comparison of uncorrected kaon fractions as a function of z to the TPC Coherent Fit. In addition, the MC prediction from PYTHIA Perugia-0 is overlaid.

Appendix H. Cross-Checks

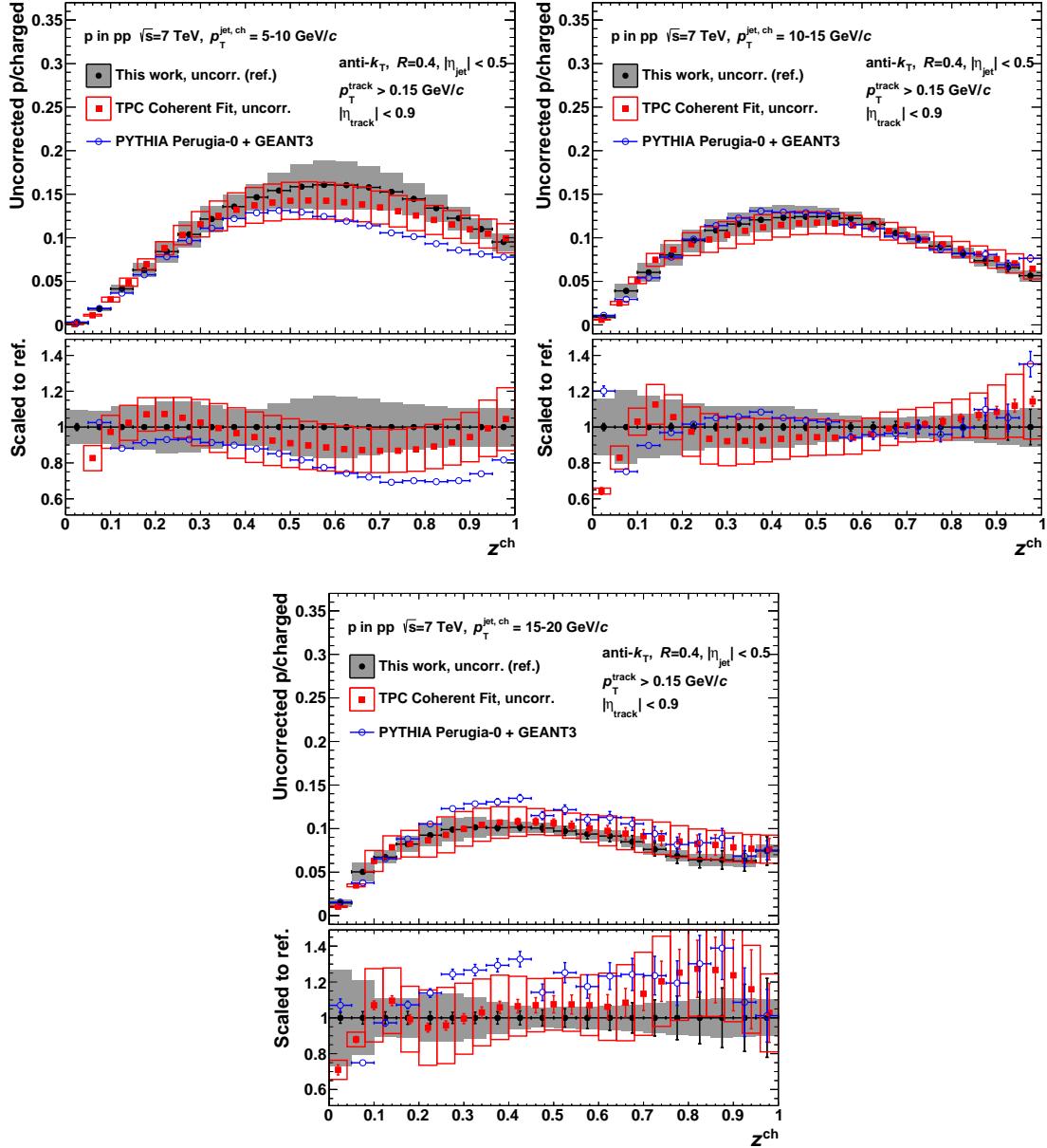


Figure H.3.: Comparison of uncorrected proton fractions as a function of z to the TPC Coherent Fit. In addition, the MC prediction from PYTHIA Perugia-0 is overlaid.

H.2. Combined PID and Lund

In Section 6.4.1, the MB yields of negatively charged particles have been compared to those obtained from the combined PID and the Lund analysis. The corresponding comparison for positively charged particles is shown in Fig. H.4. Again, good overall agreement is observed.

Appendix H. Cross-Checks

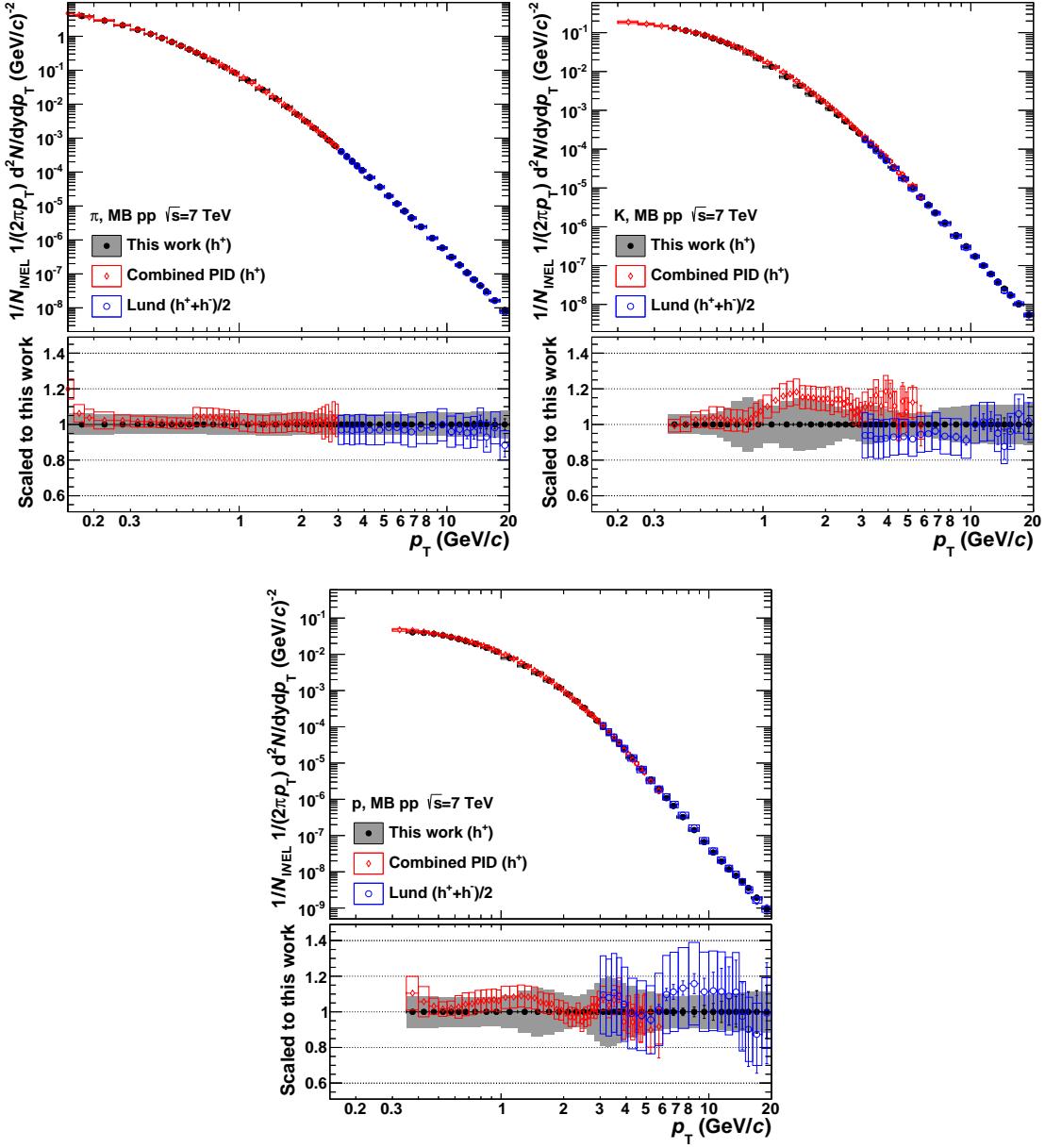


Figure H.4.: The corrected MB yields of positively charged particles are compared to the combined PID [125] and the Lund analysis [126]. The error bars indicate the statistical and the boxes the systematic uncertainties. Note that the Lund result (blue) shows the yield of both charges divided by two.

Appendix I.

Robustness of the Analysis Against Cut Variations

In this section, further plots illustrating the robustness of the analysis against cut variations as discussed in Section 6.3.10 are collected. In Figs. I.1, I.2 and I.3, the results from cut variations changing the PID quality are compared for pions, kaons and protons, respectively. Shown here are inclusive, $p_T^{\text{jet}} = 5 - 10, 15 - 20$ and $20 - 30 \text{ GeV}/c$, whereas $p_T^{\text{jet}} = 10 - 15 \text{ GeV}/c$ was already depicted in Fig. 6.17. The cut variations are based on the default hybrid cuts. As can be seen, the results are in agreement within systematic uncertainties.

The impact of cut variations changing the efficiency and secondary contamination is shown in Figs. I.4, I.5 and I.6 for pions, kaons and protons, respectively. The plots show the ratio to the modified hybrid track cuts as a function of z in jets. The corresponding plots as a function of p_T in the inclusive case are depicted in Fig. I.7. In all cases, the systematic uncertainties cover the spread of the results. The only exception are protons in the inclusive case, which has been discussed in Section 6.3.10.

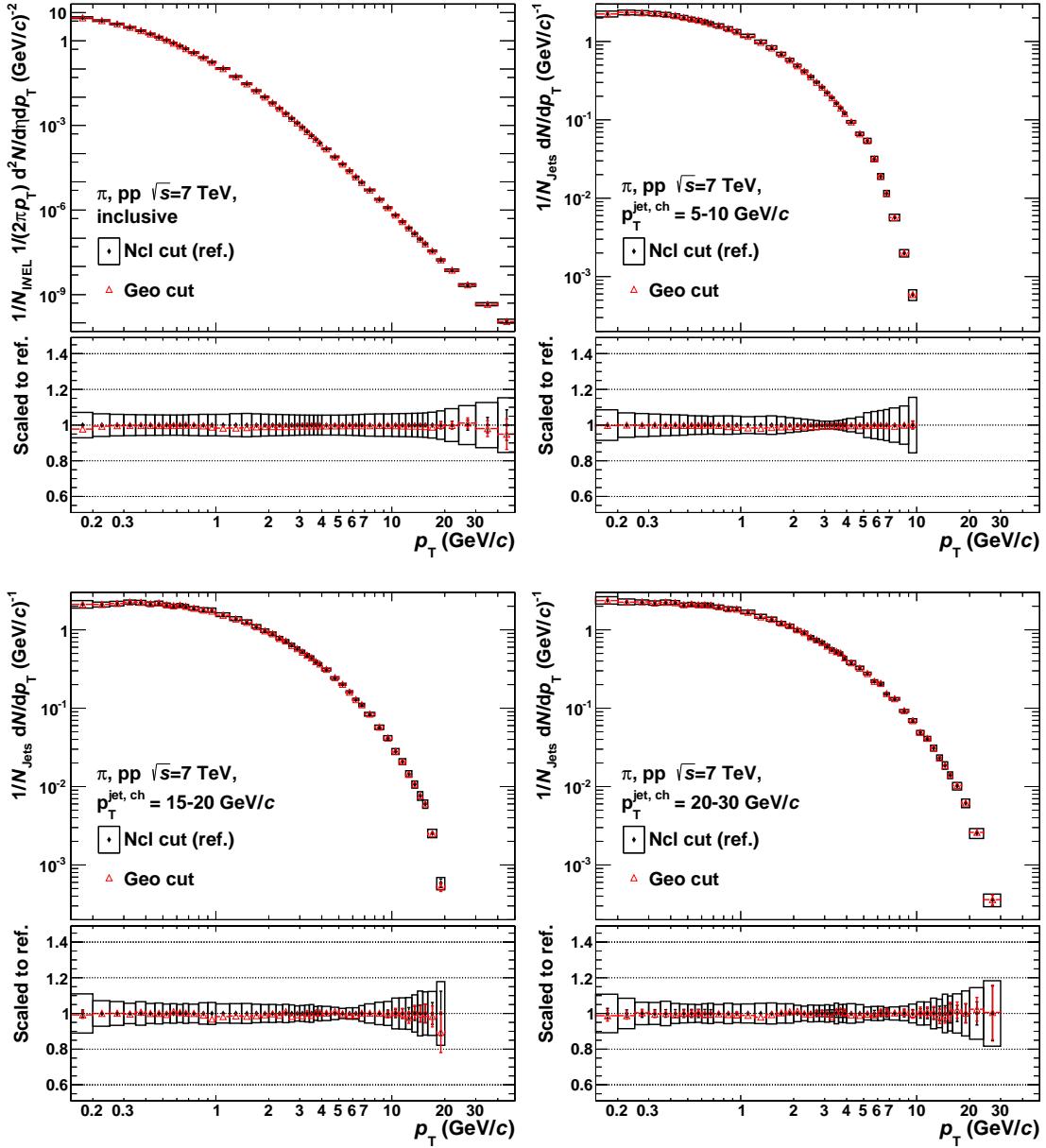


Figure I.1.: Comparison of the corrected pion yields for the TPC geometrical cut (red open triangles) and the ncl cut (black diamonds) versus p_T^{track} in data. Each set of cuts is corrected with its own MC correction factors. The boxes correspond to the systematic errors for the ncl cut, including the uncertainties of the corresponding MC correction factors.

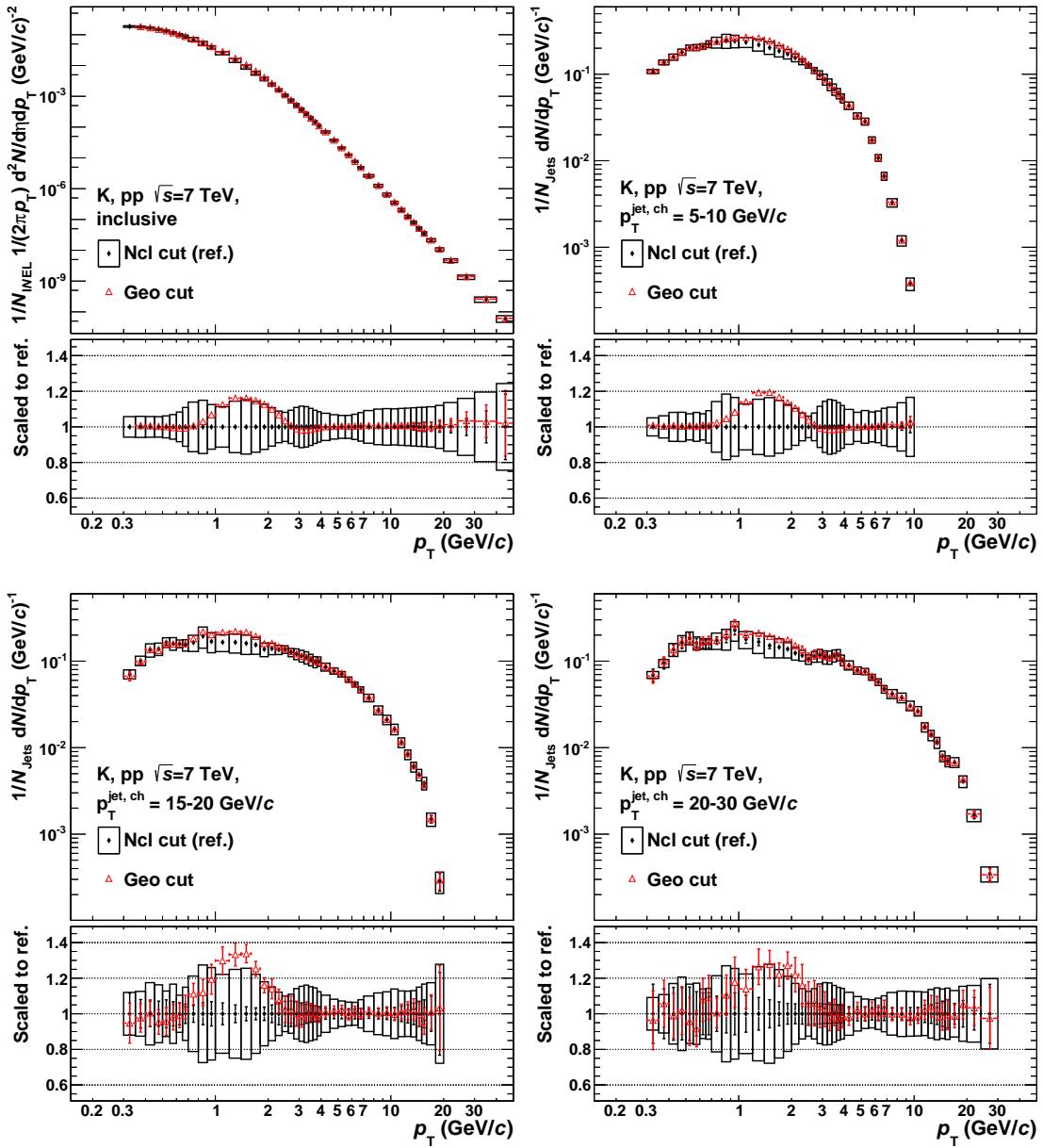


Figure I.2.: Comparison of the corrected kaon yields for the TPC geometrical cut (red open triangles) and the ncl cut (black diamonds) versus p_T^{track} in data. Each set of cuts is corrected with its own MC correction factors. The boxes correspond to the systematic errors for the ncl cut, including the uncertainties of the corresponding MC correction factors.

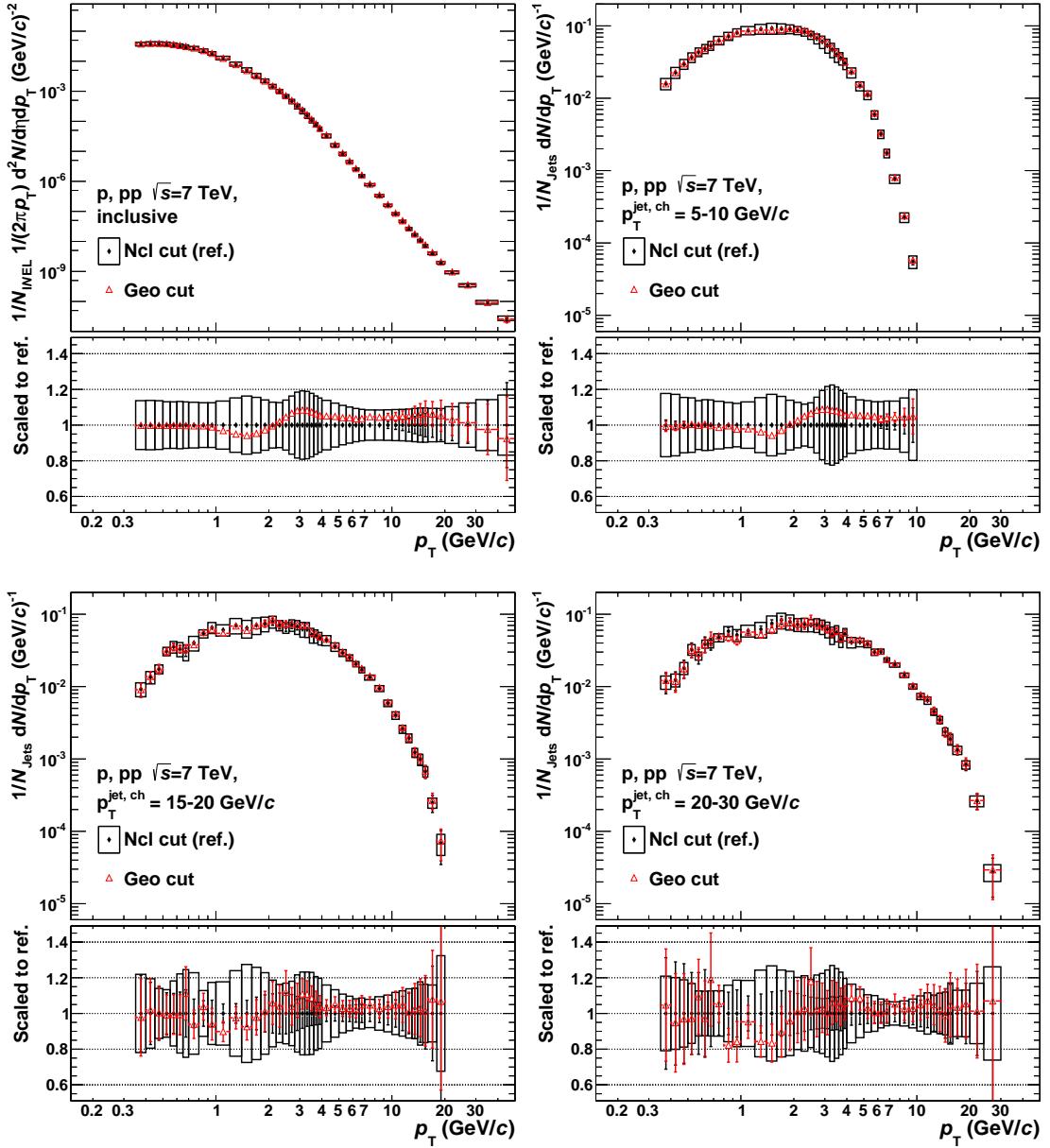


Figure I.3.: Comparison of the corrected proton yields for the TPC geometrical cut (red open triangles) and the ncl cut (black diamonds) versus p_T^{track} in data. Each set of cuts is corrected with its own MC correction factors. The boxes correspond to the systematic errors for the ncl cut, including the uncertainties of the corresponding MC correction factors.

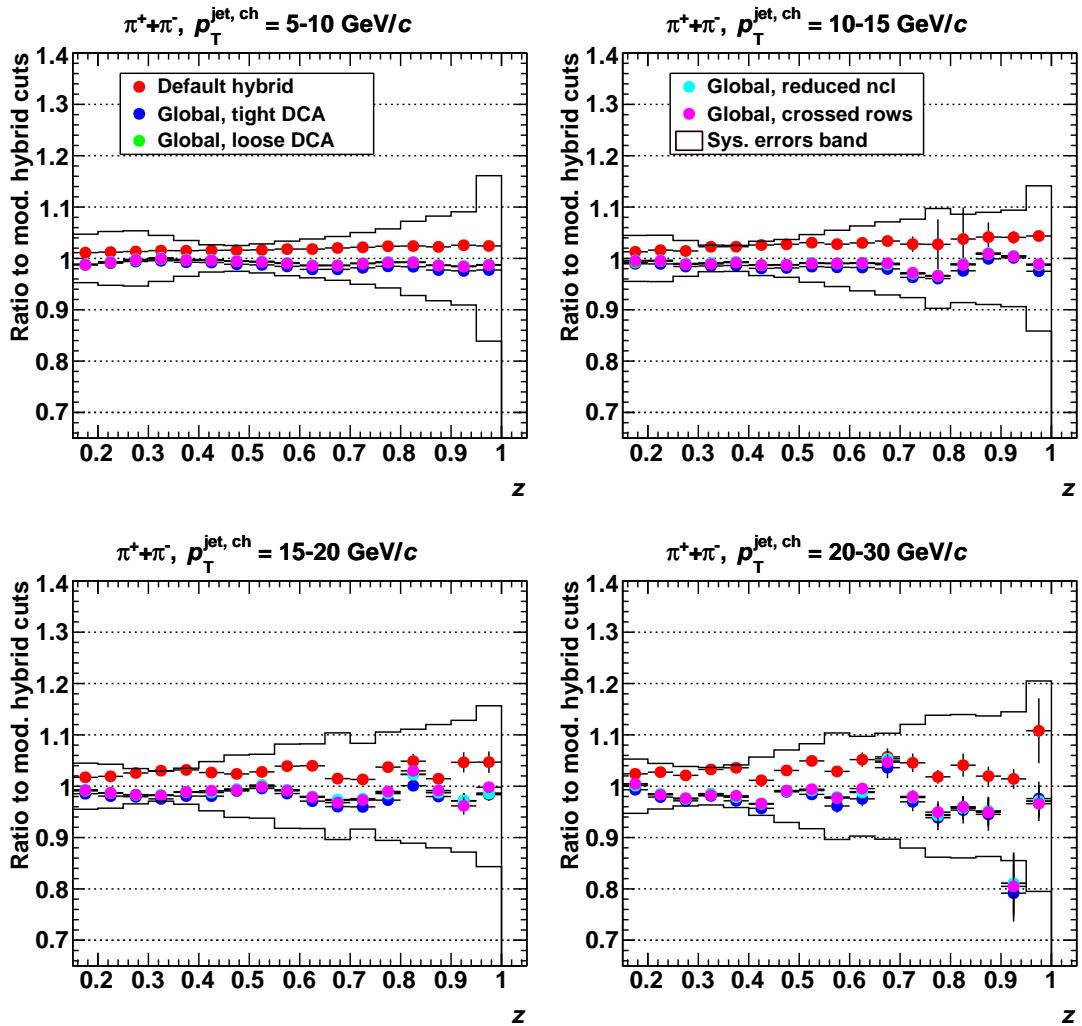


Figure I.4.: Cut variations for pions in jets as a function of z . The ratio to the modified hybrid track cuts is shown. Each set of cuts is corrected with its own MC correction factors. The systematic error band shows the total uncertainty from fit and MC corrections.

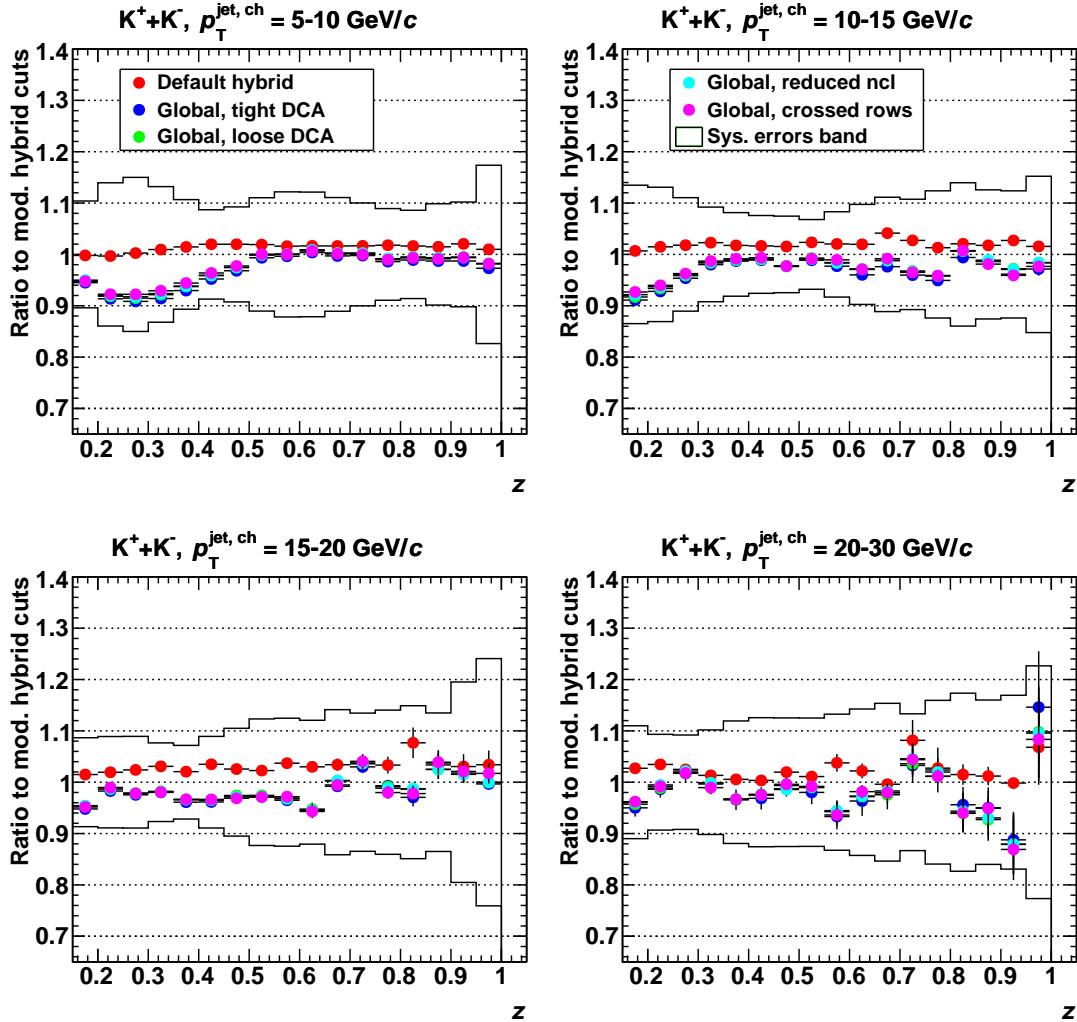


Figure I.5.: Cut variations for kaons in jets as a function of z . The ratio to the modified hybrid track cuts is shown. Each set of cuts is corrected with its own MC correction factors. The systematic error band shows the total uncertainty from fit and MC corrections.

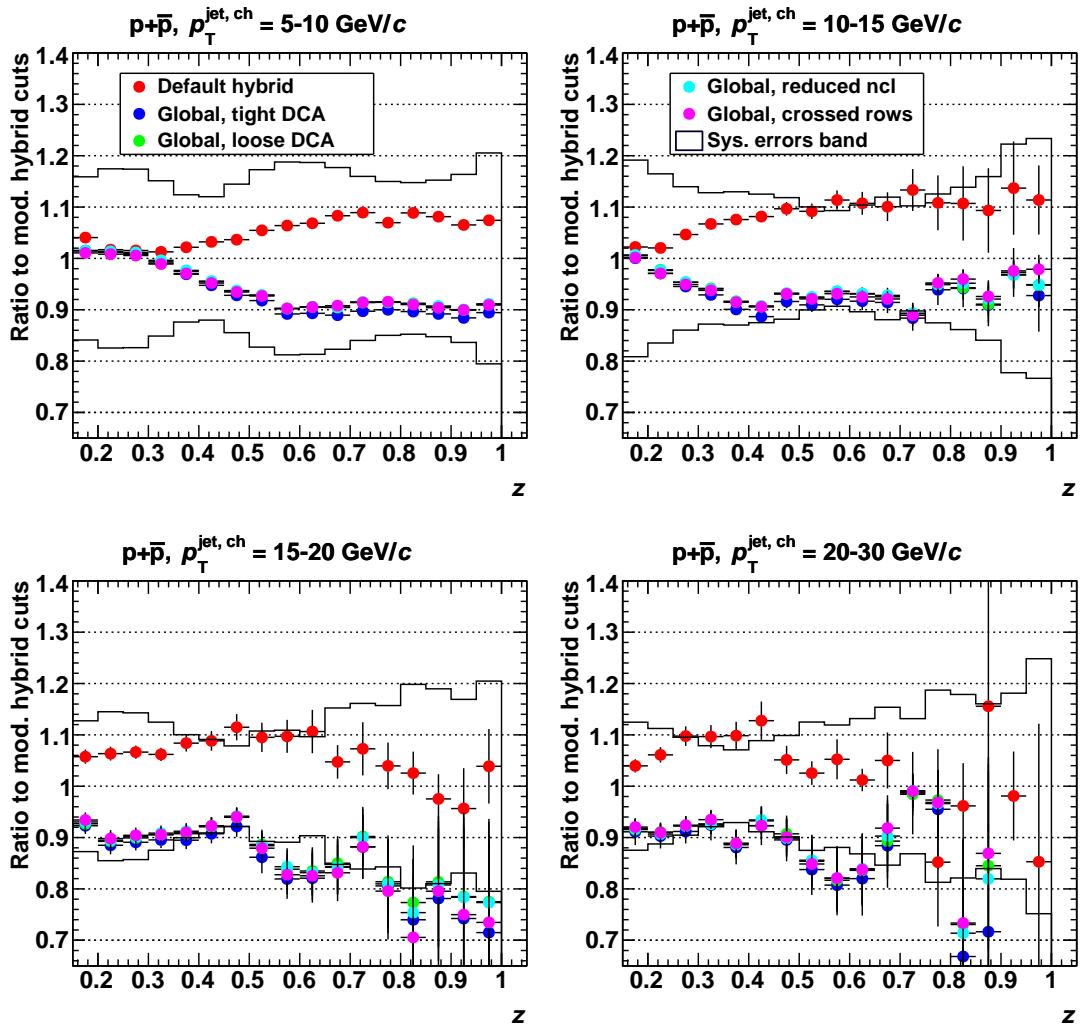


Figure I.6.: Cut variations for protons in jets as a function of z . The ratio to the modified hybrid track cuts is shown. Each set of cuts is corrected with its own MC correction factors. The systematic error band shows the total uncertainty from fit and MC corrections.

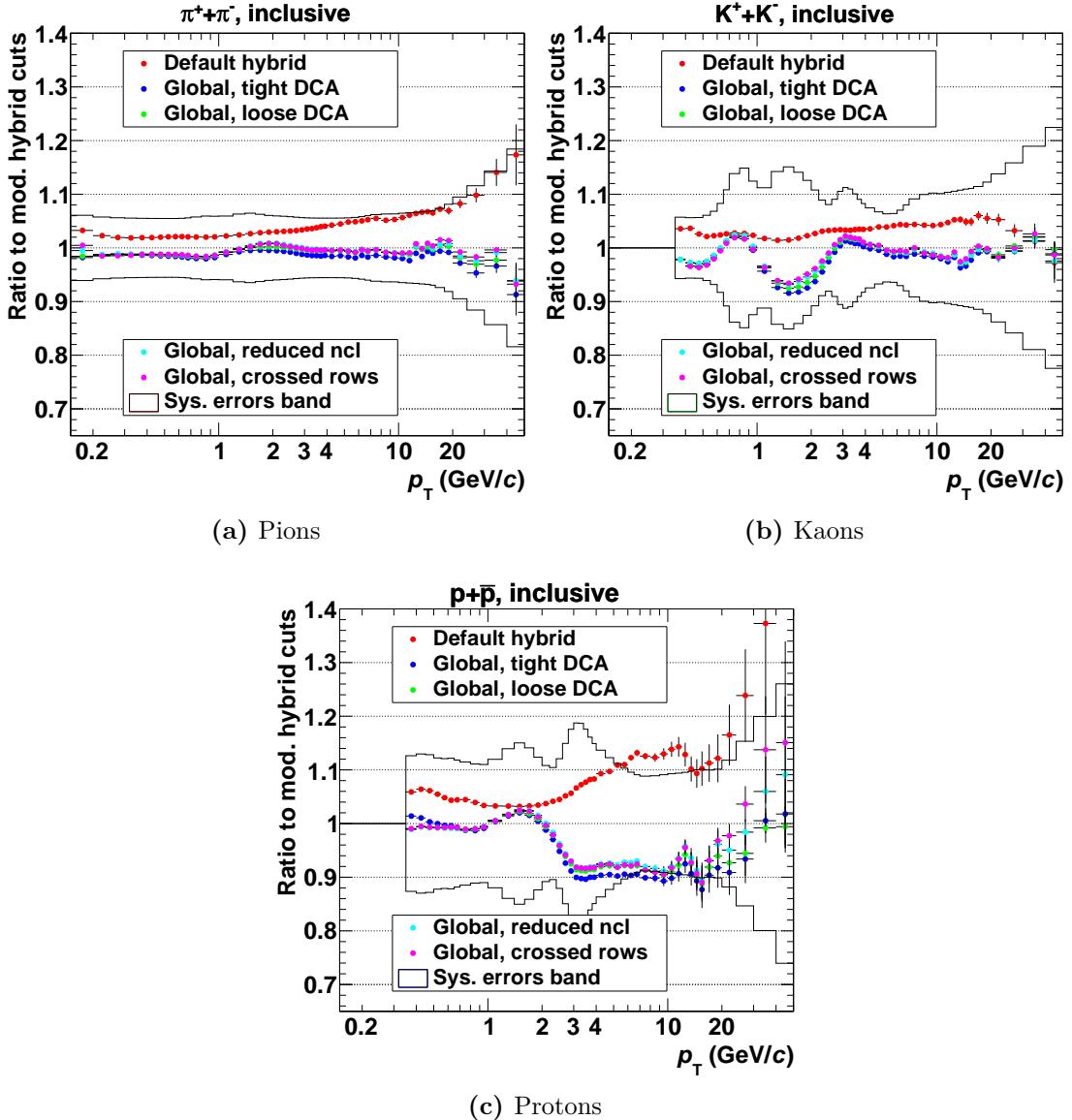


Figure I.7.: Cut variations for pions (a), kaons (b) and protons (c) in the inclusive case. The ratio to the modified hybrid track cuts is shown. Each set of cuts is corrected with its own MC correction factors. The systematic error band shows the total uncertainty from fit and MC corrections.

Appendix J.

Charge Differential Analysis Results

To judge a possible charge dependence of the particle production at mid-rapidity, the analysis is run separately on positively and negatively charged particles. Further results, supplementing the discussion in Chapter 6, are presented in this appendix.

J.1. Multiplicity-Binned Analysis

In Section 6.5.4, it was demonstrated that the yield ratio of positively and negatively charged particles does not depend on the considered bin of V0M multiplicity. The observations are basically identical for the reference multiplicity bins 4 – 6 and 40 – 49, as shown in Fig. J.1, as well as for other bins of this multiplicity estimator.

Appendix J. Charge Differential Analysis Results

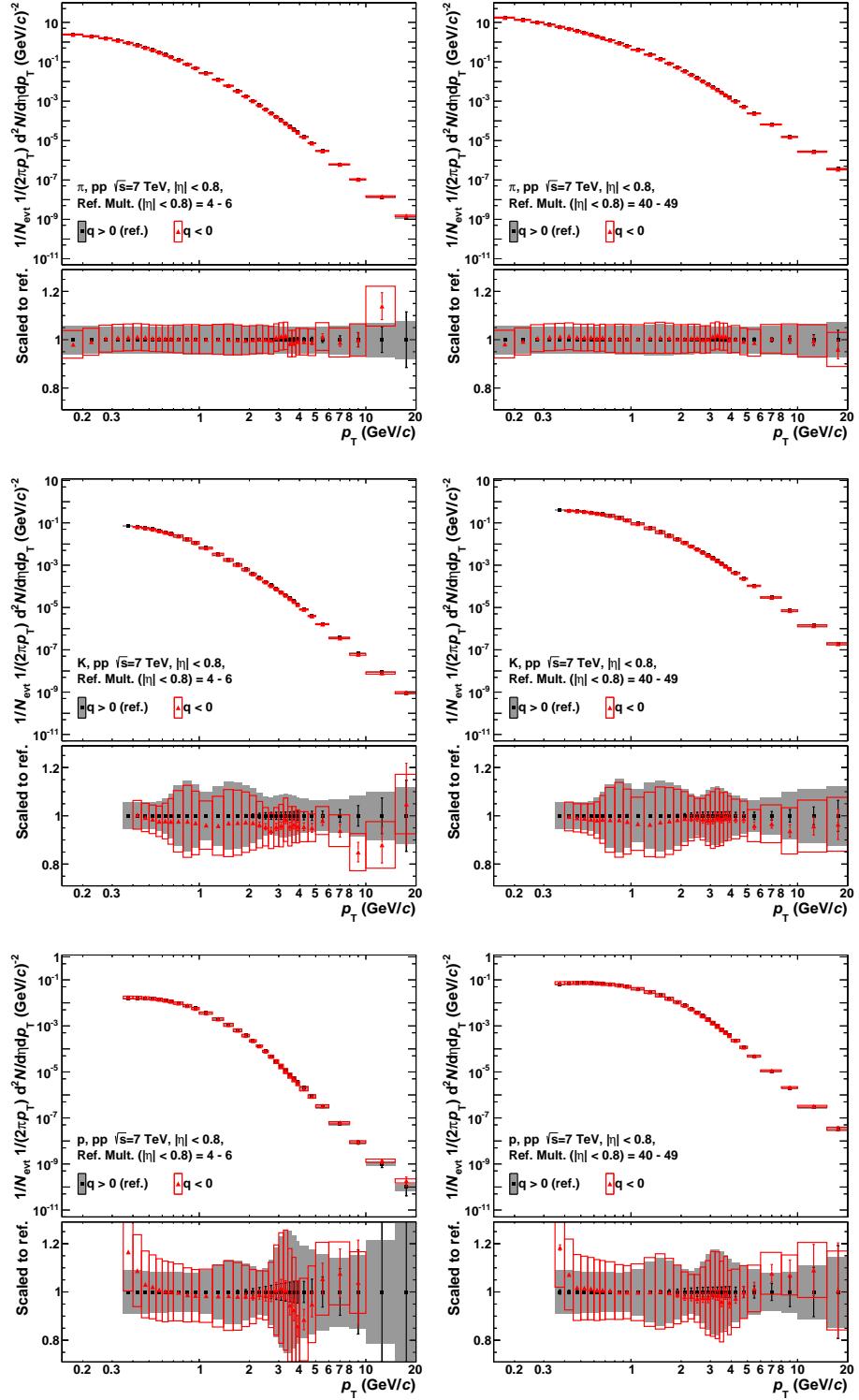


Figure J.1.: Comparison of positively ($q > 0$) and negatively ($q < 0$) charged particles for the reference multiplicity bins 4 – 6 (left) and 40 – 49 (right). The rows show (from top to bottom): π , K and p.

J.2. Jet Analysis

In addition to the plots for lower p_T^{jet} shown in Section 6.6.3, the Fig. J.2 compares positively and negatively charged particles in jets with $p_T^{\text{jet}} = 15 - 20 \text{ GeV}/c$. The results are similar to the lower jet momenta, but the large statistical uncertainties — which are uncorrelated for the different charges — make a more quantitative statement difficult.

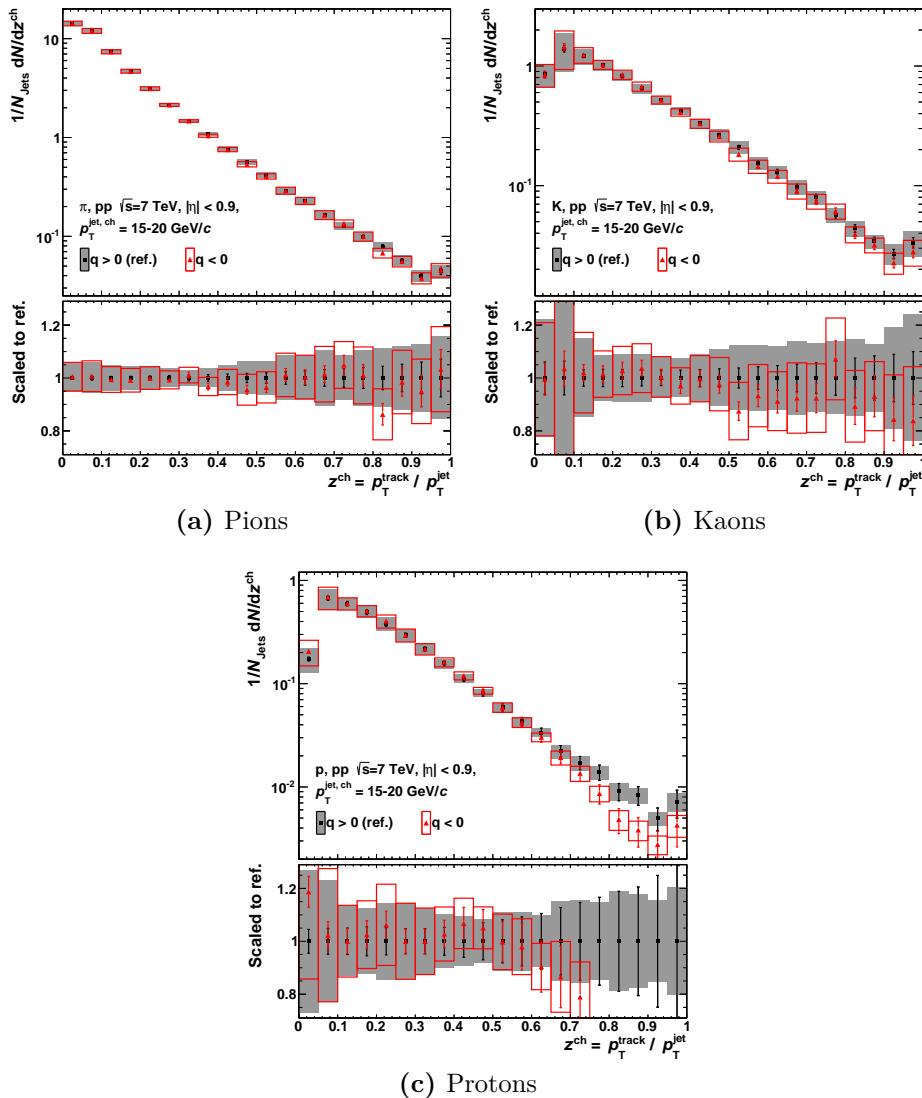


Figure J.2.: Comparison of positively ($q > 0$) and negatively ($q < 0$) charged particles in jets with $p_T^{\text{jet}} = 15 - 20 \text{ GeV}/c$. The results are shown for pions (a), kaons (b) and protons (c).

Appendix K.

Comparison of Corrected Particle Yields in Jets with MC

In Section 6.6.2, the fully corrected identified particle yields were compared to predictions from various PYTHIA Perugia tunes. The corresponding plots for $\pi/K/p$ as a function of z are shown in Figs. K.1, K.2 and K.3. The features are the same as for p_T binning, but the results get smeared along the horizontal axis and this axis is also scaled. The underestimation of the pion yield at low p_T is visible for $z \lesssim 0.1$ (the precise value of the threshold depends on p_T^{jet}). To focus on the high- p_T and high- z region, the plots are shown again with a logarithmic scale for the vertical axis in Figs. K.4, K.5 and K.6 for p_T binning and in Figs. K.7, K.8 and K.9 for z binning. In particular, the wrongly predicted high- p_T slope for the protons is visible in all p_T^{jet} bins.

Finally, the K/π and p/π ratios are shown as a function of z in Figs. K.10 and K.11, respectively. As for the p_T binning, the trends observed for the yields are just propagated to the to-pion ratios.

Appendix K. Comparison of Corrected Particle Yields in Jets with MC

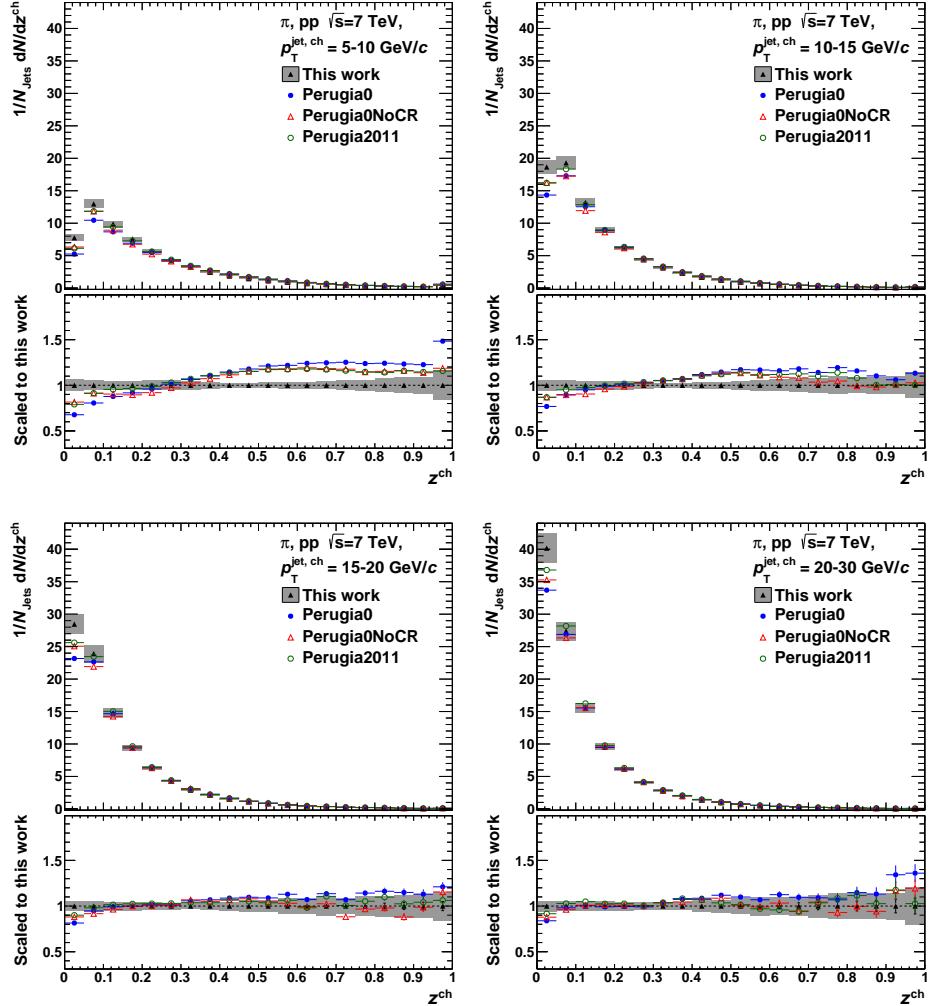


Figure K.1.: Comparison of corrected pion yields in jets to different MC predictions as a function of z . The p_T^{jet} increases from the upper left to the lower right panel.

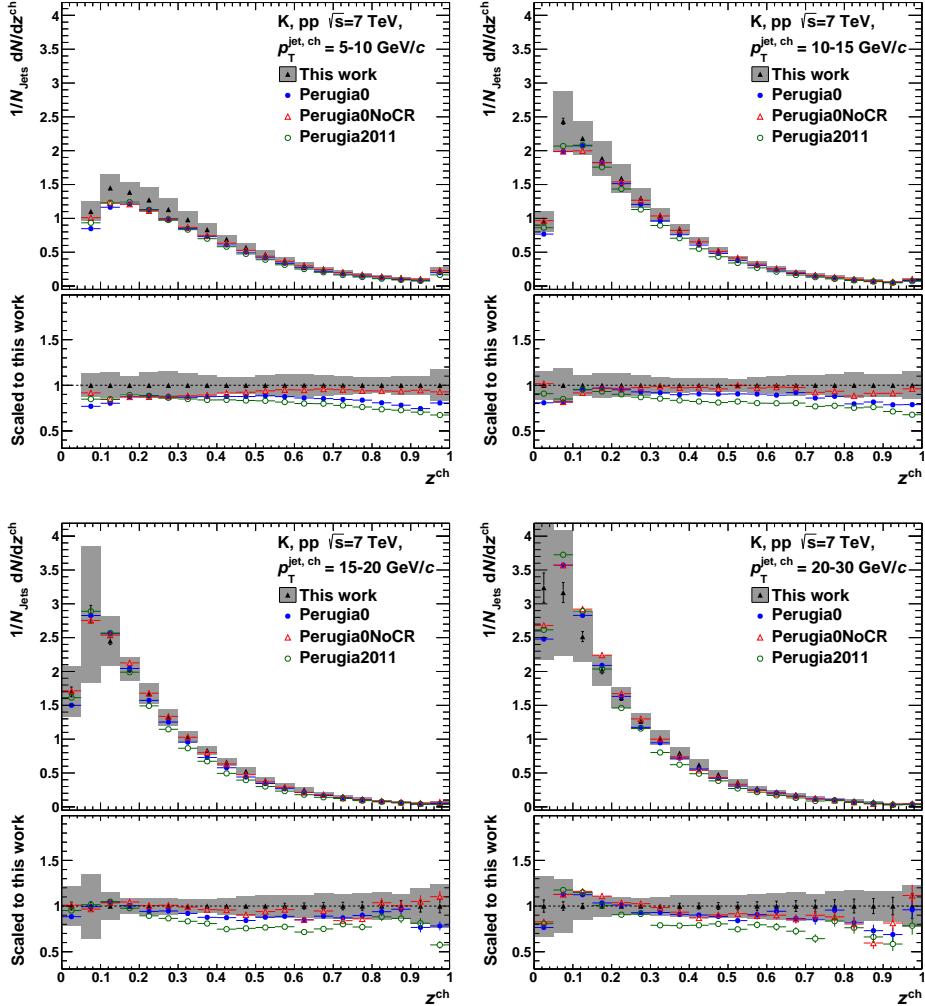


Figure K.2.: Comparison of corrected kaon yields in jets to different MC predictions as a function of z . The p_T^{jet} increases from the upper left to the lower right panel.

Appendix K. Comparison of Corrected Particle Yields in Jets with MC

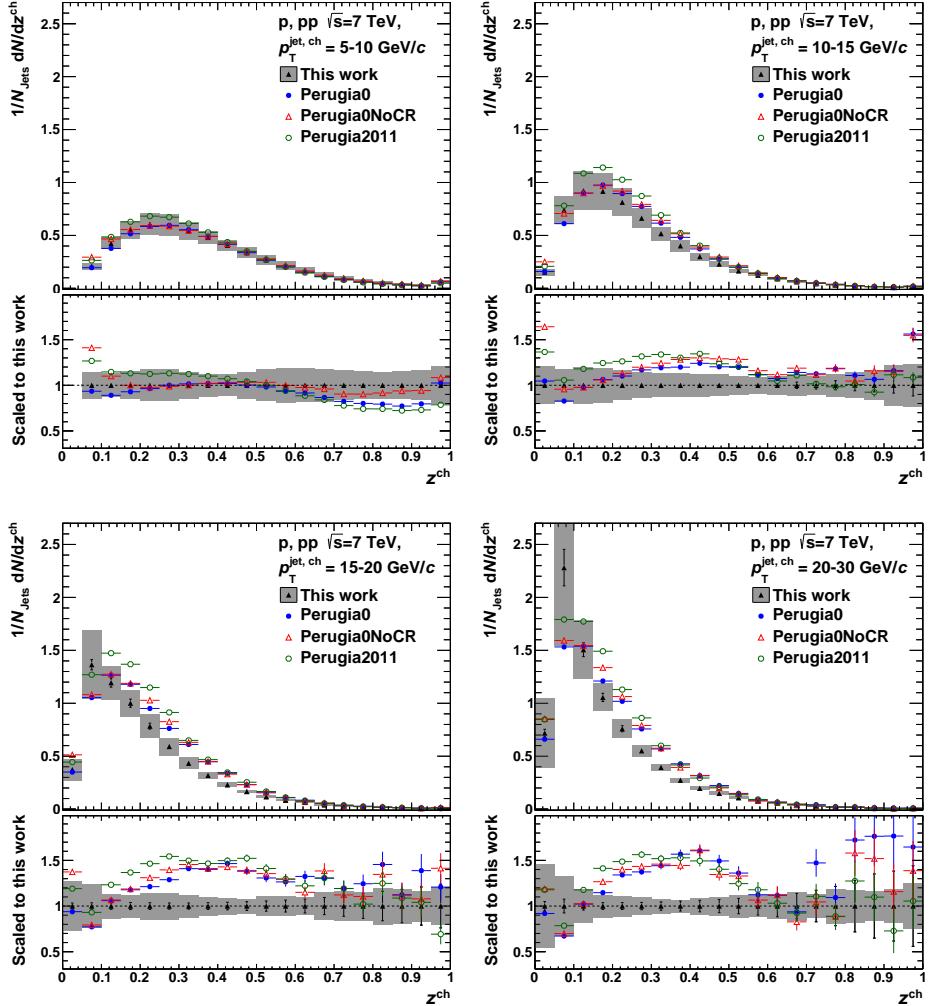


Figure K.3.: Comparison of corrected proton yields in jets to different MC predictions as a function of z . The $p_{\text{T}}^{\text{jet}}$ increases from the upper left to the lower right panel.

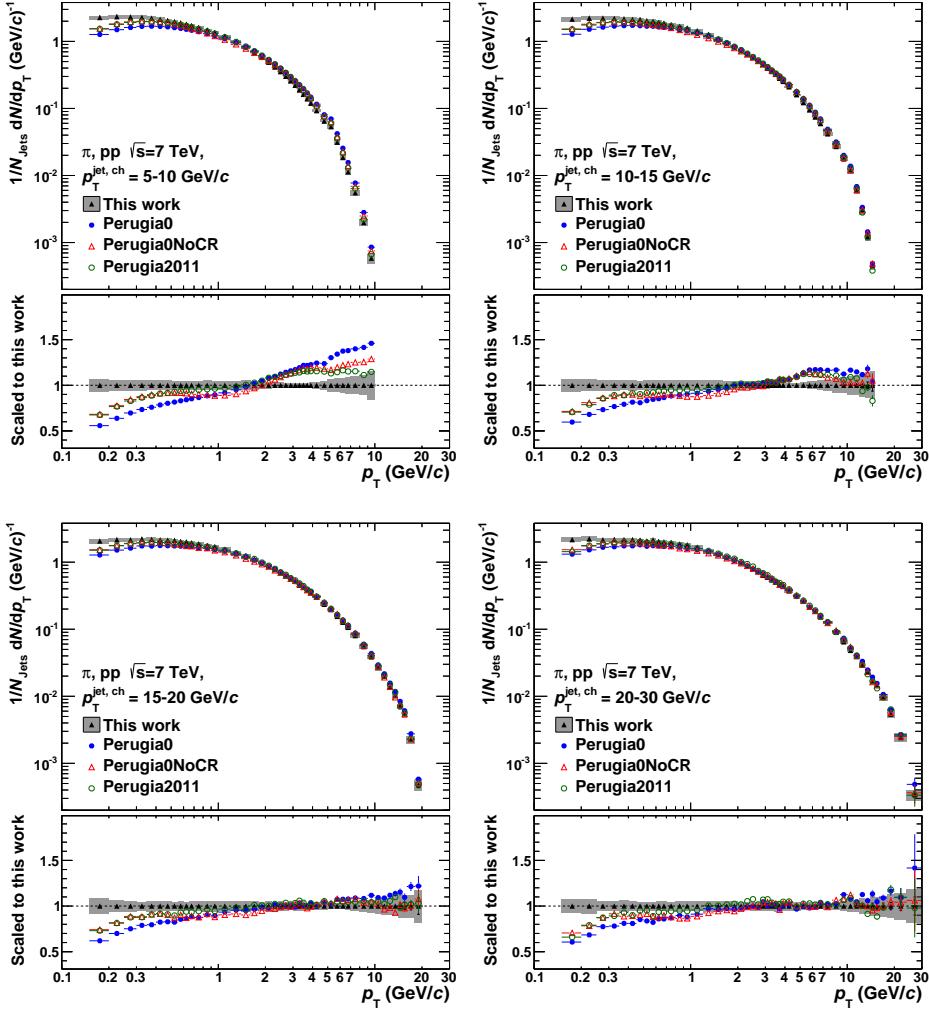


Figure K.4.: Comparison of corrected pion yields in jets to different MC predictions as a function of p_T . The p_T^{jet} increases from the upper left to the lower right panel.

Appendix K. Comparison of Corrected Particle Yields in Jets with MC

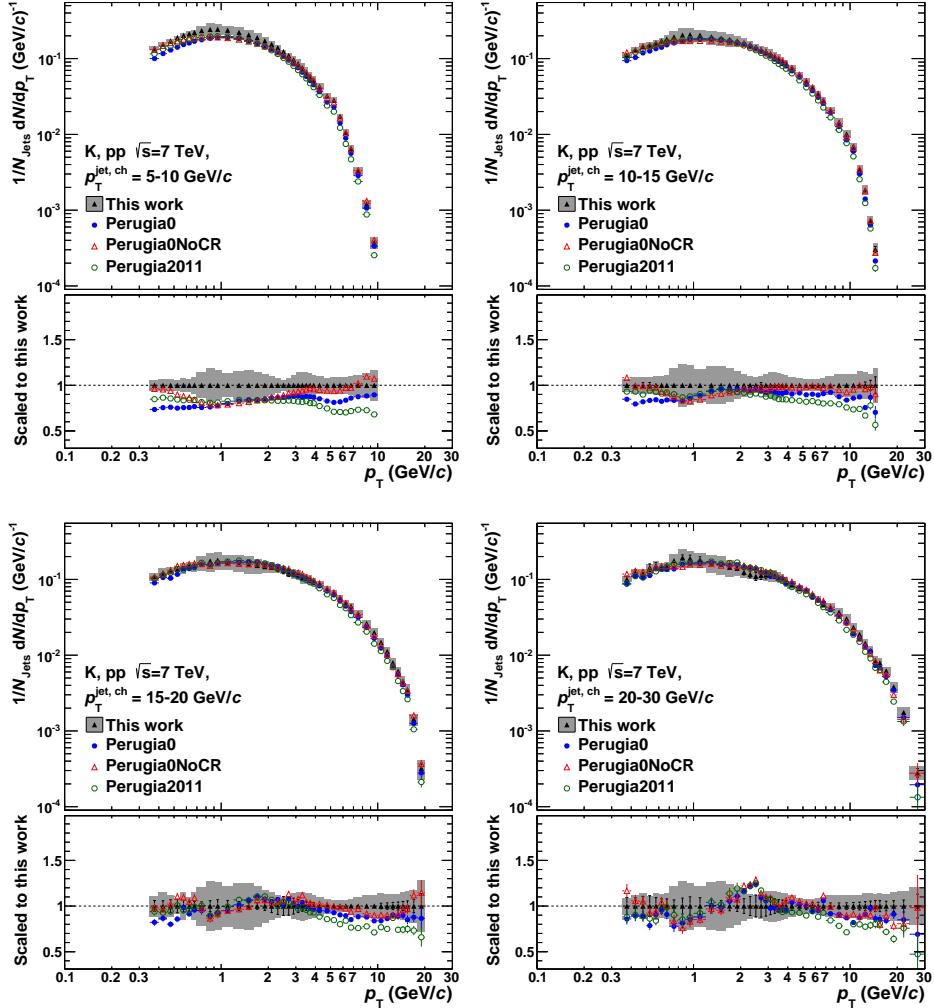


Figure K.5.: Comparison of corrected kaon yields in jets to different MC predictions as a function of p_T . The p_T^{jet} increases from the upper left to the lower right panel.

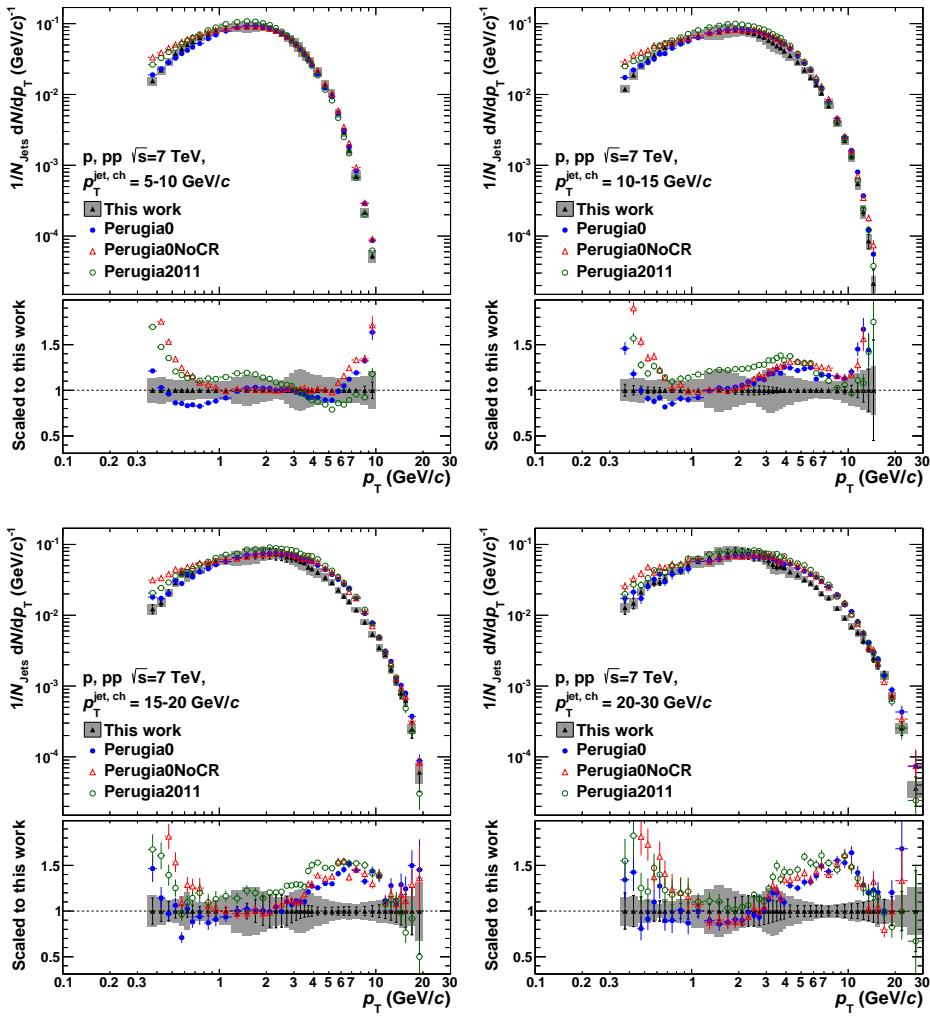


Figure K.6.: Comparison of corrected proton yields in jets to different MC predictions as a function of p_T . The p_T^{jet} increases from the upper left to the lower right panel.

Appendix K. Comparison of Corrected Particle Yields in Jets with MC

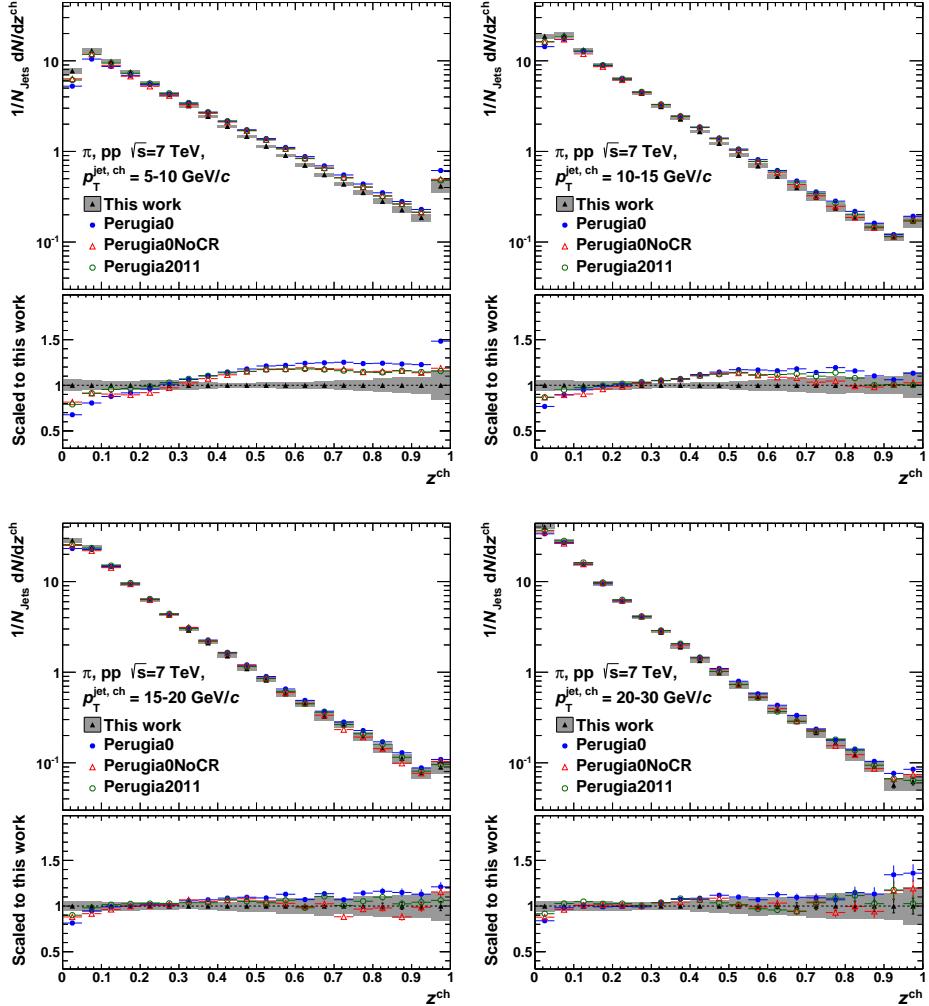


Figure K.7.: Comparison of corrected pion yields in jets to different MC predictions as a function of z . The p_T^{jet} increases from the upper left to the lower right panel.

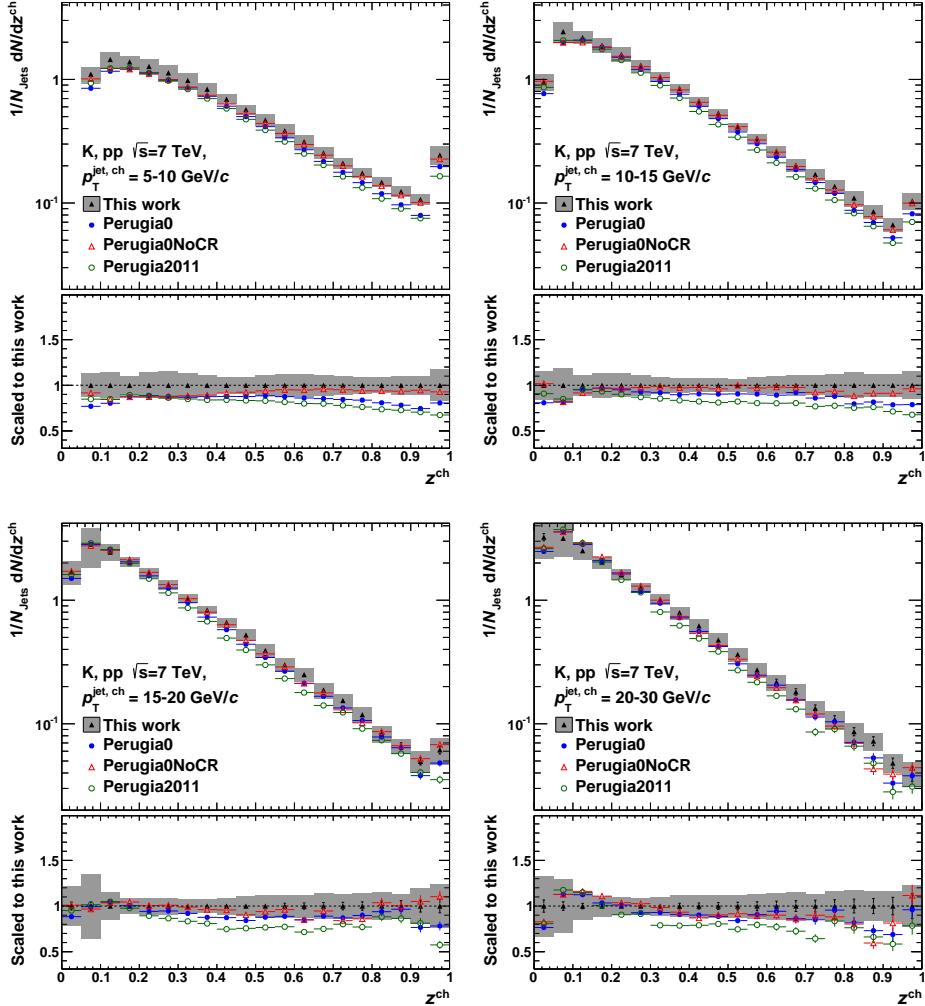


Figure K.8.: Comparison of corrected kaon yields in jets to different MC predictions as a function of z . The p_T^{jet} increases from the upper left to the lower right panel.

Appendix K. Comparison of Corrected Particle Yields in Jets with MC

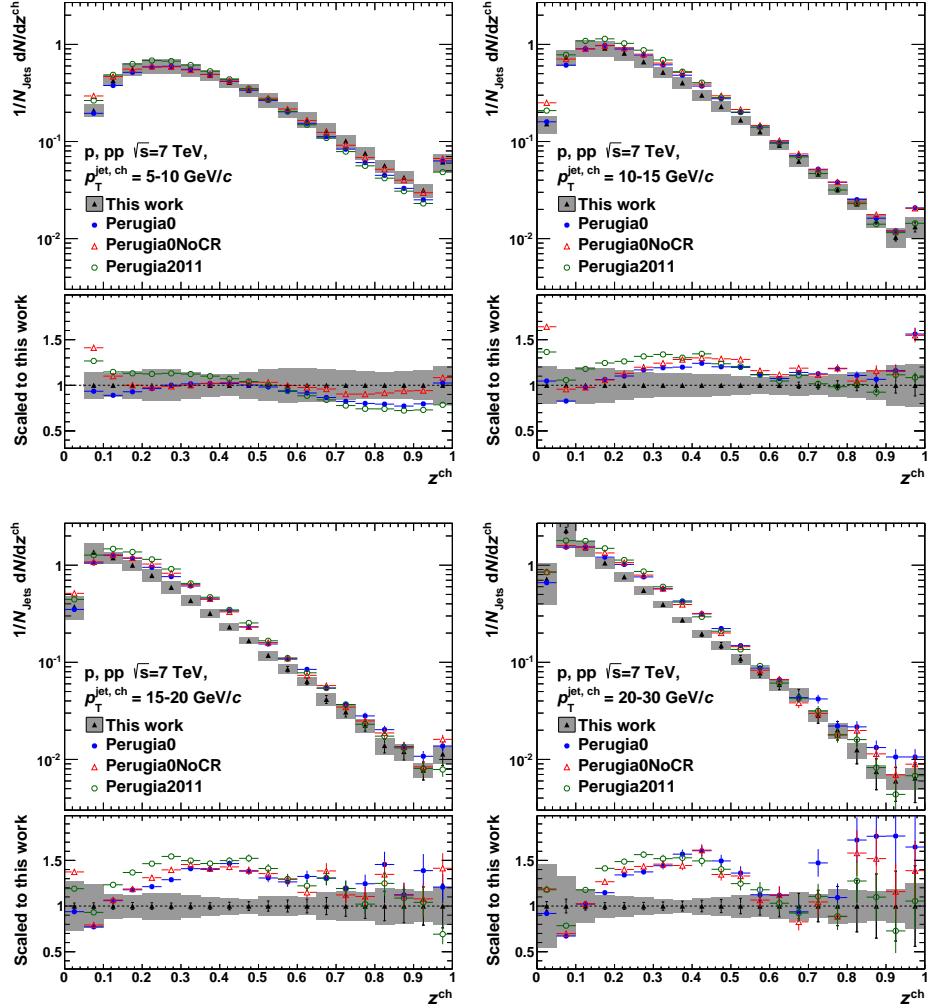


Figure K.9.: Comparison of corrected proton yields in jets to different MC predictions as a function of z . The $p_{\text{T}}^{\text{jet}}$ increases from the upper left to the lower right panel.

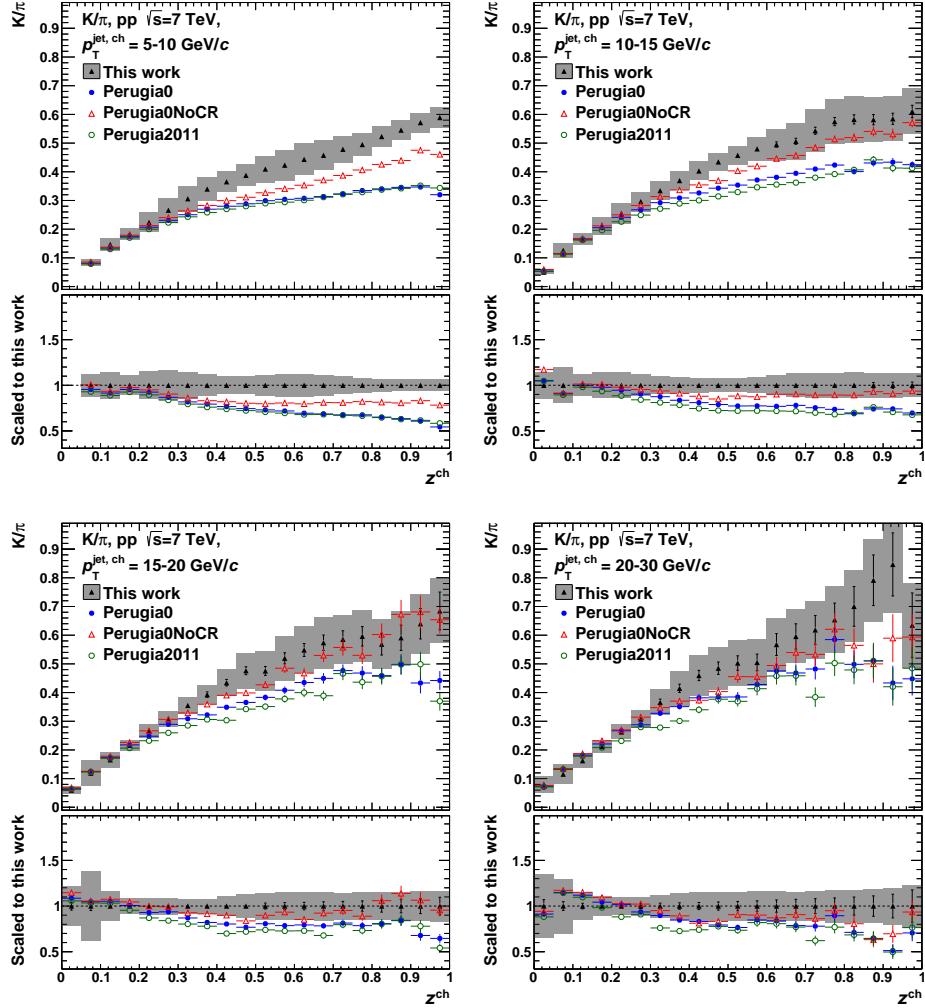


Figure K.10.: Comparison of corrected K/π ratios in jets to different MC predictions as a function of z . The p_T^{jet} increases from the upper left to the lower right panel.

Appendix K. Comparison of Corrected Particle Yields in Jets with MC

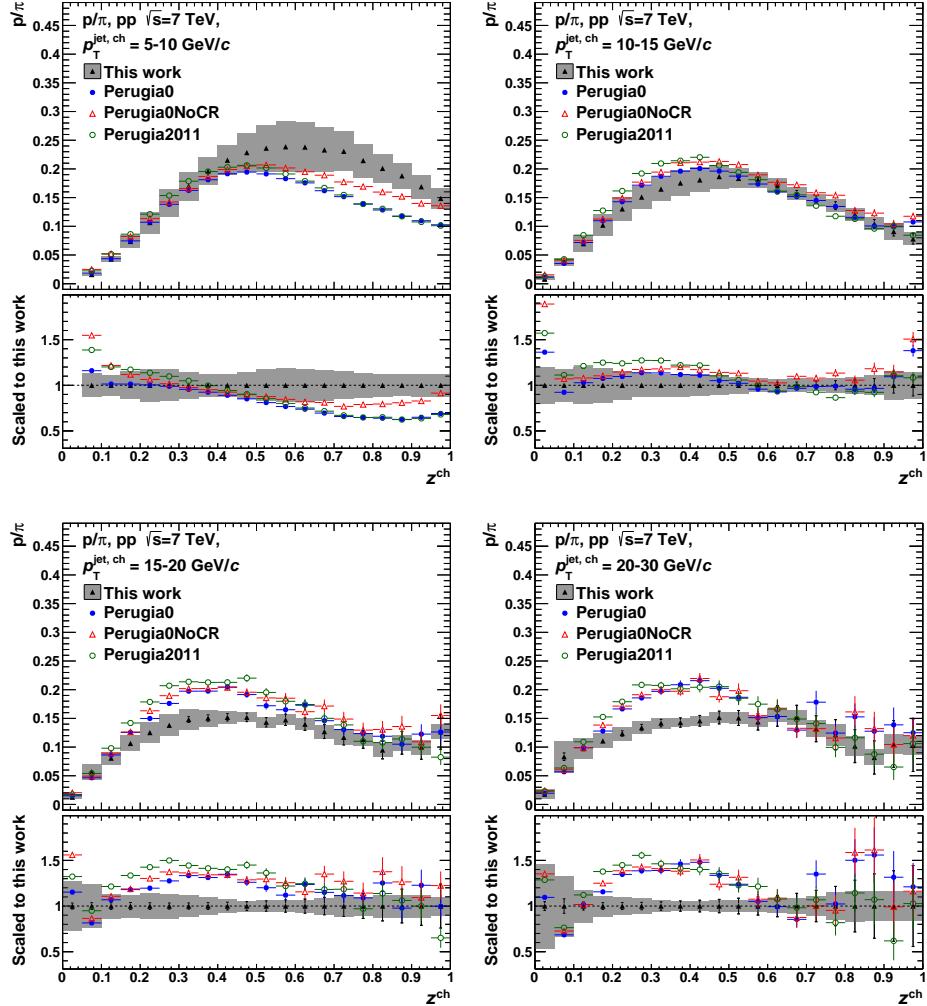


Figure K.11.: Comparison of corrected p/π ratios in jets to different MC predictions as a function of z . The $p_T^{\text{jet}, \text{ch}}$ increases from the upper left to the lower right panel.

Bibliography

- [1] A. Calaprice. *The New Quotable Einstein*. Princeton University Press, Princeton NJ, 2005.
- [2] D. J. Schwarz. The first second of the universe. *Annalen Phys.*, 12:220–270, 2003. arXiv:astro-ph/0303574, <http://dx.doi.org/10.1002/andp.200310010>.
- [3] K. Olive et al. (Particle Data Group). Review of Particle Physics. *Chin.Phys.*, C38:090001, 2014. <http://dx.doi.org/10.1088/1674-1137/38/9/090001>.
- [4] I. J. R. Aitchison, A. J. G. Hey. *Gauge Theories in Particle Physics, Volume I: From Relativistic Quantum Mechanics to QED*. Graduate student series in physics. Taylor & Francis Group, New York London, 3rd edition, 2003.
- [5] I. J. R. Aitchison, A. J. G. Hey. *Gauge Theories in Particle Physics, Volume II: Non-Abelian Gauge Theories: QCD and the Electroweak Theory*. Graduate student series in physics. Taylor & Francis Group, New York London, 3rd edition, 2004.
- [6] R. Bouchendira, P. Cladé, S. Guellati-Khélifa, et al. New Determination of the Fine Structure Constant and Test of the Quantum Electrodynamics. *Phys.Rev.Lett.*, 106:080801, 2011. arXiv:1012.3627, <http://dx.doi.org/10.1103/PhysRevLett.106.080801>.
- [7] W. J. Marciano and H. Pagels. Quantum Chromodynamics: A Review. *Phys.Rept.*, 36:137, 1978. [http://dx.doi.org/10.1016/0370-1573\(78\)90208-9](http://dx.doi.org/10.1016/0370-1573(78)90208-9).
- [8] P. W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys.Lett.*, 12:132–133, 1964. [http://dx.doi.org/10.1016/0031-9163\(64\)91136-9](http://dx.doi.org/10.1016/0031-9163(64)91136-9).
- [9] P. W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys.Rev.Lett.*, 13:508–509, 1964. <http://dx.doi.org/10.1103/PhysRevLett.13.508>.
- [10] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys.Rev.Lett.*, 13:321–323, 1964. <http://dx.doi.org/10.1103/PhysRevLett.13.321>.

Bibliography

- [11] G. Guralnik, C. Hagen, and T. Kibble. Global Conservation Laws and Massless Particles. *Phys.Rev.Lett.*, 13:585–587, 1964. <http://dx.doi.org/10.1103/PhysRevLett.13.585>.
- [12] G. Aad et al. (ATLAS Collaboration). Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys.Lett.*, B716:1–29, 2012. arXiv:1207.7214, <http://dx.doi.org/10.1016/j.physletb.2012.08.020>.
- [13] S. Chatrchyan et al. (CMS Collaboration). Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716:30–61, 2012. arXiv:1207.7235, <http://dx.doi.org/10.1016/j.physletb.2012.08.021>.
- [14] Q. Ahmad et al. (SNO Collaboration). Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by 8B solar neutrinos at the Sudbury Neutrino Observatory. *Phys.Rev.Lett.*, 87:071301, 2001. arXiv:nucl-ex/0106015, <http://dx.doi.org/10.1103/PhysRevLett.87.071301>.
- [15] S. Raby and R. Slansky. Neutrino masses: How to add them to the standard model. *Los Alamos Sci.*, 25:64–71, 1997. http://theta13.lbl.gov/neutrino_book_pdfs/04_03_Neutrino_Masses.pdf.
- [16] S. Chatrchyan et al. (CMS Collaboration). Search for fractionally charged particles in pp collisions at $\sqrt{s} = 7$ TeV. *Phys.Rev.*, D87(9):092008, 2013. arXiv:1210.2311, <http://dx.doi.org/10.1103/PhysRevD.87.092008>.
- [17] S. Scherer. Introduction to chiral perturbation theory. *Adv.Nucl.Phys.*, 27:277, 2003. arXiv:hep-ph/0210398.
- [18] G. Lepage. Lattice QCD for Novices. *pre-print*, 1998. arXiv:hep-lat/0506036.
- [19] N. Cabibbo and G. Parisi. Exponential Hadronic Spectrum and Quark Liberation. *Phys.Lett.*, B59:67–69, 1975. [http://dx.doi.org/10.1016/0370-2693\(75\)90158-6](http://dx.doi.org/10.1016/0370-2693(75)90158-6).
- [20] J. C. Collins and M. Perry. Superdense Matter: Neutrons Or Asymptotically Free Quarks? *Phys.Rev.Lett.*, 34:1353, 1975. <http://dx.doi.org/10.1103/PhysRevLett.34.1353>.
- [21] R. S. Bhalerao. Relativistic heavy-ion collisions. *pre-print*, 2014. arXiv:1404.3294.
- [22] J. C. Collins, D. E. Soper, and G. F. Sterman. Factorization of Hard Processes in QCD. *Adv.Ser.Direct.High Energy Phys.*, 5:1–91, 1988. arXiv:hep-ph/0409313.

- [23] B. Jäger, A. Schäfer, M. Stratmann, et al. Next-to-leading order QCD corrections to high- p_T pion production in longitudinally polarized pp collisions. *Phys.Rev.*, D67:054005, 2003. <http://dx.doi.org/10.1103/PhysRevD.67.054005>.
- [24] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. *Nucl.Phys.*, B126(2):298–318, 1977. [http://dx.doi.org/10.1016/0550-3213\(77\)90384-4](http://dx.doi.org/10.1016/0550-3213(77)90384-4).
- [25] Y. L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov.Phys.JETP*, 46:641–653, 1977.
- [26] V. N. Gribov and L. N. Lipatov. Deep inelastic e p scattering in perturbation theory. *Sov.J.Nucl.Phys.*, 15:438–450, 1972.
- [27] Y. I. Azimov, Y. L. Dokshitzer, V. A. Khoze, et al. Humpbacked QCD Plateau in Hadron Spectra. *Z.Phys.*, C31:213, 1986. <http://dx.doi.org/10.1007/BF01479529>.
- [28] S. Albino, B. Kniehl, and G. Kramer. Low x particle spectra in the Modified Leading Logarithm Approximation. *Eur.Phys.J.*, C38:177–185, 2004. arXiv:hep-ph/0408112, <http://dx.doi.org/10.1140/epjc/s2004-02030-1>.
- [29] J. Lees et al. (BaBar Collaboration). Production of charged pions, kaons, and protons in e^+e^- annihilations into hadrons at $\sqrt{s} = 10.54$ GeV. *Phys.Rev.*, D88:032011, 2013. arXiv:1306.2895, <http://dx.doi.org/10.1103/PhysRevD.88.032011>.
- [30] M. Leitgab et al. (Belle Collaboration). Precision Measurement of Charged Pion and Kaon Differential Cross Sections in e^+e^- Annihilation at $\sqrt{s} = 10.52$ GeV. *Phys.Rev.Lett.*, 111:062002, 2013. arXiv:1301.6183, <http://dx.doi.org/10.1103/PhysRevLett.111.062002>.
- [31] P. Abreu et al. (DELPHI Collaboration). Identified charged particles in quark and gluon jets. *Eur.Phys.J.*, C17:207–222, 2000. arXiv:hep-ex/0106063, <http://dx.doi.org/10.1007/s100520000449>.
- [32] T. Aaltonen et al. (CDF Collaboration). Production of K_S^0 , $K^{*\pm}(892)$ and $\phi^0(1020)$ in minimum bias events and K_S^0 and Λ^0 in jets in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys.Rev.*, D88:092005, 2013. <http://dx.doi.org/10.1103/PhysRevD.88.092005>.
- [33] R. Sassot, P. Zurita, and M. Stratmann. Inclusive Hadron Production in the CERN-LHC Era. *Phys.Rev.*, D82:074011, 2010. arXiv:1008.0540, <http://dx.doi.org/10.1103/PhysRevD.82.074011>.

Bibliography

- [34] B. Abelev et al. (ALICE Collaboration). Charged jet cross sections and properties in proton-proton collisions at $\sqrt{s} = 7 \text{ TeV}$. *pre-print*, 2014. arXiv:1411.4969.
- [35] R. Sassot, M. Stratmann, and P. Zurita. Fragmentations Functions in Nuclear Media. *Phys.Rev.*, D81:054001, 2010. arXiv:0912.1311, <http://dx.doi.org/10.1103/PhysRevD.81.054001>.
- [36] A. Airapetian et al. (HERMES Collaboration). Transverse momentum broadening of hadrons produced in semi-inclusive deep-inelastic scattering on nuclei. *Phys.Lett.*, B684:114–118, 2010. arXiv:0906.2478, <http://dx.doi.org/10.1016/j.physletb.2010.01.020>.
- [37] J. Adams et al. (STAR Collaboration). Pion, kaon, proton and anti-proton transverse momentum distributions from p + p and d + Au collisions at $\sqrt{s_{\text{NN}}} = 200 \text{ GeV}$. *Phys.Lett.*, B616:8–16, 2005. arXiv:nucl-ex/0309012, <http://dx.doi.org/10.1016/j.physletb.2005.04.041>.
- [38] D. G. d'Enterria. Hard scattering cross-sections at LHC in the Glauber approach: From pp to p–A and A–A collisions. *pre-print*, 2003. arXiv:nucl-ex/0302016.
- [39] M. L. Miller, K. Reygers, S. J. Sanders, et al. Glauber modeling in high energy nuclear collisions. *Ann.Rev.Nucl.Part.Sci.*, 57:205–243, 2007. arXiv:nucl-ex/0701025, <http://dx.doi.org/10.1146/annurev.nucl.57.090506.123020>.
- [40] B. Abelev et al. (ALICE Collaboration). Transverse momentum distribution and nuclear modification factor of charged particles in p–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02 \text{ TeV}$. *Phys.Rev.Lett.*, 110(8):082302, 2013. arXiv:1210.4520, <http://dx.doi.org/10.1103/PhysRevLett.110.082302>.
- [41] S. Chatrchyan et al. (CMS Collaboration). Multiplicity and transverse momentum dependence of two- and four-particle correlations in p–Pb and Pb–Pb collisions. *Phys.Lett.*, B724:213–240, 2013. arXiv:1305.0609, <http://dx.doi.org/10.1016/j.physletb.2013.06.028>.
- [42] V. Khachatryan et al. (CMS Collaboration). Observation of Long-Range Near-Side Angular Correlations in Proton-Proton Collisions at the LHC. *JHEP*, 1009:091, 2010. arXiv:1009.4122, [http://dx.doi.org/10.1007/JHEP09\(2010\)091](http://dx.doi.org/10.1007/JHEP09(2010)091).
- [43] K. Aamodt et al. (ALICE Collaboration). Two-pion Bose-Einstein correlations in central Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 2.76 \text{ TeV}$. *Phys.Lett.*, B696:328–337, 2011. arXiv:1012.4035, <http://dx.doi.org/10.1016/j.physletb.2010.12.053>.

- [44] J. Adam et al. (ALICE Collaboration). Two-pion femtoscopy in p–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02 \text{ TeV}$. *pre-print*, 2015. arXiv:1502.00559.
- [45] K. Aamodt et al. (ALICE Collaboration). Femtoscopy of pp collisions at $\sqrt{s} = 0.9$ and 7 TeV at the LHC with two-pion Bose-Einstein correlations. *Phys.Rev.*, D84:112004, 2011. arXiv:1101.3665, <http://dx.doi.org/10.1103/PhysRevD.84.112004>.
- [46] W. Li. Observation of a 'Ridge' correlation structure in high multiplicity proton-proton collisions: A brief review. *Mod.Phys.Lett.*, A27:1230018, 2012. arXiv:1206.0148, <http://dx.doi.org/10.1142/S0217732312300182>.
- [47] B. Abelev et al. (ALICE Collaboration). Multiplicity Dependence of Pion, Kaon, Proton and Lambda Production in p–Pb Collisions at $\sqrt{s_{\text{NN}}} = 5.02 \text{ TeV}$. *Phys.Lett.*, B728:25–38, 2014. arXiv:1307.6796, <http://dx.doi.org/10.1016/j.physletb.2013.11.020>.
- [48] P. Bozek. Hydrodynamic flow from RHIC to LHC. *Acta Phys.Polon.*, B43:689, 2012. arXiv:1111.4398, <http://dx.doi.org/10.5506/APhysPolB.43.689>.
- [49] V. Minissale, F. Scardina, and V. Greco. Hadrons from Coalescence plus Fragmentation in A–A collisions from RHIC to LHC energy. *pre-print*, 2015. arXiv:1502.06213.
- [50] V. Greco, C. Ko, and P. Levai. Parton coalescence at RHIC. *Phys.Rev.*, C68:034904, 2003. arXiv:nucl-th/0305024, <http://dx.doi.org/10.1103/PhysRevC.68.034904>.
- [51] R. Fries, B. Müller, C. Nonaka, et al. Hadron production in heavy ion collisions: Fragmentation and recombination from a dense parton phase. *Phys.Rev.*, C68:044902, 2003. arXiv:nucl-th/0306027, <http://dx.doi.org/10.1103/PhysRevC.68.044902>.
- [52] R. J. Fries, B. Müller, C. Nonaka, et al. Hadronization in Heavy-Ion Collisions: Recombination and Fragmentation of Partons. *Phys.Rev.Lett.*, 90:202303, 2003. <http://dx.doi.org/10.1103/PhysRevLett.90.202303>.
- [53] J. Sollfrank, F. Becattini, K. Redlich, et al. Canonical strangeness enhancement. *Nucl.Phys.*, A638:399C–402C, 1998. arXiv:nucl-th/9802046, [http://dx.doi.org/10.1016/S0375-9474\(98\)00395-9](http://dx.doi.org/10.1016/S0375-9474(98)00395-9).
- [54] J. Letessier and J. Rafelski. Strangeness chemical equilibration in a quark-gluon plasma. *Phys.Rev.*, C75:014905, 2007. <http://dx.doi.org/10.1103/PhysRevC.75.014905>.

Bibliography

- [55] O. S. Brüning, P. Collier, P. Lebrun, et al. *LHC Design Report*, volume 1. CERN, Geneva, 2004.
- [56] L. Evans and P. Bryant. LHC Machine. *JINST*, 3:S08001, 2008. <http://dx.doi.org/10.1088/1748-0221/3/08/S08001>.
- [57] J. Jowett and C. Carli. The LHC as a proton nucleus collider. *Conf.Proc.*, C060626:550–552, 2006.
- [58] C. Salgado, J. Alvarez-Muniz, F. Arleo, et al. Proton-Nucleus Collisions at the LHC: Scientific Opportunities and Requirements. *J.Phys.*, G39:015010, 2012. arXiv:1105.3919, <http://dx.doi.org/10.1088/0954-3899/39/1/015010>.
- [59] M. Lamont. Status of the LHC. *J.Phys.Conf.Ser.*, 455:012001, 2013. <http://dx.doi.org/10.1088/1742-6596/455/1/012001>.
- [60] G. Anelli et al. (TOTEM Collaboration). The TOTEM experiment at the CERN Large Hadron Collider. *JINST*, 3:S08007, 2008. <http://dx.doi.org/10.1088/1748-0221/3/08/S08007>.
- [61] O. Adriani et al. (LHCf Collaboration). The LHCf detector at the CERN Large Hadron Collider. *JINST*, 3:S08006, 2008. <http://dx.doi.org/10.1088/1748-0221/3/08/S08006>.
- [62] J. Pinfold et al. (MoEDAL Collaboration). Technical Design Report of the MoEDAL Experiment. Technical Report CERN-LHCC-2009-006. MoEDAL-TDR-001, CERN, Geneva, 2009.
- [63] J.-L. Caron. Layout of the LEP tunnel including future LHC infrastructures, 1997. AC Collection. Legacy of AC. Pictures from 1992 to 2002.
- [64] G. Aad et al. (ATLAS Collaboration). The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008. <http://dx.doi.org/10.1088/1748-0221/3/08/S08003>.
- [65] S. Chatrchyan et al. (CMS Collaboration). The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008. <http://dx.doi.org/10.1088/1748-0221/3/08/S08004>.
- [66] K. Aamodt et al. (ALICE Collaboration). The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008. <http://dx.doi.org/10.1088/1748-0221/3/08/S08002>.
- [67] A. Augusto Alves Jr et al. (LHCb Collaboration). The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008. <http://dx.doi.org/10.1088/1748-0221/3/08/S08005>.

- [68] F. Carminati et al. (ALICE Collaboration). ALICE: Physics performance report, volume I. *J.Phys.*, G30:1517–1763, 2004. <http://dx.doi.org/10.1088/0954-3899/30/11/001>.
- [69] Y. Schutz. Heavy-Ion Physics at LHC. Presentation at Quark Matter in Oakland, January 2004.
- [70] B. B. Abelev et al. (ALICE Collaboration). Performance of the ALICE Experiment at the CERN LHC. *Int.J.Mod.Phys.*, A29:1430044, 2014. arXiv:1402.4476, <http://dx.doi.org/10.1142/S0217751X14300440>.
- [71] G. Dellacasa et al. (ALICE Collaboration). *ALICE Inner Tracking System (ITS): Technical Design Report*. Technical Design Report ALICE. CERN, Geneva, 1999.
- [72] G. Dellacasa et al. (ALICE Collaboration). *ALICE time projection chamber: Technical Design Report*. Technical Design Report ALICE. CERN, Geneva, 2000.
- [73] J. Alme et al. (ALICE Collaboration). The ALICE TPC, a large 3-dimensional tracking device with fast readout for ultra-high multiplicity events. *Nucl.Instrum.Meth.*, A622(1):316–367, 2010. <http://dx.doi.org/10.1016/j.nima.2010.04.042>.
- [74] D. Vranic. Private communication.
- [75] P. Cortese (ALICE Collaboration). *ALICE transition-radiation detector: Technical Design Report*. Technical Design Report ALICE. CERN, Geneva, 2001.
- [76] G. Dellacasa et al. (ALICE Collaboration). *ALICE Time-Of-Flight system (TOF): Technical Design Report*. Technical Design Report ALICE. CERN, Geneva, 2000.
- [77] F. Piuz, W. Klempt, L. Leistam, et al. (ALICE Collaboration). *ALICE high-momentum particle identification: Technical Design Report*. Technical Design Report ALICE. CERN, Geneva, 1998.
- [78] V. I. Man’ko, W. Klempt, L. Leistam, et al. (ALICE Collaboration). *ALICE Photon Spectrometer (PHOS): Technical Design Report*. Technical Design Report ALICE. CERN, Geneva, 1999.
- [79] P. Cortese et al. (ALICE Collaboration). ALICE Electromagnetic Calorimeter Technical Design Report. Technical Report CERN-LHCC-2008-014. ALICE-TDR-14, CERN, Geneva, 2008.

Bibliography

- [80] G. Dellacasa et al. (ALICE Collaboration). *ALICE dimuon forward spectrometer: Technical Design Report*. Technical Design Report ALICE. CERN, Geneva, 1999.
- [81] P. Cortese et al. (ALICE Collaboration). *ALICE forward detectors: FMD, TO and VO: Technical Design Report*. Technical Design Report ALICE. CERN, Geneva, 2004.
- [82] A. Fernandez, E. Gamez, G. Herrera, et al. ACORDE a Cosmic Ray Detector for ALICE. *Nucl.Instrum.Meth.*, A572:102–103, 2007. arXiv:physics/0606051, <http://dx.doi.org/10.1016/j.nima.2006.10.336>.
- [83] AliRoot Documentation. <http://aliweb.cern.ch/Offline/AliRoot/Manual.html>.
- [84] R. Brun and F. Rademakers. ROOT – An object oriented data analysis framework. *Nucl.Instrum.Meth.*, A389(1-2):81–86, 1997. [http://dx.doi.org/10.1016/S0168-9002\(97\)00048-X](http://dx.doi.org/10.1016/S0168-9002(97)00048-X).
- [85] I. Antcheva, M. Ballintijn, B. Bellenot, et al. ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization. *Comput.Phys.Commun.*, 180:2499–2512, 2009. <http://dx.doi.org/10.1016/j.cpc.2009.08.005>.
- [86] ROOT Documentation. <https://root.cern.ch/drupal/content/documentation>.
- [87] T. Sjostrand, S. Mrenna, and P. Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 0605:026, 2006. arXiv:hep-ph/0603175, <http://dx.doi.org/10.1088/1126-6708/2006/05/026>.
- [88] M. Gyulassy and X.-N. Wang. HIJING 1.0: A Monte Carlo program for parton and particle production in high-energy hadronic and nuclear collisions. *Comput.Phys.Commun.*, 83:307, 1994. arXiv:nucl-th/9502021, [http://dx.doi.org/10.1016/0010-4655\(94\)90057-4](http://dx.doi.org/10.1016/0010-4655(94)90057-4).
- [89] R. Brun, F. Carminati, and S. Giani. GEANT – Detector Description and Simulation Tool. *CERN Program Library Long Writeup W5013*, 1994.
- [90] S. Agostinelli et al. (GEANT4 Collaboration). GEANT4: A Simulation toolkit. *Nucl.Instrum.Meth.*, A506:250–303, 2003. [http://dx.doi.org/10.1016/S0168-9002\(03\)01368-8](http://dx.doi.org/10.1016/S0168-9002(03)01368-8).
- [91] A. Ferrari, P. R. Sala, A. Fassò, et al. *FLUKA: A multi-particle transport code (program version 2005)*. CERN, Geneva, 2005.

- [92] R. Frühwirth. Application of Kalman filtering to track and vertex fitting. *Nucl.Instrum.Meth.*, A262:444–450, 1987. [http://dx.doi.org/10.1016/0168-9002\(87\)90887-4](http://dx.doi.org/10.1016/0168-9002(87)90887-4).
- [93] G. C. Blazey, J. R. Dittmann, S. D. Ellis, et al. Run II jet physics. *pre-print*, 2000. arXiv:hep-ex/0005012.
- [94] G. P. Salam and G. Soyez. A Practical Seedless Infrared-Safe Cone jet algorithm. *JHEP*, 0705:086, 2007. arXiv:0704.0292, <http://dx.doi.org/10.1088/1126-6708/2007/05/086>.
- [95] S. D. Ellis and D. E. Soper. Successive combination jet algorithm for hadron collisions. *Phys.Rev.*, D48:3160–3166, 1993. <http://dx.doi.org/10.1103/PhysRevD.48.3160>.
- [96] S. Catani, Y. Dokshitzer, M. Seymour, et al. Longitudinally-invariant k_\perp -clustering algorithms for hadron-hadron collisions. *Nucl.Phys.*, B406(1–2):187–224, 1993. [http://dx.doi.org/10.1016/0550-3213\(93\)90166-M](http://dx.doi.org/10.1016/0550-3213(93)90166-M).
- [97] Y. L. Dokshitzer, G. Leder, S. Moretti, et al. Better jet clustering algorithms. *JHEP*, 9708:001, 1997. arXiv:hep-ph/9707323, <http://dx.doi.org/10.1088/1126-6708/1997/08/001>.
- [98] M. Cacciari, G. P. Salam, and G. Soyez. The Anti- k_T jet clustering algorithm. *JHEP*, 0804:063, 2008. arXiv:0802.1189, <http://dx.doi.org/10.1088/1126-6708/2008/04/063>.
- [99] G. Corcella, I. Knowles, G. Marchesini, et al. HERWIG 6.5 release note. *pre-print*, 2002. arXiv:hep-ph/0210213.
- [100] J. Allen et al. (ALICE Collaboration). ALICE DCal: An Addendum to the EMCAL Technical Design Report Di-Jet and Hadron-Jet correlation measurements in ALICE. Technical Report CERN-LHCC-2010-011. ALICE-TDR-14-add-1, CERN, Geneva, 2010.
- [101] H. Bethe. Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie. *Annalen der Physik*, 397(3):325–400, 1930. <http://dx.doi.org/10.1002/andp.19303970303>.
- [102] E. Fermi. The Ionization Loss of Energy in Gases and in Condensed Materials. *Phys.Rev.*, 57:485–493, 1940. <http://dx.doi.org/10.1103/PhysRev.57.485>.
- [103] U. Fano. Penetration of protons, alpha particles, and mesons. *Ann.Rev.Nucl.Part.Sci.*, 13:1–66, 1963. <http://dx.doi.org/10.1146/annurev.ns.13.120163.000245>.

Bibliography

- [104] W. Blum, W. Riegler, and L. Rolandi. *Particle detection with drift chambers; 2nd ed.*. Springer, Berlin, 2008.
- [105] A. Kalweit. *Production of light flavor hadrons and anti-nuclei at the LHC*. Doctoral thesis, Darmstadt University of Technology, 2012.
- [106] P. Christiansen (for the ALICE Collaboration). High p_T identified particle production in ALICE. *Nucl.Phys.*, A910-911:20–26, 2013. arXiv:1208.5368, <http://dx.doi.org/10.1016/j.nuclphysa.2012.12.012>.
- [107] J. Podolanski and R. Armenteros. III. Analysis of V-events. *Phil.Mag.*, 45(360):13–30, 1954. <http://dx.doi.org/10.1080/14786440108520416>.
- [108] D. Décamp et al. (ALEPH Collaboration). ALEPH: A detector for electron-positron annihilations at LEP. *Nucl.Instrum.Meth.*, A294:121–178, 1990. Erratum-ibid. A303:393, 1991, [http://dx.doi.org/10.1016/0168-9002\(90\)91831-U](http://dx.doi.org/10.1016/0168-9002(90)91831-U).
- [109] S. Baker and R. D. Cousins. Clarification of the use of Chi-square and likelihood functions in fits to histograms. *Nucl.Instrum.Meth.*, 221(2):437–442, 1984. [http://dx.doi.org/10.1016/0167-5087\(84\)90016-4](http://dx.doi.org/10.1016/0167-5087(84)90016-4).
- [110] F. James. MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1. *CERN Program Library Long Writeup D506*, 1994.
- [111] W. T. Eadie, D. Drijard, F. E. James, et al. *Statistical Methods in Experimental Physics*. North Holland, 1971.
- [112] X.-G. Lu. *Exploring the performance limits of the ALICE Time Projection Chamber and Transition Radiation Detector for measuring identified hadron production at the LHC*. Doctoral thesis, University of Heidelberg, 2013.
- [113] W. H. Press, S. A. Teukolsky, W. T. Vetterling, et al. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, New York, NY, USA, 2nd edition, 1992.
- [114] P. Z. Skands. Tuning Monte Carlo Generators: The Perugia Tunes. *Phys.Rev.*, D82:074018, 2010. arXiv:1005.3457, <http://dx.doi.org/10.1103/PhysRevD.82.074018>.
- [115] V. Khachatryan et al. (CMS Collaboration). Strange Particle Production in pp Collisions at $\sqrt{s} = 0.9$ and 7 TeV. *JHEP*, 1105:064, 2011. arXiv:1102.4282, [http://dx.doi.org/10.1007/JHEP05\(2011\)064](http://dx.doi.org/10.1007/JHEP05(2011)064).
- [116] A. Karneyeu, L. Mijovic, S. Prestel, et al. MCPLOTS: a particle physics resource based on volunteer computing. *Eur.Phys.J.*, C74:2714, 2014. arXiv:1306.3436, <http://dx.doi.org/10.1140/epjc/s10052-014-2714-9>.

- [117] E. Abbas et al. (ALICE Collaboration). Mid-rapidity anti-baryon to baryon ratios in pp collisions at $\sqrt{s} = 0.9, 2.76$ and 7 TeV measured by ALICE. *Eur.Phys.J.*, C73:2496, 2013. arXiv:1305.1562, <http://dx.doi.org/10.1140/epjc/s10052-013-2496-5>.
- [118] K. Aamodt et al. (ALICE). Charged-particle multiplicity measurement in proton-proton collisions at $\sqrt{s} = 0.9$ and 2.36 TeV with ALICE at LHC. *Eur.Phys.J.*, C68:89–108, 2010. arXiv:1004.3034, <http://dx.doi.org/10.1140/epjc/s10052-010-1339-x>.
- [119] B. Abelev et al. (ALICE Collaboration). Measurement of the inclusive differential jet cross section in pp collisions at $\sqrt{s} = 2.76\text{ TeV}$. *Phys.Lett.*, B722:262–272, 2013. arXiv:1301.3475, <http://dx.doi.org/10.1016/j.physletb.2013.04.026>.
- [120] X.-G. Lu. Private communication.
- [121] X.-G. Lu (for the ALICE Collaboration). Measurement of hadron composition in charged jets from pp collisions with the ALICE experiment. *Nucl.Phys.*, A, 2014. arXiv:1407.8385, <http://dx.doi.org/10.1016/j.nuclphysa.2014.08.003>.
- [122] L. Lyons. On estimating systematic errors from repeated measurements. *J.Phys.A:Math.Gen.*, 25(7):1967, 1992. <http://dx.doi.org/10.1088/0305-4470/25/7/035>.
- [123] D. D. Chinellato. Private communication.
- [124] O. Busch. Private communication.
- [125] J. Adam et al. (ALICE Collaboration). Measurement of pion, kaon and proton production in proton-proton collisions at $\sqrt{s} = 7\text{ TeV}$. *pre-print*, 2015. arXiv:1504.00024.
- [126] A. Ortiz Velasquez (for the ALICE Collaboration). Production of pions, kaons and protons at high p_T in $\sqrt{s_{NN}} = 2.76\text{ TeV}$ Pb-Pb collisions. *Nucl.Phys.*, A904-905:763c–766c, 2013. arXiv:1210.6995, <http://dx.doi.org/10.1016/j.nuclphysa.2013.02.129>.
- [127] B. Abelev et al. (ALICE Collaboration). Measurement of inelastic, single- and double-diffraction cross sections in proton–proton collisions at the LHC with ALICE. *Eur.Phys.J.*, C73:2456, 2013. arXiv:1208.4968, <http://dx.doi.org/10.1140/epjc/s10052-013-2456-0>.
- [128] F.-H. Liu, Y.-H. Chen, Y.-Q. Gao, et al. On Current Conversion between Particle Rapidity and Pseudorapidity Distributions in High Energy Collisions. *Adv.High Energy Phys.*, 2013:710534, 2013. <http://dx.doi.org/10.1155/2013/710534>.

Bibliography

- [129] N. Jacazio. Private communication.
- [130] V. Vislavicius. Private communication.
- [131] M. Chojnacki. Private communication.
- [132] K. Aamodt et al. (ALICE Collaboration). Midrapidity antiproton-to-proton ratio in pp collisions at $\sqrt{s} = 0.9$ and 7 TeV measured by the ALICE experiment. *Phys.Rev.Lett.*, 105:072002, 2010. arXiv:1006.5432, <http://dx.doi.org/10.1103/PhysRevLett.105.072002>.
- [133] M. Cacciari, G. P. Salam, and G. Soyez. FastJet User Manual. *Eur.Phys.J.*, C72:1896, 2012. arXiv:1111.6097, <http://dx.doi.org/10.1140/epjc/s10052-012-1896-2>.
- [134] T. Kaufmann, A. Mukherjee, and W. Vogelsang. Hadron Fragmentation Inside Jets in Hadronic Collisions. *pre-print*, 2015. arXiv:1506.01415.
- [135] T. Kaufmann and W. Vogelsang. Private communication.
- [136] P. Bartalini, E. Berger, B. Blok, et al. Multi-Parton Interactions at the LHC. *pre-print*, 2011. arXiv:1111.0469.
- [137] C. Andrei (for the ALICE Collaboration). Light flavor hadron spectra at low p_T and search for collective phenomena in high multiplicity pp, p-Pb and Pb-Pb collisions measured with the ALICE experiment. *Nucl.Phys.*, A, 2014. arXiv:1408.0093, <http://dx.doi.org/10.1016/j.nuclphysa.2014.08.002>.
- [138] P. Z. Skands and D. Wicke. Non-perturbative QCD effects and the top mass at the Tevatron. *Eur.Phys.J.*, C52:133–140, 2007. arXiv:hep-ph/0703081, <http://dx.doi.org/10.1140/epjc/s10052-007-0352-1>.
- [139] A. Ortiz Velasquez, P. Christiansen, E. Cuautle Flores, et al. Color Reconnection and Flowlike Patterns in pp Collisions. *Phys.Rev.Lett.*, 111(4):042001, 2013. arXiv:1303.6326, <http://dx.doi.org/10.1103/PhysRevLett.111.042001>.
- [140] F. Gelis. Color Glass Condensate and Glasma. *Int.J.Mod.Phys.*, A28:1330001, 2013. arXiv:1211.3327, <http://dx.doi.org/10.1142/S0217751X13300019>.
- [141] K. Dusling and R. Venugopalan. Comparison of the color glass condensate to dihadron correlations in proton-proton and proton-nucleus collisions. *Phys.Rev.*, D87(9):094034, 2013. arXiv:1302.7018, <http://dx.doi.org/10.1103/PhysRevD.87.094034>.
- [142] M. Ivanov. Private communication.

Acknowledgements

First of all, I want to thank Prof. Dr. Hans Rudolf Schmidt for the outstanding opportunity to work in this comprehensive and interesting field of research, accompanied by very exciting times at CERN. Working in his group also gave me the opportunity to gain insights into the CBM experiment.

Furthermore, I want to thank Prof. Dr. Werner Vogelsang for supervising my thesis in view of the theoretical aspects. We had many exciting discussions about the theoretical implications of my results.

I am deeply indebted to Dr. Jens Wiechula, who introduced me to the (TPC) PID in ALICE and the PID framework. I learnt a lot from him during our collaboration in the last years. We spent quite some time together to debug and improve the PID framework.

Many special thanks go to Prof. Dr. Oliver Busch, with whom I collaborated on the jet analysis. We had countless technical discussions about the MC correction procedure and estimation of the systematic uncertainties. I benefited a lot from his experience in these procedures from an earlier analysis.

For many discussions around the mathematics behind the fitting procedure and about the jet analysis, I want to express my gratitude to Xianguo Lu.

I owe many thanks to Prof. Dr. Peter Christiansen for the introduction to particle identification with the TPC in the relativistic rise and many lively debates about the interpretation of my results.

I greatly appreciate the (sometimes very technical) discussions about the multiplicity dependent analysis with Prof. Dr. David Dobrigkeit Chinellato, Dr. Alexander Kalweit, Marek Chojnacki, Nicolo Jacazio and Vytautas Vislavicius.

I want to express my gratitude to Dr. Jochen Klein for many discussions about the jet analysis and to Tom Kaufmann, who carried out next-to-leading order calculations, which we started to compare to my results — it was mesmerising to experience the duality of experiment and theory.

For proof-reading (parts of) my thesis, I'm indebted to Dr. Jens Wiechula, Prof. Dr. Oliver Busch, Prof. Dr. Peter Christiansen and Tom Kaufmann.

Further special thanks go to Dr. Dariusz Miskowiec, who introduced me to femtoscopy, and to Dr. Marian Ivanov, with whom I discussed a lot the technicalities of the TPC signal.

Last but not least, I would like to thank the GSI IT department for providing an excellent environment for data analysis with plenty of computer power.

Many more people deserve to be thanked and I sincerely apologise that it is impossible to list all of them here.